Generating SQL for SQLite using Ollama, ChromaDB

This notebook runs through the process of using the `vanna` Python package to generate SQL using AI (RAG + LLMs) including connecting to a database and training. If you're not ready to train on your own database, you can still try it using a sample SQLite database.

Which LLM do you want to use?

- OpenAI via Vanna.AI (Recommended)

  Use Vanna.AI for free to generate your queries
- OpenAI

  Use OpenAI with your own API key
- Azure OpenAI

  If you have OpenAI models deployed on Azure
- [Selected] Ollama

  Use Ollama locally for free. Requires additional setup.
- Mistral via Mistral API

  If you have a Mistral API key
- Other LLM

  If you have a different LLM model

Where do you want to store the 'training' data?

- Vanna Hosted Vector DB (Recommended)

  Use Vanna.AIs hosted vector database (pgvector) for free. This is usable across machines with no additional setup.
- [Selected] ChromaDB

  Use ChromaDBs open-source vector database for free locally. No additional setup is necessary -- all database files will be created and stored locally.
- Marqo

  Use Marqo locally for free. Requires additional setup. Or use their hosted option.
- Other VectorDB

  Use any other vector database. Requires additional setup.

Setup

!pwd!pip install vanna!pip install 'vanna[chromadb]'!pip install ollama!pip show vanna # 0.5.5, 0.2.1!pip show ollama # 0.2.0

In [1]:
```
import warnings
import re
```

```
warnings.filterwarnings('ignore', category=DeprecationWarning, message='^Num
# warnings.filterwarnings('ignore', category=DeprecationWarning, message=re.

import os

import re
from time import time

from vanna.ollama import Ollama
from vanna.chromadb.chromadb_vector import ChromaDB_VectorStore
```

In [2]:
```
class MyVanna(ChromaDB_VectorStore, Ollama):
    def __init__(self, config=None):
        ChromaDB_VectorStore.__init__(self, config=config)
        Ollama.__init__(self, config=config)
```

In [3]:
```
file_db = "~/Downloads/chinook.sqlite"
model_name = "phi3:14b" # 'llama3'

clean_and_train = True  # False
```

In [4]:
```
config = {
    'model': model_name,   # 'mistral' # "starcoder2"
}
vn = MyVanna(config=config)
```

In [5]:
```
hostname = os.uname().nodename
print("Hostname:", hostname)
```

```
Hostname: ducklover1
```

In [6]:
```
file_db = os.path.abspath(os.path.expanduser(file_db))
vn.connect_to_sqlite(file_db)
```

In [7]:
```
vn.run_sql_is_set
```

Out[7]:
```
True
```

In [8]:
```
def remove_collections(collection_name=None, ACCEPTED_TYPES = ["sql", "ddl",
    if not collection_name:
        collections = ACCEPTED_TYPES
    elif isinstance(collection_name, str):
        collections = [collection_name]
    elif isinstance(collection_name, list):
        collections = collection_name
    else:
        print(f"\t{collection_name} is unknown: Skipped")
        return

    for c in collections:
        if not c in ACCEPTED_TYPES:
            print(f"\t{c} is unknown: Skipped")
            continue
        # print(f"vn.remove_collection('{c}')")
        vn.remove_collection(c)
```

In [9]:

```
def strip_brackets(ddl):
    """
    This function removes square brackets from table and column names in a D

    Args:
        ddl (str): The DDL script containing square brackets.

    Returns:
        str: The DDL script with square brackets removed.
    """
    # Use regular expressions to match and replace square brackets
    pattern = r"\[([^\]]+)]"  # Match any character except ] within square b
    return re.sub(pattern, r"\1", ddl)
```

In [10]:

```
if clean_and_train:
    remove_collections()
```

Training

SQLite sample database

You only need to train once. Do not train again unless you want to add more training data.

In [11]:

```
df_ddl = vn.run_sql("SELECT type, sql FROM sqlite_master WHERE sql is not nu
```

In [12]:

```
df_ddl
```

Out[12]:

	type	sql
0	table	CREATE TABLE "albums"\r\n(\r\n [AlbumId] IN...
1	table	CREATE TABLE sqlite_sequence(name,seq)
2	table	CREATE TABLE "artists"\r\n(\r\n [ArtistId] ...
3	table	CREATE TABLE "customers"\r\n(\r\n [Customer...
4	table	CREATE TABLE "employees"\r\n(\r\n [Employee...
5	table	CREATE TABLE "genres"\r\n(\r\n [GenreId] IN...
6	table	CREATE TABLE "invoices"\r\n(\r\n [InvoiceId...
7	table	CREATE TABLE "invoice_items"\r\n(\r\n [Invo...
8	table	CREATE TABLE "media_types"\r\n(\r\n [MediaT...
9	table	CREATE TABLE "playlists"\r\n(\r\n [Playlist...
10	table	CREATE TABLE "playlist_track"\r\n(\r\n [Pla...
11	table	CREATE TABLE "tracks"\r\n(\r\n [TrackId] IN...
12	index	CREATE INDEX [IFK_AlbumArtistId] ON "albums" (...
13	index	CREATE INDEX [IFK_CustomerSupportRepId] ON "cu...
14	index	CREATE INDEX [IFK_EmployeeReportsTo] ON "emplo...
15	index	CREATE INDEX [IFK_InvoiceCustomerId] ON "invoi...
16	index	CREATE INDEX [IFK_InvoiceLineInvoiceId] ON "in...
17	index	CREATE INDEX [IFK_InvoiceLineTrackId] ON "invo...
18	index	CREATE INDEX [IFK_PlaylistTrackTrackId] ON "pl...
19	index	CREATE INDEX [IFK_TrackAlbumId] ON "tracks" ([...
20	index	CREATE INDEX [IFK_TrackGenreId] ON "tracks" ([...
21	index	CREATE INDEX [IFK_TrackMediaTypeId] ON "tracks...
22	table	CREATE TABLE sqlite_stat1(tbl,idx,stat)

In [13]:

```
if clean_and_train:
    for ddl in df_ddl['sql'].to_list():
        ddl = strip_brackets(ddl)
        vn.train(ddl=ddl)

    # Sometimes you may want to add documentation about your business termir
    vn.train(documentation="In the chinook database invoice means order")
```

```
Adding ddl: CREATE TABLE "albums"
(
    AlbumId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,
    Title NVARCHAR(160)  NOT NULL,
    ArtistId INTEGER  NOT NULL,
    FOREIGN KEY (ArtistId) REFERENCES "artists" (ArtistId)
                ON DELETE NO ACTION ON UPDATE NO ACTION
)
Adding ddl: CREATE TABLE sqlite_sequence(name,seq)
Adding ddl: CREATE TABLE "artists"
(
    ArtistId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,
    Name NVARCHAR(120)
)
Adding ddl: CREATE TABLE "customers"
(
    CustomerId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,
    FirstName NVARCHAR(40)  NOT NULL,
    LastName NVARCHAR(20)  NOT NULL,
    Company NVARCHAR(80),
    Address NVARCHAR(70),
    City NVARCHAR(40),
    State NVARCHAR(40),
    Country NVARCHAR(40),
    PostalCode NVARCHAR(10),
    Phone NVARCHAR(24),
    Fax NVARCHAR(24),
    Email NVARCHAR(60)  NOT NULL,
    SupportRepId INTEGER,
    FOREIGN KEY (SupportRepId) REFERENCES "employees" (EmployeeId)
                ON DELETE NO ACTION ON UPDATE NO ACTION
)
Adding ddl: CREATE TABLE "employees"
(
    EmployeeId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,
    LastName NVARCHAR(20)  NOT NULL,
    FirstName NVARCHAR(20)  NOT NULL,
    Title NVARCHAR(30),
    ReportsTo INTEGER,
    BirthDate DATETIME,
    HireDate DATETIME,
    Address NVARCHAR(70),
    City NVARCHAR(40),
    State NVARCHAR(40),
    Country NVARCHAR(40),
    PostalCode NVARCHAR(10),
    Phone NVARCHAR(24),
    Fax NVARCHAR(24),
    Email NVARCHAR(60),
    FOREIGN KEY (ReportsTo) REFERENCES "employees" (EmployeeId)
                ON DELETE NO ACTION ON UPDATE NO ACTION
)
Adding ddl: CREATE TABLE "genres"
(
    GenreId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,
    Name NVARCHAR(120)
```

```
)
Adding ddl: CREATE TABLE "invoices"
(
    InvoiceId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,
    CustomerId INTEGER  NOT NULL,
    InvoiceDate DATETIME  NOT NULL,
    BillingAddress NVARCHAR(70),
    BillingCity NVARCHAR(40),
    BillingState NVARCHAR(40),
    BillingCountry NVARCHAR(40),
    BillingPostalCode NVARCHAR(10),
    Total NUMERIC(10,2)  NOT NULL,
    FOREIGN KEY (CustomerId) REFERENCES "customers" (CustomerId)
                ON DELETE NO ACTION ON UPDATE NO ACTION
)
Adding ddl: CREATE TABLE "invoice_items"
(
    InvoiceLineId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,
    InvoiceId INTEGER  NOT NULL,
    TrackId INTEGER  NOT NULL,
    UnitPrice NUMERIC(10,2)  NOT NULL,
    Quantity INTEGER  NOT NULL,
    FOREIGN KEY (InvoiceId) REFERENCES "invoices" (InvoiceId)
                ON DELETE NO ACTION ON UPDATE NO ACTION,
    FOREIGN KEY (TrackId) REFERENCES "tracks" (TrackId)
                ON DELETE NO ACTION ON UPDATE NO ACTION
)
Adding ddl: CREATE TABLE "media_types"
(
    MediaTypeId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,
    Name NVARCHAR(120)
)
Adding ddl: CREATE TABLE "playlists"
(
    PlaylistId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,
    Name NVARCHAR(120)
)
Adding ddl: CREATE TABLE "playlist_track"
(
    PlaylistId INTEGER  NOT NULL,
    TrackId INTEGER  NOT NULL,
    CONSTRAINT PK_PlaylistTrack PRIMARY KEY  (PlaylistId, TrackId),
    FOREIGN KEY (PlaylistId) REFERENCES "playlists" (PlaylistId)
                ON DELETE NO ACTION ON UPDATE NO ACTION,
    FOREIGN KEY (TrackId) REFERENCES "tracks" (TrackId)
                ON DELETE NO ACTION ON UPDATE NO ACTION
)
Adding ddl: CREATE TABLE "tracks"
(
    TrackId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,
    Name NVARCHAR(200)  NOT NULL,
    AlbumId INTEGER,
    MediaTypeId INTEGER  NOT NULL,
    GenreId INTEGER,
    Composer NVARCHAR(220),
    Milliseconds INTEGER  NOT NULL,
    Bytes INTEGER,
    UnitPrice NUMERIC(10,2)  NOT NULL,
    FOREIGN KEY (AlbumId) REFERENCES "albums" (AlbumId)
                ON DELETE NO ACTION ON UPDATE NO ACTION,
    FOREIGN KEY (GenreId) REFERENCES "genres" (GenreId)
                ON DELETE NO ACTION ON UPDATE NO ACTION,
    FOREIGN KEY (MediaTypeId) REFERENCES "media_types" (MediaTypeId)
                ON DELETE NO ACTION ON UPDATE NO ACTION
)
Adding ddl: CREATE INDEX IFK_AlbumArtistId ON "albums" (ArtistId)
Adding ddl: CREATE INDEX IFK_CustomerSupportRepId ON "customers" (SupportRep
Id)
Adding ddl: CREATE INDEX IFK_EmployeeReportsTo ON "employees" (ReportsTo)
Adding ddl: CREATE INDEX IFK_InvoiceCustomerId ON "invoices" (CustomerId)
Adding ddl: CREATE INDEX IFK_InvoiceLineInvoiceId ON "invoice_items" (Invoic
eId)
Adding ddl: CREATE INDEX IFK_InvoiceLineTrackId ON "invoice_items" (TrackId)
Adding ddl: CREATE INDEX IFK_PlaylistTrackTrackId ON "playlist_track" (Track
Id)
Adding ddl: CREATE INDEX IFK_TrackAlbumId ON "tracks" (AlbumId)
Adding ddl: CREATE INDEX IFK_TrackGenreId ON "tracks" (GenreId)
Adding ddl: CREATE INDEX IFK_TrackMediaTypeId ON "tracks" (MediaTypeId)
Adding ddl: CREATE TABLE sqlite_stat1(tbl,idx,stat)
Adding documentation....
```

In [14]:

```
# show training data
training_data = vn.get_training_data()
training_data
```

Out[14]:

	id	question	content	training_data_type
0	039f9d54-59f7-5f29-8c04-14dbc3e95671-ddl	None	CREATE TABLE "artists"\r\n(\r\n ArtistId IN...	ddl
1	0db84e3d-ef41-563c-803e-21c1b985dc19-ddl	None	CREATE TABLE "invoices"\r\n(\r\n InvoiceId ...	ddl
2	10cba811-ddba-5042-9e90-d764dfcd1629-ddl	None	CREATE INDEX IFK_InvoiceCustomerId ON "invoice...	ddl
3	2c711317-b93d-5f60-a728-cb1c6fcbc040-ddl	None	CREATE INDEX IFK_CustomerSupportRepId ON "cust...	ddl
4	37319c81-65f7-50ee-956b-795de244bee5-ddl	None	CREATE TABLE sqlite_stat1(tbl,idx,stat)	ddl
5	40bd77cd-e1de-5872-8693-624117ff413c-ddl	None	CREATE INDEX IFK_InvoiceLineInvoiceId ON "invo...	ddl
6	41130543-7164-562a-90a7-0fd0a409c154-ddl	None	CREATE TABLE "albums"\r\n(\r\n AlbumId INTE...	ddl
7	458debc8-8082-5450-a17a-66028bd55ace-ddl	None	CREATE TABLE "playlists"\r\n(\r\n PlaylistI...	ddl
8	4815f3fd-925b-53ce-9dfa-0e4285d5abd3-ddl	None	CREATE TABLE "invoice_items"\r\n(\r\n Invoi...	ddl
9	48d484e9-984c-58ff-b391-75521c69d486-ddl	None	CREATE INDEX IFK_PlaylistTrackTrackId ON "play...	ddl
10	551e1120-a6ee-554f-8b8a-ccf4f22d3636-ddl	None	CREATE INDEX IFK_AlbumArtistId ON "albums" (Ar...	ddl
11	5ff4911e-45c1-5a59-9566-243a9b6a3320-ddl	None	CREATE TABLE "employees"\r\n(\r\n EmployeeI...	ddl
12	65df0648-bf05-5f75-9365-c21f54b2302d-ddl	None	CREATE TABLE "media_types"\r\n(\r\n MediaTy...	ddl
13	6b585176-e66d-5b23-8d86-ca8a80e3af3d-ddl	None	CREATE INDEX IFK_EmployeeReportsTo ON "employe...	ddl
14	868758b8-e018-55e7-8cc3-75c0e6d211c8-ddl	None	CREATE INDEX IFK_TrackAlbumId ON "tracks" (Alb...	ddl
15	9ea4613d-c1be-5a77-ada9-c54ee3f0cab7-ddl	None	CREATE INDEX IFK_TrackMediaTypeId ON "tracks" ...	ddl
16	a9c9a852-608d-5ef2-aede-26ba098d83d1-	None	CREATE INDEX IFK_TrackGenreId ON "tracks" (Gen...	ddl

	id	question	content	training_data_type
	ddl			
17	b42cc9e1-9219-5a42-9a06-de906f76239e-ddl	None	CREATE TABLE "tracks"\r\n(\r\n TrackId INTE...	ddl
18	c387b9d2-5ff4-5a07-8364-f5dab45bb2a9-ddl	None	CREATE TABLE "genres"\r\n(\r\n GenreId INTE...	ddl
19	d654f328-dc36-549e-84c3-06ee0db7e0f7-ddl	None	CREATE TABLE "playlist_track"\r\n(\r\n Play...	ddl
20	d93f0d68-023d-5afb-8121-ba346699d318-ddl	None	CREATE TABLE "customers"\r\n(\r\n CustomerI...	ddl
21	e5879308-329e-543f-a693-0c14e2f9972e-ddl	None	CREATE INDEX IFK_InvoiceLineTrackId ON "invoic...	ddl
22	ea84418b-1a28-59b4-a1f4-2fb674208adc-ddl	None	CREATE TABLE sqlite_sequence(name,seq)	ddl
0	2b4dda0a-a6ac-5e34-8f76-e41c0734d55e-doc	None	In the chinook database invoice means order	documentation

Asking the AI

Whenever you ask a new question, it will find the 10 most relevant pieces of training data and use it as part of the LLM prompt to generate the SQL.

In [15]:

```
ts_start = time()
```

```
SELECT name FROM sqlite_master WHERE type = 'table';
```

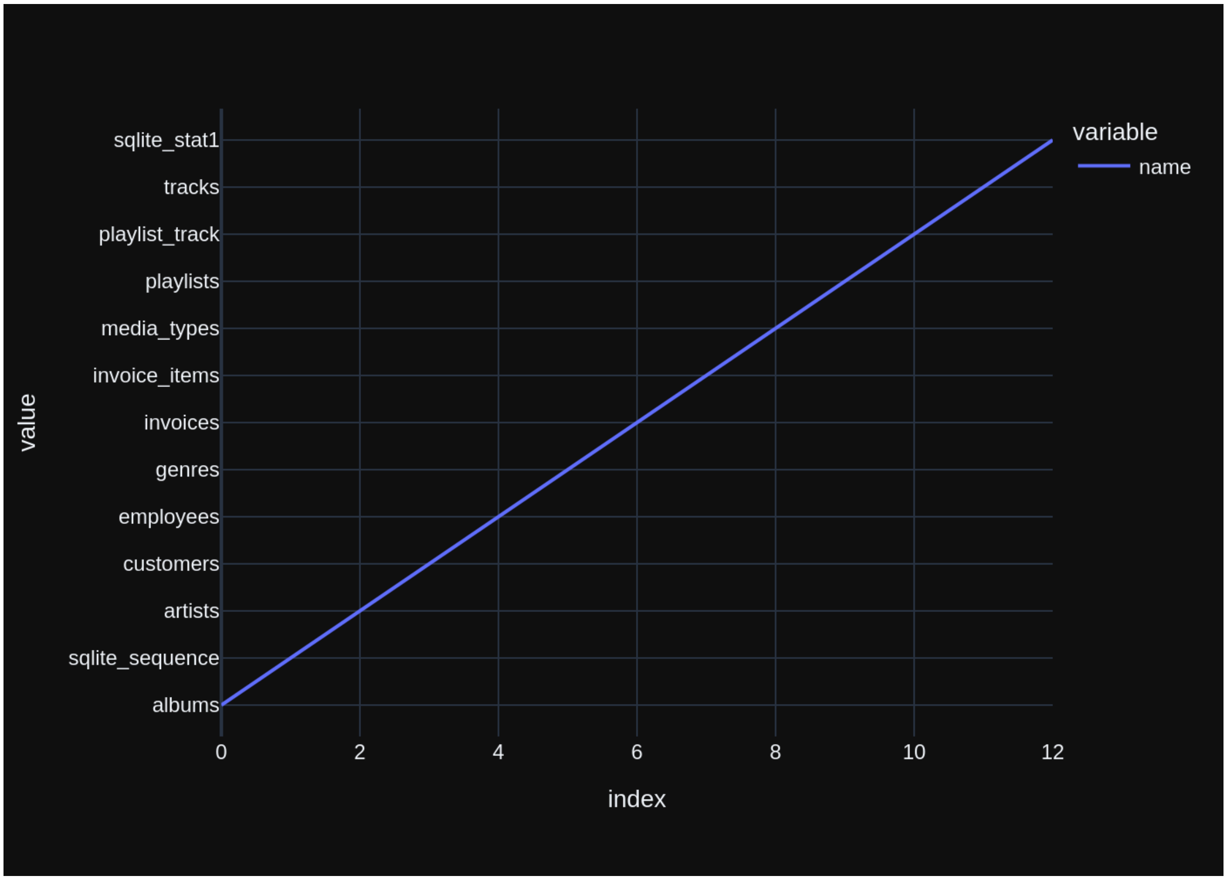
In [16]:

```
vn.ask(question="Can you list all tables in the SQLite database catalog?")
```

```
Number of requested results 10 is greater than number of elements in index
1, updating n_results = 1
```

```
SQL Prompt: [{'role': 'system', 'content': 'You are a SQLite expert. Please
help to generate a SQL query to answer the question. Your response should ON
LY be based on the given context and follow the response guidelines and form
at instructions. \n===Tables \nCREATE TABLE sqlite_stat1(tbl,idx,stat)\n\nCR
EATE TABLE sqlite_sequence(name,seq)\n\nCREATE TABLE "playlists"\r\n(\r\n
PlaylistId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR
(120)\r\n)\n\nCREATE TABLE "genres"\r\n(\r\n    GenreId INTEGER PRIMARY KEY
AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(120)\r\n)\n\nCREATE TABLE "trac
ks"\r\n(\r\n    TrackId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    N
ame NVARCHAR(200)  NOT NULL,\r\n    AlbumId INTEGER,\r\n    MediaTypeId INTE
GER  NOT NULL,\r\n    GenreId INTEGER,\r\n    Composer NVARCHAR(220),\r\n
Milliseconds INTEGER  NOT NULL,\r\n    Bytes INTEGER,\r\n    UnitPrice NUMER
IC(10,2)  NOT NULL,\r\n    FOREIGN KEY (AlbumId) REFERENCES "albums" (AlbumI
d) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (Genr
eId) REFERENCES "genres" (GenreId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO
ACTION,\r\n    FOREIGN KEY (MediaTypeId) REFERENCES "media_types" (MediaType
Id) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE "me
dia_types"\r\n(\r\n    MediaTypeId INTEGER PRIMARY KEY AUTOINCREMENT NOT NUL
L,\r\n    Name NVARCHAR(120)\r\n)\n\nCREATE TABLE "artists"\r\n(\r\n    Arti
stId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(120)\r
\n)\n\nCREATE TABLE "invoice_items"\r\n(\r\n    InvoiceLineId INTEGER PRIMAR
Y KEY AUTOINCREMENT NOT NULL,\r\n    InvoiceId INTEGER  NOT NULL,\r\n    Tra
ckId INTEGER  NOT NULL,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    Qua
ntity INTEGER  NOT NULL,\r\n    FOREIGN KEY (InvoiceId) REFERENCES "invoice
s" (InvoiceId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREI
GN KEY (TrackId) REFERENCES "tracks" (TrackId) \r\n\t\tON DELETE NO ACTION O
N UPDATE NO ACTION\r\n)\n\nCREATE TABLE "playlist_track"\r\n(\r\n    Playlis
tId INTEGER  NOT NULL,\r\n    TrackId INTEGER  NOT NULL,\r\n    CONSTRAINT P
K_PlaylistTrack PRIMARY KEY  (PlaylistId, TrackId),\r\n    FOREIGN KEY (Play
listId) REFERENCES "playlists" (PlaylistId) \r\n\t\tON DELETE NO ACTION ON U
PDATE NO ACTION,\r\n    FOREIGN KEY (TrackId) REFERENCES "tracks" (TrackId)
\r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE "album
s"\r\n(\r\n    AlbumId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Ti
tle NVARCHAR(160)  NOT NULL,\r\n    ArtistId INTEGER  NOT NULL,\r\n    FOREI
GN KEY (ArtistId) REFERENCES "artists" (ArtistId) \r\n\t\tON DELETE NO ACTIO
N ON UPDATE NO ACTION\r\n)\n\n\n===Additional Context \n\nIn the chinook dat
abase invoice means order\n\n===Response Guidelines \n1. If the provided con
text is sufficient, please generate a valid SQL query without any explanatio
ns for the question. \n2. If the provided context is almost sufficient but r
equires knowledge of a specific string in a particular column, please genera
te an intermediate SQL query to find the distinct strings in that column. Pr
epend the query with a comment saying intermediate_sql \n3. If the provided
context is insufficient, please explain why it can\'t be generated. \n4. Ple
ase use the most relevant table(s). \n5. If the question has been asked and
answered before, please repeat the answer exactly as it was given before.
\n'}, {'role': 'user', 'content': 'Can you list all tables in the SQLite dat
abase catalog?'}]
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "You are a SQLite expert. Please help to gene
rate a SQL query to answer the question. Your response should ONLY be based
on the given context and follow the response guidelines and format instructi
ons. \n===Tables \nCREATE TABLE sqlite_stat1(tbl,idx,stat)\n\nCREATE TABLE s
qlite_sequence(name,seq)\n\nCREATE TABLE \"playlists\"\r\n(\r\n    PlaylistI
d INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(120)\r\n)
\n\nCREATE TABLE \"genres\"\r\n(\r\n    GenreId INTEGER PRIMARY KEY AUTOINCR
EMENT NOT NULL,\r\n    Name NVARCHAR(120)\r\n)\n\nCREATE TABLE \"tracks\"\r
\n(\r\n    TrackId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name N
VARCHAR(200)  NOT NULL,\r\n    AlbumId INTEGER,\r\n    MediaTypeId INTEGER 
NOT NULL,\r\n    GenreId INTEGER,\r\n    Composer NVARCHAR(220),\r\n    Mill
iseconds INTEGER  NOT NULL,\r\n    Bytes INTEGER,\r\n    UnitPrice NUMERIC(1
0,2)  NOT NULL,\r\n    FOREIGN KEY (AlbumId) REFERENCES \"albums\" (AlbumId)
\r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (GenreI
d) REFERENCES \"genres\" (GenreId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO 
ACTION,\r\n    FOREIGN KEY (MediaTypeId) REFERENCES \"media_types\" (MediaTy
peId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE 
\"media_types\"\r\n(\r\n    MediaTypeId INTEGER PRIMARY KEY AUTOINCREMENT NO
T NULL,\r\n    Name NVARCHAR(120)\r\n)\n\nCREATE TABLE \"artists\"\r\n(\r\n 
ArtistId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(12
0)\r\n)\n\nCREATE TABLE \"invoice_items\"\r\n(\r\n    InvoiceLineId INTEGER 
PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    InvoiceId INTEGER  NOT NULL,\r\n 
TrackId INTEGER  NOT NULL,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    
Quantity INTEGER  NOT NULL,\r\n    FOREIGN KEY (InvoiceId) REFERENCES \"invo
ices\" (InvoiceId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    F
OREIGN KEY (TrackId) REFERENCES \"tracks\" (TrackId) \r\n\t\tON DELETE NO AC
TION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE \"playlist_track\"\r\n(\r\n   
PlaylistId INTEGER  NOT NULL,\r\n    TrackId INTEGER  NOT NULL,\r\n    CONST
RAINT PK_PlaylistTrack PRIMARY KEY  (PlaylistId, TrackId),\r\n    FOREIGN KE
Y (PlaylistId) REFERENCES \"playlists\" (PlaylistId) \r\n\t\tON DELETE NO AC
TION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackId) REFERENCES \"tracks\"
(TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TAB
LE \"albums\"\r\n(\r\n    AlbumId INTEGER PRIMARY KEY AUTOINCREMENT NOT NUL
L,\r\n    Title NVARCHAR(160)  NOT NULL,\r\n    ArtistId INTEGER  NOT NUL
L,\r\n    FOREIGN KEY (ArtistId) REFERENCES \"artists\" (ArtistId) \r\n\t\tO
N DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\n\n===Additional Context \n\nI
n the chinook database invoice means order\n\n===Response Guidelines \n1. If
the provided context is sufficient, please generate a valid SQL query withou
t any explanations for the question. \n2. If the provided context is almost 
sufficient but requires knowledge of a specific string in a particular colum
n, please generate an intermediate SQL query to find the distinct strings in
that column. Prepend the query with a comment saying intermediate_sql \n3. I
f the provided context is insufficient, please explain why it can't be gener
ated. \n4. Please use the most relevant table(s). \n5. If the question has b
een asked and answered before, please repeat the answer exactly as it was gi
ven before. \n"}, {"role": "user", "content": "Can you list all tables in th
e SQLite database catalog?"}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T21:28:39.653335518Z', 'messa
ge': {'role': 'assistant', 'content': "SELECT name FROM sqlite_master WHERE 
type = 'table';"}, 'done_reason': 'stop', 'done': True, 'total_duration': 74
227580811, 'load_duration': 3773367453, 'prompt_eval_count': 1076, 'prompt_e
val_duration': 66748813000, 'eval_count': 13, 'eval_duration': 3604572000}
LLM Response: SELECT name FROM sqlite_master WHERE type = 'table';
Info: Output from LLM: SELECT name FROM sqlite_master WHERE type = 'table';
Extracted SQL: SELECT name FROM sqlite_master WHERE type = 'table'
SELECT name FROM sqlite_master WHERE type = 'table'
              name
0           albums
1   sqlite_sequence
```

```
2           artists
3         customers
4         employees
5            genres
6          invoices
7     invoice_items
8       media_types
9         playlists
10   playlist_track
11           tracks
12     sqlite_stat1
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "The following is a pandas DataFrame that con
tains the results of the query that answers the question the user asked: 'Ca
n you list all tables in the SQLite database catalog?'\n\nThe DataFrame was 
produced using this query: SELECT name FROM sqlite_master WHERE type = 'tabl
e'\n\nThe following is information about the resulting pandas DataFrame 'd
f': \nRunning df.dtypes gives:\n name    object\ndtype: object"}, {"role": 
"user", "content": "Can you generate the Python plotly code to chart the res
ults of the dataframe? Assume the data is in a pandas dataframe called 'df'. 
If there is only one value in the dataframe, use an Indicator. Respond with 
only Python code. Do not answer with any explanations -- just the code."}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T21:29:37.669496643Z', 'messa
ge': {'role': 'assistant', 'content': '```python\nimport plotly.graph_object
s as go\n\n# Assuming df has a column named "name" which contains table name
s\nif len(df) == 1:\n    # Use an Indicator for single value dataframe\n
fig = go.Figure(data=[go.Indicator(value=df["name"][0], mode="number+rang
e")])\nelse:\n    # Create a Bar chart for multiple values in the dataframe
\n    fig = go\n    trace = go.Bar(x=df[\'name\'], y=[1]*len(df))  # Assumin
g each table is represented by one unit (e.g., count of rows)\n    fig.add_t
race(trace)\n\nfig.show()\n```'}, 'done_reason': 'stop', 'done': True, 'tota
l_duration': 57981657875, 'load_duration': 2990616, 'prompt_eval_count': 16
7, 'prompt_eval_duration': 10092785000, 'eval_count': 161, 'eval_duration': 
47793436000}
```


sqlite_stat1
tracks
playlist_track
playlists
media_types
invoice_items
invoices
genres
employees
customers
artists
sqlite_sequence
albums
value
0
2
4
6
8
10
12
index
variable
name

Out[16]:

```
("SELECT name FROM sqlite_master WHERE type = 'table'",
                 name
 0             albums
 1    sqlite_sequence
 2            artists
 3          customers
 4          employees
 5             genres
 6           invoices
 7      invoice_items
 8        media_types
 9          playlists
 10    playlist_track
 11            tracks
 12      sqlite_stat1,
 Figure({
     'data': [{'hovertemplate': 'variable=name<br>index=%{x}<br>value=%{y}<
extra></extra>',
               'legendgroup': 'name',
               'line': {'color': '#636efa', 'dash': 'solid'},
               'marker': {'symbol': 'circle'},
               'mode': 'lines',
               'name': 'name',
               'orientation': 'v',
               'showlegend': True,
               'type': 'scatter',
               'x': array([ 0,  1,  2,  3,  4,  5,  6,  7,  8,  9, 10, 11,
12]),
               'xaxis': 'x',
               'y': array(['albums', 'sqlite_sequence', 'artists', 'custome
rs', 'employees',
                           'genres', 'invoices', 'invoice_items', 'media_ty
pes', 'playlists',
                           'playlist_track', 'tracks', 'sqlite_stat1'], dty
pe=object),
               'yaxis': 'y'}],
     'layout': {'legend': {'title': {'text': 'variable'}, 'tracegroupgap':
0},
                'margin': {'t': 60},
                'template': '...',
                'xaxis': {'anchor': 'y', 'domain': [0.0, 1.0], 'title': {'t
ext': 'index'}},
                'yaxis': {'anchor': 'x', 'domain': [0.0, 1.0], 'title': {'t
ext': 'value'}}}
 }))
```

In [17]:

```
vn.ask(question="which table stores customer's orders")
```

```
Number of requested results 10 is greater than number of elements in index
1, updating n_results = 1
Number of requested results 10 is greater than number of elements in index
1, updating n_results = 1
```

```
SQL Prompt: [{'role': 'system', 'content': 'You are a SQLite expert. Please
help to generate a SQL query to answer the question. Your response should ON
LY be based on the given context and follow the response guidelines and form
at instructions. \n===Tables \nCREATE TABLE "invoices"\r\n(\r\n    InvoiceId
INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    CustomerId INTEGER  NOT N
ULL,\r\n    InvoiceDate DATETIME  NOT NULL,\r\n    BillingAddress NVARCHAR(7
0),\r\n    BillingCity NVARCHAR(40),\r\n    BillingState NVARCHAR(40),\r\n
BillingCountry NVARCHAR(40),\r\n    BillingPostalCode NVARCHAR(10),\r\n    T
otal NUMERIC(10,2)  NOT NULL,\r\n    FOREIGN KEY (CustomerId) REFERENCES "cu
stomers" (CustomerId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n
\nCREATE TABLE "invoice_items"\r\n(\r\n    InvoiceLineId INTEGER PRIMARY KEY
AUTOINCREMENT NOT NULL,\r\n    InvoiceId INTEGER  NOT NULL,\r\n    TrackId I
NTEGER  NOT NULL,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    Quantity
INTEGER  NOT NULL,\r\n    FOREIGN KEY (InvoiceId) REFERENCES "invoices" (Inv
oiceId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY
(TrackId) REFERENCES "tracks" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDAT
E NO ACTION\r\n)\n\nCREATE TABLE "customers"\r\n(\r\n    CustomerId INTEGER
PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    FirstName NVARCHAR(40)  NOT NUL
L,\r\n    LastName NVARCHAR(20)  NOT NULL,\r\n    Company NVARCHAR(80),\r\n
Address NVARCHAR(70),\r\n    City NVARCHAR(40),\r\n    State NVARCHAR(40),\r
\n    Country NVARCHAR(40),\r\n    PostalCode NVARCHAR(10),\r\n    Phone NVA
RCHAR(24),\r\n    Fax NVARCHAR(24),\r\n    Email NVARCHAR(60)  NOT NULL,\r\n
SupportRepId INTEGER,\r\n    FOREIGN KEY (SupportRepId) REFERENCES "employee
s" (EmployeeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREA
TE TABLE "employees"\r\n(\r\n    EmployeeId INTEGER PRIMARY KEY AUTOINCREMEN
T NOT NULL,\r\n    LastName NVARCHAR(20)  NOT NULL,\r\n    FirstName NVARCHA
R(20)  NOT NULL,\r\n    Title NVARCHAR(30),\r\n    ReportsTo INTEGER,\r\n
BirthDate DATETIME,\r\n    HireDate DATETIME,\r\n    Address NVARCHAR(70),\r
\n    City NVARCHAR(40),\r\n    State NVARCHAR(40),\r\n    Country NVARCHAR
(40),\r\n    PostalCode NVARCHAR(10),\r\n    Phone NVARCHAR(24),\r\n    Fax
NVARCHAR(24),\r\n    Email NVARCHAR(60),\r\n    FOREIGN KEY (ReportsTo) REFE
RENCES "employees" (EmployeeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACT
ION\r\n)\n\nCREATE TABLE sqlite_sequence(name,seq)\n\nCREATE TABLE "playlist
s"\r\n(\r\n    PlaylistId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n
Name NVARCHAR(120)\r\n)\n\nCREATE TABLE sqlite_stat1(tbl,idx,stat)\n\nCREATE
TABLE "albums"\r\n(\r\n    AlbumId INTEGER PRIMARY KEY AUTOINCREMENT NOT NUL
L,\r\n    Title NVARCHAR(160)  NOT NULL,\r\n    ArtistId INTEGER  NOT NUL
L,\r\n    FOREIGN KEY (ArtistId) REFERENCES "artists" (ArtistId) \r\n\t\tON
DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE "playlist_track"\r
\n(\r\n    PlaylistId INTEGER  NOT NULL,\r\n    TrackId INTEGER  NOT NULL,\r
\n    CONSTRAINT PK_PlaylistTrack PRIMARY KEY  (PlaylistId, TrackId),\r\n
FOREIGN KEY (PlaylistId) REFERENCES "playlists" (PlaylistId) \r\n\t\tON DELE
TE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackId) REFERENCES "t
racks" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCRE
ATE TABLE "media_types"\r\n(\r\n    MediaTypeId INTEGER PRIMARY KEY AUTOINCR
EMENT NOT NULL,\r\n    Name NVARCHAR(120)\r\n)\n\n\n===Additional Context \n
\nIn the chinook database invoice means order\n\n===Response Guidelines \n1.
If the provided context is sufficient, please generate a valid SQL query wit
hout any explanations for the question. \n2. If the provided context is almo
st sufficient but requires knowledge of a specific string in a particular co
lumn, please generate an intermediate SQL query to find the distinct strings
in that column. Prepend the query with a comment saying intermediate_sql \n
3. If the provided context is insufficient, please explain why it can\'t be
generated. \n4. Please use the most relevant table(s). \n5. If the question
has been asked and answered before, please repeat the answer exactly as it w
as given before. \n'}, {'role': 'user', 'content': 'Can you list all tables
```

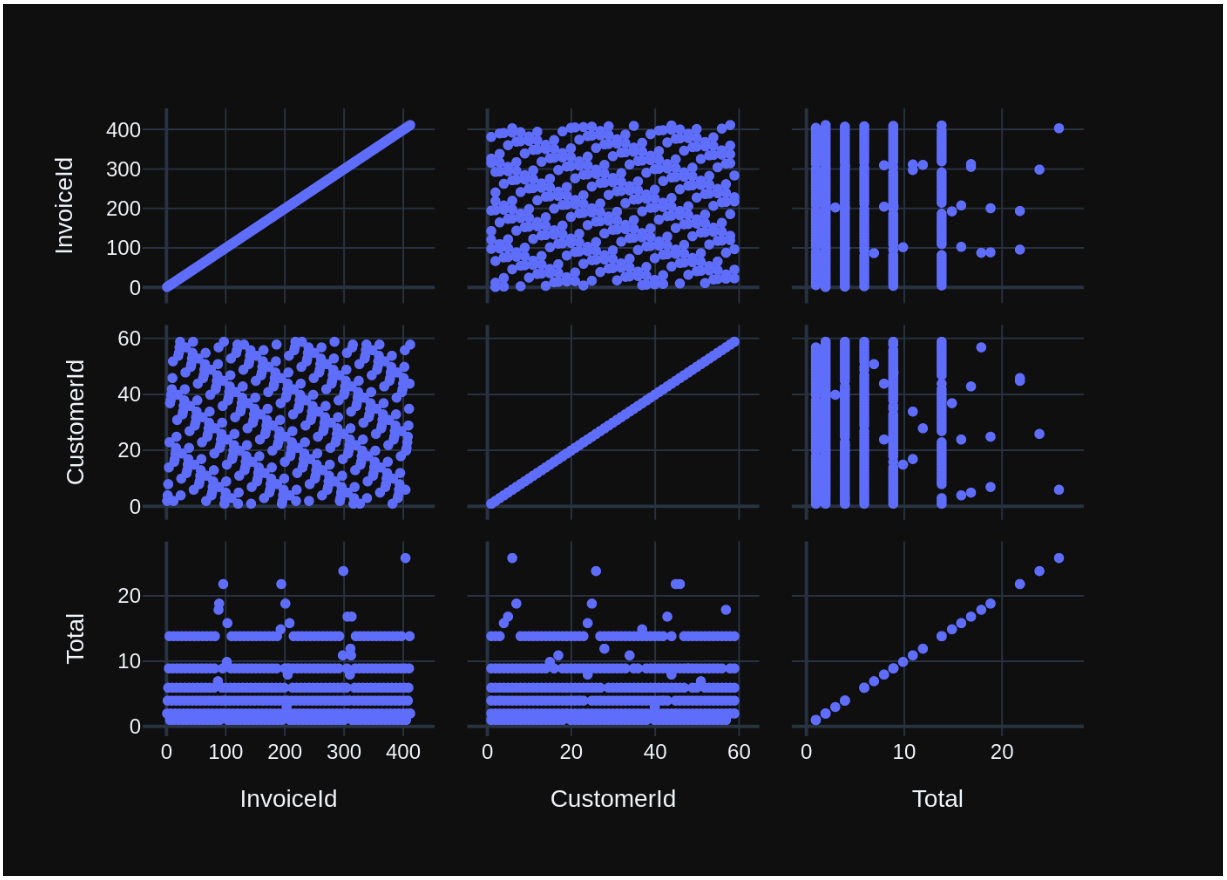
```
in the SQLite database catalog?'}, {'role': 'assistant', 'content': "SELECT
name FROM sqlite_master WHERE type = 'table'"}, {'role': 'user', 'content':
"which table stores customer's orders"}]
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "You are a SQLite expert. Please help to gene
rate a SQL query to answer the question. Your response should ONLY be based
on the given context and follow the response guidelines and format instructi
ons. \n===Tables \nCREATE TABLE \"invoices\"\r\n(\r\n    InvoiceId INTEGER P
RIMARY KEY AUTOINCREMENT NOT NULL,\r\n    CustomerId INTEGER  NOT NULL,\r\n
InvoiceDate DATETIME  NOT NULL,\r\n    BillingAddress NVARCHAR(70),\r\n    B
illingCity NVARCHAR(40),\r\n    BillingState NVARCHAR(40),\r\n    BillingCou
ntry NVARCHAR(40),\r\n    BillingPostalCode NVARCHAR(10),\r\n    Total NUMER
IC(10,2)  NOT NULL,\r\n    FOREIGN KEY (CustomerId) REFERENCES \"customers\"
(CustomerId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE
TABLE \"invoice_items\"\r\n(\r\n    InvoiceLineId INTEGER PRIMARY KEY AUTOIN
CREMENT NOT NULL,\r\n    InvoiceId INTEGER  NOT NULL,\r\n    TrackId INTEGER
NOT NULL,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    Quantity INTEGER
NOT NULL,\r\n    FOREIGN KEY (InvoiceId) REFERENCES \"invoices\" (InvoiceId)
\r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackI
d) REFERENCES \"tracks\" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO
ACTION\r\n)\n\nCREATE TABLE \"customers\"\r\n(\r\n    CustomerId INTEGER PRI
MARY KEY AUTOINCREMENT NOT NULL,\r\n    FirstName NVARCHAR(40)  NOT NULL,\r
\n    LastName NVARCHAR(20)  NOT NULL,\r\n    Company NVARCHAR(80),\r\n    A
ddress NVARCHAR(70),\r\n    City NVARCHAR(40),\r\n    State NVARCHAR(40),\r
\n    Country NVARCHAR(40),\r\n    PostalCode NVARCHAR(10),\r\n    Phone NVA
RCHAR(24),\r\n    Fax NVARCHAR(24),\r\n    Email NVARCHAR(60)  NOT NULL,\r\n
SupportRepId INTEGER,\r\n    FOREIGN KEY (SupportRepId) REFERENCES \"employe
es\" (EmployeeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCR
EATE TABLE \"employees\"\r\n(\r\n    EmployeeId INTEGER PRIMARY KEY AUTOINCR
EMENT NOT NULL,\r\n    LastName NVARCHAR(20)  NOT NULL,\r\n    FirstName NVA
RCHAR(20)  NOT NULL,\r\n    Title NVARCHAR(30),\r\n    ReportsTo INTEGER,\r
\n    BirthDate DATETIME,\r\n    HireDate DATETIME,\r\n    Address NVARCHAR
(70),\r\n    City NVARCHAR(40),\r\n    State NVARCHAR(40),\r\n    Country NV
ARCHAR(40),\r\n    PostalCode NVARCHAR(10),\r\n    Phone NVARCHAR(24),\r\n
Fax NVARCHAR(24),\r\n    Email NVARCHAR(60),\r\n    FOREIGN KEY (ReportsTo)
REFERENCES \"employees\" (EmployeeId) \r\n\t\tON DELETE NO ACTION ON UPDATE
NO ACTION\r\n)\n\nCREATE TABLE sqlite_sequence(name,seq)\n\nCREATE TABLE \"p
laylists\"\r\n(\r\n    PlaylistId INTEGER PRIMARY KEY AUTOINCREMENT NOT NUL
L,\r\n    Name NVARCHAR(120)\r\n)\n\nCREATE TABLE sqlite_stat1(tbl,idx,stat)
\n\nCREATE TABLE \"albums\"\r\n(\r\n    AlbumId INTEGER PRIMARY KEY AUTOINCR
EMENT NOT NULL,\r\n    Title NVARCHAR(160)  NOT NULL,\r\n    ArtistId INTEGE
R  NOT NULL,\r\n    FOREIGN KEY (ArtistId) REFERENCES \"artists\" (ArtistId)
\r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE \"playl
ist_track\"\r\n(\r\n    PlaylistId INTEGER  NOT NULL,\r\n    TrackId INTEGER
NOT NULL,\r\n    CONSTRAINT PK_PlaylistTrack PRIMARY KEY  (PlaylistId, Track
Id),\r\n    FOREIGN KEY (PlaylistId) REFERENCES \"playlists\" (PlaylistId)
\r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackI
d) REFERENCES \"tracks\" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO
ACTION\r\n)\n\nCREATE TABLE \"media_types\"\r\n(\r\n    MediaTypeId INTEGER
PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(120)\r\n)\n\n\n===A
dditional Context \n\nIn the chinook database invoice means order\n\n===Resp
onse Guidelines \n1. If the provided context is sufficient, please generate
```

~~~
a valid SQL query without any explanations for the question. \n2. If the pro
vided context is almost sufficient but requires knowledge of a specific stri
ng in a particular column, please generate an intermediate SQL query to find
the distinct strings in that column. Prepend the query with a comment saying
intermediate_sql \n3. If the provided context is insufficient, please explai
n why it can't be generated. \n4. Please use the most relevant table(s). \n
5. If the question has been asked and answered before, please repeat the ans
wer exactly as it was given before. \n"}, {"role": "user", "content": "Can y
ou list all tables in the SQLite database catalog?"}, {"role": "assistant",
"content": "SELECT name FROM sqlite_master WHERE type = 'table'"}, {"role":
"user", "content": "which table stores customer's orders"}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T21:31:28.031565534Z', 'messa
ge': {'role': 'assistant', 'content': 'The "invoices" table stores customers
\' orders. The context mentions that in the chinook database, an invoice mea
ns an order. Here is a SQL query to select all columns from the "invoices" t
able:\n\n```sql\nSELECT * FROM invoices;\n```'}, 'done_reason': 'stop', 'don
e': True, 'total_duration': 109932628453, 'load_duration': 2575132, 'prompt_
eval_count': 1397, 'prompt_eval_duration': 90649200000, 'eval_count': 63, 'e
val_duration': 19089563000}
LLM Response: The "invoices" table stores customers' orders. The context men
tions that in the chinook database, an invoice means an order. Here is a SQL
query to select all columns from the "invoices" table:

```sql
SELECT * FROM invoices;
~~~
Info: Output from LLM: The "invoices" table stores customers' orders. The co
ntext mentions that in the chinook database, an invoice means an order. Here
is a SQL query to select all columns from the "invoices" table:

```sql
SELECT * FROM invoices;
```
Extracted SQL: SELECT * FROM invoices
SELECT * FROM invoices
     InvoiceId CustomerId   InvoiceDate \
0   1   2 2009-01-01 00:00:00
1   2   4 2009-01-02 00:00:00
2   3   8 2009-01-03 00:00:00
3   4   14 2009-01-06 00:00:00
4   5   23 2009-01-11 00:00:00
..   ..
407   408   25 2013-12-05 00:00:00
408   409   29 2013-12-06 00:00:00
409   410   35 2013-12-09 00:00:00
410   411   44 2013-12-14 00:00:00
411   412   58 2013-12-22 00:00:00

                               BillingAddress BillingCity BillingState \
0   Theodor-Heuss-Straße 34   Stuttgart   None
1   Ullevålsveien 14   Oslo   None
2   Grétrystraat 63   Brussels   None
3   8210 111 ST NW   Edmonton   AB
4   69 Salem Street   Boston   MA
..   ..
```
```

```
407                      319 N. Frances Street     Madison           WI
408                     796 Dundas Street West     Toronto           ON
409  Rua dos Campeões Europeus de Viena, 4350       Porto         None
410                            Porthaninkatu 9    Helsinki         None
411                        12,Community Centre       Delhi         None

    BillingCountry BillingPostalCode  Total
0          Germany             70174   1.98
1           Norway              0171   3.96
2          Belgium              1000   5.94
3           Canada           T6G 2C7   8.91
4              USA              2113  13.86
..             ...               ...    ...
407            USA             53703   3.96
408         Canada           M6J 1V1   5.94
409       Portugal              None   8.91
410        Finland             00530  13.86
411          India            110017   1.99

[412 rows x 9 columns]
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "The following is a pandas DataFrame that con
tains the results of the query that answers the question the user asked: 'wh
ich table stores customer's orders'\n\nThe DataFrame was produced using this
query: SELECT * FROM invoices\n\nThe following is information about the resu
lting pandas DataFrame 'df': \nRunning df.dtypes gives:\n InvoiceId
int64\nCustomerId             int64\nInvoiceDate           object\nBillingAd
dress        object\nBillingCity           object\nBillingState          obj
ect\nBillingCountry        object\nBillingPostalCode     object\nTotal
float64\ndtype: object"}, {"role": "user", "content": "Can you generate the
Python plotly code to chart the results of the dataframe? Assume the data is
in a pandas dataframe called 'df'. If there is only one value in the datafra
me, use an Indicator. Respond with only Python code. Do not answer with any
explanations -- just the code."}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T21:32:10.123346138Z', 'messa
ge': {'role': 'assistant', 'content': 'Here\'s your Plotly code:\n\n```pytho
n\nimport plotly.express as px\n\n# Assume df is a DataFrame containing the
data you want to visualize\nif len(df) > 1:\n    fig = px.scatter_matrix(df,
dimensions=["InvoiceId", "CustomerId", "Total"])\nelse:\n    fig = px.indica
tor()\n    \nfig.show()\n```'}, 'done_reason': 'stop', 'done': True, 'total_
duration': 42062216722, 'load_duration': 45667486, 'prompt_eval_count': 213,
'prompt_eval_duration': 12740386000, 'eval_count': 99, 'eval_duration': 2923
1081000}
```


InvoiceId
CustomerId
Total
0
100
200
300
400
0
20
40
60
0
10
20
InvoiceId
CustomerId
Total

```
Out[17]: ('SELECT * FROM invoices',
               InvoiceId  CustomerId          InvoiceDate  \
          0            1           2  2009-01-01 00:00:00
          1            2           4  2009-01-02 00:00:00
          2            3           8  2009-01-03 00:00:00
          3            4          14  2009-01-06 00:00:00
          4            5          23  2009-01-11 00:00:00
          ..         ...         ...                  ...
          407        408          25  2013-12-05 00:00:00
          408        409          29  2013-12-06 00:00:00
          409        410          35  2013-12-09 00:00:00
          410        411          44  2013-12-14 00:00:00
          411        412          58  2013-12-22 00:00:00

                                    BillingAddress BillingCity BillingState  \
          0                Theodor-Heuss-Straße 34   Stuttgart         None
          1                       Ullevålsveien 14        Oslo         None
          2                        Grétrystraat 63    Brussels         None
          3                         8210 111 ST NW    Edmonton           AB
          4                        69 Salem Street      Boston           MA
          ..                                   ...         ...          ...
          407                 319 N. Frances Street     Madison           WI
          408                796 Dundas Street West     Toronto           ON
          409  Rua dos Campeões Europeus de Viena, 4350       Porto         None
          410                      Porthaninkatu 9    Helsinki         None
          411                  12,Community Centre       Delhi         None

              BillingCountry BillingPostalCode  Total
          0          Germany             70174   1.98
          1           Norway              0171   3.96
          2          Belgium              1000   5.94
          3           Canada           T6G 2C7   8.91
          4              USA              2113  13.86
          ..             ...               ...    ...
          407            USA             53703   3.96
          408         Canada           M6J 1V1   5.94
          409       Portugal              None   8.91
          410        Finland             00530  13.86
          411          India            110017   1.99

          [412 rows x 9 columns],
          Figure({
              'data': [{'dimensions': [{'axis': {'matches': True},
                                        'label': 'InvoiceId',
                                        'values': array([  1,   2,   3, ..., 410, 41
         1, 412])},
                                       {'axis': {'matches': True},
                                        'label': 'CustomerId',
                                        'values': array([ 2,  4,  8, ..., 35, 44, 5
         8])},
                                       {'axis': {'matches': True},
                                        'label': 'Total',
                                        'values': array([ 1.98,  3.96,  5.94, ...,
         8.91, 13.86,  1.99])}],
                        'hovertemplate': '%{xaxis.title.text}=%{x}<br>%{yaxis.title.
         text}=%{y}<extra></extra>',
                        'legendgroup': '',
                        'marker': {'color': '#636efa', 'symbol': 'circle'},
                        'name': '',
                        'showlegend': False,
                        'type': 'splom'}],
              'layout': {'dragmode': 'select', 'legend': {'tracegroupgap': 0}, 'marg
         in': {'t': 60}, 'template': '...'}
          }))
```

In [18]:

```
vn.ask(question="How many customers are there")
```

```
Number of requested results 10 is greater than number of elements in index
2, updating n_results = 2
Number of requested results 10 is greater than number of elements in index
1, updating n_results = 1
```

```
SQL Prompt: [{'role': 'system', 'content': 'You are a SQLite expert. Please
help to generate a SQL query to answer the question. Your response should ON
LY be based on the given context and follow the response guidelines and form
at instructions. \n===Tables \nCREATE TABLE "invoices"\r\n(\r\n    InvoiceId
INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    CustomerId INTEGER  NOT N
ULL,\r\n    InvoiceDate DATETIME  NOT NULL,\r\n    BillingAddress NVARCHAR(7
0),\r\n    BillingCity NVARCHAR(40),\r\n    BillingState NVARCHAR(40),\r\n
BillingCountry NVARCHAR(40),\r\n    BillingPostalCode NVARCHAR(10),\r\n    T
otal NUMERIC(10,2)  NOT NULL,\r\n    FOREIGN KEY (CustomerId) REFERENCES "cu
stomers" (CustomerId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n
\nCREATE INDEX IFK_CustomerSupportRepId ON "customers" (SupportRepId)\n\nCRE
ATE TABLE "customers"\r\n(\r\n    CustomerId INTEGER PRIMARY KEY AUTOINCREME
NT NOT NULL,\r\n    FirstName NVARCHAR(40)  NOT NULL,\r\n    LastName NVARCH
AR(20)  NOT NULL,\r\n    Company NVARCHAR(80),\r\n    Address NVARCHAR(7
0),\r\n    City NVARCHAR(40),\r\n    State NVARCHAR(40),\r\n    Country NVAR
CHAR(40),\r\n    PostalCode NVARCHAR(10),\r\n    Phone NVARCHAR(24),\r\n
Fax NVARCHAR(24),\r\n    Email NVARCHAR(60)  NOT NULL,\r\n    SupportRepId I
NTEGER,\r\n    FOREIGN KEY (SupportRepId) REFERENCES "employees" (EmployeeI
d) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_
InvoiceCustomerId ON "invoices" (CustomerId)\n\nCREATE TABLE "invoice_item
s"\r\n(\r\n    InvoiceLineId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n
InvoiceId INTEGER  NOT NULL,\r\n    TrackId INTEGER  NOT NULL,\r\n    UnitPr
ice NUMERIC(10,2)  NOT NULL,\r\n    Quantity INTEGER  NOT NULL,\r\n    FOREI
GN KEY (InvoiceId) REFERENCES "invoices" (InvoiceId) \r\n\t\tON DELETE NO AC
TION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackId) REFERENCES "tracks"
(TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE IND
EX IFK_InvoiceLineInvoiceId ON "invoice_items" (InvoiceId)\n\nCREATE TABLE
"albums"\r\n(\r\n    AlbumId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n
Title NVARCHAR(160)  NOT NULL,\r\n    ArtistId INTEGER  NOT NULL,\r\n    FOR
EIGN KEY (ArtistId) REFERENCES "artists" (ArtistId) \r\n\t\tON DELETE NO ACT
ION ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_InvoiceLineTrackId ON "invo
ice_items" (TrackId)\n\nCREATE TABLE "employees"\r\n(\r\n    EmployeeId INTE
GER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    LastName NVARCHAR(20)  NOT NU
LL,\r\n    FirstName NVARCHAR(20)  NOT NULL,\r\n    Title NVARCHAR(30),\r\n
ReportsTo INTEGER,\r\n    BirthDate DATETIME,\r\n    HireDate DATETIME,\r\n
Address NVARCHAR(70),\r\n    City NVARCHAR(40),\r\n    State NVARCHAR(40),\r
\n    Country NVARCHAR(40),\r\n    PostalCode NVARCHAR(10),\r\n    Phone NVA
RCHAR(24),\r\n    Fax NVARCHAR(24),\r\n    Email NVARCHAR(60),\r\n    FOREIG
N KEY (ReportsTo) REFERENCES "employees" (EmployeeId) \r\n\t\tON DELETE NO A
CTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE "playlists"\r\n(\r\n    Playl
istId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(120)
\r\n)\n\n\n===Additional Context \n\nIn the chinook database invoice means o
rder\n\n===Response Guidelines \n1. If the provided context is sufficient, p
lease generate a valid SQL query without any explanations for the question.
\n2. If the provided context is almost sufficient but requires knowledge of
a specific string in a particular column, please generate an intermediate SQ
L query to find the distinct strings in that column. Prepend the query with
a comment saying intermediate_sql \n3. If the provided context is insufficie
nt, please explain why it can\'t be generated. \n4. Please use the most rele
vant table(s). \n5. If the question has been asked and answered before, plea
se repeat the answer exactly as it was given before. \n'}, {'role': 'user',
'content': "which table stores customer's orders"}, {'role': 'assistant', 'c
ontent': 'SELECT * FROM invoices'}, {'role': 'user', 'content': 'Can you lis
t all tables in the SQLite database catalog?'}, {'role': 'assistant', 'conte
nt': "SELECT name FROM sqlite_master WHERE type = 'table'"}, {'role': 'use
r', 'content': 'How many customers are there'}]
```

```
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "You are a SQLite expert. Please help to gene
rate a SQL query to answer the question. Your response should ONLY be based
on the given context and follow the response guidelines and format instructi
ons. \n===Tables \nCREATE TABLE \"invoices\"\r\n(\r\n    InvoiceId INTEGER P
RIMARY KEY AUTOINCREMENT NOT NULL,\r\n    CustomerId INTEGER  NOT NULL,\r\n
InvoiceDate DATETIME  NOT NULL,\r\n    BillingAddress NVARCHAR(70),\r\n    B
illingCity NVARCHAR(40),\r\n    BillingState NVARCHAR(40),\r\n    BillingCou
ntry NVARCHAR(40),\r\n    BillingPostalCode NVARCHAR(10),\r\n    Total NUMER
IC(10,2)  NOT NULL,\r\n    FOREIGN KEY (CustomerId) REFERENCES \"customers\"
(CustomerId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE
INDEX IFK_CustomerSupportRepId ON \"customers\" (SupportRepId)\n\nCREATE TAB
LE \"customers\"\r\n(\r\n    CustomerId INTEGER PRIMARY KEY AUTOINCREMENT NO
T NULL,\r\n    FirstName NVARCHAR(40)  NOT NULL,\r\n    LastName NVARCHAR(2
0)  NOT NULL,\r\n    Company NVARCHAR(80),\r\n    Address NVARCHAR(70),\r\n
City NVARCHAR(40),\r\n    State NVARCHAR(40),\r\n    Country NVARCHAR(40),\r
\n    PostalCode NVARCHAR(10),\r\n    Phone NVARCHAR(24),\r\n    Fax NVARCHA
R(24),\r\n    Email NVARCHAR(60)  NOT NULL,\r\n    SupportRepId INTEGER,\r\n
FOREIGN KEY (SupportRepId) REFERENCES \"employees\" (EmployeeId) \r\n\t\tON
DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_InvoiceCustome
rId ON \"invoices\" (CustomerId)\n\nCREATE TABLE \"invoice_items\"\r\n(\r\n
InvoiceLineId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    InvoiceId I
NTEGER  NOT NULL,\r\n    TrackId INTEGER  NOT NULL,\r\n    UnitPrice NUMERIC
(10,2)  NOT NULL,\r\n    Quantity INTEGER  NOT NULL,\r\n    FOREIGN KEY (Inv
oiceId) REFERENCES \"invoices\" (InvoiceId) \r\n\t\tON DELETE NO ACTION ON U
PDATE NO ACTION,\r\n    FOREIGN KEY (TrackId) REFERENCES \"tracks\" (TrackI
d) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_
InvoiceLineInvoiceId ON \"invoice_items\" (InvoiceId)\n\nCREATE TABLE \"albu
ms\"\r\n(\r\n    AlbumId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n
Title NVARCHAR(160)  NOT NULL,\r\n    ArtistId INTEGER  NOT NULL,\r\n    FOR
EIGN KEY (ArtistId) REFERENCES \"artists\" (ArtistId) \r\n\t\tON DELETE NO A
CTION ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_InvoiceLineTrackId ON \"i
nvoice_items\" (TrackId)\n\nCREATE TABLE \"employees\"\r\n(\r\n    EmployeeI
d INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    LastName NVARCHAR(20)
NOT NULL,\r\n    FirstName NVARCHAR(20)  NOT NULL,\r\n    Title NVARCHAR(3
0),\r\n    ReportsTo INTEGER,\r\n    BirthDate DATETIME,\r\n    HireDate DAT
ETIME,\r\n    Address NVARCHAR(70),\r\n    City NVARCHAR(40),\r\n    State N
VARCHAR(40),\r\n    Country NVARCHAR(40),\r\n    PostalCode NVARCHAR(10),\r
\n    Phone NVARCHAR(24),\r\n    Fax NVARCHAR(24),\r\n    Email NVARCHAR(6
0),\r\n    FOREIGN KEY (ReportsTo) REFERENCES \"employees\" (EmployeeId) \r
\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE \"playlis
ts\"\r\n(\r\n    PlaylistId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n
Name NVARCHAR(120)\r\n)\n\n\n===Additional Context \n\nIn the chinook databa
se invoice means order\n\n===Response Guidelines \n1. If the provided contex
t is sufficient, please generate a valid SQL query without any explanations
for the question. \n2. If the provided context is almost sufficient but requ
ires knowledge of a specific string in a particular column, please generate
an intermediate SQL query to find the distinct strings in that column. Prepe
nd the query with a comment saying intermediate_sql \n3. If the provided con
text is insufficient, please explain why it can't be generated. \n4. Please
use the most relevant table(s). \n5. If the question has been asked and answ
ered before, please repeat the answer exactly as it was given before. \n"},
```

~~~
{"role": "user", "content": "which table stores customer's orders"}, {"rol
e": "assistant", "content": "SELECT * FROM invoices"}, {"role": "user", "con
tent": "Can you list all tables in the SQLite database catalog?"}, {"role":
"assistant", "content": "SELECT name FROM sqlite_master WHERE type = 'tabl
e'"}, {"role": "user", "content": "How many customers are there"}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T21:34:17.392972861Z', 'messa
ge': {'role': 'assistant', 'content': 'To determine how many customers are i
n the "customers" table, we can use a SQL query that counts the number of ro
ws. Here is the appropriate SQL statement:\n\n```sql\nSELECT COUNT(*) AS Num
berOfCustomers FROM customers;\n```\n\nThis query will return a single row w
ith a column named \'NumberOfCustomers\' containing the total count of all c
ustomer records in the "customers" table. The `COUNT(*)` function is used to
count the number of rows, and by assigning this result to an alias (in this
case, \'NumberOfCustomers\'), we can easily reference the resulting value wi
thin our application or further query processing.'}, 'done_reason': 'stop',
'done': True, 'total_duration': 126421814089, 'load_duration': 3114010, 'pro
mpt_eval_count': 1302, 'prompt_eval_duration': 84180661000, 'eval_count': 13
7, 'eval_duration': 41938219000}
LLM Response: To determine how many customers are in the "customers" table,
we can use a SQL query that counts the number of rows. Here is the appropria
te SQL statement:

```sql
SELECT COUNT(*) AS NumberOfCustomers FROM customers;
```

This query will return a single row with a column named 'NumberOfCustomers'
containing the total count of all customer records in the "customers" table.
The `COUNT(*)` function is used to count the number of rows, and by assignin
g this result to an alias (in this case, 'NumberOfCustomers'), we can easily
reference the resulting value within our application or further query proces
sing.
Info: Output from LLM: To determine how many customers are in the "customer
s" table, we can use a SQL query that counts the number of rows. Here is the
appropriate SQL statement:

```sql
SELECT COUNT(*) AS NumberOfCustomers FROM customers;
```

This query will return a single row with a column named 'NumberOfCustomers'
containing the total count of all customer records in the "customers" table.
The `COUNT(*)` function is used to count the number of rows, and by assignin
g this result to an alias (in this case, 'NumberOfCustomers'), we can easily
reference the resulting value within our application or further query proces
sing.
Extracted SQL: SELECT COUNT(*) AS NumberOfCustomers FROM customers
SELECT COUNT(*) AS NumberOfCustomers FROM customers
   NumberOfCustomers
0                 59
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
~~~

```
[{"role": "system", "content": "The following is a pandas DataFrame that con
tains the results of the query that answers the question the user asked: 'Ho
w many customers are there'\n\nThe DataFrame was produced using this query:
SELECT COUNT(*) AS NumberOfCustomers FROM customers\n\nThe following is info
rmation about the resulting pandas DataFrame 'df': \nRunning df.dtypes give
s:\n NumberOfCustomers    int64\ndtype: object"}, {"role": "user", "conten
t": "Can you generate the Python plotly code to chart the results of the dat
aframe? Assume the data is in a pandas dataframe called 'df'. If there is on
ly one value in the dataframe, use an Indicator. Respond with only Python co
de. Do not answer with any explanations -- just the code."}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T21:34:42.096569586Z', 'messa
ge': {'role': 'assistant', 'content': "'''python\nimport plotly.express as p
x\n\n# Assuming df contains our data from the SQL query:\nfig = px.bar(df, x
=['NumberOfCustomers'])\nfig.show()\n'''"}, 'done_reason': 'stop', 'done': T
rue, 'total_duration': 24676846651, 'load_duration': 47149579, 'prompt_eval_
count': 165, 'prompt_eval_duration': 9976887000, 'eval_count': 50, 'eval_dur
ation': 14562265000}
Couldn't run plotly code:  'NoneType' object has no attribute 'show'
```

```
Traceback (most recent call last):
  File "/home/gongai/anaconda3/envs/vanna/lib/python3.11/site-packages/vann
a/base/base.py", line 1684, in ask
    img_bytes = fig.to_image(format="png", scale=2)
                ^^^^^^^^^^^^
AttributeError: 'NoneType' object has no attribute 'to_image'

During handling of the above exception, another exception occurred:

Traceback (most recent call last):
  File "/home/gongai/anaconda3/envs/vanna/lib/python3.11/site-packages/vann
a/base/base.py", line 1687, in ask
    fig.show()
    ^^^^^^^^
AttributeError: 'NoneType' object has no attribute 'show'
```

In [ ]:

In [19]:

```
vn.ask(question="what are the top 5 countries that customers come from?")
```

```
Number of requested results 10 is greater than number of elements in index
3, updating n_results = 3
Number of requested results 10 is greater than number of elements in index
1, updating n_results = 1
```

```
SQL Prompt: [{'role': 'system', 'content': 'You are a SQLite expert. Please
help to generate a SQL query to answer the question. Your response should ON
LY be based on the given context and follow the response guidelines and form
at instructions. \n===Tables \nCREATE TABLE "invoices"\r\n(\r\n    InvoiceId
INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    CustomerId INTEGER  NOT N
ULL,\r\n    InvoiceDate DATETIME  NOT NULL,\r\n    BillingAddress NVARCHAR(7
0),\r\n    BillingCity NVARCHAR(40),\r\n    BillingState NVARCHAR(40),\r\n
BillingCountry NVARCHAR(40),\r\n    BillingPostalCode NVARCHAR(10),\r\n    T
otal NUMERIC(10,2)  NOT NULL,\r\n    FOREIGN KEY (CustomerId) REFERENCES "cu
stomers" (CustomerId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n
\nCREATE TABLE "customers"\r\n(\r\n    CustomerId INTEGER PRIMARY KEY AUTOIN
CREMENT NOT NULL,\r\n    FirstName NVARCHAR(40)  NOT NULL,\r\n    LastName N
VARCHAR(20)  NOT NULL,\r\n    Company NVARCHAR(80),\r\n    Address NVARCHAR
(70),\r\n    City NVARCHAR(40),\r\n    State NVARCHAR(40),\r\n    Country NV
ARCHAR(40),\r\n    PostalCode NVARCHAR(10),\r\n    Phone NVARCHAR(24),\r\n
Fax NVARCHAR(24),\r\n    Email NVARCHAR(60)  NOT NULL,\r\n    SupportRepId I
NTEGER,\r\n    FOREIGN KEY (SupportRepId) REFERENCES "employees" (EmployeeI
d) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE "inv
oice_items"\r\n(\r\n    InvoiceLineId INTEGER PRIMARY KEY AUTOINCREMENT NOT
NULL,\r\n    InvoiceId INTEGER  NOT NULL,\r\n    TrackId INTEGER  NOT NUL
L,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    Quantity INTEGER  NOT NU
LL,\r\n    FOREIGN KEY (InvoiceId) REFERENCES "invoices" (InvoiceId) \r\n\t
\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackId) REFE
RENCES "tracks" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r
\n)\n\nCREATE TABLE "media_types"\r\n(\r\n    MediaTypeId INTEGER PRIMARY KE
Y AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(120)\r\n)\n\nCREATE INDEX IFK
_CustomerSupportRepId ON "customers" (SupportRepId)\n\nCREATE TABLE "employe
es"\r\n(\r\n    EmployeeId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n
LastName NVARCHAR(20)  NOT NULL,\r\n    FirstName NVARCHAR(20)  NOT NULL,\r
\n    Title NVARCHAR(30),\r\n    ReportsTo INTEGER,\r\n    BirthDate DATETIM
E,\r\n    HireDate DATETIME,\r\n    Address NVARCHAR(70),\r\n    City NVARCH
AR(40),\r\n    State NVARCHAR(40),\r\n    Country NVARCHAR(40),\r\n    Posta
lCode NVARCHAR(10),\r\n    Phone NVARCHAR(24),\r\n    Fax NVARCHAR(24),\r\n
Email NVARCHAR(60),\r\n    FOREIGN KEY (ReportsTo) REFERENCES "employees" (E
mployeeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TA
BLE "albums"\r\n(\r\n    AlbumId INTEGER PRIMARY KEY AUTOINCREMENT NOT NUL
L,\r\n    Title NVARCHAR(160)  NOT NULL,\r\n    ArtistId INTEGER  NOT NUL
L,\r\n    FOREIGN KEY (ArtistId) REFERENCES "artists" (ArtistId) \r\n\t\tON
DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE "playlist_track"\r
\n(\r\n    PlaylistId INTEGER  NOT NULL,\r\n    TrackId INTEGER  NOT NULL,\r
\n    CONSTRAINT PK_PlaylistTrack PRIMARY KEY  (PlaylistId, TrackId),\r\n
FOREIGN KEY (PlaylistId) REFERENCES "playlists" (PlaylistId) \r\n\t\tON DELE
TE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackId) REFERENCES "t
racks" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCRE
ATE TABLE sqlite_sequence(name,seq)\n\nCREATE TABLE "tracks"\r\n(\r\n    Tra
ckId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(200)
NOT NULL,\r\n    AlbumId INTEGER,\r\n    MediaTypeId INTEGER  NOT NULL,\r\n
GenreId INTEGER,\r\n    Composer NVARCHAR(220),\r\n    Milliseconds INTEGER
NOT NULL,\r\n    Bytes INTEGER,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r
\n    FOREIGN KEY (AlbumId) REFERENCES "albums" (AlbumId) \r\n\t\tON DELETE
NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (GenreId) REFERENCES "genr
es" (GenreId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIG
N KEY (MediaTypeId) REFERENCES "media_types" (MediaTypeId) \r\n\t\tON DELETE
NO ACTION ON UPDATE NO ACTION\r\n)\n\n\n===Additional Context \n\nIn the chi
nook database invoice means order\n\n===Response Guidelines \n1. If the prov
ided context is sufficient, please generate a valid SQL query without any ex
```

```
planations for the question. \n2. If the provided context is almost sufficie
nt but requires knowledge of a specific string in a particular column, pleas
e generate an intermediate SQL query to find the distinct strings in that co
lumn. Prepend the query with a comment saying intermediate_sql \n3. If the p
rovided context is insufficient, please explain why it can\'t be generated.
\n4. Please use the most relevant table(s). \n5. If the question has been as
ked and answered before, please repeat the answer exactly as it was given be
fore. \n'}, {'role': 'user', 'content': "which table stores customer's order
s"}, {'role': 'assistant', 'content': 'SELECT * FROM invoices'}, {'role': 'u
ser', 'content': 'How many customers are there'}, {'role': 'assistant', 'con
tent': 'SELECT COUNT(*) AS NumberOfCustomers FROM customers'}, {'role': 'use
r', 'content': 'Can you list all tables in the SQLite database catalog?'},
{'role': 'assistant', 'content': "SELECT name FROM sqlite_master WHERE type
= 'table'"}, {'role': 'user', 'content': 'what are the top 5 countries that
customers come from?'}]
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "You are a SQLite expert. Please help to gene
rate a SQL query to answer the question. Your response should ONLY be based
on the given context and follow the response guidelines and format instructi
ons. \n===Tables \nCREATE TABLE \"invoices\"\r\n(\r\n    InvoiceId INTEGER P
RIMARY KEY AUTOINCREMENT NOT NULL,\r\n    CustomerId INTEGER  NOT NULL,\r\n
InvoiceDate DATETIME  NOT NULL,\r\n    BillingAddress NVARCHAR(70),\r\n    B
illingCity NVARCHAR(40),\r\n    BillingState NVARCHAR(40),\r\n    BillingCou
ntry NVARCHAR(40),\r\n    BillingPostalCode NVARCHAR(10),\r\n    Total NUMER
IC(10,2)  NOT NULL,\r\n    FOREIGN KEY (CustomerId) REFERENCES \"customers\"
(CustomerId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE
TABLE \"customers\"\r\n(\r\n    CustomerId INTEGER PRIMARY KEY AUTOINCREMENT
NOT NULL,\r\n    FirstName NVARCHAR(40)  NOT NULL,\r\n    LastName NVARCHAR
(20)  NOT NULL,\r\n    Company NVARCHAR(80),\r\n    Address NVARCHAR(70),\r
\n    City NVARCHAR(40),\r\n    State NVARCHAR(40),\r\n    Country NVARCHAR
(40),\r\n    PostalCode NVARCHAR(10),\r\n    Phone NVARCHAR(24),\r\n    Fax
NVARCHAR(24),\r\n    Email NVARCHAR(60)  NOT NULL,\r\n    SupportRepId INTEG
ER,\r\n    FOREIGN KEY (SupportRepId) REFERENCES \"employees\" (EmployeeId)
\r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE \"invoi
ce_items\"\r\n(\r\n    InvoiceLineId INTEGER PRIMARY KEY AUTOINCREMENT NOT N
ULL,\r\n    InvoiceId INTEGER  NOT NULL,\r\n    TrackId INTEGER  NOT NULL,\r
\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    Quantity INTEGER  NOT NUL
L,\r\n    FOREIGN KEY (InvoiceId) REFERENCES \"invoices\" (InvoiceId) \r\n\t
\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackId) REFE
RENCES \"tracks\" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION
\r\n)\n\nCREATE TABLE \"media_types\"\r\n(\r\n    MediaTypeId INTEGER PRIMAR
Y KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(120)\r\n)\n\nCREATE INDEX
IFK_CustomerSupportRepId ON \"customers\" (SupportRepId)\n\nCREATE TABLE \"e
mployees\"\r\n(\r\n    EmployeeId INTEGER PRIMARY KEY AUTOINCREMENT NOT NUL
L,\r\n    LastName NVARCHAR(20)  NOT NULL,\r\n    FirstName NVARCHAR(20)  NO
T NULL,\r\n    Title NVARCHAR(30),\r\n    ReportsTo INTEGER,\r\n    BirthDat
e DATETIME,\r\n    HireDate DATETIME,\r\n    Address NVARCHAR(70),\r\n    Ci
ty NVARCHAR(40),\r\n    State NVARCHAR(40),\r\n    Country NVARCHAR(40),\r\n
PostalCode NVARCHAR(10),\r\n    Phone NVARCHAR(24),\r\n    Fax NVARCHAR(2
4),\r\n    Email NVARCHAR(60),\r\n    FOREIGN KEY (ReportsTo) REFERENCES \"e
mployees\" (EmployeeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)
\n\nCREATE TABLE \"albums\"\r\n(\r\n    AlbumId INTEGER PRIMARY KEY AUTOINCR
```

~~~
EMENT NOT NULL,\r\n    Title NVARCHAR(160)  NOT NULL,\r\n    ArtistId INTEGE
R  NOT NULL,\r\n    FOREIGN KEY (ArtistId) REFERENCES \"artists\" (ArtistId)
\r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE \"playl
ist_track\"\r\n(\r\n    PlaylistId INTEGER  NOT NULL,\r\n    TrackId INTEGER
NOT NULL,\r\n    CONSTRAINT PK_PlaylistTrack PRIMARY KEY  (PlaylistId, Track
Id),\r\n    FOREIGN KEY (PlaylistId) REFERENCES \"playlists\" (PlaylistId)
\r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackI
d) REFERENCES \"tracks\" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO
ACTION\r\n)\n\nCREATE TABLE sqlite_sequence(name,seq)\n\nCREATE TABLE \"trac
ks\"\r\n(\r\n    TrackId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n
Name NVARCHAR(200)  NOT NULL,\r\n    AlbumId INTEGER,\r\n    MediaTypeId INT
EGER  NOT NULL,\r\n    GenreId INTEGER,\r\n    Composer NVARCHAR(220),\r\n
Milliseconds INTEGER  NOT NULL,\r\n    Bytes INTEGER,\r\n    UnitPrice NUMER
IC(10,2)  NOT NULL,\r\n    FOREIGN KEY (AlbumId) REFERENCES \"albums\" (Albu
mId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (Ge
nreId) REFERENCES \"genres\" (GenreId) \r\n\t\tON DELETE NO ACTION ON UPDATE
NO ACTION,\r\n    FOREIGN KEY (MediaTypeId) REFERENCES \"media_types\" (Medi
aTypeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\n\n===Additi
onal Context \n\nIn the chinook database invoice means order\n\n===Response
Guidelines \n1. If the provided context is sufficient, please generate a val
id SQL query without any explanations for the question. \n2. If the provided
context is almost sufficient but requires knowledge of a specific string in
a particular column, please generate an intermediate SQL query to find the d
istinct strings in that column. Prepend the query with a comment saying inte
rmediate_sql \n3. If the provided context is insufficient, please explain wh
y it can't be generated. \n4. Please use the most relevant table(s). \n5. If
the question has been asked and answered before, please repeat the answer ex
actly as it was given before. \n"}, {"role": "user", "content": "which table
stores customer's orders"}, {"role": "assistant", "content": "SELECT * FROM
invoices"}, {"role": "user", "content": "How many customers are there"}, {"r
ole": "assistant", "content": "SELECT COUNT(*) AS NumberOfCustomers FROM cus
tomers"}, {"role": "user", "content": "Can you list all tables in the SQLite
database catalog?"}, {"role": "assistant", "content": "SELECT name FROM sqli
te_master WHERE type = 'table'"}, {"role": "user", "content": "what are the
top 5 countries that customers come from?"}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T21:37:19.513010015Z', 'messa
ge': {'role': 'assistant', 'content': 'To generate an intermediate SQL query
to find distinct strings (countries) in a particular column of the "customer
s" table:\n\nintermediate_sql - Find distinct values in the Country colum
n.\n\n```sql\nSELECT DISTINCT Country FROM customers;\n```\nNow, using this
intermediate query as part of the main query to count and sort the number of
customers from each country:\n\nmain_sql - Count customers by their countrie
s and get top 5 countries with most customers.\n\n```sql\nWITH CountryCounts
AS (\n    SELECT Country, COUNT(*) AS NumberOfCustomers FROM (\n        SELE
CT DISTINCT Country FROM customers\n    ) GROUP BY Country\n)\nSELECT Countr
y, NumberOfCustomers FROM CountryCounts ORDER BY NumberOfCustomers DESC LIMI
T 5;\n```'}, 'done_reason': 'stop', 'done': True, 'total_duration': 15737444
6996, 'load_duration': 3840778, 'prompt_eval_count': 1662, 'prompt_eval_dura
tion': 105242593000, 'eval_count': 171, 'eval_duration': 51678606000}
LLM Response: To generate an intermediate SQL query to find distinct strings
(countries) in a particular column of the "customers" table:

intermediate_sql - Find distinct values in the Country column.

```sql
~~~

~~~
SELECT DISTINCT Country FROM customers;
```
Now, using this intermediate query as part of the main query to count and so
rt the number of customers from each country:

main_sql - Count customers by their countries and get top 5 countries with m
ost customers.

```sql
WITH CountryCounts AS (
    SELECT Country, COUNT(*) AS NumberOfCustomers FROM (
        SELECT DISTINCT Country FROM customers
    ) GROUP BY Country
)
SELECT Country, NumberOfCustomers FROM CountryCounts ORDER BY NumberOfCustom
ers DESC LIMIT 5;
```
The LLM is not allowed to see the data in your database. Your question requi
res database introspection to generate the necessary SQL. Please set allow_l
lm_to_see_data=True to enable this.
Couldn't run sql:  Execution failed on sql 'The LLM is not allowed to see th
e data in your database. Your question requires database introspection to ge
nerate the necessary SQL. Please set allow_llm_to_see_data=True to enable th
is.': near "The": syntax error
~~~

More SQL questions

see `sample-sql-queries-sqlite-chinook.ipynb`

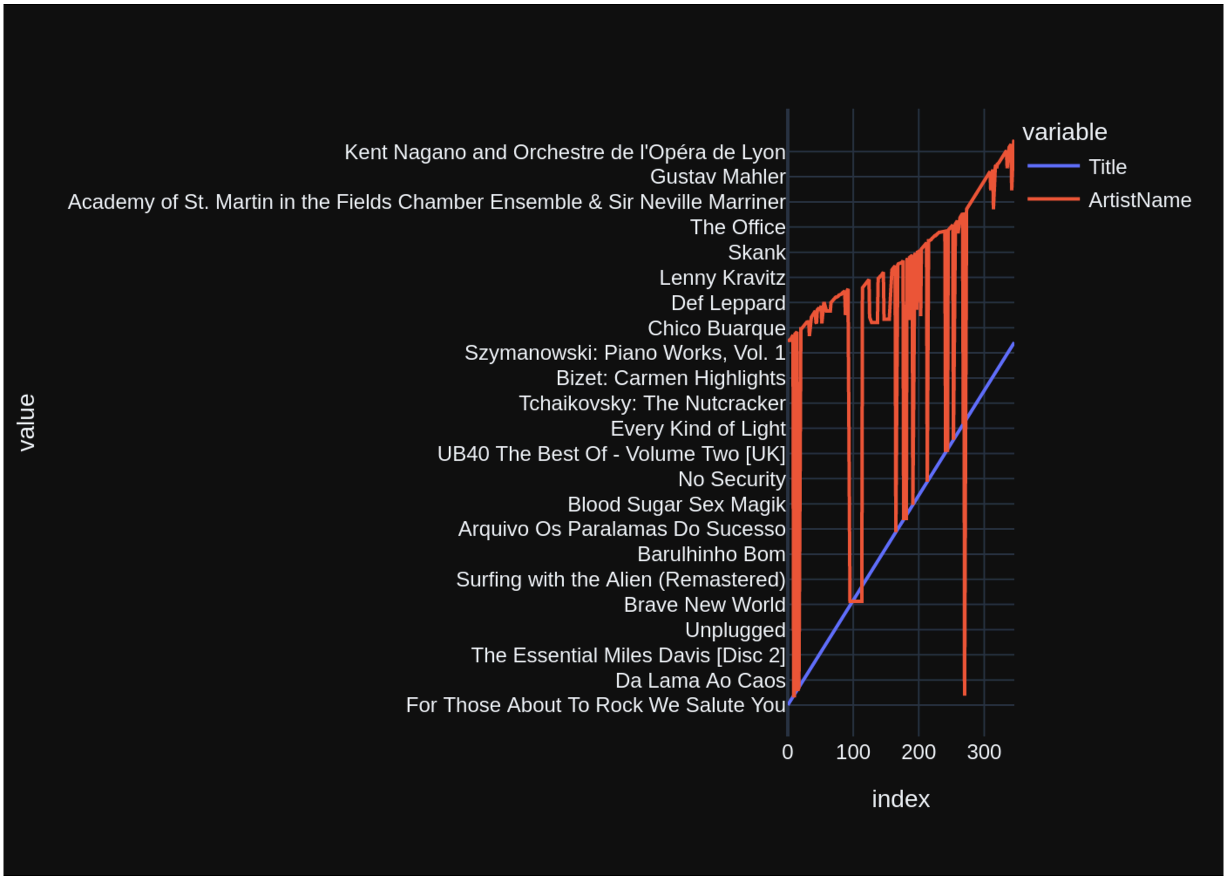
In [20]:

```
question = """
    List all albums and their corresponding artist names
"""

vn.ask(question=question)
```

```
Number of requested results 10 is greater than number of elements in index
3, updating n_results = 3
Number of requested results 10 is greater than number of elements in index
1, updating n_results = 1
```

```
SQL Prompt: [{'role': 'system', 'content': 'You are a SQLite expert. Please
help to generate a SQL query to answer the question. Your response should ON
LY be based on the given context and follow the response guidelines and form
at instructions. \n===Tables \nCREATE INDEX IFK_AlbumArtistId ON "albums" (A
rtistId)\n\nCREATE TABLE "albums"\r\n(\r\n    AlbumId INTEGER PRIMARY KEY AU
TOINCREMENT NOT NULL,\r\n    Title NVARCHAR(160)  NOT NULL,\r\n    ArtistId
INTEGER  NOT NULL,\r\n    FOREIGN KEY (ArtistId) REFERENCES "artists" (Artis
tId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE "t
racks"\r\n(\r\n    TrackId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n
Name NVARCHAR(200)  NOT NULL,\r\n    AlbumId INTEGER,\r\n    MediaTypeId INT
EGER  NOT NULL,\r\n    GenreId INTEGER,\r\n    Composer NVARCHAR(220),\r\n
Milliseconds INTEGER  NOT NULL,\r\n    Bytes INTEGER,\r\n    UnitPrice NUMER
IC(10,2)  NOT NULL,\r\n    FOREIGN KEY (AlbumId) REFERENCES "albums" (AlbumI
d) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (Genr
eId) REFERENCES "genres" (GenreId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO
ACTION,\r\n    FOREIGN KEY (MediaTypeId) REFERENCES "media_types" (MediaType
Id) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK
_TrackAlbumId ON "tracks" (AlbumId)\n\nCREATE TABLE "artists"\r\n(\r\n    Ar
tistId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(120)
\r\n)\n\nCREATE INDEX IFK_TrackGenreId ON "tracks" (GenreId)\n\nCREATE INDEX
IFK_PlaylistTrackTrackId ON "playlist_track" (TrackId)\n\nCREATE TABLE "play
lists"\r\n(\r\n    PlaylistId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r
\n    Name NVARCHAR(120)\r\n)\n\nCREATE TABLE "genres"\r\n(\r\n    GenreId I
NTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(120)\r\n)\n
\nCREATE INDEX IFK_TrackMediaTypeId ON "tracks" (MediaTypeId)\n\n\n===Additi
onal Context \n\nIn the chinook database invoice means order\n\n===Response
Guidelines \n1. If the provided context is sufficient, please generate a val
id SQL query without any explanations for the question. \n2. If the provided
context is almost sufficient but requires knowledge of a specific string in
a particular column, please generate an intermediate SQL query to find the d
istinct strings in that column. Prepend the query with a comment saying inte
rmediate_sql \n3. If the provided context is insufficient, please explain wh
y it can\'t be generated. \n4. Please use the most relevant table(s). \n5. I
f the question has been asked and answered before, please repeat the answer
exactly as it was given before. \n'}, {'role': 'user', 'content': 'Can you l
ist all tables in the SQLite database catalog?'}, {'role': 'assistant', 'con
tent': "SELECT name FROM sqlite_master WHERE type = 'table'"}, {'role': 'use
r', 'content': 'How many customers are there'}, {'role': 'assistant', 'conte
nt': 'SELECT COUNT(*) AS NumberOfCustomers FROM customers'}, {'role': 'use
r', 'content': "which table stores customer's orders"}, {'role': 'assistan
t', 'content': 'SELECT * FROM invoices'}, {'role': 'user', 'content': '  \n
List all albums and their corresponding artist names  \n'}]
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "You are a SQLite expert. Please help to gene
rate a SQL query to answer the question. Your response should ONLY be based
on the given context and follow the response guidelines and format instructi
ons. \n===Tables \nCREATE INDEX IFK_AlbumArtistId ON \"albums\" (ArtistId)\n
\nCREATE TABLE \"albums\"\r\n(\r\n    AlbumId INTEGER PRIMARY KEY AUTOINCREM
ENT NOT NULL,\r\n    Title NVARCHAR(160)  NOT NULL,\r\n    ArtistId INTEGER
NOT NULL,\r\n    FOREIGN KEY (ArtistId) REFERENCES \"artists\" (ArtistId) \r
\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE \"tracks
\"\r\n(\r\n    TrackId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Na
me NVARCHAR(200)  NOT NULL,\r\n    AlbumId INTEGER,\r\n    MediaTypeId INTEG
ER  NOT NULL,\r\n    GenreId INTEGER,\r\n    Composer NVARCHAR(220),\r\n    
Milliseconds INTEGER  NOT NULL,\r\n    Bytes INTEGER,\r\n    UnitPrice NUMER
IC(10,2)  NOT NULL,\r\n    FOREIGN KEY (AlbumId) REFERENCES \"albums\" (Albu
mId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (Ge
nreId) REFERENCES \"genres\" (GenreId) \r\n\t\tON DELETE NO ACTION ON UPDATE
NO ACTION,\r\n    FOREIGN KEY (MediaTypeId) REFERENCES \"media_types\" (Medi
aTypeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE INDE
X IFK_TrackAlbumId ON \"tracks\" (AlbumId)\n\nCREATE TABLE \"artists\"\r\n
(\r\n    ArtistId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NV
ARCHAR(120)\r\n)\n\nCREATE INDEX IFK_TrackGenreId ON \"tracks\" (GenreId)\n
\nCREATE INDEX IFK_PlaylistTrackTrackId ON \"playlist_track\" (TrackId)\n\nC
REATE TABLE \"playlists\"\r\n(\r\n    PlaylistId INTEGER PRIMARY KEY AUTOINC
REMENT NOT NULL,\r\n    Name NVARCHAR(120)\r\n)\n\nCREATE TABLE \"genres\"\r
\n(\r\n    GenreId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name N
VARCHAR(120)\r\n)\n\nCREATE INDEX IFK_TrackMediaTypeId ON \"tracks\" (MediaT
ypeId)\n\n\n===Additional Context \n\nIn the chinook database invoice means
order\n\n===Response Guidelines \n1. If the provided context is sufficient,
please generate a valid SQL query without any explanations for the question.
\n2. If the provided context is almost sufficient but requires knowledge of
a specific string in a particular column, please generate an intermediate SQ
L query to find the distinct strings in that column. Prepend the query with
a comment saying intermediate_sql \n3. If the provided context is insufficie
nt, please explain why it can't be generated. \n4. Please use the most relev
ant table(s). \n5. If the question has been asked and answered before, pleas
e repeat the answer exactly as it was given before. \n"}, {"role": "user",
"content": "Can you list all tables in the SQLite database catalog?"}, {"rol
e": "assistant", "content": "SELECT name FROM sqlite_master WHERE type = 'ta
ble'"}, {"role": "user", "content": "How many customers are there"}, {"rol
e": "assistant", "content": "SELECT COUNT(*) AS NumberOfCustomers FROM custo
mers"}, {"role": "user", "content": "which table stores customer's orders"},
{"role": "assistant", "content": "SELECT * FROM invoices"}, {"role": "user",
"content": "  \n    List all albums and their corresponding artist names
\n"}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T21:38:24.231141678Z', 'messa
ge': {'role': 'assistant', 'content': '```sql\nSELECT a.Title, ar.Name AS Ar
tistName\nFROM albums a\nJOIN artists ar ON a.ArtistId = ar.ArtistId;\n``
`'}, 'done_reason': 'stop', 'done': True, 'total_duration': 64675963328, 'lo
ad_duration': 3855346, 'prompt_eval_count': 897, 'prompt_eval_duration': 526
04359000, 'eval_count': 40, 'eval_duration': 11646860000}
LLM Response: ```sql
SELECT a.Title, ar.Name AS ArtistName
FROM albums a
JOIN artists ar ON a.ArtistId = ar.ArtistId;
```
Info: Output from LLM: ```sql
SELECT a.Title, ar.Name AS ArtistName
FROM albums a
JOIN artists ar ON a.ArtistId = ar.ArtistId;
```
Extracted SQL: SELECT a.Title, ar.Name AS ArtistName
FROM albums a
JOIN artists ar ON a.ArtistId = ar.ArtistId
SELECT a.Title, ar.Name AS ArtistName
FROM albums a
```

```
JOIN artists ar ON a.ArtistId = ar.ArtistId
                                                 Title  \
0                For Those About To Rock We Salute You   
1                                    Balls to the Wall   
2                                    Restless and Wild   
3                                    Let There Be Rock   
4                                             Big Ones   
..                                                 ...   
342                             Respighi:Pines of Rome   
343  Schubert: The Late String Quartets & String Qu...   
344                                Monteverdi: L'Orfeo   
345                              Mozart: Chamber Music   
346  Koyaanisqatsi (Soundtrack from the Motion Pict...   

                                            ArtistName  
0                                                AC/DC  
1                                               Accept  
2                                               Accept  
3                                                AC/DC  
4                                            Aerosmith  
..                                                 ...  
342                                     Eugene Ormandy  
343                             Emerson String Quartet  
344  C. Monteverdi, Nigel Rogers - Chiaroscuro; Lon...  
345                                      Nash Ensemble  
346                              Philip Glass Ensemble  

[347 rows x 2 columns]
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "The following is a pandas DataFrame that con
tains the results of the query that answers the question the user asked: '
\n    List all albums and their corresponding artist names  \n'\n\nThe DataF
rame was produced using this query: SELECT a.Title, ar.Name AS ArtistName\nF
ROM albums a\nJOIN artists ar ON a.ArtistId = ar.ArtistId\n\nThe following i
s information about the resulting pandas DataFrame 'df': \nRunning df.dtypes
gives:\n Title         object\nArtistName    object\ndtype: object"}, {"rol
e": "user", "content": "Can you generate the Python plotly code to chart the
results of the dataframe? Assume the data is in a pandas dataframe called 'd
f'. If there is only one value in the dataframe, use an Indicator. Respond w
ith only Python code. Do not answer with any explanations -- just the cod
e."}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T21:39:04.930800429Z', 'messa
ge': {'role': 'assistant', 'content': '```python\nimport plotly.express as p
x\n\n# Check if there\'s only one row or multiple rows\nif df.shape[0] ==  v
ice versa, "No data to show") # For a single value use an Indicator\n    fig
= px.indicator(df)\nelse:\n    fig = px.bar(df, x=\'ArtistName\', y=[\'Title
\'])\n    \nfig.show()\n```'}, 'done_reason': 'stop', 'done': True, 'total_d
uration': 40672077783, 'load_duration': 3543389, 'prompt_eval_count': 197,
'prompt_eval_duration': 11384762000, 'eval_count': 101, 'eval_duration': 291
50286000}
```


value
Kent Nagano and Orchestre de l'Opéra de Lyon
Gustav Mahler
Academy of St. Martin in the Fields Chamber Ensemble & Sir Neville Marriner
The Office
Skank
Lenny Kravitz
Def Leppard
Chico Buarque
Szymanowski: Piano Works, Vol. 1
Bizet: Carmen Highlights
Tchaikovsky: The Nutcracker
Every Kind of Light
UB40 The Best Of - Volume Two [UK]
No Security
Blood Sugar Sex Magik
Arquivo Os Paralamas Do Sucesso
Barulhinho Bom
Surfing with the Alien (Remastered)
Brave New World
Unplugged
The Essential Miles Davis [Disc 2]
Da Lama Ao Caos
For Those About To Rock We Salute You
0
100
200
300
index
variable
Title
ArtistName

```
Out[20]: ('SELECT a.Title, ar.Name AS ArtistName\nFROM albums a\nJOIN artists ar ON
         a.ArtistId = ar.ArtistId',
                                                         Title  \
          0               For Those About To Rock We Salute You   
          1                                   Balls to the Wall   
          2                                   Restless and Wild   
          3                                   Let There Be Rock   
          4                                            Big Ones   
          ..                                                ...   
          342                            Respighi:Pines of Rome   
          343  Schubert: The Late String Quartets & String Qu...   
          344                               Monteverdi: L'Orfeo   
          345                             Mozart: Chamber Music   
          346  Koyaanisqatsi (Soundtrack from the Motion Pict...   
         
                                                      ArtistName  
          0                                                AC/DC  
          1                                               Accept  
          2                                               Accept  
          3                                                AC/DC  
          4                                            Aerosmith  
          ..                                                 ...  
          342                                     Eugene Ormandy  
          343                             Emerson String Quartet  
          344  C. Monteverdi, Nigel Rogers - Chiaroscuro; Lon...  
          345                                      Nash Ensemble  
          346                              Philip Glass Ensemble  
         
          [347 rows x 2 columns],
          Figure({
              'data': [{'hovertemplate': 'variable=Title<br>index=%{x}<br>value=%{y}
         <extra></extra>',
                        'legendgroup': 'Title',
                        'line': {'color': '#636efa', 'dash': 'solid'},
                        'marker': {'symbol': 'circle'},
                        'mode': 'lines',
                        'name': 'Title',
                        'orientation': 'v',
                        'showlegend': True,
                        'type': 'scatter',
                        'x': array([  0,   1,   2, ..., 344, 345, 346]),
                        'xaxis': 'x',
                        'y': array(['For Those About To Rock We Salute You', 'Balls
         to the Wall',
                                    'Restless and Wild', ..., "Monteverdi: L'Orfeo",
                                    'Mozart: Chamber Music',
                                    'Koyaanisqatsi (Soundtrack from the Motion Pictu
         re)'], dtype=object),
                        'yaxis': 'y'},
                       {'hovertemplate': 'variable=ArtistName<br>index=%{x}<br>value
         =%{y}<extra></extra>',
                        'legendgroup': 'ArtistName',
                        'line': {'color': '#EF553B', 'dash': 'solid'},
                        'marker': {'symbol': 'circle'},
                        'mode': 'lines',
                        'name': 'ArtistName',
                        'orientation': 'v',
                        'showlegend': True,
                        'type': 'scatter',
                        'x': array([  0,   1,   2, ..., 344, 345, 346]),
                        'xaxis': 'x',
                        'y': array(['AC/DC', 'Accept', 'Accept', ...,
                                    'C. Monteverdi, Nigel Rogers - Chiaroscuro; Lond
         on Baroque; London Cornett & Sackbu',
                                    'Nash Ensemble', 'Philip Glass Ensemble'], dtype
         =object),
                        'yaxis': 'y'}],
              'layout': {'legend': {'title': {'text': 'variable'}, 'tracegroupgap':
         0},
                         'margin': {'t': 60},
                         'template': '...',
                         'xaxis': {'anchor': 'y', 'domain': [0.0, 1.0], 'title': {'t
         ext': 'index'}},
                         'yaxis': {'anchor': 'x', 'domain': [0.0, 1.0], 'title': {'t
         ext': 'value'}}}
          }))
```

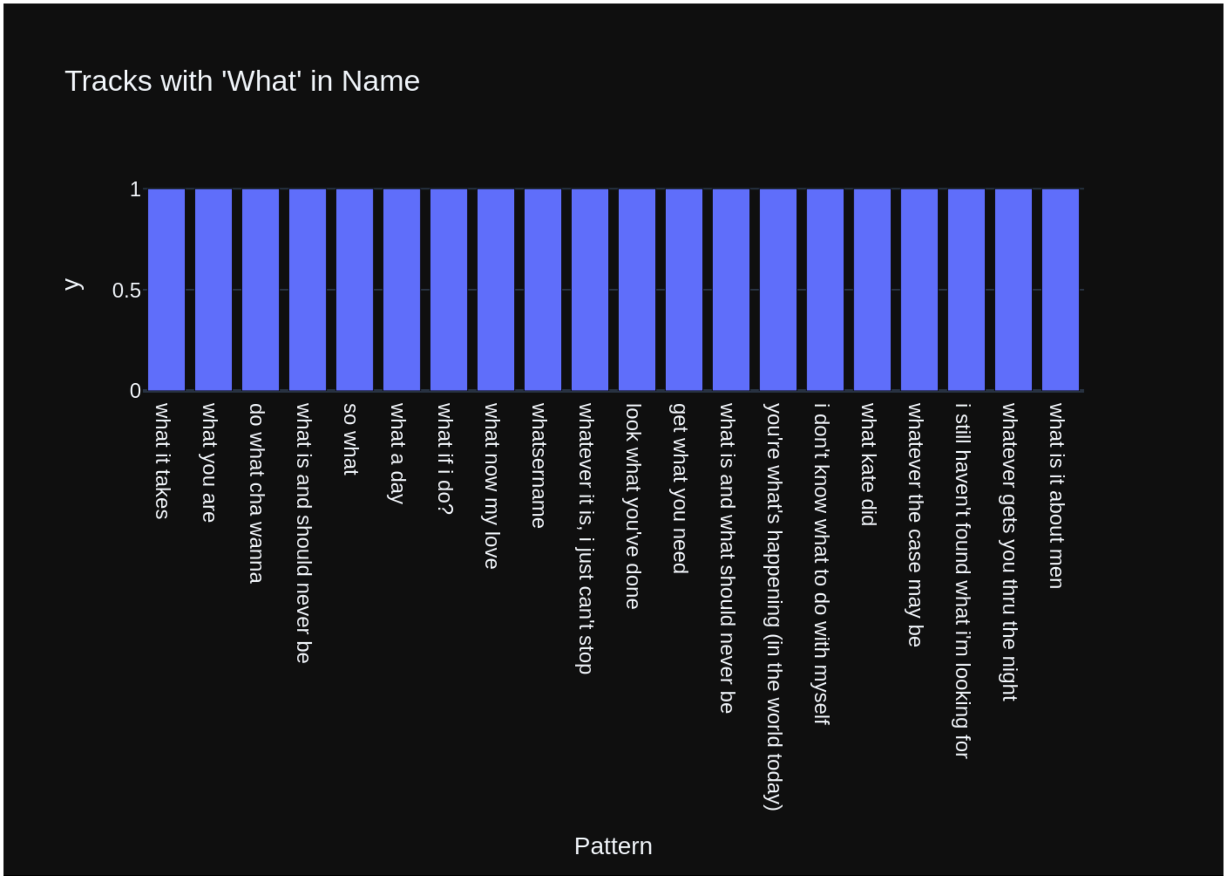
In [21]:

```
question = """
    Find all tracks with a name containing "What" (case-insensitive)
"""

vn.ask(question=question)
```

```
Number of requested results 10 is greater than number of elements in index
4, updating n_results = 4
Number of requested results 10 is greater than number of elements in index
1, updating n_results = 1
```

```
SQL Prompt: [{'role': 'system', 'content': 'You are a SQLite expert. Please
help to generate a SQL query to answer the question. Your response should ON
LY be based on the given context and follow the response guidelines and form
at instructions. \n===Tables \nCREATE INDEX IFK_TrackGenreId ON "tracks" (Ge
nreId)\n\nCREATE INDEX IFK_PlaylistTrackTrackId ON "playlist_track" (TrackI
d)\n\nCREATE TABLE "tracks"\r\n(\r\n    TrackId INTEGER PRIMARY KEY AUTOINCR
EMENT NOT NULL,\r\n    Name NVARCHAR(200)  NOT NULL,\r\n    AlbumId INTEGE
R,\r\n    MediaTypeId INTEGER  NOT NULL,\r\n    GenreId INTEGER,\r\n    Comp
oser NVARCHAR(220),\r\n    Milliseconds INTEGER  NOT NULL,\r\n    Bytes INTE
GER,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    FOREIGN KEY (AlbumId)
REFERENCES "albums" (AlbumId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTIO
N,\r\n    FOREIGN KEY (GenreId) REFERENCES "genres" (GenreId) \r\n\t\tON DEL
ETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (MediaTypeId) REFERENC
ES "media_types" (MediaTypeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTI
ON\r\n)\n\nCREATE INDEX IFK_TrackAlbumId ON "tracks" (AlbumId)\n\nCREATE IND
EX IFK_TrackMediaTypeId ON "tracks" (MediaTypeId)\n\nCREATE TABLE "playlist_
track"\r\n(\r\n    PlaylistId INTEGER  NOT NULL,\r\n    TrackId INTEGER  NOT
NULL,\r\n    CONSTRAINT PK_PlaylistTrack PRIMARY KEY  (PlaylistId, TrackI
d),\r\n    FOREIGN KEY (PlaylistId) REFERENCES "playlists" (PlaylistId) \r\n
\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackId) RE
FERENCES "tracks" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION
\r\n)\n\nCREATE INDEX IFK_InvoiceLineTrackId ON "invoice_items" (TrackId)\n
\nCREATE INDEX IFK_AlbumArtistId ON "albums" (ArtistId)\n\nCREATE TABLE "pla
ylists"\r\n(\r\n    PlaylistId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r
\n    Name NVARCHAR(120)\r\n)\n\nCREATE TABLE "genres"\r\n(\r\n    GenreId I
NTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(120)\r\n)\n
\n\n===Additional Context \n\nIn the chinook database invoice means order\n
\n===Response Guidelines \n1. If the provided context is sufficient, please
generate a valid SQL query without any explanations for the question. \n2. I
f the provided context is almost sufficient but requires knowledge of a spec
ific string in a particular column, please generate an intermediate SQL quer
y to find the distinct strings in that column. Prepend the query with a comm
ent saying intermediate_sql \n3. If the provided context is insufficient, pl
ease explain why it can\'t be generated. \n4. Please use the most relevant t
able(s). \n5. If the question has been asked and answered before, please rep
eat the answer exactly as it was given before. \n'}, {'role': 'user', 'conte
nt': '  \n    List all albums and their corresponding artist names  \n'},
{'role': 'assistant', 'content': 'SELECT a.Title, ar.Name AS ArtistName\nFRO
M albums a\nJOIN artists ar ON a.ArtistId = ar.ArtistId'}, {'role': 'user',
'content': 'Can you list all tables in the SQLite database catalog?'}, {'rol
e': 'assistant', 'content': "SELECT name FROM sqlite_master WHERE type = 'ta
ble'"}, {'role': 'user', 'content': "which table stores customer's orders"},
{'role': 'assistant', 'content': 'SELECT * FROM invoices'}, {'role': 'user',
'content': 'How many customers are there'}, {'role': 'assistant', 'content':
'SELECT COUNT(*) AS NumberOfCustomers FROM customers'}, {'role': 'user', 'co
ntent': '  \n    Find all tracks with a name containing "What" (case-insensi
tive)\n'}]
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "You are a SQLite expert. Please help to gene
rate a SQL query to answer the question. Your response should ONLY be based
on the given context and follow the response guidelines and format instructi
ons. \n===Tables \nCREATE INDEX IFK_TrackGenreId ON \"tracks\" (GenreId)\n\n
CREATE INDEX IFK_PlaylistTrackTrackId ON \"playlist_track\" (TrackId)\n\nCRE
ATE TABLE \"tracks\"\r\n(\r\n    TrackId INTEGER PRIMARY KEY AUTOINCREMENT N
OT NULL,\r\n    Name NVARCHAR(200)  NOT NULL,\r\n    AlbumId INTEGER,\r\n
MediaTypeId INTEGER  NOT NULL,\r\n    GenreId INTEGER,\r\n    Composer NVARC
HAR(220),\r\n    Milliseconds INTEGER  NOT NULL,\r\n    Bytes INTEGER,\r\n
UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    FOREIGN KEY (AlbumId) REFERENCES
\"albums\" (AlbumId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n
FOREIGN KEY (GenreId) REFERENCES \"genres\" (GenreId) \r\n\t\tON DELETE NO A
CTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (MediaTypeId) REFERENCES \"med
ia_types\" (MediaTypeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r
\n)\n\nCREATE INDEX IFK_TrackAlbumId ON \"tracks\" (AlbumId)\n\nCREATE INDEX
IFK_TrackMediaTypeId ON \"tracks\" (MediaTypeId)\n\nCREATE TABLE \"playlist_
track\"\r\n(\r\n    PlaylistId INTEGER  NOT NULL,\r\n    TrackId INTEGER  NO
T NULL,\r\n    CONSTRAINT PK_PlaylistTrack PRIMARY KEY  (PlaylistId, TrackI
d),\r\n    FOREIGN KEY (PlaylistId) REFERENCES \"playlists\" (PlaylistId) \r
\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackId)
REFERENCES \"tracks\" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACT
ION\r\n)\n\nCREATE INDEX IFK_InvoiceLineTrackId ON \"invoice_items\" (TrackI
d)\n\nCREATE INDEX IFK_AlbumArtistId ON \"albums\" (ArtistId)\n\nCREATE TABL
E \"playlists\"\r\n(\r\n    PlaylistId INTEGER PRIMARY KEY AUTOINCREMENT NOT
NULL,\r\n    Name NVARCHAR(120)\r\n)\n\nCREATE TABLE \"genres\"\r\n(\r\n
GenreId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(12
0)\r\n)\n\n\n===Additional Context \n\nIn the chinook database invoice means
order\n\n===Response Guidelines \n1. If the provided context is sufficient,
please generate a valid SQL query without any explanations for the question.
\n2. If the provided context is almost sufficient but requires knowledge of
a specific string in a particular column, please generate an intermediate SQ
L query to find the distinct strings in that column. Prepend the query with
a comment saying intermediate_sql \n3. If the provided context is insufficie
nt, please explain why it can't be generated. \n4. Please use the most relev
ant table(s). \n5. If the question has been asked and answered before, pleas
e repeat the answer exactly as it was given before. \n"}, {"role": "user",
"content": "  \n    List all albums and their corresponding artist names
\n"}, {"role": "assistant", "content": "SELECT a.Title, ar.Name AS ArtistNam
e\nFROM albums a\nJOIN artists ar ON a.ArtistId = ar.ArtistId"}, {"role": "u
ser", "content": "Can you list all tables in the SQLite database catalog?"},
{"role": "assistant", "content": "SELECT name FROM sqlite_master WHERE type
= 'table'"}, {"role": "user", "content": "which table stores customer's orde
rs"}, {"role": "assistant", "content": "SELECT * FROM invoices"}, {"role":
"user", "content": "How many customers are there"}, {"role": "assistant", "c
ontent": "SELECT COUNT(*) AS NumberOfCustomers FROM customers"}, {"role": "u
ser", "content": "  \n    Find all tracks with a name containing \"What\" (c
ase-insensitive)\n"}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T21:40:36.541197554Z', 'messa
ge': {'role': 'assistant', 'content': '-- Intermediate_sql: To find the dist
inct case-sensitive and insensitive patterns for \'what\' in track names.\nS
ELECT DISTINCT LOWER(name) AS Pattern FROM tracks WHERE lower(Name) LIKE \'%
what%\';\n\n-- Main SQL query to list all tracks with a name containing "Wha
t" (case-insensitive).\nSELECT TrackId, Name\nFROM tracks\nWHERE Name COLLAT
E NOCASE LIKE \'%what%\';'}, 'done_reason': 'stop', 'done': True, 'total_dur
ation': 91487256123, 'load_duration': 3053524, 'prompt_eval_count': 988, 'pr
ompt_eval_duration': 61613375000, 'eval_count': 99, 'eval_duration': 2929772
1000}
LLM Response: -- Intermediate_sql: To find the distinct case-sensitive and i
nsensitive patterns for 'what' in track names.
```

```
SELECT DISTINCT LOWER(name) AS Pattern FROM tracks WHERE lower(Name) LIKE '%
what%';

-- Main SQL query to list all tracks with a name containing "What" (case-ins
ensitive).
SELECT TrackId, Name
FROM tracks
WHERE Name COLLATE NOCASE LIKE '%what%';
Info: Output from LLM: -- Intermediate_sql: To find the distinct case-sensit
ive and insensitive patterns for 'what' in track names.
SELECT DISTINCT LOWER(name) AS Pattern FROM tracks WHERE lower(Name) LIKE '%
what%';

-- Main SQL query to list all tracks with a name containing "What" (case-ins
ensitive).
SELECT TrackId, Name
FROM tracks
WHERE Name COLLATE NOCASE LIKE '%what%';
Extracted SQL: SELECT DISTINCT LOWER(name) AS Pattern FROM tracks WHERE lowe
r(Name) LIKE '%what%'
SELECT DISTINCT LOWER(name) AS Pattern FROM tracks WHERE lower(Name) LIKE '%
what%'
                                         Pattern
0                                  what it takes
1                                   what you are
2                              do what cha wanna
3                    what is and should never be
4                                        so what
5                                     what a day
6                                  what if i do?
7                               what now my love
8                                    whatsername
9              whatever it is, i just can't stop
10                         look what you've done
11                             get what you need
12              what is and what should never be
13  you're what's happening (in the world today)
14           i don't know what to do with myself
15                                 what kate did
16                      whatever the case may be
17    i still haven't found what i'm looking for
18              whatever gets you thru the night
19                          what is it about men
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "The following is a pandas DataFrame that con
tains the results of the query that answers the question the user asked: '
\n    Find all tracks with a name containing \"What\" (case-insensitive)
\n'\n\nThe DataFrame was produced using this query: SELECT DISTINCT LOWER(na
me) AS Pattern FROM tracks WHERE lower(Name) LIKE '%what%'\n\nThe following
is information about the resulting pandas DataFrame 'df': \nRunning df.dtype
s gives:\n Pattern    object\ndtype: object"}, {"role": "user", "content":
"Can you generate the Python plotly code to chart the results of the datafra
me? Assume the data is in a pandas dataframe called 'df'. If there is only o
ne value in the dataframe, use an Indicator. Respond with only Python code.
Do not answer with any explanations -- just the code."}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T21:41:24.127722189Z', 'messa
ge': {'role': 'assistant', 'content': '```python\nimport plotly.express as p
x\n\n# Check if df has more than one unique value\nif len(df[\'Pattern\'].un
ique()) > 1:\n    fig = px.bar(df, x=\'Pattern\', y=df.groupby(\'Pattern\').
size(), title="Tracks with \'What\' in Name")\nelse:\n    fig = px.indicator
(data_frame=[{\'value\': df[\'Pattern\'].iloc[0]}] + [{"text": "Only one tra
ck found"}] )\n    \nfig.show()\n```'}, 'done_reason': 'stop', 'done': True,
'total_duration': 47566857411, 'load_duration': 45981487, 'prompt_eval_coun
t': 189, 'prompt_eval_duration': 11050529000, 'eval_count': 126, 'eval_durat
ion': 36379553000}
```



```
Out[21]:  ("SELECT DISTINCT LOWER(name) AS Pattern FROM tracks WHERE lower(Name) LIKE
          '%what%'",
                                                    Pattern
           0                               what it takes
           1                                what you are
           2                           do what cha wanna
           3                 what is and should never be
           4                                     so what
           5                                  what a day
           6                               what if i do?
           7                            what now my love
           8                                 whatsername
           9           whatever it is, i just can't stop
           10                      look what you've done
           11                          get what you need
           12           what is and what should never be
           13  you're what's happening (in the world today)
           14        i don't know what to do with myself
           15                              what kate did
           16                   whatever the case may be
           17    i still haven't found what i'm looking for
           18           whatever gets you thru the night
           19                       what is it about men,
           Figure({
               'data': [{'alignmentgroup': 'True',
                         'hovertemplate': 'Pattern=%{x}<br>y=%{y}<extra></extra>',
                         'legendgroup': '',
                         'marker': {'color': '#636efa', 'pattern': {'shape': ''}},
                         'name': '',
                         'offsetgroup': '',
                         'orientation': 'v',
                         'showlegend': False,
                         'textposition': 'auto',
                         'type': 'bar',
                         'x': array(['what it takes', 'what you are', 'do what cha wa
          nna',
                                     'what is and should never be', 'so what', 'what
          a day', 'what if i do?',
                                     'what now my love', 'whatsername', "whatever it
          is, i just can't stop",
                                     "look what you've done", 'get what you need',
                                     'what is and what should never be',
                                     "you're what's happening (in the world today)",
                                     "i don't know what to do with myself", 'what kat
          e did',
                                     'whatever the case may be',
                                     "i still haven't found what i'm looking for",
                                     'whatever gets you thru the night', 'what is it
          about men'],
                                    dtype=object),
                         'xaxis': 'x',
                         'y': array([1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
          1, 1, 1, 1]),
                         'yaxis': 'y'}],
               'layout': {'barmode': 'relative',
                          'legend': {'tracegroupgap': 0},
                    'template': '...',
                    'title': {'text': "Tracks with 'What' in Name"},
                    'xaxis': {'anchor': 'y', 'domain': [0.0, 1.0], 'title': {'t
ext': 'Pattern'}},
                    'yaxis': {'anchor': 'x', 'domain': [0.0, 1.0], 'title': {'t
ext': 'y'}}}
 }))
```

In [22]:

```
question = """
    Get the total number of invoices for each customer
"""

vn.ask(question=question)
```

```
Number of requested results 10 is greater than number of elements in index
5, updating n_results = 5
Number of requested results 10 is greater than number of elements in index
1, updating n_results = 1
```

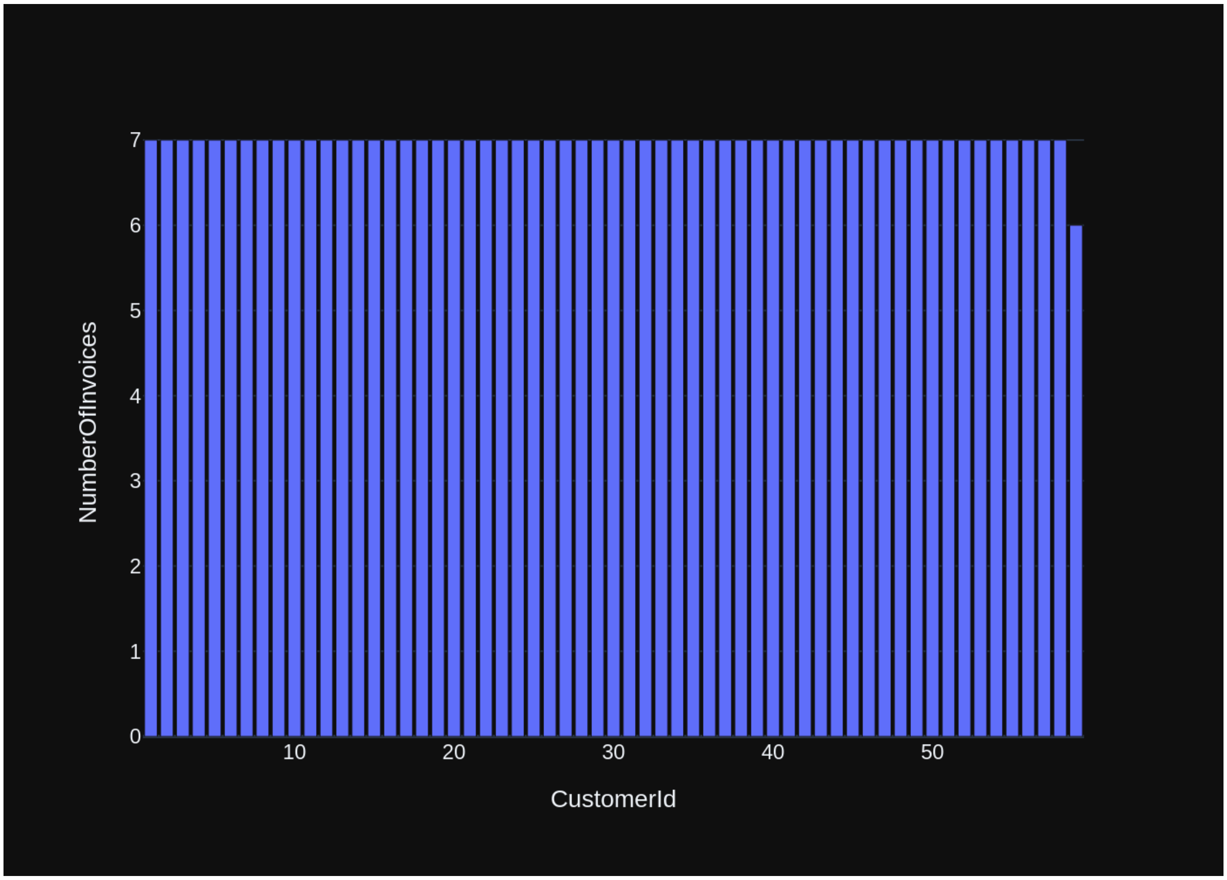
```
SQL Prompt: [{'role': 'system', 'content': 'You are a SQLite expert. Please
help to generate a SQL query to answer the question. Your response should ON
LY be based on the given context and follow the response guidelines and form
at instructions. \n===Tables \nCREATE TABLE "invoices"\r\n(\r\n    InvoiceId
INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    CustomerId INTEGER  NOT N
ULL,\r\n    InvoiceDate DATETIME  NOT NULL,\r\n    BillingAddress NVARCHAR(7
0),\r\n    BillingCity NVARCHAR(40),\r\n    BillingState NVARCHAR(40),\r\n
BillingCountry NVARCHAR(40),\r\n    BillingPostalCode NVARCHAR(10),\r\n    T
otal NUMERIC(10,2)  NOT NULL,\r\n    FOREIGN KEY (CustomerId) REFERENCES "cu
stomers" (CustomerId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n
\nCREATE INDEX IFK_InvoiceCustomerId ON "invoices" (CustomerId)\n\nCREATE IN
DEX IFK_InvoiceLineInvoiceId ON "invoice_items" (InvoiceId)\n\nCREATE TABLE
"invoice_items"\r\n(\r\n    InvoiceLineId INTEGER PRIMARY KEY AUTOINCREMENT
NOT NULL,\r\n    InvoiceId INTEGER  NOT NULL,\r\n    TrackId INTEGER  NOT NU
LL,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    Quantity INTEGER  NOT N
ULL,\r\n    FOREIGN KEY (InvoiceId) REFERENCES "invoices" (InvoiceId) \r\n\t
\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackId) REFE
RENCES "tracks" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r
\n)\n\nCREATE INDEX IFK_InvoiceLineTrackId ON "invoice_items" (TrackId)\n\nC
REATE TABLE "customers"\r\n(\r\n    CustomerId INTEGER PRIMARY KEY AUTOINCRE
MENT NOT NULL,\r\n    FirstName NVARCHAR(40)  NOT NULL,\r\n    LastName NVAR
CHAR(20)  NOT NULL,\r\n    Company NVARCHAR(80),\r\n    Address NVARCHAR(7
0),\r\n    City NVARCHAR(40),\r\n    State NVARCHAR(40),\r\n    Country NVAR
CHAR(40),\r\n    PostalCode NVARCHAR(10),\r\n    Phone NVARCHAR(24),\r\n
Fax NVARCHAR(24),\r\n    Email NVARCHAR(60)  NOT NULL,\r\n    SupportRepId I
NTEGER,\r\n    FOREIGN KEY (SupportRepId) REFERENCES "employees" (EmployeeI
d) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_
CustomerSupportRepId ON "customers" (SupportRepId)\n\nCREATE TABLE "employee
s"\r\n(\r\n    EmployeeId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n
LastName NVARCHAR(20)  NOT NULL,\r\n    FirstName NVARCHAR(20)  NOT NULL,\r
\n    Title NVARCHAR(30),\r\n    ReportsTo INTEGER,\r\n    BirthDate DATETIM
E,\r\n    HireDate DATETIME,\r\n    Address NVARCHAR(70),\r\n    City NVARCH
AR(40),\r\n    State NVARCHAR(40),\r\n    Country NVARCHAR(40),\r\n    Posta
lCode NVARCHAR(10),\r\n    Phone NVARCHAR(24),\r\n    Fax NVARCHAR(24),\r\n
Email NVARCHAR(60),\r\n    FOREIGN KEY (ReportsTo) REFERENCES "employees" (E
mployeeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE IN
DEX IFK_EmployeeReportsTo ON "employees" (ReportsTo)\n\nCREATE TABLE "track
s"\r\n(\r\n    TrackId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Na
me NVARCHAR(200)  NOT NULL,\r\n    AlbumId INTEGER,\r\n    MediaTypeId INTEG
ER  NOT NULL,\r\n    GenreId INTEGER,\r\n    Composer NVARCHAR(220),\r\n
Milliseconds INTEGER  NOT NULL,\r\n    Bytes INTEGER,\r\n    UnitPrice NUMER
IC(10,2)  NOT NULL,\r\n    FOREIGN KEY (AlbumId) REFERENCES "albums" (AlbumI
d) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (Genr
eId) REFERENCES "genres" (GenreId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO
ACTION,\r\n    FOREIGN KEY (MediaTypeId) REFERENCES "media_types" (MediaType
Id) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\n\n===Additional
Context \n\nIn the chinook database invoice means order\n\n===Response Guide
lines \n1. If the provided context is sufficient, please generate a valid SQ
L query without any explanations for the question. \n2. If the provided cont
ext is almost sufficient but requires knowledge of a specific string in a pa
rticular column, please generate an intermediate SQL query to find the disti
nct strings in that column. Prepend the query with a comment saying intermed
iate_sql \n3. If the provided context is insufficient, please explain why it
can\'t be generated. \n4. Please use the most relevant table(s). \n5. If the
question has been asked and answered before, please repeat the answer exactl
y as it was given before. \n'}, {'role': 'user', 'content': "which table sto
```

```
res customer's orders"}, {'role': 'assistant', 'content': 'SELECT * FROM inv
oices'}, {'role': 'user', 'content': 'How many customers are there'}, {'rol
e': 'assistant', 'content': 'SELECT COUNT(*) AS NumberOfCustomers FROM custo
mers'}, {'role': 'user', 'content': '  \n    List all albums and their corre
sponding artist names  \n'}, {'role': 'assistant', 'content': 'SELECT a.Titl
e, ar.Name AS ArtistName\nFROM albums a\nJOIN artists ar ON a.ArtistId = ar.
ArtistId'}, {'role': 'user', 'content': '  \n    Find all tracks with a name
containing "What" (case-insensitive)\n'}, {'role': 'assistant', 'content':
"SELECT DISTINCT LOWER(name) AS Pattern FROM tracks WHERE lower(Name) LIKE
'%what%'"}, {'role': 'user', 'content': 'Can you list all tables in the SQLi
te database catalog?'}, {'role': 'assistant', 'content': "SELECT name FROM s
qlite_master WHERE type = 'table'"}, {'role': 'user', 'content': '  \n    Ge
t the total number of invoices for each customer\n'}]
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "You are a SQLite expert. Please help to gene
rate a SQL query to answer the question. Your response should ONLY be based
on the given context and follow the response guidelines and format instructi
ons. \n===Tables \nCREATE TABLE \"invoices\"\r\n(\r\n    InvoiceId INTEGER P
RIMARY KEY AUTOINCREMENT NOT NULL,\r\n    CustomerId INTEGER  NOT NULL,\r\n
InvoiceDate DATETIME  NOT NULL,\r\n    BillingAddress NVARCHAR(70),\r\n    B
illingCity NVARCHAR(40),\r\n    BillingState NVARCHAR(40),\r\n    BillingCou
ntry NVARCHAR(40),\r\n    BillingPostalCode NVARCHAR(10),\r\n    Total NUMER
IC(10,2)  NOT NULL,\r\n    FOREIGN KEY (CustomerId) REFERENCES \"customers\"
(CustomerId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE
INDEX IFK_InvoiceCustomerId ON \"invoices\" (CustomerId)\n\nCREATE INDEX IFK
_InvoiceLineInvoiceId ON \"invoice_items\" (InvoiceId)\n\nCREATE TABLE \"inv
oice_items\"\r\n(\r\n    InvoiceLineId INTEGER PRIMARY KEY AUTOINCREMENT NOT
NULL,\r\n    InvoiceId INTEGER  NOT NULL,\r\n    TrackId INTEGER  NOT NUL
L,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    Quantity INTEGER  NOT NU
LL,\r\n    FOREIGN KEY (InvoiceId) REFERENCES \"invoices\" (InvoiceId) \r\n
\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackId) RE
FERENCES \"tracks\" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTIO
N\r\n)\n\nCREATE INDEX IFK_InvoiceLineTrackId ON \"invoice_items\" (TrackId)
\n\nCREATE TABLE \"customers\"\r\n(\r\n    CustomerId INTEGER PRIMARY KEY AU
TOINCREMENT NOT NULL,\r\n    FirstName NVARCHAR(40)  NOT NULL,\r\n    LastNa
me NVARCHAR(20)  NOT NULL,\r\n    Company NVARCHAR(80),\r\n    Address NVARC
HAR(70),\r\n    City NVARCHAR(40),\r\n    State NVARCHAR(40),\r\n    Country
NVARCHAR(40),\r\n    PostalCode NVARCHAR(10),\r\n    Phone NVARCHAR(24),\r\n
Fax NVARCHAR(24),\r\n    Email NVARCHAR(60)  NOT NULL,\r\n    SupportRepId I
NTEGER,\r\n    FOREIGN KEY (SupportRepId) REFERENCES \"employees\" (Employee
Id) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK
_CustomerSupportRepId ON \"customers\" (SupportRepId)\n\nCREATE TABLE \"empl
oyees\"\r\n(\r\n    EmployeeId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r
\n    LastName NVARCHAR(20)  NOT NULL,\r\n    FirstName NVARCHAR(20)  NOT NU
LL,\r\n    Title NVARCHAR(30),\r\n    ReportsTo INTEGER,\r\n    BirthDate DA
TETIME,\r\n    HireDate DATETIME,\r\n    Address NVARCHAR(70),\r\n    City N
VARCHAR(40),\r\n    State NVARCHAR(40),\r\n    Country NVARCHAR(40),\r\n
PostalCode NVARCHAR(10),\r\n    Phone NVARCHAR(24),\r\n    Fax NVARCHAR(2
4),\r\n    Email NVARCHAR(60),\r\n    FOREIGN KEY (ReportsTo) REFERENCES \"e
mployees\" (EmployeeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)
\n\nCREATE INDEX IFK_EmployeeReportsTo ON \"employees\" (ReportsTo)\n\nCREAT
E TABLE \"tracks\"\r\n(\r\n    TrackId INTEGER PRIMARY KEY AUTOINCREMENT NOT
```

~~~
NULL,\r\n    Name NVARCHAR(200)  NOT NULL,\r\n    AlbumId INTEGER,\r\n    Me
diaTypeId INTEGER  NOT NULL,\r\n    GenreId INTEGER,\r\n    Composer NVARCHA
R(220),\r\n    Milliseconds INTEGER  NOT NULL,\r\n    Bytes INTEGER,\r\n    
UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    FOREIGN KEY (AlbumId) REFERENCES 
\"albums\" (AlbumId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    
FOREIGN KEY (GenreId) REFERENCES \"genres\" (GenreId) \r\n\t\tON DELETE NO A
CTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (MediaTypeId) REFERENCES \"med
ia_types\" (MediaTypeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r
\n)\n\n\n===Additional Context \n\nIn the chinook database invoice means ord
er\n\n===Response Guidelines \n1. If the provided context is sufficient, ple
ase generate a valid SQL query without any explanations for the question. \n
2. If the provided context is almost sufficient but requires knowledge of a 
specific string in a particular column, please generate an intermediate SQL 
query to find the distinct strings in that column. Prepend the query with a 
comment saying intermediate_sql \n3. If the provided context is insufficien
t, please explain why it can't be generated. \n4. Please use the most releva
nt table(s). \n5. If the question has been asked and answered before, please
repeat the answer exactly as it was given before. \n"}, {"role": "user", "co
ntent": "which table stores customer's orders"}, {"role": "assistant", "cont
ent": "SELECT * FROM invoices"}, {"role": "user", "content": "How many custo
mers are there"}, {"role": "assistant", "content": "SELECT COUNT(*) AS Numbe
rOfCustomers FROM customers"}, {"role": "user", "content": "  \n    List all
albums and their corresponding artist names  \n"}, {"role": "assistant", "co
ntent": "SELECT a.Title, ar.Name AS ArtistName\nFROM albums a\nJOIN artists 
ar ON a.ArtistId = ar.ArtistId"}, {"role": "user", "content": "  \n    Find 
all tracks with a name containing \"What\" (case-insensitive)\n"}, {"role": 
"assistant", "content": "SELECT DISTINCT LOWER(name) AS Pattern FROM tracks 
WHERE lower(Name) LIKE '%what%'"}, {"role": "user", "content": "Can you list
all tables in the SQLite database catalog?"}, {"role": "assistant", "conten
t": "SELECT name FROM sqlite_master WHERE type = 'table'"}, {"role": "user",
"content": "  \n    Get the total number of invoices for each customer\n"}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T21:43:39.888719806Z', 'messa
ge': {'role': 'assistant', 'content': "To get the total number of invoices f
or each customer, you can perform an aggregation on the `invoices` table gro
uped by the `CustomerId`. The SQL query would look like this:\n\n```sql\nSEL
ECT CustomerId, COUNT(*) AS NumberOfInvoices\nFROM invoices\nGROUP BY Custom
erId;\n```\nThis will return a result set with two columns - 'CustomerId' an
d 'NumberOfInvoices', where each row represents a unique customer and the co
rresponding count of their respective invoices."}, 'done_reason': 'stop', 'd
one': True, 'total_duration': 135620929487, 'load_duration': 2994027, 'promp
t_eval_count': 1578, 'prompt_eval_duration': 100742595000, 'eval_count': 11
4, 'eval_duration': 34202360000}
LLM Response: To get the total number of invoices for each customer, you can
perform an aggregation on the `invoices` table grouped by the `CustomerId`.
The SQL query would look like this:

```sql
SELECT CustomerId, COUNT(*) AS NumberOfInvoices
FROM invoices
GROUP BY CustomerId;
```
This will return a result set with two columns - 'CustomerId' and 'NumberOfI
nvoices', where each row represents a unique customer and the corresponding 
count of their respective invoices.
Info: Output from LLM: To get the total number of invoices for each custome
~~~

~~~
r, you can perform an aggregation on the `invoices` table grouped by the `Cu
stomerId`. The SQL query would look like this:

```sql
SELECT CustomerId, COUNT(*) AS NumberOfInvoices
FROM invoices
GROUP BY CustomerId;
```
This will return a result set with two columns - 'CustomerId' and 'NumberOfI
nvoices', where each row represents a unique customer and the corresponding
count of their respective invoices.
Extracted SQL: SELECT CustomerId, COUNT(*) AS NumberOfInvoices
FROM invoices
GROUP BY CustomerId
SELECT CustomerId, COUNT(*) AS NumberOfInvoices
FROM invoices
GROUP BY CustomerId
    CustomerId  NumberOfInvoices
0            1                 7
1            2                 7
2            3                 7
3            4                 7
4            5                 7
5            6                 7
6            7                 7
7            8                 7
8            9                 7
9           10                 7
10          11                 7
11          12                 7
12          13                 7
13          14                 7
14          15                 7
15          16                 7
16          17                 7
17          18                 7
18          19                 7
19          20                 7
20          21                 7
21          22                 7
22          23                 7
23          24                 7
24          25                 7
25          26                 7
26          27                 7
27          28                 7
28          29                 7
29          30                 7
30          31                 7
31          32                 7
32          33                 7
33          34                 7
34          35                 7
35          36                 7
36          37                 7
37          38                 7
~~~

```
38          39                 7
39          40                 7
40          41                 7
41          42                 7
42          43                 7
43          44                 7
44          45                 7
45          46                 7
46          47                 7
47          48                 7
48          49                 7
49          50                 7
50          51                 7
51          52                 7
52          53                 7
53          54                 7
54          55                 7
55          56                 7
56          57                 7
57          58                 7
58          59                 6
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "The following is a pandas DataFrame that con
tains the results of the query that answers the question the user asked: '
\n    Get the total number of invoices for each customer\n'\n\nThe DataFrame
was produced using this query: SELECT CustomerId, COUNT(*) AS NumberOfInvoic
es\nFROM invoices\nGROUP BY CustomerId\n\nThe following is information about
the resulting pandas DataFrame 'df': \nRunning df.dtypes gives:\n CustomerId
int64\nNumberOfInvoices    int64\ndtype: object"}, {"role": "user", "conten
t": "Can you generate the Python plotly code to chart the results of the dat
aframe? Assume the data is in a pandas dataframe called 'df'. If there is on
ly one value in the dataframe, use an Indicator. Respond with only Python co
de. Do not answer with any explanations -- just the code."}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T21:44:14.436194466Z', 'messa
ge': {'role': 'assistant', 'content': "```python\nimport plotly.express as p
x\n\nif df['NumberOfInvoices'].nunique() > 1:\n    fig = px.bar(df, x='Custo
merId', y='NumberOfInvoices')\nelse:\n    fig = px.indicator(data_frame=df)
\nfig.show()\n```"}, 'done_reason': 'stop', 'done': True, 'total_duration':
34520389249, 'load_duration': 51108033, 'prompt_eval_count': 195, 'prompt_ev
al_duration': 11264678000, 'eval_count': 80, 'eval_duration': 23113977000}
```


NumberOfInvoices
0
1
2
3
4
5
6
7
10
20
30
40
50
CustomerId

```
Out[22]: ('SELECT CustomerId, COUNT(*) AS NumberOfInvoices\nFROM invoices\nGROUP BY
         CustomerId',
              CustomerId  NumberOfInvoices
          0            1                 7
          1            2                 7
          2            3                 7
          3            4                 7
          4            5                 7
          5            6                 7
          6            7                 7
          7            8                 7
          8            9                 7
          9           10                 7
          10          11                 7
          11          12                 7
          12          13                 7
          13          14                 7
          14          15                 7
          15          16                 7
          16          17                 7
          17          18                 7
          18          19                 7
          19          20                 7
          20          21                 7
          21          22                 7
          22          23                 7
          23          24                 7
          24          25                 7
          25          26                 7
          26          27                 7
          27          28                 7
          28          29                 7
          29          30                 7
          30          31                 7
          31          32                 7
          32          33                 7
          33          34                 7
          34          35                 7
          35          36                 7
          36          37                 7
          37          38                 7
          38          39                 7
          39          40                 7
          40          41                 7
          41          42                 7
          42          43                 7
          43          44                 7
          44          45                 7
          45          46                 7
          46          47                 7
          47          48                 7
          48          49                 7
          49          50                 7
          50          51                 7
          51          52                 7
          52          53                 7
```

```
 53          54                  7
 54          55                  7
 55          56                  7
 56          57                  7
 57          58                  7
 58          59                  6,
 Figure({
     'data': [{'alignmentgroup': 'True',
               'hovertemplate': 'CustomerId=%{x}<br>NumberOfInvoices=%{y}<e
xtra></extra>',
               'legendgroup': '',
               'marker': {'color': '#636efa', 'pattern': {'shape': ''}},
               'name': '',
               'offsetgroup': '',
               'orientation': 'v',
               'showlegend': False,
               'textposition': 'auto',
               'type': 'bar',
               'x': array([ 1,  2,  3,  4,  5,  6,  7,  8,  9, 10, 11, 12,
13, 14, 15, 16, 17, 18,
                           19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30,
31, 32, 33, 34, 35, 36,
                           37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48,
49, 50, 51, 52, 53, 54,
                           55, 56, 57, 58, 59]),
               'xaxis': 'x',
               'y': array([7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7,
7, 7, 7, 7, 7, 7, 7, 7,
                           7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7,
7, 7, 7, 7, 7, 7, 7, 7,
                           7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 6]),
               'yaxis': 'y'}],
     'layout': {'barmode': 'relative',
                'legend': {'tracegroupgap': 0},
                'margin': {'t': 60},
                'template': '...',
                'xaxis': {'anchor': 'y', 'domain': [0.0, 1.0], 'title': {'t
ext': 'CustomerId'}},
                'yaxis': {'anchor': 'x', 'domain': [0.0, 1.0], 'title': {'t
ext': 'NumberOfInvoices'}}}
 }))
```

In [23]:

```
question = """
    Find the total number of invoices per country:
"""

vn.ask(question=question)
```

```
Number of requested results 10 is greater than number of elements in index
6, updating n_results = 6
Number of requested results 10 is greater than number of elements in index
1, updating n_results = 1
```

```
SQL Prompt: [{'role': 'system', 'content': 'You are a SQLite expert. Please
help to generate a SQL query to answer the question. Your response should ON
LY be based on the given context and follow the response guidelines and form
at instructions. \n===Tables \nCREATE TABLE "invoices"\r\n(\r\n    InvoiceId
INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    CustomerId INTEGER  NOT N
ULL,\r\n    InvoiceDate DATETIME  NOT NULL,\r\n    BillingAddress NVARCHAR(7
0),\r\n    BillingCity NVARCHAR(40),\r\n    BillingState NVARCHAR(40),\r\n
BillingCountry NVARCHAR(40),\r\n    BillingPostalCode NVARCHAR(10),\r\n    T
otal NUMERIC(10,2)  NOT NULL,\r\n    FOREIGN KEY (CustomerId) REFERENCES "cu
stomers" (CustomerId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n
\nCREATE TABLE "invoice_items"\r\n(\r\n    InvoiceLineId INTEGER PRIMARY KEY
AUTOINCREMENT NOT NULL,\r\n    InvoiceId INTEGER  NOT NULL,\r\n    TrackId I
NTEGER  NOT NULL,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    Quantity
INTEGER  NOT NULL,\r\n    FOREIGN KEY (InvoiceId) REFERENCES "invoices" (Inv
oiceId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY
(TrackId) REFERENCES "tracks" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDAT
E NO ACTION\r\n)\n\nCREATE INDEX IFK_InvoiceCustomerId ON "invoices" (Custom
erId)\n\nCREATE INDEX IFK_InvoiceLineInvoiceId ON "invoice_items" (InvoiceI
d)\n\nCREATE INDEX IFK_InvoiceLineTrackId ON "invoice_items" (TrackId)\n\nCR
EATE TABLE "employees"\r\n(\r\n    EmployeeId INTEGER PRIMARY KEY AUTOINCREM
ENT NOT NULL,\r\n    LastName NVARCHAR(20)  NOT NULL,\r\n    FirstName NVARC
HAR(20)  NOT NULL,\r\n    Title NVARCHAR(30),\r\n    ReportsTo INTEGER,\r\n
BirthDate DATETIME,\r\n    HireDate DATETIME,\r\n    Address NVARCHAR(70),\r
\n    City NVARCHAR(40),\r\n    State NVARCHAR(40),\r\n    Country NVARCHAR
(40),\r\n    PostalCode NVARCHAR(10),\r\n    Phone NVARCHAR(24),\r\n    Fax
NVARCHAR(24),\r\n    Email NVARCHAR(60),\r\n    FOREIGN KEY (ReportsTo) REFE
RENCES "employees" (EmployeeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACT
ION\r\n)\n\nCREATE TABLE "customers"\r\n(\r\n    CustomerId INTEGER PRIMARY
KEY AUTOINCREMENT NOT NULL,\r\n    FirstName NVARCHAR(40)  NOT NULL,\r\n
LastName NVARCHAR(20)  NOT NULL,\r\n    Company NVARCHAR(80),\r\n    Address
NVARCHAR(70),\r\n    City NVARCHAR(40),\r\n    State NVARCHAR(40),\r\n    Co
untry NVARCHAR(40),\r\n    PostalCode NVARCHAR(10),\r\n    Phone NVARCHAR(2
4),\r\n    Fax NVARCHAR(24),\r\n    Email NVARCHAR(60)  NOT NULL,\r\n    Sup
portRepId INTEGER,\r\n    FOREIGN KEY (SupportRepId) REFERENCES "employees"
(EmployeeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE
TABLE "albums"\r\n(\r\n    AlbumId INTEGER PRIMARY KEY AUTOINCREMENT NOT NUL
L,\r\n    Title NVARCHAR(160)  NOT NULL,\r\n    ArtistId INTEGER  NOT NUL
L,\r\n    FOREIGN KEY (ArtistId) REFERENCES "artists" (ArtistId) \r\n\t\tON
DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE "tracks"\r\n(\r\n
TrackId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(20
0)  NOT NULL,\r\n    AlbumId INTEGER,\r\n    MediaTypeId INTEGER  NOT NUL
L,\r\n    GenreId INTEGER,\r\n    Composer NVARCHAR(220),\r\n    Millisecond
s INTEGER  NOT NULL,\r\n    Bytes INTEGER,\r\n    UnitPrice NUMERIC(10,2)  N
OT NULL,\r\n    FOREIGN KEY (AlbumId) REFERENCES "albums" (AlbumId) \r\n\t\t
ON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (GenreId) REFERE
NCES "genres" (GenreId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n
FOREIGN KEY (MediaTypeId) REFERENCES "media_types" (MediaTypeId) \r\n\t\tON
DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_EmployeeReport
sTo ON "employees" (ReportsTo)\n\n\n===Additional Context \n\nIn the chinook
database invoice means order\n\n===Response Guidelines \n1. If the provided
context is sufficient, please generate a valid SQL query without any explana
tions for the question. \n2. If the provided context is almost sufficient bu
t requires knowledge of a specific string in a particular column, please gen
erate an intermediate SQL query to find the distinct strings in that column.
Prepend the query with a comment saying intermediate_sql \n3. If the provide
d context is insufficient, please explain why it can\'t be generated. \n4. P
```

```
lease use the most relevant table(s). \n5. If the question has been asked an
d answered before, please repeat the answer exactly as it was given before.
\n'}, {'role': 'user', 'content': '  \n    Get the total number of invoices
for each customer\n'}, {'role': 'assistant', 'content': 'SELECT CustomerId,
COUNT(*) AS NumberOfInvoices\nFROM invoices\nGROUP BY CustomerId'}, {'role':
'user', 'content': "which table stores customer's orders"}, {'role': 'assist
ant', 'content': 'SELECT * FROM invoices'}, {'role': 'user', 'content': 'How
many customers are there'}, {'role': 'assistant', 'content': 'SELECT COUNT
(*) AS NumberOfCustomers FROM customers'}, {'role': 'user', 'content': '  \n
List all albums and their corresponding artist names  \n'}, {'role': 'assist
ant', 'content': 'SELECT a.Title, ar.Name AS ArtistName\nFROM albums a\nJOIN
artists ar ON a.ArtistId = ar.ArtistId'}, {'role': 'user', 'content': '  \n
Find all tracks with a name containing "What" (case-insensitive)\n'}, {'rol
e': 'assistant', 'content': "SELECT DISTINCT LOWER(name) AS Pattern FROM tra
cks WHERE lower(Name) LIKE '%what%'"}, {'role': 'user', 'content': 'Can you
list all tables in the SQLite database catalog?'}, {'role': 'assistant', 'co
ntent': "SELECT name FROM sqlite_master WHERE type = 'table'"}, {'role': 'us
er', 'content': '  \n    Find the total number of invoices per country:\n'}]
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "You are a SQLite expert. Please help to gene
rate a SQL query to answer the question. Your response should ONLY be based
on the given context and follow the response guidelines and format instructi
ons. \n===Tables \nCREATE TABLE \"invoices\"\r\n(\r\n    InvoiceId INTEGER P
RIMARY KEY AUTOINCREMENT NOT NULL,\r\n    CustomerId INTEGER  NOT NULL,\r\n
InvoiceDate DATETIME  NOT NULL,\r\n    BillingAddress NVARCHAR(70),\r\n    B
illingCity NVARCHAR(40),\r\n    BillingState NVARCHAR(40),\r\n    BillingCou
ntry NVARCHAR(40),\r\n    BillingPostalCode NVARCHAR(10),\r\n    Total NUMER
IC(10,2)  NOT NULL,\r\n    FOREIGN KEY (CustomerId) REFERENCES \"customers\"
(CustomerId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE
TABLE \"invoice_items\"\r\n(\r\n    InvoiceLineId INTEGER PRIMARY KEY AUTOIN
CREMENT NOT NULL,\r\n    InvoiceId INTEGER  NOT NULL,\r\n    TrackId INTEGER
NOT NULL,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    Quantity INTEGER
NOT NULL,\r\n    FOREIGN KEY (InvoiceId) REFERENCES \"invoices\" (InvoiceId)
\r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackI
d) REFERENCES \"tracks\" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO
ACTION\r\n)\n\nCREATE INDEX IFK_InvoiceCustomerId ON \"invoices\" (CustomerI
d)\n\nCREATE INDEX IFK_InvoiceLineInvoiceId ON \"invoice_items\" (InvoiceId)
\n\nCREATE INDEX IFK_InvoiceLineTrackId ON \"invoice_items\" (TrackId)\n\nCR
EATE TABLE \"employees\"\r\n(\r\n    EmployeeId INTEGER PRIMARY KEY AUTOINCR
EMENT NOT NULL,\r\n    LastName NVARCHAR(20)  NOT NULL,\r\n    FirstName NVA
RCHAR(20)  NOT NULL,\r\n    Title NVARCHAR(30),\r\n    ReportsTo INTEGER,\r
\n    BirthDate DATETIME,\r\n    HireDate DATETIME,\r\n    Address NVARCHAR
(70),\r\n    City NVARCHAR(40),\r\n    State NVARCHAR(40),\r\n    Country NV
ARCHAR(40),\r\n    PostalCode NVARCHAR(10),\r\n    Phone NVARCHAR(24),\r\n
Fax NVARCHAR(24),\r\n    Email NVARCHAR(60),\r\n    FOREIGN KEY (ReportsTo)
REFERENCES \"employees\" (EmployeeId) \r\n\t\tON DELETE NO ACTION ON UPDATE
NO ACTION\r\n)\n\nCREATE TABLE \"customers\"\r\n(\r\n    CustomerId INTEGER
PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    FirstName NVARCHAR(40)  NOT NUL
L,\r\n    LastName NVARCHAR(20)  NOT NULL,\r\n    Company NVARCHAR(80),\r\n
Address NVARCHAR(70),\r\n    City NVARCHAR(40),\r\n    State NVARCHAR(40),\r
\n    Country NVARCHAR(40),\r\n    PostalCode NVARCHAR(10),\r\n    Phone NVA
RCHAR(24),\r\n    Fax NVARCHAR(24),\r\n    Email NVARCHAR(60)  NOT NULL,\r\n
SupportRepId INTEGER,\r\n    FOREIGN KEY (SupportRepId) REFERENCES \"employe
es\" (EmployeeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCR
EATE TABLE \"albums\"\r\n(\r\n    AlbumId INTEGER PRIMARY KEY AUTOINCREMENT 
NOT NULL,\r\n    Title NVARCHAR(160)  NOT NULL,\r\n    ArtistId INTEGER  NOT
NULL,\r\n    FOREIGN KEY (ArtistId) REFERENCES \"artists\" (ArtistId) \r\n\t
\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE \"tracks\"\r
\n(\r\n    TrackId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name N
VARCHAR(200)  NOT NULL,\r\n    AlbumId INTEGER,\r\n    MediaTypeId INTEGER 
NOT NULL,\r\n    GenreId INTEGER,\r\n    Composer NVARCHAR(220),\r\n    Mill
iseconds INTEGER  NOT NULL,\r\n    Bytes INTEGER,\r\n    UnitPrice NUMERIC(1
0,2)  NOT NULL,\r\n    FOREIGN KEY (AlbumId) REFERENCES \"albums\" (AlbumId)
\r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (GenreI
d) REFERENCES \"genres\" (GenreId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO 
ACTION,\r\n    FOREIGN KEY (MediaTypeId) REFERENCES \"media_types\" (MediaTy
peId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX I
FK_EmployeeReportsTo ON \"employees\" (ReportsTo)\n\n\n===Additional Context
\n\nIn the chinook database invoice means order\n\n===Response Guidelines \n
1. If the provided context is sufficient, please generate a valid SQL query 
without any explanations for the question. \n2. If the provided context is a
lmost sufficient but requires knowledge of a specific string in a particular
column, please generate an intermediate SQL query to find the distinct strin
gs in that column. Prepend the query with a comment saying intermediate_sql 
\n3. If the provided context is insufficient, please explain why it can't be
generated. \n4. Please use the most relevant table(s). \n5. If the question 
has been asked and answered before, please repeat the answer exactly as it w
as given before. \n"}, {"role": "user", "content": "  \n    Get the total nu
mber of invoices for each customer\n"}, {"role": "assistant", "content": "SE
LECT CustomerId, COUNT(*) AS NumberOfInvoices\nFROM invoices\nGROUP BY Custo
merId"}, {"role": "user", "content": "which table stores customer's order
s"}, {"role": "assistant", "content": "SELECT * FROM invoices"}, {"role": "u
ser", "content": "How many customers are there"}, {"role": "assistant", "con
tent": "SELECT COUNT(*) AS NumberOfCustomers FROM customers"}, {"role": "use
r", "content": "  \n    List all albums and their corresponding artist names
\n"}, {"role": "assistant", "content": "SELECT a.Title, ar.Name AS ArtistNam
e\nFROM albums a\nJOIN artists ar ON a.ArtistId = ar.ArtistId"}, {"role": "u
ser", "content": "  \n    Find all tracks with a name containing \"What\" (c
ase-insensitive)\n"}, {"role": "assistant", "content": "SELECT DISTINCT LOWE
R(name) AS Pattern FROM tracks WHERE lower(Name) LIKE '%what%'"}, {"role": 
"user", "content": "Can you list all tables in the SQLite database catalo
g?"}, {"role": "assistant", "content": "SELECT name FROM sqlite_master WHERE
type = 'table'"}, {"role": "user", "content": "  \n    Find the total number
of invoices per country:\n"}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T21:46:16.122537048Z', 'messa
ge': {'role': 'assistant', 'content': 'SELECT b.Country, COUNT(*) AS NumberO
fInvoices\nFROM invoices i\nJOIN customers c ON i.CustomerId = c.CustomerId
\nGROUP BY b.Country'}, 'done_reason': 'stop', 'done': True, 'total_duratio
n': 121591877069, 'load_duration': 3202556, 'prompt_eval_count': 1700, 'prom
pt_eval_duration': 108517056000, 'eval_count': 41, 'eval_duration': 12237290
000}
LLM Response: SELECT b.Country, COUNT(*) AS NumberOfInvoices
FROM invoices i
JOIN customers c ON i.CustomerId = c.CustomerId
GROUP BY b.Country
SELECT b.Country, COUNT(*) AS NumberOfInvoices
FROM invoices i
```

```
JOIN customers c ON i.CustomerId = c.CustomerId
GROUP BY b.Country
Couldn't run sql:  Execution failed on sql 'SELECT b.Country, COUNT(*) AS Nu
mberOfInvoices
FROM invoices i
JOIN customers c ON i.CustomerId = c.CustomerId
GROUP BY b.Country': no such column: b.Country
```

In [24]:

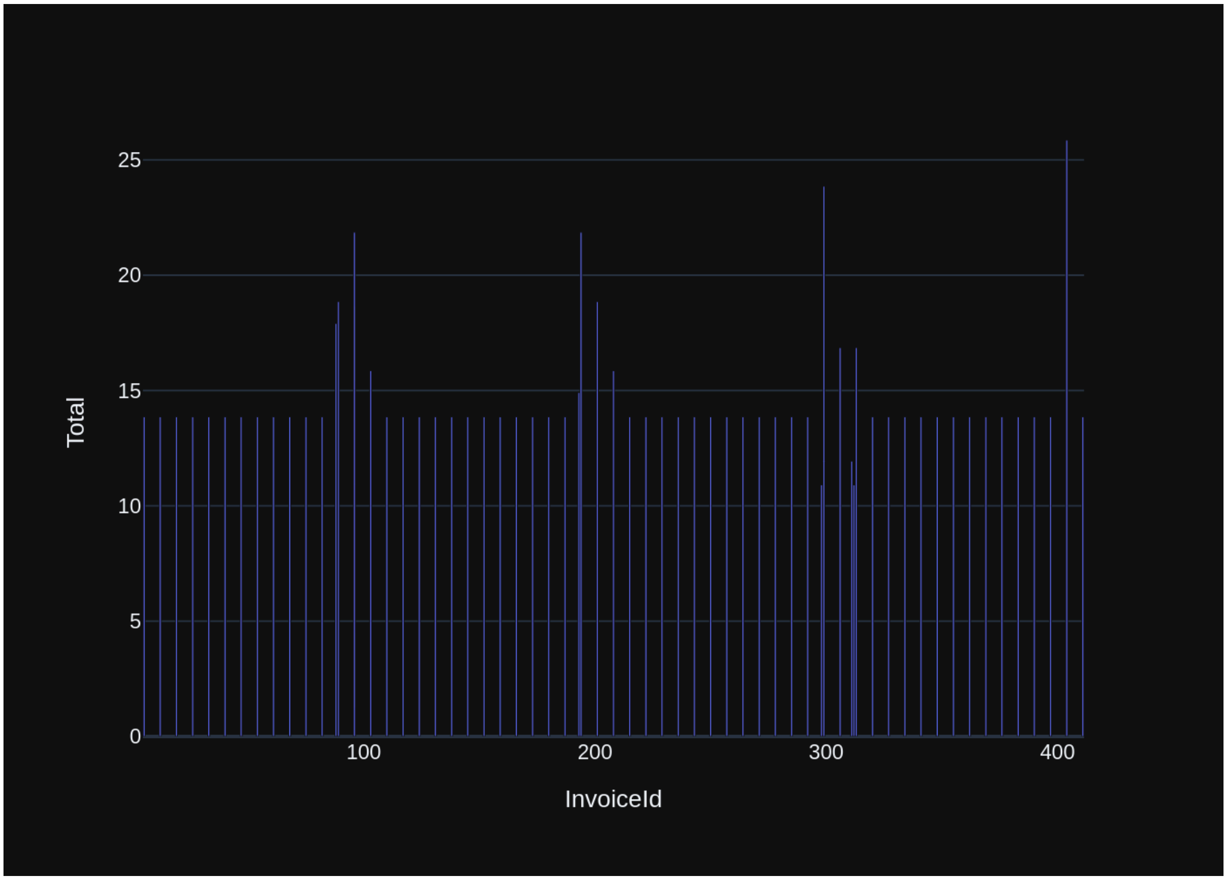
```
question = """
    List all invoices with a total exceeding $10:
"""

vn.ask(question=question)
```

```
Number of requested results 10 is greater than number of elements in index
6, updating n_results = 6
Number of requested results 10 is greater than number of elements in index
1, updating n_results = 1
```

```
SQL Prompt: [{'role': 'system', 'content': 'You are a SQLite expert. Please
help to generate a SQL query to answer the question. Your response should ON
LY be based on the given context and follow the response guidelines and form
at instructions. \n===Tables \nCREATE TABLE "invoice_items"\r\n(\r\n    Invo
iceLineId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    InvoiceId INTEG
ER  NOT NULL,\r\n    TrackId INTEGER  NOT NULL,\r\n    UnitPrice NUMERIC(10,
2)  NOT NULL,\r\n    Quantity INTEGER  NOT NULL,\r\n    FOREIGN KEY (Invoice
Id) REFERENCES "invoices" (InvoiceId) \r\n\t\tON DELETE NO ACTION ON UPDATE
NO ACTION,\r\n    FOREIGN KEY (TrackId) REFERENCES "tracks" (TrackId) \r\n\t
\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_InvoiceLi
neInvoiceId ON "invoice_items" (InvoiceId)\n\nCREATE TABLE "invoices"\r\n(\r
\n    InvoiceId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    CustomerI
d INTEGER  NOT NULL,\r\n    InvoiceDate DATETIME  NOT NULL,\r\n    BillingAd
dress NVARCHAR(70),\r\n    BillingCity NVARCHAR(40),\r\n    BillingState NVA
RCHAR(40),\r\n    BillingCountry NVARCHAR(40),\r\n    BillingPostalCode NVAR
CHAR(10),\r\n    Total NUMERIC(10,2)  NOT NULL,\r\n    FOREIGN KEY (Customer
Id) REFERENCES "customers" (CustomerId) \r\n\t\tON DELETE NO ACTION ON UPDAT
E NO ACTION\r\n)\n\nCREATE INDEX IFK_InvoiceLineTrackId ON "invoice_items"
(TrackId)\n\nCREATE INDEX IFK_InvoiceCustomerId ON "invoices" (CustomerId)\n
\nCREATE TABLE "tracks"\r\n(\r\n    TrackId INTEGER PRIMARY KEY AUTOINCREMEN
T NOT NULL,\r\n    Name NVARCHAR(200)  NOT NULL,\r\n    AlbumId INTEGER,\r\n
MediaTypeId INTEGER  NOT NULL,\r\n    GenreId INTEGER,\r\n    Composer NVARC
HAR(220),\r\n    Milliseconds INTEGER  NOT NULL,\r\n    Bytes INTEGER,\r\n
UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    FOREIGN KEY (AlbumId) REFERENCES
"albums" (AlbumId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    F
OREIGN KEY (GenreId) REFERENCES "genres" (GenreId) \r\n\t\tON DELETE NO ACTI
ON ON UPDATE NO ACTION,\r\n    FOREIGN KEY (MediaTypeId) REFERENCES "media_t
ypes" (MediaTypeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\n
CREATE INDEX IFK_EmployeeReportsTo ON "employees" (ReportsTo)\n\nCREATE TABL
E "customers"\r\n(\r\n    CustomerId INTEGER PRIMARY KEY AUTOINCREMENT NOT N
ULL,\r\n    FirstName NVARCHAR(40)  NOT NULL,\r\n    LastName NVARCHAR(20)
NOT NULL,\r\n    Company NVARCHAR(80),\r\n    Address NVARCHAR(70),\r\n    C
ity NVARCHAR(40),\r\n    State NVARCHAR(40),\r\n    Country NVARCHAR(40),\r
\n    PostalCode NVARCHAR(10),\r\n    Phone NVARCHAR(24),\r\n    Fax NVARCHA
R(24),\r\n    Email NVARCHAR(60)  NOT NULL,\r\n    SupportRepId INTEGER,\r\n
FOREIGN KEY (SupportRepId) REFERENCES "employees" (EmployeeId) \r\n\t\tON DE
LETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE "employees"\r\n(\r\n
EmployeeId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    LastName NVARC
HAR(20)  NOT NULL,\r\n    FirstName NVARCHAR(20)  NOT NULL,\r\n    Title NVA
RCHAR(30),\r\n    ReportsTo INTEGER,\r\n    BirthDate DATETIME,\r\n    HireD
ate DATETIME,\r\n    Address NVARCHAR(70),\r\n    City NVARCHAR(40),\r\n
State NVARCHAR(40),\r\n    Country NVARCHAR(40),\r\n    PostalCode NVARCHAR
(10),\r\n    Phone NVARCHAR(24),\r\n    Fax NVARCHAR(24),\r\n    Email NVARC
HAR(60),\r\n    FOREIGN KEY (ReportsTo) REFERENCES "employees" (EmployeeId)
\r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_Cus
tomerSupportRepId ON "customers" (SupportRepId)\n\n\n===Additional Context
\n\nIn the chinook database invoice means order\n\n===Response Guidelines \n
1. If the provided context is sufficient, please generate a valid SQL query
without any explanations for the question. \n2. If the provided context is a
lmost sufficient but requires knowledge of a specific string in a particular
column, please generate an intermediate SQL query to find the distinct strin
gs in that column. Prepend the query with a comment saying intermediate_sql
\n3. If the provided context is insufficient, please explain why it can\'t b
e generated. \n4. Please use the most relevant table(s). \n5. If the questio
n has been asked and answered before, please repeat the answer exactly as it
was given before. \n'}, {'role': 'user', 'content': '  \n    Get the total n
```

```
umber of invoices for each customer\n'}, {'role': 'assistant', 'content': 'S
ELECT CustomerId, COUNT(*) AS NumberOfInvoices\nFROM invoices\nGROUP BY Cust
omerId'}, {'role': 'user', 'content': "which table stores customer's order
s"}, {'role': 'assistant', 'content': 'SELECT * FROM invoices'}, {'role': 'u
ser', 'content': 'How many customers are there'}, {'role': 'assistant', 'con
tent': 'SELECT COUNT(*) AS NumberOfCustomers FROM customers'}, {'role': 'use
r', 'content': '  \n    List all albums and their corresponding artist names
\n'}, {'role': 'assistant', 'content': 'SELECT a.Title, ar.Name AS ArtistNam
e\nFROM albums a\nJOIN artists ar ON a.ArtistId = ar.ArtistId'}, {'role': 'u
ser', 'content': '  \n    Find all tracks with a name containing "What" (cas
e-insensitive)\n'}, {'role': 'assistant', 'content': "SELECT DISTINCT LOWER
(name) AS Pattern FROM tracks WHERE lower(Name) LIKE '%what%'"}, {'role': 'u
ser', 'content': 'Can you list all tables in the SQLite database catalog?'},
{'role': 'assistant', 'content': "SELECT name FROM sqlite_master WHERE type
= 'table'"}, {'role': 'user', 'content': '  \n    List all invoices with a t
otal exceeding $10:\n'}]
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "You are a SQLite expert. Please help to gene
rate a SQL query to answer the question. Your response should ONLY be based
on the given context and follow the response guidelines and format instructi
ons. \n===Tables \nCREATE TABLE \"invoice_items\"\r\n(\r\n    InvoiceLineId
INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    InvoiceId INTEGER  NOT NU
LL,\r\n    TrackId INTEGER  NOT NULL,\r\n    UnitPrice NUMERIC(10,2)  NOT NU
LL,\r\n    Quantity INTEGER  NOT NULL,\r\n    FOREIGN KEY (InvoiceId) REFERE
NCES \"invoices\" (InvoiceId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTIO
N,\r\n    FOREIGN KEY (TrackId) REFERENCES \"tracks\" (TrackId) \r\n\t\tON D
ELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_InvoiceLineInvo
iceId ON \"invoice_items\" (InvoiceId)\n\nCREATE TABLE \"invoices\"\r\n(\r\n
InvoiceId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    CustomerId INTE
GER  NOT NULL,\r\n    InvoiceDate DATETIME  NOT NULL,\r\n    BillingAddress
NVARCHAR(70),\r\n    BillingCity NVARCHAR(40),\r\n    BillingState NVARCHAR
(40),\r\n    BillingCountry NVARCHAR(40),\r\n    BillingPostalCode NVARCHAR
(10),\r\n    Total NUMERIC(10,2)  NOT NULL,\r\n    FOREIGN KEY (CustomerId)
REFERENCES \"customers\" (CustomerId) \r\n\t\tON DELETE NO ACTION ON UPDATE
NO ACTION\r\n)\n\nCREATE INDEX IFK_InvoiceLineTrackId ON \"invoice_items\"
(TrackId)\n\nCREATE INDEX IFK_InvoiceCustomerId ON \"invoices\" (CustomerId)
\n\nCREATE TABLE \"tracks\"\r\n(\r\n    TrackId INTEGER PRIMARY KEY AUTOINCR
EMENT NOT NULL,\r\n    Name NVARCHAR(200)  NOT NULL,\r\n    AlbumId INTEGE
R,\r\n    MediaTypeId INTEGER  NOT NULL,\r\n    GenreId INTEGER,\r\n    Comp
oser NVARCHAR(220),\r\n    Milliseconds INTEGER  NOT NULL,\r\n    Bytes INTE
GER,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    FOREIGN KEY (AlbumId)
REFERENCES \"albums\" (AlbumId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACT
ION,\r\n    FOREIGN KEY (GenreId) REFERENCES \"genres\" (GenreId) \r\n\t\tON
DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (MediaTypeId) REFER
ENCES \"media_types\" (MediaTypeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO
ACTION\r\n)\n\nCREATE INDEX IFK_EmployeeReportsTo ON \"employees\" (ReportsT
o)\n\nCREATE TABLE \"customers\"\r\n(\r\n    CustomerId INTEGER PRIMARY KEY
AUTOINCREMENT NOT NULL,\r\n    FirstName NVARCHAR(40)  NOT NULL,\r\n    Last
Name NVARCHAR(20)  NOT NULL,\r\n    Company NVARCHAR(80),\r\n    Address NVA
RCHAR(70),\r\n    City NVARCHAR(40),\r\n    State NVARCHAR(40),\r\n    Count
ry NVARCHAR(40),\r\n    PostalCode NVARCHAR(10),\r\n    Phone NVARCHAR(2
4),\r\n    Fax NVARCHAR(24),\r\n    Email NVARCHAR(60)  NOT NULL,\r\n    Sup
portRepId INTEGER,\r\n    FOREIGN KEY (SupportRepId) REFERENCES \"employees
\" (EmployeeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREA
TE TABLE \"employees\"\r\n(\r\n    EmployeeId INTEGER PRIMARY KEY AUTOINCREM
ENT NOT NULL,\r\n    LastName NVARCHAR(20)  NOT NULL,\r\n    FirstName NVARC
HAR(20)  NOT NULL,\r\n    Title NVARCHAR(30),\r\n    ReportsTo INTEGER,\r\n
BirthDate DATETIME,\r\n    HireDate DATETIME,\r\n    Address NVARCHAR(70),\r
\n    City NVARCHAR(40),\r\n    State NVARCHAR(40),\r\n    Country NVARCHAR
(40),\r\n    PostalCode NVARCHAR(10),\r\n    Phone NVARCHAR(24),\r\n    Fax
NVARCHAR(24),\r\n    Email NVARCHAR(60),\r\n    FOREIGN KEY (ReportsTo) REFE
RENCES \"employees\" (EmployeeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO A
CTION\r\n)\n\nCREATE INDEX IFK_CustomerSupportRepId ON \"customers\" (Suppor
tRepId)\n\n\n===Additional Context \n\nIn the chinook database invoice means
order\n\n===Response Guidelines \n1. If the provided context is sufficient,
please generate a valid SQL query without any explanations for the question.
\n2. If the provided context is almost sufficient but requires knowledge of
a specific string in a particular column, please generate an intermediate SQ
L query to find the distinct strings in that column. Prepend the query with
a comment saying intermediate_sql \n3. If the provided context is insufficie
nt, please explain why it can't be generated. \n4. Please use the most relev
ant table(s). \n5. If the question has been asked and answered before, pleas
e repeat the answer exactly as it was given before. \n"}, {"role": "user",
"content": "  \n    Get the total number of invoices for each customer\n"},
{"role": "assistant", "content": "SELECT CustomerId, COUNT(*) AS NumberOfInv
oices\nFROM invoices\nGROUP BY CustomerId"}, {"role": "user", "content": "wh
ich table stores customer's orders"}, {"role": "assistant", "content": "SELE
CT * FROM invoices"}, {"role": "user", "content": "How many customers are th
ere"}, {"role": "assistant", "content": "SELECT COUNT(*) AS NumberOfCustomer
s FROM customers"}, {"role": "user", "content": "  \n    List all albums and
their corresponding artist names  \n"}, {"role": "assistant", "content": "SE
LECT a.Title, ar.Name AS ArtistName\nFROM albums a\nJOIN artists ar ON a.Art
istId = ar.ArtistId"}, {"role": "user", "content": "  \n    Find all tracks
with a name containing \"What\" (case-insensitive)\n"}, {"role": "assistan
t", "content": "SELECT DISTINCT LOWER(name) AS Pattern FROM tracks WHERE low
er(Name) LIKE '%what%'"}, {"role": "user", "content": "Can you list all tabl
es in the SQLite database catalog?"}, {"role": "assistant", "content": "SELE
CT name FROM sqlite_master WHERE type = 'table'"}, {"role": "user", "conten
t": "  \n    List all invoices with a total exceeding $10:\n"}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T21:48:01.845900169Z', 'messa
ge': {'role': 'assistant', 'content': 'SELECT * FROM invoices WHERE Total >
10.0;'}, 'done_reason': 'stop', 'done': True, 'total_duration': 10567991999
7, 'load_duration': 3537115, 'prompt_eval_count': 1628, 'prompt_eval_duratio
n': 100325690000, 'eval_count': 16, 'eval_duration': 4512850000}
LLM Response: SELECT * FROM invoices WHERE Total > 10.0;
Info: Output from LLM: SELECT * FROM invoices WHERE Total > 10.0;
Extracted SQL: SELECT * FROM invoices WHERE Total > 10.0
SELECT * FROM invoices WHERE Total > 10.0
    InvoiceId  CustomerId          InvoiceDate             BillingAddress  \
0           5          23  2009-01-11 00:00:00            69 Salem Street
1          12           2  2009-02-11 00:00:00    Theodor-Heuss-Straße 34
2          19          40  2009-03-14 00:00:00             8, Rue Hanovre
3          26          19  2009-04-14 00:00:00            1 Infinite Loop
4          33          57  2009-05-15 00:00:00            Calle Lira, 198
..        ...         ...                  ...                        ...
59        383          10  2013-08-12 00:00:00  Rua Dr. Falcão Filho, 155
60        390          48  2013-09-12 00:00:00      Lijnbaansgracht 120bg
```

```
61        397          27  2013-10-13 00:00:00            1033 N Park Ave
62        404           6  2013-11-13 00:00:00              Rilská 3174/6
63        411          44  2013-12-14 00:00:00            Porthaninkatu 9

   BillingCity BillingState  BillingCountry BillingPostalCode  Total
0       Boston           MA             USA              2113  13.86
1    Stuttgart         None         Germany             70174  13.86
2        Paris         None          France             75002  13.86
3    Cupertino           CA             USA             95014  13.86
4     Santiago         None           Chile              None  13.86
..         ...          ...             ...               ...    ...
59   São Paulo           SP          Brazil         01007-010  13.86
60   Amsterdam           VV     Netherlands              1016  13.86
61      Tucson           AZ             USA             85719  13.86
62      Prague         None  Czech Republic             14300  25.86
63    Helsinki         None         Finland             00530  13.86

[64 rows x 9 columns]
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "The following is a pandas DataFrame that con
tains the results of the query that answers the question the user asked: '
\n    List all invoices with a total exceeding $10:\n'\n\nThe DataFrame was
produced using this query: SELECT * FROM invoices WHERE Total > 10.0\n\nThe
following is information about the resulting pandas DataFrame 'df': \nRunnin
g df.dtypes gives:\n InvoiceId              int64\nCustomerId             in
t64\nInvoiceDate            object\nBillingAddress        object\nBillingCity
object\nBillingState          object\nBillingCountry        object\nBillingP
ostalCode     object\nTotal                float64\ndtype: object"}, {"rol
e": "user", "content": "Can you generate the Python plotly code to chart the
results of the dataframe? Assume the data is in a pandas dataframe called 'd
f'. If there is only one value in the dataframe, use an Indicator. Respond w
ith only Python code. Do not answer with any explanations -- just the cod
e."}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T21:48:44.27763644Z', 'messag
e': {'role': 'assistant', 'content': '```python\nimport plotly.express as px
\n\nif len(set(df[\'Total\'])) > 1:\n    fig = px.bar(df, x=\'InvoiceId\', y
=\'Total\')\nelse:\n    fig = px.indicator(data_frame=[{"value": df["Tota
l"].values[0], "intervals": [{\'stop\': len(df)}]}])\n    \nfig.show()\n``
`'}, 'done_reason': 'stop', 'done': True, 'total_duration': 42403167350, 'lo
ad_duration': 2987779, 'prompt_eval_count': 232, 'prompt_eval_duration': 136
94174000, 'eval_count': 99, 'eval_duration': 28618800000}
```


Total
25
20
15
10
5
0
100
200
300
400
InvoiceId

```
Out[24]: ('SELECT * FROM invoices WHERE Total > 10.0',
              InvoiceId  CustomerId          InvoiceDate             BillingAddress
         \
          0           5          23  2009-01-11 00:00:00            69 Salem Street
          1          12           2  2009-02-11 00:00:00    Theodor-Heuss-Straße 34
          2          19          40  2009-03-14 00:00:00             8, Rue Hanovre
          3          26          19  2009-04-14 00:00:00            1 Infinite Loop
          4          33          57  2009-05-15 00:00:00            Calle Lira, 198
          ..        ...         ...                  ...                        ...
          59        383          10  2013-08-12 00:00:00  Rua Dr. Falcão Filho, 155
          60        390          48  2013-09-12 00:00:00     Lijnbaansgracht 120bg
          61        397          27  2013-10-13 00:00:00            1033 N Park Ave
          62        404           6  2013-11-13 00:00:00              Rilská 3174/6
          63        411          44  2013-12-14 00:00:00            Porthaninkatu 9

             BillingCity BillingState  BillingCountry BillingPostalCode  Total
          0       Boston           MA             USA              2113  13.86
          1    Stuttgart         None         Germany             70174  13.86
          2        Paris         None          France             75002  13.86
          3    Cupertino           CA             USA             95014  13.86
          4     Santiago         None           Chile              None  13.86
          ..         ...          ...             ...               ...    ...
          59   São Paulo           SP          Brazil         01007-010  13.86
          60   Amsterdam           VV     Netherlands              1016  13.86
          61      Tucson           AZ             USA             85719  13.86
          62      Prague         None  Czech Republic             14300  25.86
          63    Helsinki         None         Finland             00530  13.86

          [64 rows x 9 columns],
          Figure({
              'data': [{'alignmentgroup': 'True',
                        'hovertemplate': 'InvoiceId=%{x}<br>Total=%{y}<extra></extra
         >',
                        'legendgroup': '',
                        'marker': {'color': '#636efa', 'pattern': {'shape': ''}},
                        'name': '',
                        'offsetgroup': '',
                        'orientation': 'v',
                        'showlegend': False,
                        'textposition': 'auto',
                        'type': 'bar',
                        'x': array([  5,  12,  19,  26,  33,  40,  47,  54,  61,  6
         8,  75,  82,  88,  89,
                                     96, 103, 110, 117, 124, 131, 138, 145, 152, 15
         9, 166, 173, 180, 187,
                                    193, 194, 201, 208, 215, 222, 229, 236, 243, 25
         0, 257, 264, 271, 278,
                                    285, 292, 298, 299, 306, 311, 312, 313, 320, 32
         7, 334, 341, 348, 355,
                                    362, 369, 376, 383, 390, 397, 404, 411]),
                        'xaxis': 'x',
                        'y': array([13.86, 13.86, 13.86, 13.86, 13.86, 13.86, 13.86,
         13.86, 13.86, 13.86,
                                    13.86, 13.86, 17.91, 18.86, 21.86, 15.86, 13.86,
         13.86, 13.86, 13.86,
                                    13.86, 13.86, 13.86, 13.86, 13.86, 13.86, 13.86,
         13.86, 14.91, 21.86,
                                    18.86, 15.86, 13.86, 13.86, 13.86, 13.86, 13.86,
         13.86, 13.86, 13.86,
                                    13.86, 13.86, 13.86, 13.86, 10.91, 23.86, 16.86,
         11.94, 10.91, 16.86,
                                    13.86, 13.86, 13.86, 13.86, 13.86, 13.86, 13.86,
         13.86, 13.86, 13.86,
                                    13.86, 13.86, 25.86, 13.86]),
                        'yaxis': 'y'}],
              'layout': {'barmode': 'relative',
                         'legend': {'tracegroupgap': 0},
                         'margin': {'t': 60},
                         'template': '...',
                         'xaxis': {'anchor': 'y', 'domain': [0.0, 1.0], 'title': {'t
         ext': 'InvoiceId'}},
                         'yaxis': {'anchor': 'x', 'domain': [0.0, 1.0], 'title': {'t
         ext': 'Total'}}}
          }))
```

In [25]:

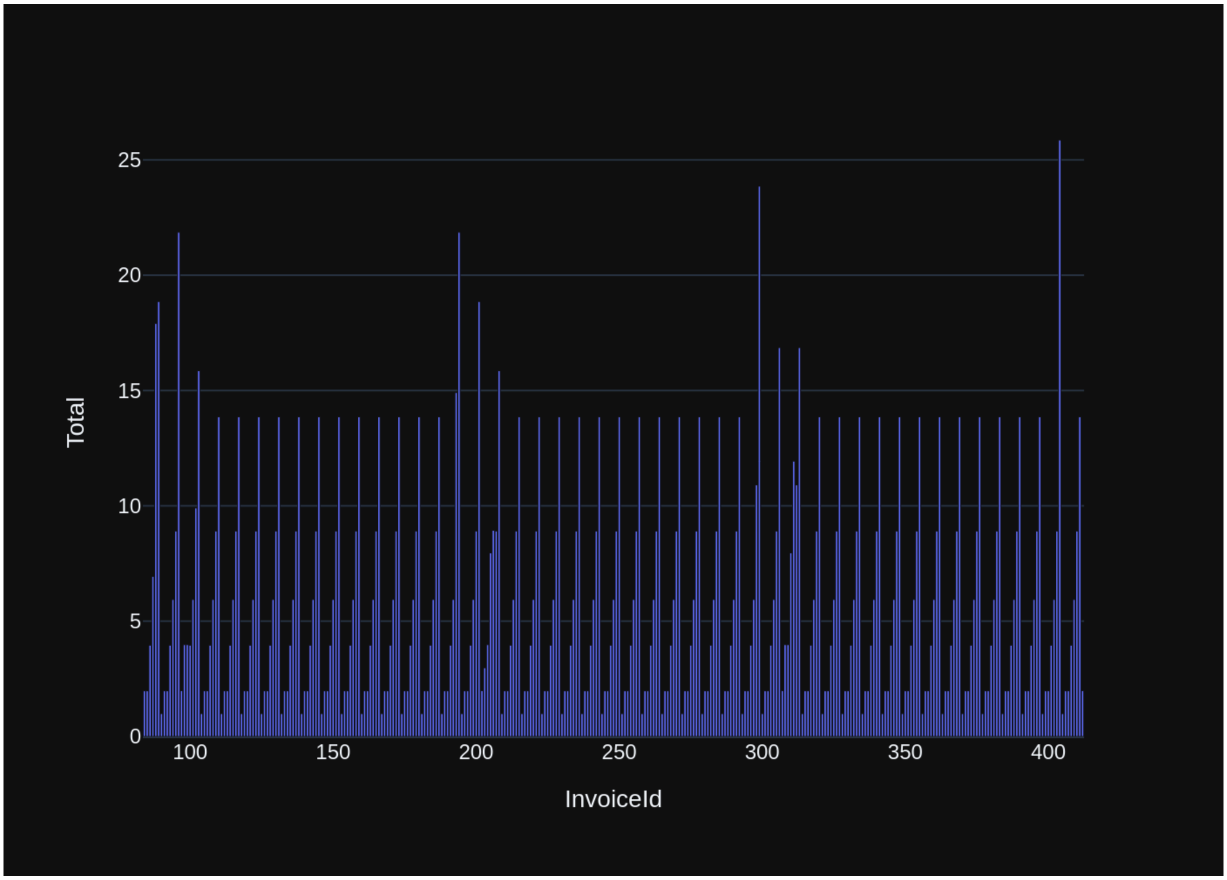
```
question = """
    Find all invoices since 2010 and the total amount invoiced:
"""

vn.ask(question=question)
```

```
Number of requested results 10 is greater than number of elements in index
7, updating n_results = 7
Number of requested results 10 is greater than number of elements in index
1, updating n_results = 1
```

```
SQL Prompt: [{'role': 'system', 'content': 'You are a SQLite expert. Please
help to generate a SQL query to answer the question. Your response should ON
LY be based on the given context and follow the response guidelines and form
at instructions. \n===Tables \nCREATE TABLE "invoices"\r\n(\r\n    InvoiceId
INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    CustomerId INTEGER  NOT N
ULL,\r\n    InvoiceDate DATETIME  NOT NULL,\r\n    BillingAddress NVARCHAR(7
0),\r\n    BillingCity NVARCHAR(40),\r\n    BillingState NVARCHAR(40),\r\n
BillingCountry NVARCHAR(40),\r\n    BillingPostalCode NVARCHAR(10),\r\n    T
otal NUMERIC(10,2)  NOT NULL,\r\n    FOREIGN KEY (CustomerId) REFERENCES "cu
stomers" (CustomerId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n
\nCREATE TABLE "invoice_items"\r\n(\r\n    InvoiceLineId INTEGER PRIMARY KEY
AUTOINCREMENT NOT NULL,\r\n    InvoiceId INTEGER  NOT NULL,\r\n    TrackId I
NTEGER  NOT NULL,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    Quantity
INTEGER  NOT NULL,\r\n    FOREIGN KEY (InvoiceId) REFERENCES "invoices" (Inv
oiceId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY
(TrackId) REFERENCES "tracks" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDAT
E NO ACTION\r\n)\n\nCREATE INDEX IFK_InvoiceLineInvoiceId ON "invoice_items"
(InvoiceId)\n\nCREATE INDEX IFK_InvoiceCustomerId ON "invoices" (CustomerId)
\n\nCREATE INDEX IFK_InvoiceLineTrackId ON "invoice_items" (TrackId)\n\nCREA
TE TABLE "employees"\r\n(\r\n    EmployeeId INTEGER PRIMARY KEY AUTOINCREMEN
T NOT NULL,\r\n    LastName NVARCHAR(20)  NOT NULL,\r\n    FirstName NVARCHA
R(20)  NOT NULL,\r\n    Title NVARCHAR(30),\r\n    ReportsTo INTEGER,\r\n
BirthDate DATETIME,\r\n    HireDate DATETIME,\r\n    Address NVARCHAR(70),\r
\n    City NVARCHAR(40),\r\n    State NVARCHAR(40),\r\n    Country NVARCHAR
(40),\r\n    PostalCode NVARCHAR(10),\r\n    Phone NVARCHAR(24),\r\n    Fax
NVARCHAR(24),\r\n    Email NVARCHAR(60),\r\n    FOREIGN KEY (ReportsTo) REFE
RENCES "employees" (EmployeeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACT
ION\r\n)\n\nCREATE TABLE "customers"\r\n(\r\n    CustomerId INTEGER PRIMARY
KEY AUTOINCREMENT NOT NULL,\r\n    FirstName NVARCHAR(40)  NOT NULL,\r\n
LastName NVARCHAR(20)  NOT NULL,\r\n    Company NVARCHAR(80),\r\n    Address
NVARCHAR(70),\r\n    City NVARCHAR(40),\r\n    State NVARCHAR(40),\r\n    Co
untry NVARCHAR(40),\r\n    PostalCode NVARCHAR(10),\r\n    Phone NVARCHAR(2
4),\r\n    Fax NVARCHAR(24),\r\n    Email NVARCHAR(60)  NOT NULL,\r\n    Sup
portRepId INTEGER,\r\n    FOREIGN KEY (SupportRepId) REFERENCES "employees"
(EmployeeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE
TABLE "tracks"\r\n(\r\n    TrackId INTEGER PRIMARY KEY AUTOINCREMENT NOT NUL
L,\r\n    Name NVARCHAR(200)  NOT NULL,\r\n    AlbumId INTEGER,\r\n    Media
TypeId INTEGER  NOT NULL,\r\n    GenreId INTEGER,\r\n    Composer NVARCHAR(2
20),\r\n    Milliseconds INTEGER  NOT NULL,\r\n    Bytes INTEGER,\r\n    Uni
tPrice NUMERIC(10,2)  NOT NULL,\r\n    FOREIGN KEY (AlbumId) REFERENCES "alb
ums" (AlbumId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREI
GN KEY (GenreId) REFERENCES "genres" (GenreId) \r\n\t\tON DELETE NO ACTION O
N UPDATE NO ACTION,\r\n    FOREIGN KEY (MediaTypeId) REFERENCES "media_type
s" (MediaTypeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCRE
ATE TABLE "albums"\r\n(\r\n    AlbumId INTEGER PRIMARY KEY AUTOINCREMENT NOT
NULL,\r\n    Title NVARCHAR(160)  NOT NULL,\r\n    ArtistId INTEGER  NOT NUL
L,\r\n    FOREIGN KEY (ArtistId) REFERENCES "artists" (ArtistId) \r\n\t\tON
DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE "playlist_track"\r
\n(\r\n    PlaylistId INTEGER  NOT NULL,\r\n    TrackId INTEGER  NOT NULL,\r
\n    CONSTRAINT PK_PlaylistTrack PRIMARY KEY  (PlaylistId, TrackId),\r\n
FOREIGN KEY (PlaylistId) REFERENCES "playlists" (PlaylistId) \r\n\t\tON DELE
TE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackId) REFERENCES "t
racks" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\n\n=
==Additional Context \n\nIn the chinook database invoice means order\n\n===R
esponse Guidelines \n1. If the provided context is sufficient, please genera
te a valid SQL query without any explanations for the question. \n2. If the
```

```
provided context is almost sufficient but requires knowledge of a specific s
tring in a particular column, please generate an intermediate SQL query to f
ind the distinct strings in that column. Prepend the query with a comment sa
ying intermediate_sql \n3. If the provided context is insufficient, please e
xplain why it can\'t be generated. \n4. Please use the most relevant table
(s). \n5. If the question has been asked and answered before, please repeat
the answer exactly as it was given before. \n'}, {'role': 'user', 'content':
'  \n    List all invoices with a total exceeding $10:\n'}, {'role': 'assist
ant', 'content': 'SELECT * FROM invoices WHERE Total > 10.0'}, {'role': 'use
r', 'content': '  \n    Get the total number of invoices for each customer
\n'}, {'role': 'assistant', 'content': 'SELECT CustomerId, COUNT(*) AS Numbe
rOfInvoices\nFROM invoices\nGROUP BY CustomerId'}, {'role': 'user', 'conten
t': "which table stores customer's orders"}, {'role': 'assistant', 'conten
t': 'SELECT * FROM invoices'}, {'role': 'user', 'content': 'How many custome
rs are there'}, {'role': 'assistant', 'content': 'SELECT COUNT(*) AS NumberO
fCustomers FROM customers'}, {'role': 'user', 'content': '  \n    List all a
lbums and their corresponding artist names  \n'}, {'role': 'assistant', 'con
tent': 'SELECT a.Title, ar.Name AS ArtistName\nFROM albums a\nJOIN artists a
r ON a.ArtistId = ar.ArtistId'}, {'role': 'user', 'content': '  \n    Find a
ll tracks with a name containing "What" (case-insensitive)\n'}, {'role': 'as
sistant', 'content': "SELECT DISTINCT LOWER(name) AS Pattern FROM tracks WHE
RE lower(Name) LIKE '%what%'"}, {'role': 'user', 'content': 'Can you list al
l tables in the SQLite database catalog?'}, {'role': 'assistant', 'content':
"SELECT name FROM sqlite_master WHERE type = 'table'"}, {'role': 'user', 'co
ntent': '  \n    Find all invoices since 2010 and the total amount invoice
d:\n'}]
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "You are a SQLite expert. Please help to gene
rate a SQL query to answer the question. Your response should ONLY be based
on the given context and follow the response guidelines and format instructi
ons. \n===Tables \nCREATE TABLE \"invoices\"\r\n(\r\n    InvoiceId INTEGER P
RIMARY KEY AUTOINCREMENT NOT NULL,\r\n    CustomerId INTEGER  NOT NULL,\r\n
InvoiceDate DATETIME  NOT NULL,\r\n    BillingAddress NVARCHAR(70),\r\n    B
illingCity NVARCHAR(40),\r\n    BillingState NVARCHAR(40),\r\n    BillingCou
ntry NVARCHAR(40),\r\n    BillingPostalCode NVARCHAR(10),\r\n    Total NUMER
IC(10,2)  NOT NULL,\r\n    FOREIGN KEY (CustomerId) REFERENCES \"customers\"
(CustomerId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE
TABLE \"invoice_items\"\r\n(\r\n    InvoiceLineId INTEGER PRIMARY KEY AUTOIN
CREMENT NOT NULL,\r\n    InvoiceId INTEGER  NOT NULL,\r\n    TrackId INTEGER
NOT NULL,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    Quantity INTEGER
NOT NULL,\r\n    FOREIGN KEY (InvoiceId) REFERENCES \"invoices\" (InvoiceId)
\r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackI
d) REFERENCES \"tracks\" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO
ACTION\r\n)\n\nCREATE INDEX IFK_InvoiceLineInvoiceId ON \"invoice_items\" (I
nvoiceId)\n\nCREATE INDEX IFK_InvoiceCustomerId ON \"invoices\" (CustomerId)
\n\nCREATE INDEX IFK_InvoiceLineTrackId ON \"invoice_items\" (TrackId)\n\nCR
EATE TABLE \"employees\"\r\n(\r\n    EmployeeId INTEGER PRIMARY KEY AUTOINCR
EMENT NOT NULL,\r\n    LastName NVARCHAR(20)  NOT NULL,\r\n    FirstName NVA
RCHAR(20)  NOT NULL,\r\n    Title NVARCHAR(30),\r\n    ReportsTo INTEGER,\r
\n    BirthDate DATETIME,\r\n    HireDate DATETIME,\r\n    Address NVARCHAR
(70),\r\n    City NVARCHAR(40),\r\n    State NVARCHAR(40),\r\n    Country NV
ARCHAR(40),\r\n    PostalCode NVARCHAR(10),\r\n    Phone NVARCHAR(24),\r\n
Fax NVARCHAR(24),\r\n    Email NVARCHAR(60),\r\n    FOREIGN KEY (ReportsTo)
REFERENCES \"employees\" (EmployeeId) \r\n\t\tON DELETE NO ACTION ON UPDATE
NO ACTION\r\n)\n\nCREATE TABLE \"customers\"\r\n(\r\n    CustomerId INTEGER
PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    FirstName NVARCHAR(40)  NOT NUL
L,\r\n    LastName NVARCHAR(20)  NOT NULL,\r\n    Company NVARCHAR(80),\r\n
Address NVARCHAR(70),\r\n    City NVARCHAR(40),\r\n    State NVARCHAR(40),\r
\n    Country NVARCHAR(40),\r\n    PostalCode NVARCHAR(10),\r\n    Phone NVA
RCHAR(24),\r\n    Fax NVARCHAR(24),\r\n    Email NVARCHAR(60)  NOT NULL,\r\n
SupportRepId INTEGER,\r\n    FOREIGN KEY (SupportRepId) REFERENCES \"employe
es\" (EmployeeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCR
EATE TABLE \"tracks\"\r\n(\r\n    TrackId INTEGER PRIMARY KEY AUTOINCREMENT
NOT NULL,\r\n    Name NVARCHAR(200)  NOT NULL,\r\n    AlbumId INTEGER,\r\n
MediaTypeId INTEGER  NOT NULL,\r\n    GenreId INTEGER,\r\n    Composer NVARC
HAR(220),\r\n    Milliseconds INTEGER  NOT NULL,\r\n    Bytes INTEGER,\r\n
UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    FOREIGN KEY (AlbumId) REFERENCES
\"albums\" (AlbumId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n
FOREIGN KEY (GenreId) REFERENCES \"genres\" (GenreId) \r\n\t\tON DELETE NO A
CTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (MediaTypeId) REFERENCES \"med
ia_types\" (MediaTypeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r
\n)\n\nCREATE TABLE \"albums\"\r\n(\r\n    AlbumId INTEGER PRIMARY KEY AUTOI
NCREMENT NOT NULL,\r\n    Title NVARCHAR(160)  NOT NULL,\r\n    ArtistId INT
EGER  NOT NULL,\r\n    FOREIGN KEY (ArtistId) REFERENCES \"artists\" (Artist
Id) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE \"p
laylist_track\"\r\n(\r\n    PlaylistId INTEGER  NOT NULL,\r\n    TrackId INT
EGER  NOT NULL,\r\n    CONSTRAINT PK_PlaylistTrack PRIMARY KEY  (PlaylistId,
TrackId),\r\n    FOREIGN KEY (PlaylistId) REFERENCES \"playlists\" (Playlist
Id) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (Tra
ckId) REFERENCES \"tracks\" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE
NO ACTION\r\n)\n\n\n===Additional Context \n\nIn the chinook database invoic
e means order\n\n===Response Guidelines \n1. If the provided context is suff
icient, please generate a valid SQL query without any explanations for the q
uestion. \n2. If the provided context is almost sufficient but requires know
ledge of a specific string in a particular column, please generate an interm
ediate SQL query to find the distinct strings in that column. Prepend the qu
ery with a comment saying intermediate_sql \n3. If the provided context is i
nsufficient, please explain why it can't be generated. \n4. Please use the m
ost relevant table(s). \n5. If the question has been asked and answered befo
re, please repeat the answer exactly as it was given before. \n"}, {"role":
"user", "content": "  \n    List all invoices with a total exceeding $1
0:\n"}, {"role": "assistant", "content": "SELECT * FROM invoices WHERE Total
> 10.0"}, {"role": "user", "content": "  \n    Get the total number of invoi
ces for each customer\n"}, {"role": "assistant", "content": "SELECT Customer
Id, COUNT(*) AS NumberOfInvoices\nFROM invoices\nGROUP BY CustomerId"}, {"ro
le": "user", "content": "which table stores customer's orders"}, {"role": "a
ssistant", "content": "SELECT * FROM invoices"}, {"role": "user", "content":
"How many customers are there"}, {"role": "assistant", "content": "SELECT CO
UNT(*) AS NumberOfCustomers FROM customers"}, {"role": "user", "content": "
\n    List all albums and their corresponding artist names  \n"}, {"role":
"assistant", "content": "SELECT a.Title, ar.Name AS ArtistName\nFROM albums
a\nJOIN artists ar ON a.ArtistId = ar.ArtistId"}, {"role": "user", "conten
t": "  \n    Find all tracks with a name containing \"What\" (case-insensiti
ve)\n"}, {"role": "assistant", "content": "SELECT DISTINCT LOWER(name) AS Pa
ttern FROM tracks WHERE lower(Name) LIKE '%what%'"}, {"role": "user", "conte
nt": "Can you list all tables in the SQLite database catalog?"}, {"role": "a
ssistant", "content": "SELECT name FROM sqlite_master WHERE type = 'tabl
e'"}, {"role": "user", "content": "  \n    Find all invoices since 2010 and
```

```
the total amount invoiced:\n"}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T21:50:53.918467509Z', 'messa
ge': {'role': 'assistant', 'content': "SELECT InvoiceId, Total\nFROM invoice
s\nWHERE strftime('%Y', InvoiceDate) >= '2010'"}, 'done_reason': 'stop', 'do
ne': True, 'total_duration': 129546295408, 'load_duration': 3736907, 'prompt
_eval_count': 1868, 'prompt_eval_duration': 119209447000, 'eval_count': 32,
'eval_duration': 9461308000}
LLM Response: SELECT InvoiceId, Total
FROM invoices
WHERE strftime('%Y', InvoiceDate) >= '2010'
SELECT InvoiceId, Total
FROM invoices
WHERE strftime('%Y', InvoiceDate) >= '2010'
     InvoiceId  Total
0           84   1.98
1           85   1.98
2           86   3.96
3           87   6.94
4           88  17.91
..         ...    ...
324        408   3.96
325        409   5.94
326        410   8.91
327        411  13.86
328        412   1.99

[329 rows x 2 columns]
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "The following is a pandas DataFrame that con
tains the results of the query that answers the question the user asked: '
\n    Find all invoices since 2010 and the total amount invoiced:\n'\n\nThe
DataFrame was produced using this query: SELECT InvoiceId, Total\nFROM invoi
ces\nWHERE strftime('%Y', InvoiceDate) >= '2010'\n\nThe following is informa
tion about the resulting pandas DataFrame 'df': \nRunning df.dtypes gives:\n
InvoiceId      int64\nTotal        float64\ndtype: object"}, {"role": "use
r", "content": "Can you generate the Python plotly code to chart the results
of the dataframe? Assume the data is in a pandas dataframe called 'df'. If t
here is only one value in the dataframe, use an Indicator. Respond with only
Python code. Do not answer with any explanations -- just the code."}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T21:51:30.570410048Z', 'messa
ge': {'role': 'assistant', 'content': "```python\nimport plotly.express as p
x\n\n# Check if dataframe has more than one row\nif len(df) > 1:\n    fig =
px.bar(df, x='InvoiceId', y='Total')\nelse:\n    fig = px.indicator(values=
[df['Total'].item()])\n\nfig.show()\n```"}, 'done_reason': 'stop', 'done': T
rue, 'total_duration': 36624804026, 'load_duration': 45167322, 'prompt_eval_
count': 209, 'prompt_eval_duration': 12260096000, 'eval_count': 84, 'eval_du
ration': 24269827000}
```


Total
InvoiceId
0
5
10
15
20
25
100
150
200
250
300
350
400

```
Out[25]: ("SELECT InvoiceId, Total\nFROM invoices\nWHERE strftime('%Y', InvoiceDate)
         >= '2010'",
               InvoiceId  Total
          0           84   1.98
          1           85   1.98
          2           86   3.96
          3           87   6.94
          4           88  17.91
          ..         ...    ...
          324        408   3.96
          325        409   5.94
          326        410   8.91
          327        411  13.86
          328        412   1.99

          [329 rows x 2 columns],
          Figure({
              'data': [{'alignmentgroup': 'True',
                        'hovertemplate': 'InvoiceId=%{x}<br>Total=%{y}<extra></extra
         >',
                        'legendgroup': '',
                        'marker': {'color': '#636efa', 'pattern': {'shape': ''}},
                        'name': '',
                        'offsetgroup': '',
                        'orientation': 'v',
                        'showlegend': False,
                        'textposition': 'auto',
                        'type': 'bar',
                        'x': array([ 84,  85,  86, ..., 410, 411, 412]),
                        'xaxis': 'x',
                        'y': array([ 1.98,  1.98,  3.96, ...,  8.91, 13.86,  1.99]),
                        'yaxis': 'y'}],
              'layout': {'barmode': 'relative',
                         'legend': {'tracegroupgap': 0},
                         'margin': {'t': 60},
                         'template': '...',
                         'xaxis': {'anchor': 'y', 'domain': [0.0, 1.0], 'title': {'t
         ext': 'InvoiceId'}},
                         'yaxis': {'anchor': 'x', 'domain': [0.0, 1.0], 'title': {'t
         ext': 'Total'}}}
          }))
```

```
In [26]: question = """
             List all employees and their reporting manager's name (if any):
         """

         vn.ask(question=question)
```

```
Number of requested results 10 is greater than number of elements in index
8, updating n_results = 8
Number of requested results 10 is greater than number of elements in index
1, updating n_results = 1
```

```
SQL Prompt: [{'role': 'system', 'content': 'You are a SQLite expert. Please
help to generate a SQL query to answer the question. Your response should ON
LY be based on the given context and follow the response guidelines and form
at instructions. \n===Tables \nCREATE INDEX IFK_EmployeeReportsTo ON "employ
ees" (ReportsTo)\n\nCREATE TABLE "employees"\r\n(\r\n    EmployeeId INTEGER
PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    LastName NVARCHAR(20)  NOT NUL
L,\r\n    FirstName NVARCHAR(20)  NOT NULL,\r\n    Title NVARCHAR(30),\r\n
ReportsTo INTEGER,\r\n    BirthDate DATETIME,\r\n    HireDate DATETIME,\r\n
Address NVARCHAR(70),\r\n    City NVARCHAR(40),\r\n    State NVARCHAR(40),\r
\n    Country NVARCHAR(40),\r\n    PostalCode NVARCHAR(10),\r\n    Phone NVA
RCHAR(24),\r\n    Fax NVARCHAR(24),\r\n    Email NVARCHAR(60),\r\n    FOREIG
N KEY (ReportsTo) REFERENCES "employees" (EmployeeId) \r\n\t\tON DELETE NO A
CTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE "customers"\r\n(\r\n    Custo
merId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    FirstName NVARCHAR
(40)  NOT NULL,\r\n    LastName NVARCHAR(20)  NOT NULL,\r\n    Company NVARC
HAR(80),\r\n    Address NVARCHAR(70),\r\n    City NVARCHAR(40),\r\n    State
NVARCHAR(40),\r\n    Country NVARCHAR(40),\r\n    PostalCode NVARCHAR(10),\r
\n    Phone NVARCHAR(24),\r\n    Fax NVARCHAR(24),\r\n    Email NVARCHAR(60)
NOT NULL,\r\n    SupportRepId INTEGER,\r\n    FOREIGN KEY (SupportRepId) REF
ERENCES "employees" (EmployeeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO AC
TION\r\n)\n\nCREATE INDEX IFK_CustomerSupportRepId ON "customers" (SupportRe
pId)\n\nCREATE TABLE "invoices"\r\n(\r\n    InvoiceId INTEGER PRIMARY KEY AU
TOINCREMENT NOT NULL,\r\n    CustomerId INTEGER  NOT NULL,\r\n    InvoiceDat
e DATETIME  NOT NULL,\r\n    BillingAddress NVARCHAR(70),\r\n    BillingCity
NVARCHAR(40),\r\n    BillingState NVARCHAR(40),\r\n    BillingCountry NVARCH
AR(40),\r\n    BillingPostalCode NVARCHAR(10),\r\n    Total NUMERIC(10,2)  N
OT NULL,\r\n    FOREIGN KEY (CustomerId) REFERENCES "customers" (CustomerId)
\r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE "invoic
e_items"\r\n(\r\n    InvoiceLineId INTEGER PRIMARY KEY AUTOINCREMENT NOT NUL
L,\r\n    InvoiceId INTEGER  NOT NULL,\r\n    TrackId INTEGER  NOT NULL,\r\n
UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    Quantity INTEGER  NOT NULL,\r\n
FOREIGN KEY (InvoiceId) REFERENCES "invoices" (InvoiceId) \r\n\t\tON DELETE
NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackId) REFERENCES "trac
ks" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE
TABLE "artists"\r\n(\r\n    ArtistId INTEGER PRIMARY KEY AUTOINCREMENT NOT N
ULL,\r\n    Name NVARCHAR(120)\r\n)\n\nCREATE TABLE "tracks"\r\n(\r\n    Tra
ckId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(200)
NOT NULL,\r\n    AlbumId INTEGER,\r\n    MediaTypeId INTEGER  NOT NULL,\r\n
GenreId INTEGER,\r\n    Composer NVARCHAR(220),\r\n    Milliseconds INTEGER
NOT NULL,\r\n    Bytes INTEGER,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r
\n    FOREIGN KEY (AlbumId) REFERENCES "albums" (AlbumId) \r\n\t\tON DELETE
NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (GenreId) REFERENCES "genr
es" (GenreId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIG
N KEY (MediaTypeId) REFERENCES "media_types" (MediaTypeId) \r\n\t\tON DELETE
NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE "albums"\r\n(\r\n    Albu
mId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Title NVARCHAR(160)
NOT NULL,\r\n    ArtistId INTEGER  NOT NULL,\r\n    FOREIGN KEY (ArtistId) R
EFERENCES "artists" (ArtistId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTI
ON\r\n)\n\nCREATE TABLE sqlite_stat1(tbl,idx,stat)\n\n\n===Additional Contex
t \n\nIn the chinook database invoice means order\n\n===Response Guidelines
\n1. If the provided context is sufficient, please generate a valid SQL quer
y without any explanations for the question. \n2. If the provided context is
almost sufficient but requires knowledge of a specific string in a particula
r column, please generate an intermediate SQL query to find the distinct str
ings in that column. Prepend the query with a comment saying intermediate_sq
l \n3. If the provided context is insufficient, please explain why it can\'t
```

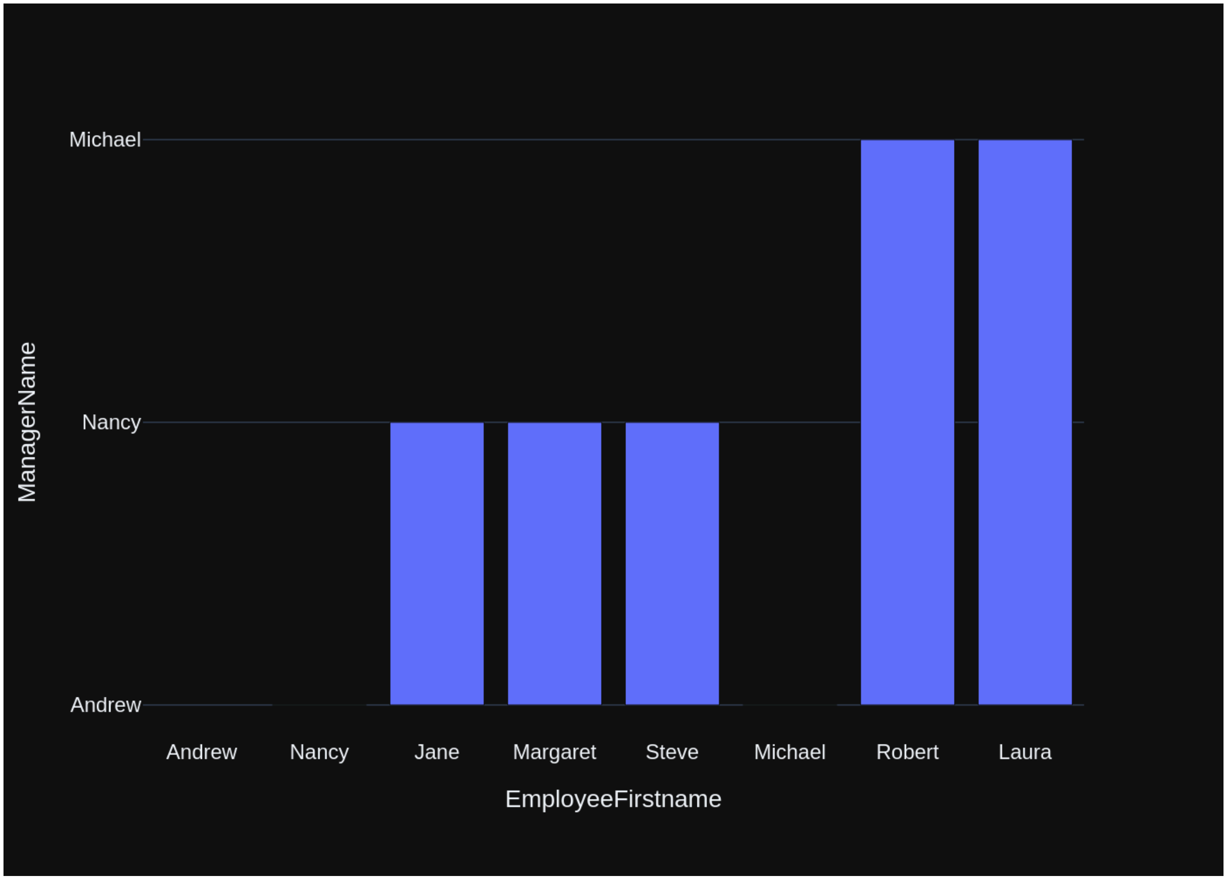
```
be generated. \n4. Please use the most relevant table(s). \n5. If the questi
on has been asked and answered before, please repeat the answer exactly as i
t was given before. \n'}, {'role': 'user', 'content': '  \n    Find all invo
ices since 2010 and the total amount invoiced:\n'}, {'role': 'assistant', 'c
ontent': "SELECT InvoiceId, Total\nFROM invoices\nWHERE strftime('%Y', Invoi
ceDate) >= '2010'"}, {'role': 'user', 'content': '  \n    Get the total numb
er of invoices for each customer\n'}, {'role': 'assistant', 'content': 'SELE
CT CustomerId, COUNT(*) AS NumberOfInvoices\nFROM invoices\nGROUP BY Custome
rId'}, {'role': 'user', 'content': "which table stores customer's orders"},
{'role': 'assistant', 'content': 'SELECT * FROM invoices'}, {'role': 'user',
'content': '  \n    List all albums and their corresponding artist names
\n'}, {'role': 'assistant', 'content': 'SELECT a.Title, ar.Name AS ArtistNam
e\nFROM albums a\nJOIN artists ar ON a.ArtistId = ar.ArtistId'}, {'role': 'u
ser', 'content': '  \n    List all invoices with a total exceeding $10:\n'},
{'role': 'assistant', 'content': 'SELECT * FROM invoices WHERE Total > 10.
0'}, {'role': 'user', 'content': 'How many customers are there'}, {'role':
'assistant', 'content': 'SELECT COUNT(*) AS NumberOfCustomers FROM customer
s'}, {'role': 'user', 'content': 'Can you list all tables in the SQLite data
base catalog?'}, {'role': 'assistant', 'content': "SELECT name FROM sqlite_m
aster WHERE type = 'table'"}, {'role': 'user', 'content': '  \n    Find all
tracks with a name containing "What" (case-insensitive)\n'}, {'role': 'assis
tant', 'content': "SELECT DISTINCT LOWER(name) AS Pattern FROM tracks WHERE
lower(Name) LIKE '%what%'"}, {'role': 'user', 'content': "  \n    List all e
mployees and their reporting manager's name (if any):\n"}]
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "You are a SQLite expert. Please help to gene
rate a SQL query to answer the question. Your response should ONLY be based
on the given context and follow the response guidelines and format instructi
ons. \n===Tables \nCREATE INDEX IFK_EmployeeReportsTo ON \"employees\" (Repo
rtsTo)\n\nCREATE TABLE \"employees\"\r\n(\r\n    EmployeeId INTEGER PRIMARY
KEY AUTOINCREMENT NOT NULL,\r\n    LastName NVARCHAR(20)  NOT NULL,\r\n    F
irstName NVARCHAR(20)  NOT NULL,\r\n    Title NVARCHAR(30),\r\n    ReportsTo
INTEGER,\r\n    BirthDate DATETIME,\r\n    HireDate DATETIME,\r\n    Address
NVARCHAR(70),\r\n    City NVARCHAR(40),\r\n    State NVARCHAR(40),\r\n    Co
untry NVARCHAR(40),\r\n    PostalCode NVARCHAR(10),\r\n    Phone NVARCHAR(2
4),\r\n    Fax NVARCHAR(24),\r\n    Email NVARCHAR(60),\r\n    FOREIGN KEY
(ReportsTo) REFERENCES \"employees\" (EmployeeId) \r\n\t\tON DELETE NO ACTIO
N ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE \"customers\"\r\n(\r\n    Custome
rId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    FirstName NVARCHAR(4
0)  NOT NULL,\r\n    LastName NVARCHAR(20)  NOT NULL,\r\n    Company NVARCHA
R(80),\r\n    Address NVARCHAR(70),\r\n    City NVARCHAR(40),\r\n    State N
VARCHAR(40),\r\n    Country NVARCHAR(40),\r\n    PostalCode NVARCHAR(10),\r
\n    Phone NVARCHAR(24),\r\n    Fax NVARCHAR(24),\r\n    Email NVARCHAR(60)
NOT NULL,\r\n    SupportRepId INTEGER,\r\n    FOREIGN KEY (SupportRepId) REF
ERENCES \"employees\" (EmployeeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO
ACTION\r\n)\n\nCREATE INDEX IFK_CustomerSupportRepId ON \"customers\" (Suppo
rtRepId)\n\nCREATE TABLE \"invoices\"\r\n(\r\n    InvoiceId INTEGER PRIMARY
KEY AUTOINCREMENT NOT NULL,\r\n    CustomerId INTEGER  NOT NULL,\r\n    Invo
iceDate DATETIME  NOT NULL,\r\n    BillingAddress NVARCHAR(70),\r\n    Billi
ngCity NVARCHAR(40),\r\n    BillingState NVARCHAR(40),\r\n    BillingCountry
NVARCHAR(40),\r\n    BillingPostalCode NVARCHAR(10),\r\n    Total NUMERIC(1
0,2)  NOT NULL,\r\n    FOREIGN KEY (CustomerId) REFERENCES \"customers\" (Cu
stomerId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TAB
LE \"invoice_items\"\r\n(\r\n    InvoiceLineId INTEGER PRIMARY KEY AUTOINCRE
MENT NOT NULL,\r\n    InvoiceId INTEGER  NOT NULL,\r\n    TrackId INTEGER  N
OT NULL,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    Quantity INTEGER  
NOT NULL,\r\n    FOREIGN KEY (InvoiceId) REFERENCES \"invoices\" (InvoiceId)
\r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackI
d) REFERENCES \"tracks\" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO 
ACTION\r\n)\n\nCREATE TABLE \"artists\"\r\n(\r\n    ArtistId INTEGER PRIMARY
KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(120)\r\n)\n\nCREATE TABLE 
\"tracks\"\r\n(\r\n    TrackId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r
\n    Name NVARCHAR(200)  NOT NULL,\r\n    AlbumId INTEGER,\r\n    MediaType
Id INTEGER  NOT NULL,\r\n    GenreId INTEGER,\r\n    Composer NVARCHAR(22
0),\r\n    Milliseconds INTEGER  NOT NULL,\r\n    Bytes INTEGER,\r\n    Unit
Price NUMERIC(10,2)  NOT NULL,\r\n    FOREIGN KEY (AlbumId) REFERENCES \"alb
ums\" (AlbumId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FORE
IGN KEY (GenreId) REFERENCES \"genres\" (GenreId) \r\n\t\tON DELETE NO ACTIO
N ON UPDATE NO ACTION,\r\n    FOREIGN KEY (MediaTypeId) REFERENCES \"media_t
ypes\" (MediaTypeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n
\nCREATE TABLE \"albums\"\r\n(\r\n    AlbumId INTEGER PRIMARY KEY AUTOINCREM
ENT NOT NULL,\r\n    Title NVARCHAR(160)  NOT NULL,\r\n    ArtistId INTEGER 
NOT NULL,\r\n    FOREIGN KEY (ArtistId) REFERENCES \"artists\" (ArtistId) \r
\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE sqlite_st
at1(tbl,idx,stat)\n\n\n===Additional Context \n\nIn the chinook database inv
oice means order\n\n===Response Guidelines \n1. If the provided context is s
ufficient, please generate a valid SQL query without any explanations for th
e question. \n2. If the provided context is almost sufficient but requires k
nowledge of a specific string in a particular column, please generate an int
ermediate SQL query to find the distinct strings in that column. Prepend the
query with a comment saying intermediate_sql \n3. If the provided context is
insufficient, please explain why it can't be generated. \n4. Please use the
most relevant table(s). \n5. If the question has been asked and answered bef
ore, please repeat the answer exactly as it was given before. \n"}, {"role":
"user", "content": "  \n    Find all invoices since 2010 and the total amoun
t invoiced:\n"}, {"role": "assistant", "content": "SELECT InvoiceId, Total\n
FROM invoices\nWHERE strftime('%Y', InvoiceDate) >= '2010'"}, {"role": "use
r", "content": "  \n    Get the total number of invoices for each customer 
\n"}, {"role": "assistant", "content": "SELECT CustomerId, COUNT(*) AS Numbe
rOfInvoices\nFROM invoices\nGROUP BY CustomerId"}, {"role": "user", "conten
t": "which table stores customer's orders"}, {"role": "assistant", "conten
t": "SELECT * FROM invoices"}, {"role": "user", "content": "  \n    List all
albums and their corresponding artist names  \n"}, {"role": "assistant", "co
ntent": "SELECT a.Title, ar.Name AS ArtistName\nFROM albums a\nJOIN artists 
ar ON a.ArtistId = ar.ArtistId"}, {"role": "user", "content": "  \n    List 
all invoices with a total exceeding $10:\n"}, {"role": "assistant", "conten
t": "SELECT * FROM invoices WHERE Total > 10.0"}, {"role": "user", "conten
t": "How many customers are there"}, {"role": "assistant", "content": "SELEC
T COUNT(*) AS NumberOfCustomers FROM customers"}, {"role": "user", "conten
t": "Can you list all tables in the SQLite database catalog?"}, {"role": "as
sistant", "content": "SELECT name FROM sqlite_master WHERE type = 'table'"},
{"role": "user", "content": "  \n    Find all tracks with a name containing 
\"What\" (case-insensitive)\n"}, {"role": "assistant", "content": "SELECT DI
STINCT LOWER(name) AS Pattern FROM tracks WHERE lower(Name) LIKE '%what%'"},
{"role": "user", "content": "  \n    List all employees and their reporting 
manager's name (if any):\n"}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T21:54:07.090723203Z', 'messa
```

~~~~
ge': {'role': 'assistant', 'content': 'This requires a self-join on the `emp
loyees` table. Here is an example SQL query:\n\n```sql\nSELECT e1.FirstName
AS EmployeeFirstname, e2.FirstName AS ManagerName\nFROM employees e1\nLEFT J
OIN employees e2 ON e1.ReportsTo = e2.EmployeeId;\n```\nThis query retrieves
all employees along with the first name of their reporting manager (if any).
The `e1` and `e2` are alias names for two instances of the same table, which
allows us to perform a self-join based on the `ReportsTo` field.'}, 'done_re
ason': 'stop', 'done': True, 'total_duration': 156408629169, 'load_duratio
n': 3002508, 'prompt_eval_count': 1815, 'prompt_eval_duration': 11540544300
0, 'eval_count': 132, 'eval_duration': 39961225000}
LLM Response: This requires a self-join on the `employees` table. Here is an
example SQL query:

```sql
SELECT e1.FirstName AS EmployeeFirstname, e2.FirstName AS ManagerName
FROM employees e1
LEFT JOIN employees e2 ON e1.ReportsTo = e2.EmployeeId;
```
This query retrieves all employees along with the first name of their report
ing manager (if any). The `e1` and `e2` are alias names for two instances of
the same table, which allows us to perform a self-join based on the `Reports
To` field.
Info: Output from LLM: This requires a self-join on the `employees` table. H
ere is an example SQL query:

```sql
SELECT e1.FirstName AS EmployeeFirstname, e2.FirstName AS ManagerName
FROM employees e1
LEFT JOIN employees e2 ON e1.ReportsTo = e2.EmployeeId;
```
This query retrieves all employees along with the first name of their report
ing manager (if any). The `e1` and `e2` are alias names for two instances of
the same table, which allows us to perform a self-join based on the `Reports
To` field.
Extracted SQL: SELECT e1.FirstName AS EmployeeFirstname, e2.FirstName AS Man
agerName
FROM employees e1
LEFT JOIN employees e2 ON e1.ReportsTo = e2.EmployeeId
SELECT e1.FirstName AS EmployeeFirstname, e2.FirstName AS ManagerName
FROM employees e1
LEFT JOIN employees e2 ON e1.ReportsTo = e2.EmployeeId
  EmployeeFirstname ManagerName
0            Andrew        None
1             Nancy      Andrew
2              Jane       Nancy
3          Margaret       Nancy
4             Steve       Nancy
5           Michael      Andrew
6            Robert     Michael
7             Laura     Michael
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "The following is a pandas DataFrame that con
~~~~

```
tains the results of the query that answers the question the user asked: '
\n    List all employees and their reporting manager's name (if any):\n'\n\n
The DataFrame was produced using this query: SELECT e1.FirstName AS Employee
Firstname, e2.FirstName AS ManagerName\nFROM employees e1\nLEFT JOIN employe
es e2 ON e1.ReportsTo = e2.EmployeeId\n\nThe following is information about
the resulting pandas DataFrame 'df': \nRunning df.dtypes gives:\n EmployeeFi
rstname    object\nManagerName          object\ndtype: object"}, {"role": "u
ser", "content": "Can you generate the Python plotly code to chart the resul
ts of the dataframe? Assume the data is in a pandas dataframe called 'df'. I
f there is only one value in the dataframe, use an Indicator. Respond with o
nly Python code. Do not answer with any explanations -- just the code."}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T21:54:42.260548012Z', 'messa
ge': {'role': 'assistant', 'content': "```python\nimport plotly.express as p
x\n\n# Assuming df is your DataFrame and it's already defined\nif len(df) ==
1:\n    fig = px.indicator()\nelse:\n    fig = px.bar(df, x='EmployeeFirstna
me', y='ManagerName')\nfig.show()\n```"}, 'done_reason': 'stop', 'done': Tru
e, 'total_duration': 35149528553, 'load_duration': 3036149, 'prompt_eval_cou
nt': 211, 'prompt_eval_duration': 12334292000, 'eval_count': 79, 'eval_durat
ion': 22718676000}
```



```
Out[26]: ('SELECT e1.FirstName AS EmployeeFirstname, e2.FirstName AS ManagerName\nFR
         OM employees e1\nLEFT JOIN employees e2 ON e1.ReportsTo = e2.EmployeeId',
            EmployeeFirstname ManagerName
          0            Andrew        None
          1             Nancy      Andrew
          2              Jane       Nancy
          3          Margaret       Nancy
          4             Steve       Nancy
          5           Michael      Andrew
          6            Robert     Michael
          7             Laura     Michael,
          Figure({
              'data': [{'alignmentgroup': 'True',
                        'hovertemplate': 'EmployeeFirstname=%{x}<br>ManagerName=%{y}
         <extra></extra>',
                        'legendgroup': '',
                        'marker': {'color': '#636efa', 'pattern': {'shape': ''}},
                        'name': '',
                        'offsetgroup': '',
                        'orientation': 'v',
                        'showlegend': False,
                        'textposition': 'auto',
                        'type': 'bar',
                        'x': array(['Andrew', 'Nancy', 'Jane', 'Margaret', 'Steve',
         'Michael', 'Robert',
                                    'Laura'], dtype=object),
                        'xaxis': 'x',
                        'y': array([None, 'Andrew', 'Nancy', 'Nancy', 'Nancy', 'Andr
         ew', 'Michael',
                                    'Michael'], dtype=object),
                        'yaxis': 'y'}],
              'layout': {'barmode': 'relative',
                         'legend': {'tracegroupgap': 0},
                         'margin': {'t': 60},
                         'template': '...',
                         'xaxis': {'anchor': 'y', 'domain': [0.0, 1.0], 'title': {'t
         ext': 'EmployeeFirstname'}},
                         'yaxis': {'anchor': 'x', 'domain': [0.0, 1.0], 'title': {'t
         ext': 'ManagerName'}}}
          }))
```

In [27]:

```
question = """
    Get the average invoice total for each customer:
"""

vn.ask(question=question)
```

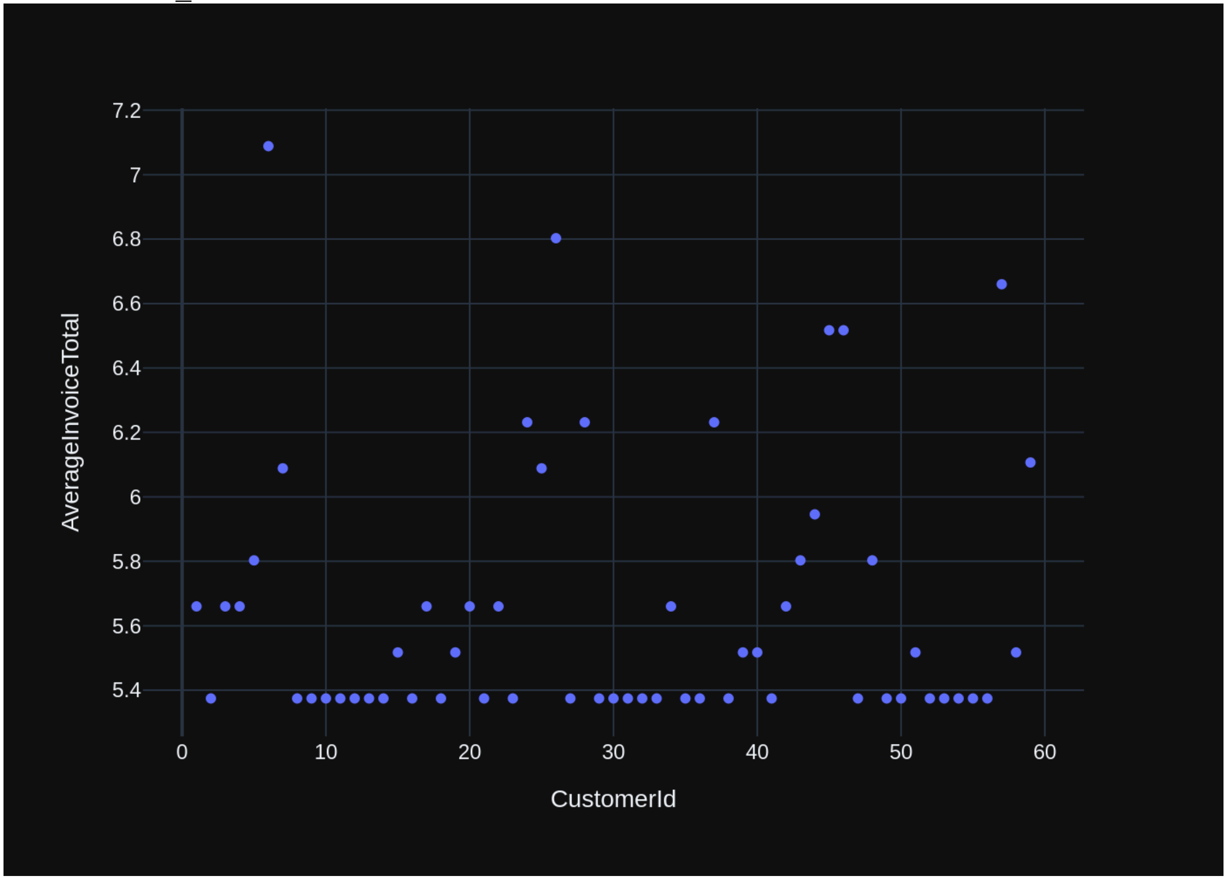
```
Number of requested results 10 is greater than number of elements in index
9, updating n_results = 9
Number of requested results 10 is greater than number of elements in index
1, updating n_results = 1
```

```
SQL Prompt: [{'role': 'system', 'content': 'You are a SQLite expert. Please
help to generate a SQL query to answer the question. Your response should ON
LY be based on the given context and follow the response guidelines and form
at instructions. \n===Tables \nCREATE TABLE "invoices"\r\n(\r\n    InvoiceId
INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    CustomerId INTEGER  NOT N
ULL,\r\n    InvoiceDate DATETIME  NOT NULL,\r\n    BillingAddress NVARCHAR(7
0),\r\n    BillingCity NVARCHAR(40),\r\n    BillingState NVARCHAR(40),\r\n
BillingCountry NVARCHAR(40),\r\n    BillingPostalCode NVARCHAR(10),\r\n    T
otal NUMERIC(10,2)  NOT NULL,\r\n    FOREIGN KEY (CustomerId) REFERENCES "cu
stomers" (CustomerId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n
\nCREATE INDEX IFK_InvoiceCustomerId ON "invoices" (CustomerId)\n\nCREATE IN
DEX IFK_InvoiceLineInvoiceId ON "invoice_items" (InvoiceId)\n\nCREATE TABLE
"invoice_items"\r\n(\r\n    InvoiceLineId INTEGER PRIMARY KEY AUTOINCREMENT
NOT NULL,\r\n    InvoiceId INTEGER  NOT NULL,\r\n    TrackId INTEGER  NOT NU
LL,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    Quantity INTEGER  NOT N
ULL,\r\n    FOREIGN KEY (InvoiceId) REFERENCES "invoices" (InvoiceId) \r\n\t
\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackId) REFE
RENCES "tracks" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r
\n)\n\nCREATE INDEX IFK_InvoiceLineTrackId ON "invoice_items" (TrackId)\n\nC
REATE TABLE sqlite_stat1(tbl,idx,stat)\n\nCREATE INDEX IFK_CustomerSupportRe
pId ON "customers" (SupportRepId)\n\nCREATE TABLE "customers"\r\n(\r\n    Cu
stomerId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    FirstName NVARCH
AR(40)  NOT NULL,\r\n    LastName NVARCHAR(20)  NOT NULL,\r\n    Company NVA
RCHAR(80),\r\n    Address NVARCHAR(70),\r\n    City NVARCHAR(40),\r\n    Sta
te NVARCHAR(40),\r\n    Country NVARCHAR(40),\r\n    PostalCode NVARCHAR(1
0),\r\n    Phone NVARCHAR(24),\r\n    Fax NVARCHAR(24),\r\n    Email NVARCHA
R(60)  NOT NULL,\r\n    SupportRepId INTEGER,\r\n    FOREIGN KEY (SupportRep
Id) REFERENCES "employees" (EmployeeId) \r\n\t\tON DELETE NO ACTION ON UPDAT
E NO ACTION\r\n)\n\nCREATE INDEX IFK_EmployeeReportsTo ON "employees" (Repor
tsTo)\n\nCREATE TABLE "employees"\r\n(\r\n    EmployeeId INTEGER PRIMARY KEY
AUTOINCREMENT NOT NULL,\r\n    LastName NVARCHAR(20)  NOT NULL,\r\n    First
Name NVARCHAR(20)  NOT NULL,\r\n    Title NVARCHAR(30),\r\n    ReportsTo INT
EGER,\r\n    BirthDate DATETIME,\r\n    HireDate DATETIME,\r\n    Address NV
ARCHAR(70),\r\n    City NVARCHAR(40),\r\n    State NVARCHAR(40),\r\n    Coun
try NVARCHAR(40),\r\n    PostalCode NVARCHAR(10),\r\n    Phone NVARCHAR(2
4),\r\n    Fax NVARCHAR(24),\r\n    Email NVARCHAR(60),\r\n    FOREIGN KEY
(ReportsTo) REFERENCES "employees" (EmployeeId) \r\n\t\tON DELETE NO ACTION
ON UPDATE NO ACTION\r\n)\n\n\n===Additional Context \n\nIn the chinook datab
ase invoice means order\n\n===Response Guidelines \n1. If the provided conte
xt is sufficient, please generate a valid SQL query without any explanations
for the question. \n2. If the provided context is almost sufficient but requ
ires knowledge of a specific string in a particular column, please generate
an intermediate SQL query to find the distinct strings in that column. Prepe
nd the query with a comment saying intermediate_sql \n3. If the provided con
text is insufficient, please explain why it can\'t be generated. \n4. Please
use the most relevant table(s). \n5. If the question has been asked and answ
ered before, please repeat the answer exactly as it was given before. \n'},
{'role': 'user', 'content': '  \n    Get the total number of invoices for ea
ch customer\n'}, {'role': 'assistant', 'content': 'SELECT CustomerId, COUNT
(*) AS NumberOfInvoices\nFROM invoices\nGROUP BY CustomerId'}, {'role': 'use
r', 'content': '  \n    Find all invoices since 2010 and the total amount in
voiced:\n'}, {'role': 'assistant', 'content': "SELECT InvoiceId, Total\nFROM
invoices\nWHERE strftime('%Y', InvoiceDate) >= '2010'"}, {'role': 'user', 'c
ontent': '  \n    List all invoices with a total exceeding $10:\n'}, {'rol
e': 'assistant', 'content': 'SELECT * FROM invoices WHERE Total > 10.0'},
{'role': 'user', 'content': "which table stores customer's orders"}, {'rol
```

```
e': 'assistant', 'content': 'SELECT * FROM invoices'}, {'role': 'user', 'con
tent': 'How many customers are there'}, {'role': 'assistant', 'content': 'SE
LECT COUNT(*) AS NumberOfCustomers FROM customers'}, {'role': 'user', 'conte
nt': "  \n    List all employees and their reporting manager's name (if an
y):\n"}, {'role': 'assistant', 'content': 'SELECT e1.FirstName AS EmployeeFi
rstname, e2.FirstName AS ManagerName\nFROM employees e1\nLEFT JOIN employees
e2 ON e1.ReportsTo = e2.EmployeeId'}, {'role': 'user', 'content': '  \n    L
ist all albums and their corresponding artist names  \n'}, {'role': 'assista
nt', 'content': 'SELECT a.Title, ar.Name AS ArtistName\nFROM albums a\nJOIN
artists ar ON a.ArtistId = ar.ArtistId'}, {'role': 'user', 'content': '  \n
Find all tracks with a name containing "What" (case-insensitive)\n'}, {'rol
e': 'assistant', 'content': "SELECT DISTINCT LOWER(name) AS Pattern FROM tra
cks WHERE lower(Name) LIKE '%what%'"}, {'role': 'user', 'content': 'Can you
list all tables in the SQLite database catalog?'}, {'role': 'assistant', 'co
ntent': "SELECT name FROM sqlite_master WHERE type = 'table'"}, {'role': 'us
er', 'content': '  \n    Get the average invoice total for each custome
r:\n'}]
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "You are a SQLite expert. Please help to gene
rate a SQL query to answer the question. Your response should ONLY be based
on the given context and follow the response guidelines and format instructi
ons. \n===Tables \nCREATE TABLE \"invoices\"\r\n(\r\n    InvoiceId INTEGER P
RIMARY KEY AUTOINCREMENT NOT NULL,\r\n    CustomerId INTEGER  NOT NULL,\r\n
InvoiceDate DATETIME  NOT NULL,\r\n    BillingAddress NVARCHAR(70),\r\n    B
illingCity NVARCHAR(40),\r\n    BillingState NVARCHAR(40),\r\n    BillingCou
ntry NVARCHAR(40),\r\n    BillingPostalCode NVARCHAR(10),\r\n    Total NUMER
IC(10,2)  NOT NULL,\r\n    FOREIGN KEY (CustomerId) REFERENCES \"customers\"
(CustomerId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE
INDEX IFK_InvoiceCustomerId ON \"invoices\" (CustomerId)\n\nCREATE INDEX IFK
_InvoiceLineInvoiceId ON \"invoice_items\" (InvoiceId)\n\nCREATE TABLE \"inv
oice_items\"\r\n(\r\n    InvoiceLineId INTEGER PRIMARY KEY AUTOINCREMENT NOT
NULL,\r\n    InvoiceId INTEGER  NOT NULL,\r\n    TrackId INTEGER  NOT NUL
L,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    Quantity INTEGER  NOT NU
LL,\r\n    FOREIGN KEY (InvoiceId) REFERENCES \"invoices\" (InvoiceId) \r\n
\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackId) RE
FERENCES \"tracks\" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTIO
N\r\n)\n\nCREATE INDEX IFK_InvoiceLineTrackId ON \"invoice_items\" (TrackId)
\n\nCREATE TABLE sqlite_stat1(tbl,idx,stat)\n\nCREATE INDEX IFK_CustomerSupp
ortRepId ON \"customers\" (SupportRepId)\n\nCREATE TABLE \"customers\"\r\n
(\r\n    CustomerId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    First
Name NVARCHAR(40)  NOT NULL,\r\n    LastName NVARCHAR(20)  NOT NULL,\r\n
Company NVARCHAR(80),\r\n    Address NVARCHAR(70),\r\n    City NVARCHAR(4
0),\r\n    State NVARCHAR(40),\r\n    Country NVARCHAR(40),\r\n    PostalCod
e NVARCHAR(10),\r\n    Phone NVARCHAR(24),\r\n    Fax NVARCHAR(24),\r\n    E
mail NVARCHAR(60)  NOT NULL,\r\n    SupportRepId INTEGER,\r\n    FOREIGN KEY
(SupportRepId) REFERENCES \"employees\" (EmployeeId) \r\n\t\tON DELETE NO AC
TION ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_EmployeeReportsTo ON \"emp
loyees\" (ReportsTo)\n\nCREATE TABLE \"employees\"\r\n(\r\n    EmployeeId IN
TEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    LastName NVARCHAR(20)  NOT
NULL,\r\n    FirstName NVARCHAR(20)  NOT NULL,\r\n    Title NVARCHAR(30),\r
\n    ReportsTo INTEGER,\r\n    BirthDate DATETIME,\r\n    HireDate DATETIM
E,\r\n    Address NVARCHAR(70),\r\n    City NVARCHAR(40),\r\n    State NVARC
HAR(40),\r\n    Country NVARCHAR(40),\r\n    PostalCode NVARCHAR(10),\r\n
Phone NVARCHAR(24),\r\n    Fax NVARCHAR(24),\r\n    Email NVARCHAR(60),\r\n
FOREIGN KEY (ReportsTo) REFERENCES \"employees\" (EmployeeId) \r\n\t\tON DEL
ETE NO ACTION ON UPDATE NO ACTION\r\n)\n\n\n===Additional Context \n\nIn the
chinook database invoice means order\n\n===Response Guidelines \n1. If the p
rovided context is sufficient, please generate a valid SQL query without any
explanations for the question. \n2. If the provided context is almost suffic
ient but requires knowledge of a specific string in a particular column, ple
ase generate an intermediate SQL query to find the distinct strings in that
column. Prepend the query with a comment saying intermediate_sql \n3. If the
provided context is insufficient, please explain why it can't be generated.
\n4. Please use the most relevant table(s). \n5. If the question has been as
ked and answered before, please repeat the answer exactly as it was given be
fore. \n"}, {"role": "user", "content": "  \n    Get the total number of inv
oices for each customer\n"}, {"role": "assistant", "content": "SELECT Custom
erId, COUNT(*) AS NumberOfInvoices\nFROM invoices\nGROUP BY CustomerId"},
{"role": "user", "content": "  \n    Find all invoices since 2010 and the to
tal amount invoiced:\n"}, {"role": "assistant", "content": "SELECT InvoiceI
d, Total\nFROM invoices\nWHERE strftime('%Y', InvoiceDate) >= '2010'"}, {"ro
le": "user", "content": "  \n    List all invoices with a total exceeding $1
0:\n"}, {"role": "assistant", "content": "SELECT * FROM invoices WHERE Total
> 10.0"}, {"role": "user", "content": "which table stores customer's order
s"}, {"role": "assistant", "content": "SELECT * FROM invoices"}, {"role": "u
ser", "content": "How many customers are there"}, {"role": "assistant", "con
tent": "SELECT COUNT(*) AS NumberOfCustomers FROM customers"}, {"role": "use
r", "content": "  \n    List all employees and their reporting manager's nam
e (if any):\n"}, {"role": "assistant", "content": "SELECT e1.FirstName AS Em
ployeeFirstname, e2.FirstName AS ManagerName\nFROM employees e1\nLEFT JOIN e
mployees e2 ON e1.ReportsTo = e2.EmployeeId"}, {"role": "user", "content": "
\n    List all albums and their corresponding artist names  \n"}, {"role":
"assistant", "content": "SELECT a.Title, ar.Name AS ArtistName\nFROM albums
a\nJOIN artists ar ON a.ArtistId = ar.ArtistId"}, {"role": "user", "conten
t": "  \n    Find all tracks with a name containing \"What\" (case-insensiti
ve)\n"}, {"role": "assistant", "content": "SELECT DISTINCT LOWER(name) AS Pa
ttern FROM tracks WHERE lower(Name) LIKE '%what%'"}, {"role": "user", "conte
nt": "Can you list all tables in the SQLite database catalog?"}, {"role": "a
ssistant", "content": "SELECT name FROM sqlite_master WHERE type = 'tabl
e'"}, {"role": "user", "content": "  \n    Get the average invoice total for
each customer:\n"}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T21:56:30.436605942Z', 'messa
ge': {'role': 'assistant', 'content': 'SELECT CustomerId, AVG(Total) AS Aver
ageInvoiceTotal\nFROM invoices\nGROUP BY CustomerId'}, 'done_reason': 'sto
p', 'done': True, 'total_duration': 108090210462, 'load_duration': 3471787,
'prompt_eval_count': 1568, 'prompt_eval_duration': 99397604000, 'eval_coun
t': 26, 'eval_duration': 7530956000}
LLM Response: SELECT CustomerId, AVG(Total) AS AverageInvoiceTotal
FROM invoices
GROUP BY CustomerId
SELECT CustomerId, AVG(Total) AS AverageInvoiceTotal
FROM invoices
GROUP BY CustomerId
    CustomerId  AverageInvoiceTotal
0            1             5.660000
1            2             5.374286
2            3             5.660000
```

```
3            4            5.660000
4            5            5.802857
5            6            7.088571
6            7            6.088571
7            8            5.374286
8            9            5.374286
9           10            5.374286
10          11            5.374286
11          12            5.374286
12          13            5.374286
13          14            5.374286
14          15            5.517143
15          16            5.374286
16          17            5.660000
17          18            5.374286
18          19            5.517143
19          20            5.660000
20          21            5.374286
21          22            5.660000
22          23            5.374286
23          24            6.231429
24          25            6.088571
25          26            6.802857
26          27            5.374286
27          28            6.231429
28          29            5.374286
29          30            5.374286
30          31            5.374286
31          32            5.374286
32          33            5.374286
33          34            5.660000
34          35            5.374286
35          36            5.374286
36          37            6.231429
37          38            5.374286
38          39            5.517143
39          40            5.517143
40          41            5.374286
41          42            5.660000
42          43            5.802857
43          44            5.945714
44          45            6.517143
45          46            6.517143
46          47            5.374286
47          48            5.802857
48          49            5.374286
49          50            5.374286
50          51            5.517143
51          52            5.374286
52          53            5.374286
53          54            5.374286
54          55            5.374286
55          56            5.374286
56          57            6.660000
57          58            5.517143
58          59            6.106667
```

```
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "The following is a pandas DataFrame that con
tains the results of the query that answers the question the user asked: '
\n    Get the average invoice total for each customer:\n'\n\nThe DataFrame w
as produced using this query: SELECT CustomerId, AVG(Total) AS AverageInvoic
eTotal\nFROM invoices\nGROUP BY CustomerId\n\nThe following is information a
bout the resulting pandas DataFrame 'df': \nRunning df.dtypes gives:\n Custo
merId               int64\nAverageInvoiceTotal    float64\ndtype: object"},
{"role": "user", "content": "Can you generate the Python plotly code to char
t the results of the dataframe? Assume the data is in a pandas dataframe cal
led 'df'. If there is only one value in the dataframe, use an Indicator. Res
pond with only Python code. Do not answer with any explanations -- just the
code."}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T21:57:32.870305226Z', 'messa
ge': {'role': 'assistant', 'content': '```python\nimport plotly.express as p
x\n\nif len(df) > 1:\n    fig = px.bar(df, x=\'CustomerId\', y=\'AverageInvo
iceTotal\')\nelse:\n    fig = px.indicator(data_frame=[{\'name\': \'Customer
Id\', \'value\': df[\'CustomerId\'][0], \'text\': f"Average invoice total fo
r customer {df[\'CustomerId\'][0]} is {df[\'AverageInvoiceTotal\'][0]}"},\n
hover_data={\'text\': True})\n    fig.update_layout(title=\'Average Invoice
Total by Customer\')\nelse:\n    print(\'There is only one value in the data
frame. Cannot generate a chart for this data.\')\n```'}, 'done_reason': 'sto
p', 'done': True, 'total_duration': 62406577794, 'load_duration': 3844246,
'prompt_eval_count': 197, 'prompt_eval_duration': 11448387000, 'eval_count':
175, 'eval_duration': 50818016000}
```



```
Out[27]: ('SELECT CustomerId, AVG(Total) AS AverageInvoiceTotal\nFROM invoices\nGROU
         P BY CustomerId',
               CustomerId  AverageInvoiceTotal
           0            1             5.660000
           1            2             5.374286
           2            3             5.660000
           3            4             5.660000
           4            5             5.802857
           5            6             7.088571
           6            7             6.088571
           7            8             5.374286
           8            9             5.374286
           9           10             5.374286
           10          11             5.374286
           11          12             5.374286
           12          13             5.374286
           13          14             5.374286
           14          15             5.517143
           15          16             5.374286
           16          17             5.660000
           17          18             5.374286
           18          19             5.517143
           19          20             5.660000
           20          21             5.374286
           21          22             5.660000
           22          23             5.374286
           23          24             6.231429
           24          25             6.088571
           25          26             6.802857
           26          27             5.374286
           27          28             6.231429
           28          29             5.374286
           29          30             5.374286
           30          31             5.374286
           31          32             5.374286
           32          33             5.374286
           33          34             5.660000
           34          35             5.374286
           35          36             5.374286
           36          37             6.231429
           37          38             5.374286
           38          39             5.517143
           39          40             5.517143
           40          41             5.374286
           41          42             5.660000
           42          43             5.802857
           43          44             5.945714
           44          45             6.517143
           45          46             6.517143
           46          47             5.374286
           47          48             5.802857
           48          49             5.374286
           49          50             5.374286
           50          51             5.517143
           51          52             5.374286
           52          53             5.374286
```

```
 53          54             5.374286
 54          55             5.374286
 55          56             5.374286
 56          57             6.660000
 57          58             5.517143
 58          59             6.106667,
 Figure({
     'data': [{'hovertemplate': 'CustomerId=%{x}<br>AverageInvoiceTotal=%
{y}<extra></extra>',
               'legendgroup': '',
               'marker': {'color': '#636efa', 'symbol': 'circle'},
               'mode': 'markers',
               'name': '',
               'orientation': 'v',
               'showlegend': False,
               'type': 'scatter',
               'x': array([ 1,  2,  3,  4,  5,  6,  7,  8,  9, 10, 11, 12,
13, 14, 15, 16, 17, 18,
                           19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30,
31, 32, 33, 34, 35, 36,
                           37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48,
49, 50, 51, 52, 53, 54,
                           55, 56, 57, 58, 59]),
               'xaxis': 'x',
               'y': array([5.66      , 5.37428571, 5.66      , 5.66      ,
5.80285714, 7.08857143,
                           6.08857143, 5.37428571, 5.37428571, 5.37428571,
5.37428571, 5.37428571,
                           5.37428571, 5.37428571, 5.51714286, 5.37428571,
5.66      , 5.37428571,
                           5.51714286, 5.66      , 5.37428571, 5.66      ,
5.37428571, 6.23142857,
                           6.08857143, 6.80285714, 5.37428571, 6.23142857,
5.37428571, 5.37428571,
                           5.37428571, 5.37428571, 5.37428571, 5.66      ,
5.37428571, 5.37428571,
                           6.23142857, 5.37428571, 5.51714286, 5.51714286,
5.37428571, 5.66      ,
                           5.80285714, 5.94571429, 6.51714286, 6.51714286,
5.37428571, 5.80285714,
                           5.37428571, 5.37428571, 5.51714286, 5.37428571,
5.37428571, 5.37428571,
                           5.37428571, 5.37428571, 6.66      , 5.51714286,
6.10666667]),
               'yaxis': 'y'}],
     'layout': {'legend': {'tracegroupgap': 0},
                'margin': {'t': 60},
                'template': '...',
                'xaxis': {'anchor': 'y', 'domain': [0.0, 1.0], 'title': {'t
ext': 'CustomerId'}},
                'yaxis': {'anchor': 'x', 'domain': [0.0, 1.0], 'title': {'t
ext': 'AverageInvoiceTotal'}}}
 }))
```

In [28]:

```
question = """
    Find the top 5 most expensive tracks (based on unit price):
```

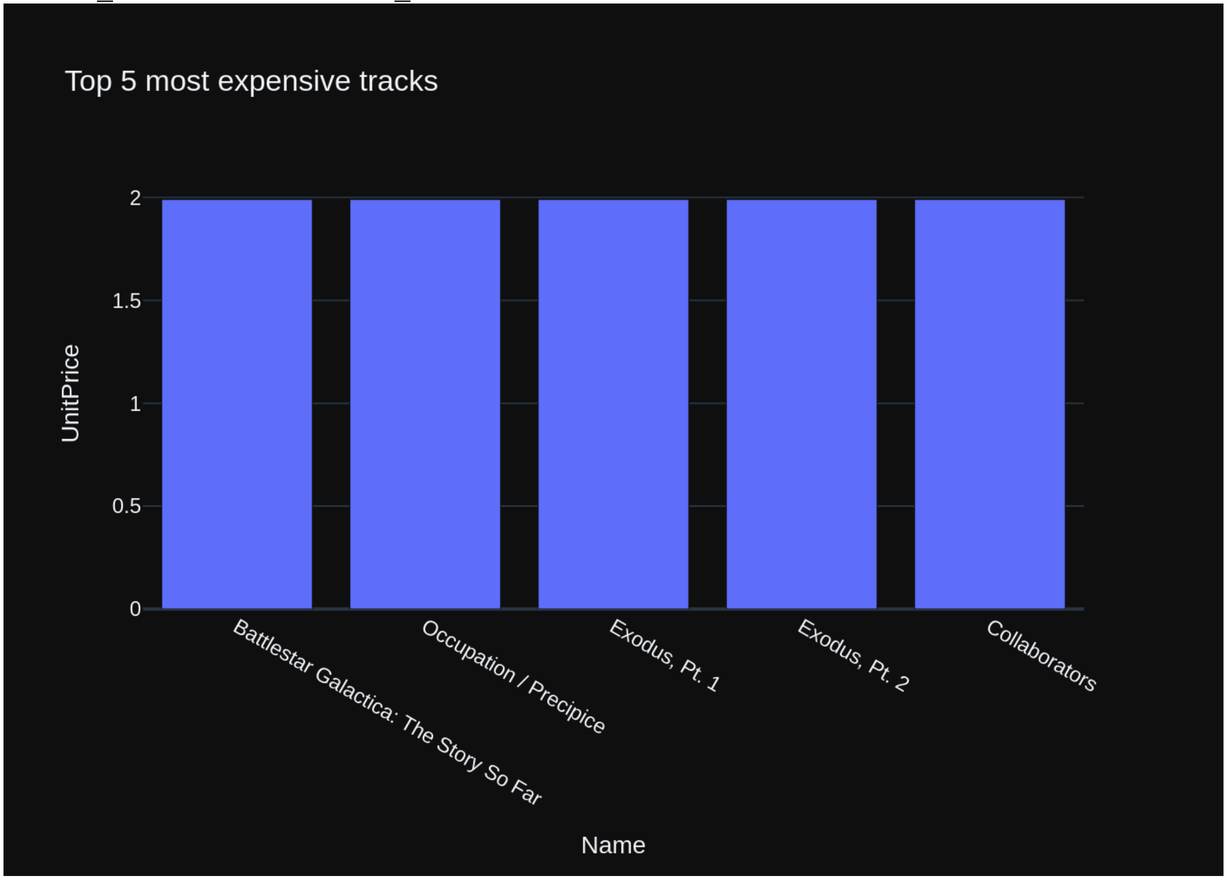
```
"""

vn.ask(question=question)
```

```
Number of requested results 10 is greater than number of elements in index
1, updating n_results = 1
```

```
SQL Prompt: [{'role': 'system', 'content': 'You are a SQLite expert. Please
help to generate a SQL query to answer the question. Your response should ON
LY be based on the given context and follow the response guidelines and form
at instructions. \n===Tables \nCREATE TABLE "tracks"\r\n(\r\n    TrackId INT
EGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(200)  NOT NUL
L,\r\n    AlbumId INTEGER,\r\n    MediaTypeId INTEGER  NOT NULL,\r\n    Genr
eId INTEGER,\r\n    Composer NVARCHAR(220),\r\n    Milliseconds INTEGER  NOT
NULL,\r\n    Bytes INTEGER,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n
FOREIGN KEY (AlbumId) REFERENCES "albums" (AlbumId) \r\n\t\tON DELETE NO ACT
ION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (GenreId) REFERENCES "genres" (G
enreId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY
(MediaTypeId) REFERENCES "media_types" (MediaTypeId) \r\n\t\tON DELETE NO AC
TION ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_TrackAlbumId ON "tracks"
(AlbumId)\n\nCREATE INDEX IFK_TrackGenreId ON "tracks" (GenreId)\n\nCREATE I
NDEX IFK_PlaylistTrackTrackId ON "playlist_track" (TrackId)\n\nCREATE INDEX
IFK_InvoiceLineTrackId ON "invoice_items" (TrackId)\n\nCREATE INDEX IFK_Trac
kMediaTypeId ON "tracks" (MediaTypeId)\n\nCREATE TABLE "invoice_items"\r\n
(\r\n    InvoiceLineId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    In
voiceId INTEGER  NOT NULL,\r\n    TrackId INTEGER  NOT NULL,\r\n    UnitPric
e NUMERIC(10,2)  NOT NULL,\r\n    Quantity INTEGER  NOT NULL,\r\n    FOREIGN
KEY (InvoiceId) REFERENCES "invoices" (InvoiceId) \r\n\t\tON DELETE NO ACTIO
N ON UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackId) REFERENCES "tracks" (Tra
ckId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE
"playlist_track"\r\n(\r\n    PlaylistId INTEGER  NOT NULL,\r\n    TrackId IN
TEGER  NOT NULL,\r\n    CONSTRAINT PK_PlaylistTrack PRIMARY KEY  (PlaylistI
d, TrackId),\r\n    FOREIGN KEY (PlaylistId) REFERENCES "playlists" (Playlis
tId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (Tr
ackId) REFERENCES "tracks" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE N
O ACTION\r\n)\n\nCREATE INDEX IFK_AlbumArtistId ON "albums" (ArtistId)\n\nCR
EATE TABLE "albums"\r\n(\r\n    AlbumId INTEGER PRIMARY KEY AUTOINCREMENT NO
T NULL,\r\n    Title NVARCHAR(160)  NOT NULL,\r\n    ArtistId INTEGER  NOT N
ULL,\r\n    FOREIGN KEY (ArtistId) REFERENCES "artists" (ArtistId) \r\n\t\tO
N DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\n\n===Additional Context \n\nI
n the chinook database invoice means order\n\n===Response Guidelines \n1. If
the provided context is sufficient, please generate a valid SQL query withou
t any explanations for the question. \n2. If the provided context is almost
sufficient but requires knowledge of a specific string in a particular colum
n, please generate an intermediate SQL query to find the distinct strings in
that column. Prepend the query with a comment saying intermediate_sql \n3. I
f the provided context is insufficient, please explain why it can\'t be gene
rated. \n4. Please use the most relevant table(s). \n5. If the question has
been asked and answered before, please repeat the answer exactly as it was g
iven before. \n'}, {'role': 'user', 'content': '  \n    Find all tracks with
a name containing "What" (case-insensitive)\n'}, {'role': 'assistant', 'cont
ent': "SELECT DISTINCT LOWER(name) AS Pattern FROM tracks WHERE lower(Name)
LIKE '%what%'"}, {'role': 'user', 'content': '  \n    List all invoices with
a total exceeding $10:\n'}, {'role': 'assistant', 'content': 'SELECT * FROM
invoices WHERE Total > 10.0'}, {'role': 'user', 'content': '  \n    List all
albums and their corresponding artist names  \n'}, {'role': 'assistant', 'co
ntent': 'SELECT a.Title, ar.Name AS ArtistName\nFROM albums a\nJOIN artists
ar ON a.ArtistId = ar.ArtistId'}, {'role': 'user', 'content': '  \n    Get t
he average invoice total for each customer:\n'}, {'role': 'assistant', 'cont
ent': 'SELECT CustomerId, AVG(Total) AS AverageInvoiceTotal\nFROM invoices\n
GROUP BY CustomerId'}, {'role': 'user', 'content': '  \n    Find all invoice
s since 2010 and the total amount invoiced:\n'}, {'role': 'assistant', 'cont
ent': "SELECT InvoiceId, Total\nFROM invoices\nWHERE strftime('%Y', InvoiceD
ate) >= '2010'"}, {'role': 'user', 'content': 'Can you list all tables in th
e SQLite database catalog?'}, {'role': 'assistant', 'content': "SELECT name
FROM sqlite_master WHERE type = 'table'"}, {'role': 'user', 'content': 'How
many customers are there'}, {'role': 'assistant', 'content': 'SELECT COUNT
(*) AS NumberOfCustomers FROM customers'}, {'role': 'user', 'content': "whic
h table stores customer's orders"}, {'role': 'assistant', 'content': 'SELECT
* FROM invoices'}, {'role': 'user', 'content': '  \n    Get the total number
of invoices for each customer\n'}, {'role': 'assistant', 'content': 'SELECT
CustomerId, COUNT(*) AS NumberOfInvoices\nFROM invoices\nGROUP BY CustomerI
d'}, {'role': 'user', 'content': "  \n    List all employees and their repor
ting manager's name (if any):\n"}, {'role': 'assistant', 'content': 'SELECT
e1.FirstName AS EmployeeFirstname, e2.FirstName AS ManagerName\nFROM employe
es e1\nLEFT JOIN employees e2 ON e1.ReportsTo = e2.EmployeeId'}, {'role': 'u
ser', 'content': '  \n    Find the top 5 most expensive tracks (based on uni
t price):\n'}]
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "You are a SQLite expert. Please help to gene
rate a SQL query to answer the question. Your response should ONLY be based
on the given context and follow the response guidelines and format instructi
ons. \n===Tables \nCREATE TABLE \"tracks\"\r\n(\r\n    TrackId INTEGER PRIMA
RY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(200)  NOT NULL,\r\n    A
lbumId INTEGER,\r\n    MediaTypeId INTEGER  NOT NULL,\r\n    GenreId INTEGE
R,\r\n    Composer NVARCHAR(220),\r\n    Milliseconds INTEGER  NOT NULL,\r\n
Bytes INTEGER,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    FOREIGN KEY
(AlbumId) REFERENCES \"albums\" (AlbumId) \r\n\t\tON DELETE NO ACTION ON UPD
ATE NO ACTION,\r\n    FOREIGN KEY (GenreId) REFERENCES \"genres\" (GenreId)
\r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (MediaTy
peId) REFERENCES \"media_types\" (MediaTypeId) \r\n\t\tON DELETE NO ACTION O
N UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_TrackAlbumId ON \"tracks\" (Albu
mId)\n\nCREATE INDEX IFK_TrackGenreId ON \"tracks\" (GenreId)\n\nCREATE INDE
X IFK_PlaylistTrackTrackId ON \"playlist_track\" (TrackId)\n\nCREATE INDEX I
FK_InvoiceLineTrackId ON \"invoice_items\" (TrackId)\n\nCREATE INDEX IFK_Tra
ckMediaTypeId ON \"tracks\" (MediaTypeId)\n\nCREATE TABLE \"invoice_items
\"\r\n(\r\n    InvoiceLineId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n
InvoiceId INTEGER  NOT NULL,\r\n    TrackId INTEGER  NOT NULL,\r\n    UnitPr
ice NUMERIC(10,2)  NOT NULL,\r\n    Quantity INTEGER  NOT NULL,\r\n    FOREI
GN KEY (InvoiceId) REFERENCES \"invoices\" (InvoiceId) \r\n\t\tON DELETE NO
ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackId) REFERENCES \"tracks
\" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE
TABLE \"playlist_track\"\r\n(\r\n    PlaylistId INTEGER  NOT NULL,\r\n    Tr
ackId INTEGER  NOT NULL,\r\n    CONSTRAINT PK_PlaylistTrack PRIMARY KEY  (Pl
aylistId, TrackId),\r\n    FOREIGN KEY (PlaylistId) REFERENCES \"playlists\"
(PlaylistId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN
KEY (TrackId) REFERENCES \"tracks\" (TrackId) \r\n\t\tON DELETE NO ACTION ON
UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_AlbumArtistId ON \"albums\" (Artis
tId)\n\nCREATE TABLE \"albums\"\r\n(\r\n    AlbumId INTEGER PRIMARY KEY AUTO
INCREMENT NOT NULL,\r\n    Title NVARCHAR(160)  NOT NULL,\r\n    ArtistId IN
TEGER  NOT NULL,\r\n    FOREIGN KEY (ArtistId) REFERENCES \"artists\" (Artis
tId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\n\n===Additional
Context \n\nIn the chinook database invoice means order\n\n===Response Guide
lines \n1. If the provided context is sufficient, please generate a valid SQ
L query without any explanations for the question. \n2. If the provided cont
```

```
ext is almost sufficient but requires knowledge of a specific string in a pa
rticular column, please generate an intermediate SQL query to find the disti
nct strings in that column. Prepend the query with a comment saying intermed
iate_sql \n3. If the provided context is insufficient, please explain why it
can't be generated. \n4. Please use the most relevant table(s). \n5. If the
question has been asked and answered before, please repeat the answer exactl
y as it was given before. \n"}, {"role": "user", "content": "  \n    Find al
l tracks with a name containing \"What\" (case-insensitive)\n"}, {"role": "a
ssistant", "content": "SELECT DISTINCT LOWER(name) AS Pattern FROM tracks WH
ERE lower(Name) LIKE '%what%'"}, {"role": "user", "content": "  \n    List a
ll invoices with a total exceeding $10:\n"}, {"role": "assistant", "conten
t": "SELECT * FROM invoices WHERE Total > 10.0"}, {"role": "user", "conten
t": "  \n    List all albums and their corresponding artist names  \n"}, {"r
ole": "assistant", "content": "SELECT a.Title, ar.Name AS ArtistName\nFROM a
lbums a\nJOIN artists ar ON a.ArtistId = ar.ArtistId"}, {"role": "user", "co
ntent": "  \n    Get the average invoice total for each customer:\n"}, {"rol
e": "assistant", "content": "SELECT CustomerId, AVG(Total) AS AverageInvoice
Total\nFROM invoices\nGROUP BY CustomerId"}, {"role": "user", "content": "
\n    Find all invoices since 2010 and the total amount invoiced:\n"}, {"rol
e": "assistant", "content": "SELECT InvoiceId, Total\nFROM invoices\nWHERE s
trftime('%Y', InvoiceDate) >= '2010'"}, {"role": "user", "content": "Can you
list all tables in the SQLite database catalog?"}, {"role": "assistant", "co
ntent": "SELECT name FROM sqlite_master WHERE type = 'table'"}, {"role": "us
er", "content": "How many customers are there"}, {"role": "assistant", "cont
ent": "SELECT COUNT(*) AS NumberOfCustomers FROM customers"}, {"role": "use
r", "content": "which table stores customer's orders"}, {"role": "assistan
t", "content": "SELECT * FROM invoices"}, {"role": "user", "content": "  \n
Get the total number of invoices for each customer\n"}, {"role": "assistan
t", "content": "SELECT CustomerId, COUNT(*) AS NumberOfInvoices\nFROM invoic
es\nGROUP BY CustomerId"}, {"role": "user", "content": "  \n    List all emp
loyees and their reporting manager's name (if any):\n"}, {"role": "assistan
t", "content": "SELECT e1.FirstName AS EmployeeFirstname, e2.FirstName AS Ma
nagerName\nFROM employees e1\nLEFT JOIN employees e2 ON e1.ReportsTo = e2.Em
ployeeId"}, {"role": "user", "content": "  \n    Find the top 5 most expensi
ve tracks (based on unit price):\n"}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T21:59:13.749626181Z', 'messa
ge': {'role': 'assistant', 'content': 'SELECT TrackId, Name, UnitPrice FROM
tracks ORDER BY UnitPrice DESC LIMIT 5'}, 'done_reason': 'stop', 'done': Tru
e, 'total_duration': 100783227978, 'load_duration': 3943169, 'prompt_eval_co
unt': 1490, 'prompt_eval_duration': 94047928000, 'eval_count': 19, 'eval_dur
ation': 5429538000}
LLM Response: SELECT TrackId, Name, UnitPrice FROM tracks ORDER BY UnitPrice
DESC LIMIT 5
SELECT TrackId, Name, UnitPrice FROM tracks ORDER BY UnitPrice DESC LIMIT 5
   TrackId                                    Name  UnitPrice
0     2819  Battlestar Galactica: The Story So Far       1.99
1     2820                  Occupation / Precipice       1.99
2     2821                           Exodus, Pt. 1       1.99
3     2822                           Exodus, Pt. 2       1.99
4     2823                           Collaborators       1.99
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
```

```
[{"role": "system", "content": "The following is a pandas DataFrame that con
tains the results of the query that answers the question the user asked: '
\n    Find the top 5 most expensive tracks (based on unit price):\n'\n\nThe
DataFrame was produced using this query: SELECT TrackId, Name, UnitPrice FRO
M tracks ORDER BY UnitPrice DESC LIMIT 5\n\nThe following is information abo
ut the resulting pandas DataFrame 'df': \nRunning df.dtypes gives:\n TrackId
int64\nName          object\nUnitPrice    float64\ndtype: object"}, {"role":
"user", "content": "Can you generate the Python plotly code to chart the res
ults of the dataframe? Assume the data is in a pandas dataframe called 'df'.
If there is only one value in the dataframe, use an Indicator. Respond with
only Python code. Do not answer with any explanations -- just the code."}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T22:00:02.001974889Z', 'messa
ge': {'role': 'assistant', 'content': 'Sure, here\'s the plotly python code
to chart the results of the dataframe:\n\n```python\nimport plotly.express a
s px\n\n# Check if there is only one row in the DataFrame\nif len(df) ==
1:\n    fig = px.indicator()\nelse:\n    # If more than one rows, create a b
ar chart for unit price of tracks\n    fig = px.bar(df, x=\'Name\', y="UnitP
rice", title=\'Top 5 most expensive tracks\')\n    \nfig.show()\n```'}, 'don
e_reason': 'stop', 'done': True, 'total_duration': 48226631212, 'load_durati
on': 3842638, 'prompt_eval_count': 195, 'prompt_eval_duration': 11524015000,
'eval_count': 126, 'eval_duration': 36566045000}
```



Out[28]:

```
('SELECT TrackId, Name, UnitPrice FROM tracks ORDER BY UnitPrice DESC LIMIT
5',
    TrackId                                    Name  UnitPrice
 0     2819  Battlestar Galactica: The Story So Far       1.99
 1     2820                   Occupation / Precipice       1.99
 2     2821                            Exodus, Pt. 1       1.99
 3     2822                            Exodus, Pt. 2       1.99
 4     2823                            Collaborators       1.99,
 Figure({
     'data': [{'alignmentgroup': 'True',
               'hovertemplate': 'Name=%{x}<br>UnitPrice=%{y}<extra></extra
>',
               'legendgroup': '',
               'marker': {'color': '#636efa', 'pattern': {'shape': ''}},
               'name': '',
               'offsetgroup': '',
               'orientation': 'v',
               'showlegend': False,
               'textposition': 'auto',
               'type': 'bar',
               'x': array(['Battlestar Galactica: The Story So Far', 'Occup
ation / Precipice',
                           'Exodus, Pt. 1', 'Exodus, Pt. 2', 'Collaborator
s'], dtype=object),
               'xaxis': 'x',
               'y': array([1.99, 1.99, 1.99, 1.99, 1.99]),
               'yaxis': 'y'}],
     'layout': {'barmode': 'relative',
                'legend': {'tracegroupgap': 0},
                'template': '...',
                'title': {'text': 'Top 5 most expensive tracks'},
                'xaxis': {'anchor': 'y', 'domain': [0.0, 1.0], 'title': {'t
ext': 'Name'}},
                'yaxis': {'anchor': 'x', 'domain': [0.0, 1.0], 'title': {'t
ext': 'UnitPrice'}}}
 }))
```

In [29]:

```
question = """
    List all genres and the number of tracks in each genre:
"""

vn.ask(question=question)
```

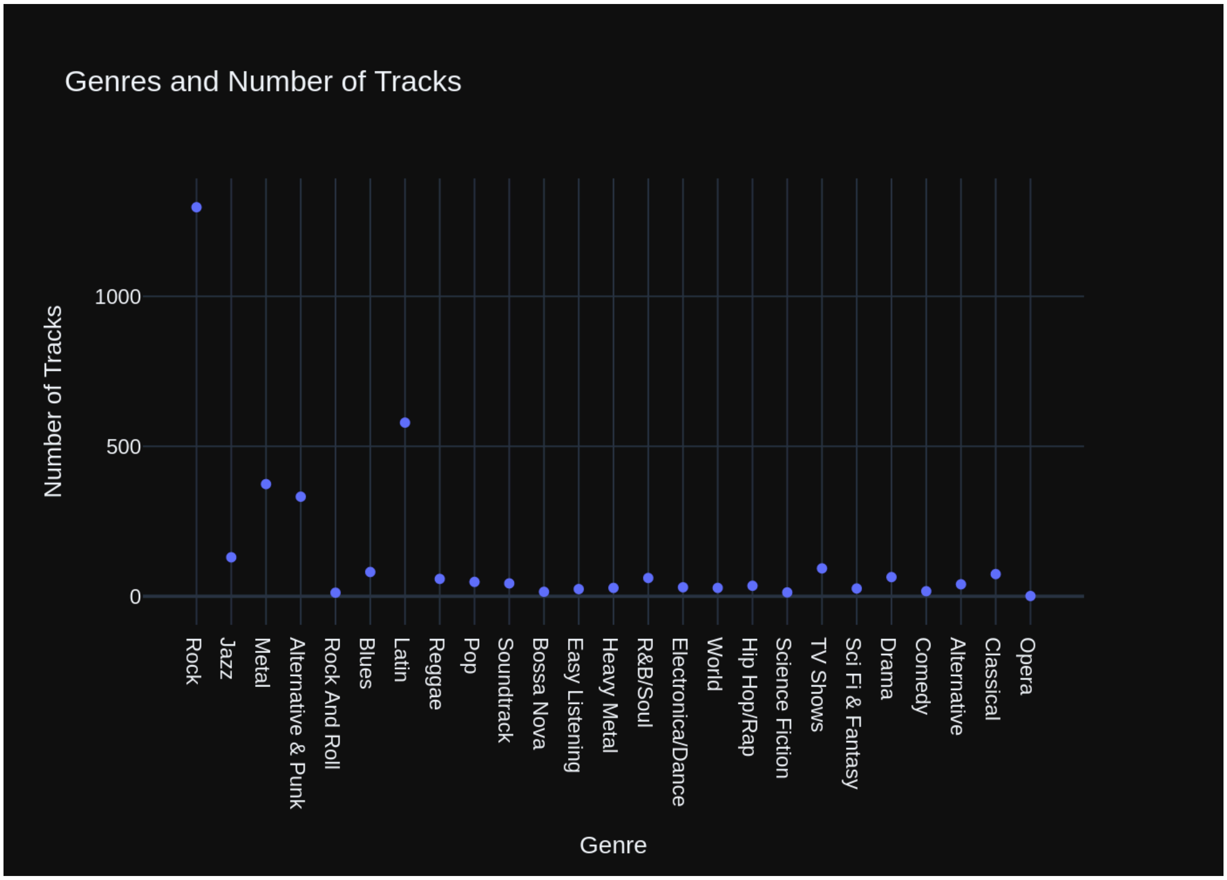
```
Number of requested results 10 is greater than number of elements in index
1, updating n_results = 1
```

```
SQL Prompt: [{'role': 'system', 'content': 'You are a SQLite expert. Please
help to generate a SQL query to answer the question. Your response should ON
LY be based on the given context and follow the response guidelines and form
at instructions. \n===Tables \nCREATE TABLE "tracks"\r\n(\r\n    TrackId INT
EGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(200)  NOT NUL
L,\r\n    AlbumId INTEGER,\r\n    MediaTypeId INTEGER  NOT NULL,\r\n    Genr
eId INTEGER,\r\n    Composer NVARCHAR(220),\r\n    Milliseconds INTEGER  NOT
NULL,\r\n    Bytes INTEGER,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n
FOREIGN KEY (AlbumId) REFERENCES "albums" (AlbumId) \r\n\t\tON DELETE NO ACT
ION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (GenreId) REFERENCES "genres" (G
enreId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY
(MediaTypeId) REFERENCES "media_types" (MediaTypeId) \r\n\t\tON DELETE NO AC
TION ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_TrackGenreId ON "tracks"
(GenreId)\n\nCREATE TABLE "genres"\r\n(\r\n    GenreId INTEGER PRIMARY KEY A
UTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(120)\r\n)\n\nCREATE INDEX IFK_Pl
aylistTrackTrackId ON "playlist_track" (TrackId)\n\nCREATE INDEX IFK_TrackAl
bumId ON "tracks" (AlbumId)\n\nCREATE TABLE "playlists"\r\n(\r\n    Playlist
Id INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(120)\r
\n)\n\nCREATE INDEX IFK_TrackMediaTypeId ON "tracks" (MediaTypeId)\n\nCREATE
TABLE "playlist_track"\r\n(\r\n    PlaylistId INTEGER  NOT NULL,\r\n    Trac
kId INTEGER  NOT NULL,\r\n    CONSTRAINT PK_PlaylistTrack PRIMARY KEY  (Play
listId, TrackId),\r\n    FOREIGN KEY (PlaylistId) REFERENCES "playlists" (Pl
aylistId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KE
Y (TrackId) REFERENCES "tracks" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPD
ATE NO ACTION\r\n)\n\nCREATE TABLE "albums"\r\n(\r\n    AlbumId INTEGER PRIM
ARY KEY AUTOINCREMENT NOT NULL,\r\n    Title NVARCHAR(160)  NOT NULL,\r\n
ArtistId INTEGER  NOT NULL,\r\n    FOREIGN KEY (ArtistId) REFERENCES "artist
s" (ArtistId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE
INDEX IFK_AlbumArtistId ON "albums" (ArtistId)\n\n\n===Additional Context \n
\nIn the chinook database invoice means order\n\n===Response Guidelines \n1.
If the provided context is sufficient, please generate a valid SQL query wit
hout any explanations for the question. \n2. If the provided context is almo
st sufficient but requires knowledge of a specific string in a particular co
lumn, please generate an intermediate SQL query to find the distinct strings
in that column. Prepend the query with a comment saying intermediate_sql \n
3. If the provided context is insufficient, please explain why it can\'t be
generated. \n4. Please use the most relevant table(s). \n5. If the question
has been asked and answered before, please repeat the answer exactly as it w
as given before. \n'}, {'role': 'user', 'content': '  \n    Find the top 5 m
ost expensive tracks (based on unit price):\n'}, {'role': 'assistant', 'cont
ent': 'SELECT TrackId, Name, UnitPrice FROM tracks ORDER BY UnitPrice DESC L
IMIT 5'}, {'role': 'user', 'content': '  \n    List all albums and their cor
responding artist names  \n'}, {'role': 'assistant', 'content': 'SELECT a.Ti
tle, ar.Name AS ArtistName\nFROM albums a\nJOIN artists ar ON a.ArtistId = a
r.ArtistId'}, {'role': 'user', 'content': '  \n    Find all tracks with a na
me containing "What" (case-insensitive)\n'}, {'role': 'assistant', 'conten
t': "SELECT DISTINCT LOWER(name) AS Pattern FROM tracks WHERE lower(Name) LI
KE '%what%'"}, {'role': 'user', 'content': 'How many customers are there'},
{'role': 'assistant', 'content': 'SELECT COUNT(*) AS NumberOfCustomers FROM
customers'}, {'role': 'user', 'content': 'Can you list all tables in the SQL
ite database catalog?'}, {'role': 'assistant', 'content': "SELECT name FROM
sqlite_master WHERE type = 'table'"}, {'role': 'user', 'content': '  \n    G
et the total number of invoices for each customer\n'}, {'role': 'assistant',
'content': 'SELECT CustomerId, COUNT(*) AS NumberOfInvoices\nFROM invoices\n
GROUP BY CustomerId'}, {'role': 'user', 'content': '  \n    List all invoice
s with a total exceeding $10:\n'}, {'role': 'assistant', 'content': 'SELECT
```

```
* FROM invoices WHERE Total > 10.0'}, {'role': 'user', 'content': '  \n    G
et the average invoice total for each customer:\n'}, {'role': 'assistant',
'content': 'SELECT CustomerId, AVG(Total) AS AverageInvoiceTotal\nFROM invoi
ces\nGROUP BY CustomerId'}, {'role': 'user', 'content': '  \n    Find all in
voices since 2010 and the total amount invoiced:\n'}, {'role': 'assistant',
'content': "SELECT InvoiceId, Total\nFROM invoices\nWHERE strftime('%Y', Inv
oiceDate) >= '2010'"}, {'role': 'user', 'content': "  \n    List all employe
es and their reporting manager's name (if any):\n"}, {'role': 'assistant',
'content': 'SELECT e1.FirstName AS EmployeeFirstname, e2.FirstName AS Manage
rName\nFROM employees e1\nLEFT JOIN employees e2 ON e1.ReportsTo = e2.Employ
eeId'}, {'role': 'user', 'content': '  \n    List all genres and the number
of tracks in each genre:\n'}]
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "You are a SQLite expert. Please help to gene
rate a SQL query to answer the question. Your response should ONLY be based
on the given context and follow the response guidelines and format instructi
ons. \n===Tables \nCREATE TABLE \"tracks\"\r\n(\r\n    TrackId INTEGER PRIMA
RY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(200)  NOT NULL,\r\n    A
lbumId INTEGER,\r\n    MediaTypeId INTEGER  NOT NULL,\r\n    GenreId INTEGE
R,\r\n    Composer NVARCHAR(220),\r\n    Milliseconds INTEGER  NOT NULL,\r\n
Bytes INTEGER,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    FOREIGN KEY
(AlbumId) REFERENCES \"albums\" (AlbumId) \r\n\t\tON DELETE NO ACTION ON UPD
ATE NO ACTION,\r\n    FOREIGN KEY (GenreId) REFERENCES \"genres\" (GenreId)
\r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (MediaTy
peId) REFERENCES \"media_types\" (MediaTypeId) \r\n\t\tON DELETE NO ACTION O
N UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_TrackGenreId ON \"tracks\" (Genr
eId)\n\nCREATE TABLE \"genres\"\r\n(\r\n    GenreId INTEGER PRIMARY KEY AUTO
INCREMENT NOT NULL,\r\n    Name NVARCHAR(120)\r\n)\n\nCREATE INDEX IFK_Playl
istTrackTrackId ON \"playlist_track\" (TrackId)\n\nCREATE INDEX IFK_TrackAlb
umId ON \"tracks\" (AlbumId)\n\nCREATE TABLE \"playlists\"\r\n(\r\n    Playl
istId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(120)
\r\n)\n\nCREATE INDEX IFK_TrackMediaTypeId ON \"tracks\" (MediaTypeId)\n\nCR
EATE TABLE \"playlist_track\"\r\n(\r\n    PlaylistId INTEGER  NOT NULL,\r\n
TrackId INTEGER  NOT NULL,\r\n    CONSTRAINT PK_PlaylistTrack PRIMARY KEY
(PlaylistId, TrackId),\r\n    FOREIGN KEY (PlaylistId) REFERENCES \"playlist
s\" (PlaylistId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOR
EIGN KEY (TrackId) REFERENCES \"tracks\" (TrackId) \r\n\t\tON DELETE NO ACTI
ON ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE \"albums\"\r\n(\r\n    AlbumId I
NTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Title NVARCHAR(160)  NOT N
ULL,\r\n    ArtistId INTEGER  NOT NULL,\r\n    FOREIGN KEY (ArtistId) REFERE
NCES \"artists\" (ArtistId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION
\r\n)\n\nCREATE INDEX IFK_AlbumArtistId ON \"albums\" (ArtistId)\n\n\n===Add
itional Context \n\nIn the chinook database invoice means order\n\n===Respon
se Guidelines \n1. If the provided context is sufficient, please generate a
valid SQL query without any explanations for the question. \n2. If the provi
ded context is almost sufficient but requires knowledge of a specific string
in a particular column, please generate an intermediate SQL query to find th
e distinct strings in that column. Prepend the query with a comment saying i
ntermediate_sql \n3. If the provided context is insufficient, please explain
why it can't be generated. \n4. Please use the most relevant table(s). \n5.
If the question has been asked and answered before, please repeat the answer
exactly as it was given before. \n"}, {"role": "user", "content": "  \n    F
```

```
ind the top 5 most expensive tracks (based on unit price):\n"}, {"role": "as
sistant", "content": "SELECT TrackId, Name, UnitPrice FROM tracks ORDER BY U
nitPrice DESC LIMIT 5"}, {"role": "user", "content": "  \n    List all album
s and their corresponding artist names  \n"}, {"role": "assistant", "conten
t": "SELECT a.Title, ar.Name AS ArtistName\nFROM albums a\nJOIN artists ar O
N a.ArtistId = ar.ArtistId"}, {"role": "user", "content": "  \n    Find all
tracks with a name containing \"What\" (case-insensitive)\n"}, {"role": "ass
istant", "content": "SELECT DISTINCT LOWER(name) AS Pattern FROM tracks WHER
E lower(Name) LIKE '%what%'"}, {"role": "user", "content": "How many custome
rs are there"}, {"role": "assistant", "content": "SELECT COUNT(*) AS NumberO
fCustomers FROM customers"}, {"role": "user", "content": "Can you list all t
ables in the SQLite database catalog?"}, {"role": "assistant", "content": "S
ELECT name FROM sqlite_master WHERE type = 'table'"}, {"role": "user", "cont
ent": "  \n    Get the total number of invoices for each customer\n"}, {"rol
e": "assistant", "content": "SELECT CustomerId, COUNT(*) AS NumberOfInvoices
\nFROM invoices\nGROUP BY CustomerId"}, {"role": "user", "content": "  \n
List all invoices with a total exceeding $10:\n"}, {"role": "assistant", "co
ntent": "SELECT * FROM invoices WHERE Total > 10.0"}, {"role": "user", "cont
ent": "  \n    Get the average invoice total for each customer:\n"}, {"rol
e": "assistant", "content": "SELECT CustomerId, AVG(Total) AS AverageInvoice
Total\nFROM invoices\nGROUP BY CustomerId"}, {"role": "user", "content": "
\n    Find all invoices since 2010 and the total amount invoiced:\n"}, {"rol
e": "assistant", "content": "SELECT InvoiceId, Total\nFROM invoices\nWHERE s
trftime('%Y', InvoiceDate) >= '2010'"}, {"role": "user", "content": "  \n
List all employees and their reporting manager's name (if any):\n"}, {"rol
e": "assistant", "content": "SELECT e1.FirstName AS EmployeeFirstname, e2.Fi
rstName AS ManagerName\nFROM employees e1\nLEFT JOIN employees e2 ON e1.Repo
rtsTo = e2.EmployeeId"}, {"role": "user", "content": "  \n    List all genre
s and the number of tracks in each genre:\n"}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T22:01:47.67316415Z', 'messag
e': {'role': 'assistant', 'content': 'SELECT g.GenreId, g.Name, COUNT(t.Trac
kId) AS NumberOfTracks\nFROM genres g\nLEFT JOIN tracks t ON g.GenreId = t.G
enreId\nGROUP BY g.GenreId'}, 'done_reason': 'stop', 'done': True, 'total_du
ration': 105573972063, 'load_duration': 3198474, 'prompt_eval_count': 1405,
'prompt_eval_duration': 88220276000, 'eval_count': 54, 'eval_duration': 1602
8877000}
LLM Response: SELECT g.GenreId, g.Name, COUNT(t.TrackId) AS NumberOfTracks
FROM genres g
LEFT JOIN tracks t ON g.GenreId = t.GenreId
GROUP BY g.GenreId
SELECT g.GenreId, g.Name, COUNT(t.TrackId) AS NumberOfTracks
FROM genres g
LEFT JOIN tracks t ON g.GenreId = t.GenreId
GROUP BY g.GenreId
    GenreId                Name  NumberOfTracks
0         1                Rock            1297
1         2                Jazz             130
2         3               Metal             374
3         4  Alternative & Punk             332
4         5       Rock And Roll              12
5         6               Blues              81
6         7               Latin             579
7         8              Reggae              58
8         9                 Pop              48
9        10          Soundtrack              43
```

```
10       11          Bossa Nova              15
11       12      Easy Listening              24
12       13         Heavy Metal              28
13       14            R&B/Soul              61
14       15   Electronica/Dance              30
15       16               World              28
16       17         Hip Hop/Rap              35
17       18     Science Fiction              13
18       19            TV Shows              93
19       20    Sci Fi & Fantasy              26
20       21               Drama              64
21       22              Comedy              17
22       23         Alternative              40
23       24           Classical              74
24       25               Opera               1
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "The following is a pandas DataFrame that con
tains the results of the query that answers the question the user asked: '
\n    List all genres and the number of tracks in each genre:\n'\n\nThe Data
Frame was produced using this query: SELECT g.GenreId, g.Name, COUNT(t.Track
Id) AS NumberOfTracks\nFROM genres g\nLEFT JOIN tracks t ON g.GenreId = t.Ge
nreId\nGROUP BY g.GenreId\n\nThe following is information about the resultin
g pandas DataFrame 'df': \nRunning df.dtypes gives:\n GenreId            int
64\nName               object\nNumberOfTracks     int64\ndtype: object"}, {"r
ole": "user", "content": "Can you generate the Python plotly code to chart t
he results of the dataframe? Assume the data is in a pandas dataframe called
'df'. If there is only one value in the dataframe, use an Indicator. Respond
with only Python code. Do not answer with any explanations -- just the cod
e."}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T22:03:16.434461592Z', 'messa
ge': {'role': 'assistant', 'content': '```python\n# Required Libraries\nimpo
rt plotly.graph_objects as go\nfrom plotly.subplots import make_subplots\n\n
if len(df) > 1:\n    fig = make_subplots()\n    scatter = go.Scatter(x=df
[\'Name\'], y=df[\'NumberOfTracks\'], mode=\'markers\')\n    fig.add_trace(s
catter, row=1, col=1)\n    \n    layout = go.Layout(title="Genres and Number
of Tracks", xaxis={"title": "Genre"}, yaxis={"title": "Number of Tracks"}, b
armode=\'stack\')\n    fig.update_layout(layout)\nelse:\n    # If only one v
alue, use an Indicator\n    indicator = go.Indicator()\n    indicator.text
="Only One Genre Found."\n    indicator.value=df[\'Name\'].values[0] if len
(df)> 1 else \'No Data\'\n    \n    fig = make_subplots()\n    fig.add_trace
(indicator, row=1, col=1)\n    \nfig.show()\n```'}, 'done_reason': 'stop',
'done': True, 'total_duration': 88734511437, 'load_duration': 3727082, 'prom
pt_eval_count': 232, 'prompt_eval_duration': 13758642000, 'eval_count': 257,
'eval_duration': 74839965000}
```


Genres and Number of Tracks
Number of Tracks
1000
500
0
Rock
Jazz
Metal
Alternative & Punk
Rock And Roll
Blues
Latin
Reggae
Pop
Soundtrack
Bossa Nova
Easy Listening
Heavy Metal
R&B/Soul
Electronica/Dance
World
Hip Hop/Rap
Science Fiction
TV Shows
Sci Fi & Fantasy
Drama
Comedy
Alternative
Classical
Opera
Genre

```
Out[29]: ('SELECT g.GenreId, g.Name, COUNT(t.TrackId) AS NumberOfTracks\nFROM genres
         g\nLEFT JOIN tracks t ON g.GenreId = t.GenreId\nGROUP BY g.GenreId',
              GenreId                Name  NumberOfTracks
          0         1                Rock            1297
          1         2                Jazz             130
          2         3               Metal             374
          3         4  Alternative & Punk             332
          4         5       Rock And Roll              12
          5         6               Blues              81
          6         7               Latin             579
          7         8              Reggae              58
          8         9                 Pop              48
          9        10          Soundtrack              43
          10       11          Bossa Nova              15
          11       12      Easy Listening              24
          12       13         Heavy Metal              28
          13       14            R&B/Soul              61
          14       15   Electronica/Dance              30
          15       16               World              28
          16       17         Hip Hop/Rap              35
          17       18     Science Fiction              13
          18       19            TV Shows              93
          19       20    Sci Fi & Fantasy              26
          20       21               Drama              64
          21       22              Comedy              17
          22       23         Alternative              40
          23       24           Classical              74
          24       25               Opera               1,
          Figure({
              'data': [{'mode': 'markers',
                        'type': 'scatter',
                        'x': array(['Rock', 'Jazz', 'Metal', 'Alternative & Punk',
         'Rock And Roll', 'Blues',
                                    'Latin', 'Reggae', 'Pop', 'Soundtrack', 'Bossa N
         ova', 'Easy Listening',
                                    'Heavy Metal', 'R&B/Soul', 'Electronica/Dance',
         'World', 'Hip Hop/Rap',
                                    'Science Fiction', 'TV Shows', 'Sci Fi & Fantas
         y', 'Drama', 'Comedy',
                                    'Alternative', 'Classical', 'Opera'], dtype=obje
         ct),
                        'xaxis': 'x',
                        'y': array([1297,  130,  374,  332,   12,   81,  579,   58,
         48,   43,   15,   24,
                                      28,   61,   30,   28,   35,   13,   93,   26,
         64,   17,   40,   74,
                                       1]),
                        'yaxis': 'y'}],
              'layout': {'barmode': 'stack',
                         'template': '...',
                         'title': {'text': 'Genres and Number of Tracks'},
                         'xaxis': {'anchor': 'y', 'domain': [0.0, 1.0], 'title': {'t
         ext': 'Genre'}},
                         'yaxis': {'anchor': 'x', 'domain': [0.0, 1.0], 'title': {'t
         ext': 'Number of Tracks'}}}
          }))
```

In [30]:

```
question = """
    Get all genres that do not have any tracks associated with them:
"""

vn.ask(question=question)
```

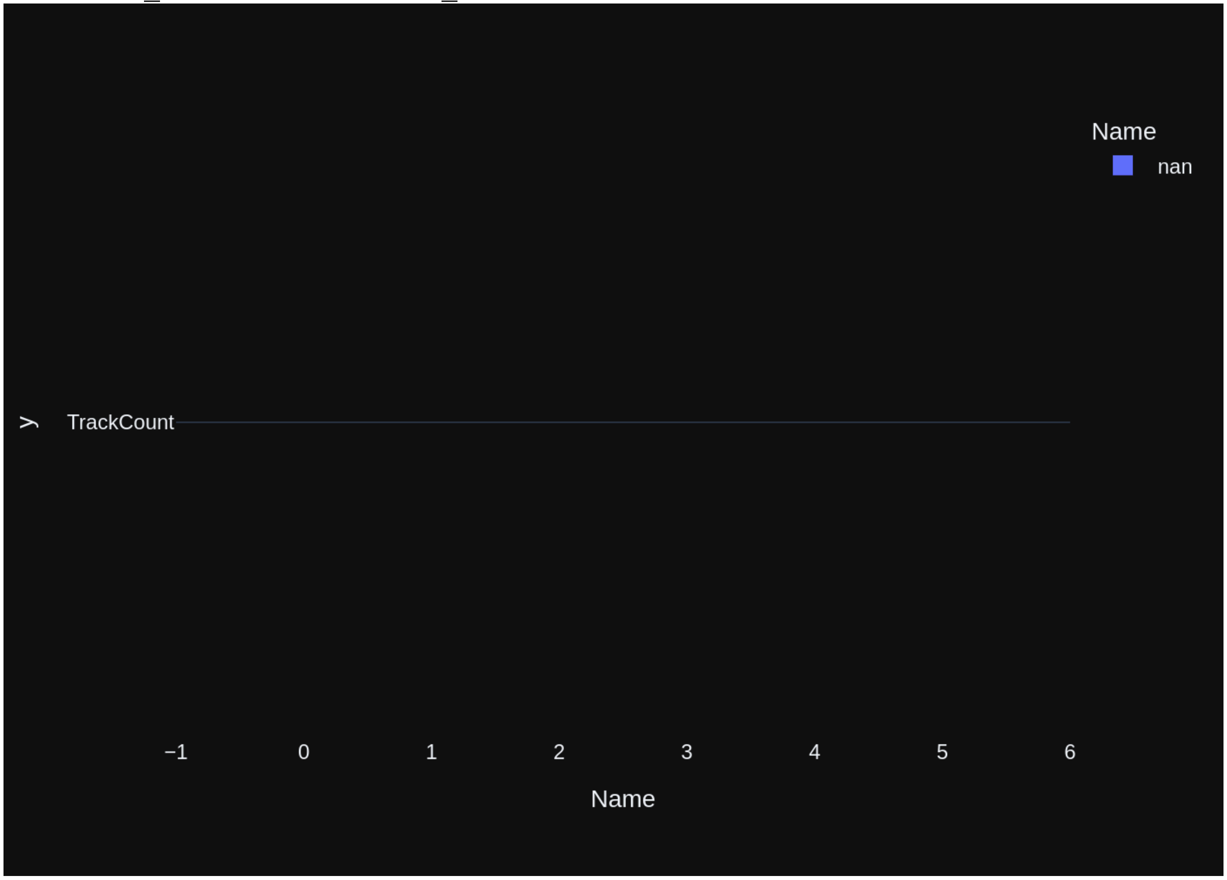
```
Number of requested results 10 is greater than number of elements in index
1, updating n_results = 1
```

```
SQL Prompt: [{'role': 'system', 'content': 'You are a SQLite expert. Please
help to generate a SQL query to answer the question. Your response should ON
LY be based on the given context and follow the response guidelines and form
at instructions. \n===Tables \nCREATE INDEX IFK_TrackGenreId ON "tracks" (Ge
nreId)\n\nCREATE TABLE "tracks"\r\n(\r\n    TrackId INTEGER PRIMARY KEY AUTO
INCREMENT NOT NULL,\r\n    Name NVARCHAR(200)  NOT NULL,\r\n    AlbumId INTE
GER,\r\n    MediaTypeId INTEGER  NOT NULL,\r\n    GenreId INTEGER,\r\n    Co
mposer NVARCHAR(220),\r\n    Milliseconds INTEGER  NOT NULL,\r\n    Bytes IN
TEGER,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    FOREIGN KEY (AlbumI
d) REFERENCES "albums" (AlbumId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO AC
TION,\r\n    FOREIGN KEY (GenreId) REFERENCES "genres" (GenreId) \r\n\t\tON
DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (MediaTypeId) REFER
ENCES "media_types" (MediaTypeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO A
CTION\r\n)\n\nCREATE INDEX IFK_PlaylistTrackTrackId ON "playlist_track" (Tra
ckId)\n\nCREATE INDEX IFK_TrackMediaTypeId ON "tracks" (MediaTypeId)\n\nCREA
TE INDEX IFK_TrackAlbumId ON "tracks" (AlbumId)\n\nCREATE TABLE "genres"\r\n
(\r\n    GenreId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVA
RCHAR(120)\r\n)\n\nCREATE TABLE "albums"\r\n(\r\n    AlbumId INTEGER PRIMARY
KEY AUTOINCREMENT NOT NULL,\r\n    Title NVARCHAR(160)  NOT NULL,\r\n    Art
istId INTEGER  NOT NULL,\r\n    FOREIGN KEY (ArtistId) REFERENCES "artists"
(ArtistId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TA
BLE "playlist_track"\r\n(\r\n    PlaylistId INTEGER  NOT NULL,\r\n    TrackI
d INTEGER  NOT NULL,\r\n    CONSTRAINT PK_PlaylistTrack PRIMARY KEY  (Playli
stId, TrackId),\r\n    FOREIGN KEY (PlaylistId) REFERENCES "playlists" (Play
listId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY
(TrackId) REFERENCES "tracks" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDAT
E NO ACTION\r\n)\n\nCREATE INDEX IFK_AlbumArtistId ON "albums" (ArtistId)\n
\nCREATE TABLE "playlists"\r\n(\r\n    PlaylistId INTEGER PRIMARY KEY AUTOIN
CREMENT NOT NULL,\r\n    Name NVARCHAR(120)\r\n)\n\n\n===Additional Context
\n\nIn the chinook database invoice means order\n\n===Response Guidelines \n
1. If the provided context is sufficient, please generate a valid SQL query
without any explanations for the question. \n2. If the provided context is a
lmost sufficient but requires knowledge of a specific string in a particular
column, please generate an intermediate SQL query to find the distinct strin
gs in that column. Prepend the query with a comment saying intermediate_sql
\n3. If the provided context is insufficient, please explain why it can\'t b
e generated. \n4. Please use the most relevant table(s). \n5. If the questio
n has been asked and answered before, please repeat the answer exactly as it
was given before. \n'}, {'role': 'user', 'content': '  \n    List all genres
and the number of tracks in each genre:\n'}, {'role': 'assistant', 'conten
t': 'SELECT g.GenreId, g.Name, COUNT(t.TrackId) AS NumberOfTracks\nFROM genr
es g\nLEFT JOIN tracks t ON g.GenreId = t.GenreId\nGROUP BY g.GenreId'}, {'r
ole': 'user', 'content': '  \n    List all albums and their corresponding ar
tist names  \n'}, {'role': 'assistant', 'content': 'SELECT a.Title, ar.Name
AS ArtistName\nFROM albums a\nJOIN artists ar ON a.ArtistId = ar.ArtistId'},
{'role': 'user', 'content': '  \n    Find all tracks with a name containing
"What" (case-insensitive)\n'}, {'role': 'assistant', 'content': "SELECT DIST
INCT LOWER(name) AS Pattern FROM tracks WHERE lower(Name) LIKE '%what%'"},
{'role': 'user', 'content': '  \n    Find the top 5 most expensive tracks (b
ased on unit price):\n'}, {'role': 'assistant', 'content': 'SELECT TrackId,
Name, UnitPrice FROM tracks ORDER BY UnitPrice DESC LIMIT 5'}, {'role': 'use
r', 'content': 'Can you list all tables in the SQLite database catalog?'},
{'role': 'assistant', 'content': "SELECT name FROM sqlite_master WHERE type
= 'table'"}, {'role': 'user', 'content': '  \n    List all invoices with a t
otal exceeding $10:\n'}, {'role': 'assistant', 'content': 'SELECT * FROM inv
oices WHERE Total > 10.0'}, {'role': 'user', 'content': '  \n    Find all in
```

```
voices since 2010 and the total amount invoiced:\n'}, {'role': 'assistant',
'content': "SELECT InvoiceId, Total\nFROM invoices\nWHERE strftime('%Y', Inv
oiceDate) >= '2010'"}, {'role': 'user', 'content': 'How many customers are t
here'}, {'role': 'assistant', 'content': 'SELECT COUNT(*) AS NumberOfCustome
rs FROM customers'}, {'role': 'user', 'content': "which table stores custome
r's orders"}, {'role': 'assistant', 'content': 'SELECT * FROM invoices'},
{'role': 'user', 'content': "  \n    List all employees and their reporting
manager's name (if any):\n"}, {'role': 'assistant', 'content': 'SELECT e1.Fi
rstName AS EmployeeFirstname, e2.FirstName AS ManagerName\nFROM employees e1
\nLEFT JOIN employees e2 ON e1.ReportsTo = e2.EmployeeId'}, {'role': 'user',
'content': '  \n    Get all genres that do not have any tracks associated wi
th them:\n'}]
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "You are a SQLite expert. Please help to gene
rate a SQL query to answer the question. Your response should ONLY be based
on the given context and follow the response guidelines and format instructi
ons. \n===Tables \nCREATE INDEX IFK_TrackGenreId ON \"tracks\" (GenreId)\n\n
CREATE TABLE \"tracks\"\r\n(\r\n    TrackId INTEGER PRIMARY KEY AUTOINCREMEN
T NOT NULL,\r\n    Name NVARCHAR(200)  NOT NULL,\r\n    AlbumId INTEGER,\r\n
MediaTypeId INTEGER  NOT NULL,\r\n    GenreId INTEGER,\r\n    Composer NVARC
HAR(220),\r\n    Milliseconds INTEGER  NOT NULL,\r\n    Bytes INTEGER,\r\n
UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    FOREIGN KEY (AlbumId) REFERENCES
\"albums\" (AlbumId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n
FOREIGN KEY (GenreId) REFERENCES \"genres\" (GenreId) \r\n\t\tON DELETE NO A
CTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (MediaTypeId) REFERENCES \"med
ia_types\" (MediaTypeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r
\n)\n\nCREATE INDEX IFK_PlaylistTrackTrackId ON \"playlist_track\" (TrackId)
\n\nCREATE INDEX IFK_TrackMediaTypeId ON \"tracks\" (MediaTypeId)\n\nCREATE
INDEX IFK_TrackAlbumId ON \"tracks\" (AlbumId)\n\nCREATE TABLE \"genres\"\r
\n(\r\n    GenreId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name N
VARCHAR(120)\r\n)\n\nCREATE TABLE \"albums\"\r\n(\r\n    AlbumId INTEGER PRI
MARY KEY AUTOINCREMENT NOT NULL,\r\n    Title NVARCHAR(160)  NOT NULL,\r\n
ArtistId INTEGER  NOT NULL,\r\n    FOREIGN KEY (ArtistId) REFERENCES \"artis
ts\" (ArtistId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREA
TE TABLE \"playlist_track\"\r\n(\r\n    PlaylistId INTEGER  NOT NULL,\r\n
TrackId INTEGER  NOT NULL,\r\n    CONSTRAINT PK_PlaylistTrack PRIMARY KEY
(PlaylistId, TrackId),\r\n    FOREIGN KEY (PlaylistId) REFERENCES \"playlist
s\" (PlaylistId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOR
EIGN KEY (TrackId) REFERENCES \"tracks\" (TrackId) \r\n\t\tON DELETE NO ACTI
ON ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_AlbumArtistId ON \"albums\"
(ArtistId)\n\nCREATE TABLE \"playlists\"\r\n(\r\n    PlaylistId INTEGER PRIM
ARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(120)\r\n)\n\n\n===Addit
ional Context \n\nIn the chinook database invoice means order\n\n===Response
Guidelines \n1. If the provided context is sufficient, please generate a val
id SQL query without any explanations for the question. \n2. If the provided
context is almost sufficient but requires knowledge of a specific string in
a particular column, please generate an intermediate SQL query to find the d
istinct strings in that column. Prepend the query with a comment saying inte
rmediate_sql \n3. If the provided context is insufficient, please explain wh
y it can't be generated. \n4. Please use the most relevant table(s). \n5. If
the question has been asked and answered before, please repeat the answer ex
actly as it was given before. \n"}, {"role": "user", "content": "  \n    Lis
```

```
t all genres and the number of tracks in each genre:\n"}, {"role": "assistan
t", "content": "SELECT g.GenreId, g.Name, COUNT(t.TrackId) AS NumberOfTracks
\nFROM genres g\nLEFT JOIN tracks t ON g.GenreId = t.GenreId\nGROUP BY g.Gen
reId"}, {"role": "user", "content": "  \n    List all albums and their corre
sponding artist names  \n"}, {"role": "assistant", "content": "SELECT a.Titl
e, ar.Name AS ArtistName\nFROM albums a\nJOIN artists ar ON a.ArtistId = ar.
ArtistId"}, {"role": "user", "content": "  \n    Find all tracks with a name
containing \"What\" (case-insensitive)\n"}, {"role": "assistant", "content":
"SELECT DISTINCT LOWER(name) AS Pattern FROM tracks WHERE lower(Name) LIKE
'%what%'"}, {"role": "user", "content": "  \n    Find the top 5 most expensi
ve tracks (based on unit price):\n"}, {"role": "assistant", "content": "SELE
CT TrackId, Name, UnitPrice FROM tracks ORDER BY UnitPrice DESC LIMIT 5"},
{"role": "user", "content": "Can you list all tables in the SQLite database
catalog?"}, {"role": "assistant", "content": "SELECT name FROM sqlite_master
WHERE type = 'table'"}, {"role": "user", "content": "  \n    List all invoic
es with a total exceeding $10:\n"}, {"role": "assistant", "content": "SELECT
* FROM invoices WHERE Total > 10.0"}, {"role": "user", "content": "  \n    F
ind all invoices since 2010 and the total amount invoiced:\n"}, {"role": "as
sistant", "content": "SELECT InvoiceId, Total\nFROM invoices\nWHERE strftime
('%Y', InvoiceDate) >= '2010'"}, {"role": "user", "content": "How many custo
mers are there"}, {"role": "assistant", "content": "SELECT COUNT(*) AS Numbe
rOfCustomers FROM customers"}, {"role": "user", "content": "which table stor
es customer's orders"}, {"role": "assistant", "content": "SELECT * FROM invo
ices"}, {"role": "user", "content": "  \n    List all employees and their re
porting manager's name (if any):\n"}, {"role": "assistant", "content": "SELE
CT e1.FirstName AS EmployeeFirstname, e2.FirstName AS ManagerName\nFROM empl
oyees e1\nLEFT JOIN employees e2 ON e1.ReportsTo = e2.EmployeeId"}, {"role":
"user", "content": "  \n    Get all genres that do not have any tracks assoc
iated with them:\n"}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T22:04:56.655305287Z', 'messa
ge': {'role': 'assistant', 'content': 'SELECT g.* FROM genres g\nLEFT JOIN t
racks t ON g.GenreId = t.GenreId\nWHERE t.TrackId IS NULL'}, 'done_reason':
'stop', 'done': True, 'total_duration': 100119743665, 'load_duration': 37330
86, 'prompt_eval_count': 1413, 'prompt_eval_duration': 89159195000, 'eval_co
unt': 33, 'eval_duration': 9656279000}
LLM Response: SELECT g.* FROM genres g
LEFT JOIN tracks t ON g.GenreId = t.GenreId
WHERE t.TrackId IS NULL
SELECT g.* FROM genres g
LEFT JOIN tracks t ON g.GenreId = t.GenreId
WHERE t.TrackId IS NULL
Empty DataFrame
Columns: [GenreId, Name]
Index: []
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "The following is a pandas DataFrame that con
tains the results of the query that answers the question the user asked: '
\n    Get all genres that do not have any tracks associated with them:\n'\n
\nThe DataFrame was produced using this query: SELECT g.* FROM genres g\nLEF
T JOIN tracks t ON g.GenreId = t.GenreId\nWHERE t.TrackId IS NULL\n\nThe fol
lowing is information about the resulting pandas DataFrame 'df': \nRunning d
```

```
f.dtypes gives:\n GenreId    object\nName       object\ndtype: object"}, {"role": "user", "content": "Can you generate the Python plotly code to chart the results of the dataframe? Assume the data is in a pandas dataframe called 'df'. If there is only one value in the dataframe, use an Indicator. Respond with only Python code. Do not answer with any explanations -- just the code."}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T22:05:38.33166581Z', 'message': {'role': 'assistant', 'content': '```python\nimport plotly.express as px\n\n# Assuming df is already defined and contains data about genres without tracks associated with them\nif len(df[\'GenreId\'].unique()) == 1:\n    fig = px.indicator(df, title=\'Single Genre Without Tracks\')\nelse:\n    fig = px.bar(df, x="Name", y=["TrackCount"], color="Name")\nfig.show()\n```'}, 'done_reason': 'stop', 'done': True, 'total_duration': 41673977394, 'load_duration': 3600425, 'prompt_eval_count': 201, 'prompt_eval_duration': 11865733000, 'eval_count': 103, 'eval_duration': 29673873000}
```



```
Out[30]: ('SELECT g.* FROM genres g\nLEFT JOIN tracks t ON g.GenreId = t.GenreId\nWH
         ERE t.TrackId IS NULL',
          Empty DataFrame
          Columns: [GenreId, Name]
          Index: [],
          Figure({
              'data': [{'alignmentgroup': 'True',
                        'hovertemplate': 'Name=%{x}<br>y=%{y}<extra></extra>',
                        'legendgroup': 'nan',
                        'marker': {'color': '#636efa', 'pattern': {'shape': ''}},
                        'name': 'nan',
                        'offsetgroup': 'nan',
                        'orientation': 'v',
                        'showlegend': True,
                        'textposition': 'auto',
                        'type': 'bar',
                        'x': array([nan], dtype=object),
                        'xaxis': 'x',
                        'y': array(['TrackCount'], dtype=object),
                        'yaxis': 'y'}],
              'layout': {'barmode': 'relative',
                         'legend': {'title': {'text': 'Name'}, 'tracegroupgap': 0},
                         'margin': {'t': 60},
                         'template': '...',
                         'xaxis': {'anchor': 'y',
                                   'categoryarray': [nan],
                                   'categoryorder': 'array',
                                   'domain': [0.0, 1.0],
                                   'title': {'text': 'Name'}},
                         'yaxis': {'anchor': 'x', 'domain': [0.0, 1.0], 'title': {'t
         ext': 'y'}}}
          }))
```

```
In [31]: question = """
             List all customers who have not placed any orders:
         """

         vn.ask(question=question)
```

```
Number of requested results 10 is greater than number of elements in index
1, updating n_results = 1
```

```
SQL Prompt: [{'role': 'system', 'content': 'You are a SQLite expert. Please
help to generate a SQL query to answer the question. Your response should ON
LY be based on the given context and follow the response guidelines and form
at instructions. \n===Tables \nCREATE TABLE "invoices"\r\n(\r\n    InvoiceId
INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    CustomerId INTEGER  NOT N
ULL,\r\n    InvoiceDate DATETIME  NOT NULL,\r\n    BillingAddress NVARCHAR(7
0),\r\n    BillingCity NVARCHAR(40),\r\n    BillingState NVARCHAR(40),\r\n
BillingCountry NVARCHAR(40),\r\n    BillingPostalCode NVARCHAR(10),\r\n    T
otal NUMERIC(10,2)  NOT NULL,\r\n    FOREIGN KEY (CustomerId) REFERENCES "cu
stomers" (CustomerId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n
\nCREATE TABLE "customers"\r\n(\r\n    CustomerId INTEGER PRIMARY KEY AUTOIN
CREMENT NOT NULL,\r\n    FirstName NVARCHAR(40)  NOT NULL,\r\n    LastName N
VARCHAR(20)  NOT NULL,\r\n    Company NVARCHAR(80),\r\n    Address NVARCHAR
(70),\r\n    City NVARCHAR(40),\r\n    State NVARCHAR(40),\r\n    Country NV
ARCHAR(40),\r\n    PostalCode NVARCHAR(10),\r\n    Phone NVARCHAR(24),\r\n
Fax NVARCHAR(24),\r\n    Email NVARCHAR(60)  NOT NULL,\r\n    SupportRepId I
NTEGER,\r\n    FOREIGN KEY (SupportRepId) REFERENCES "employees" (EmployeeI
d) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE "inv
oice_items"\r\n(\r\n    InvoiceLineId INTEGER PRIMARY KEY AUTOINCREMENT NOT
NULL,\r\n    InvoiceId INTEGER  NOT NULL,\r\n    TrackId INTEGER  NOT NUL
L,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    Quantity INTEGER  NOT NU
LL,\r\n    FOREIGN KEY (InvoiceId) REFERENCES "invoices" (InvoiceId) \r\n\t
\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackId) REFE
RENCES "tracks" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r
\n)\n\nCREATE TABLE "employees"\r\n(\r\n    EmployeeId INTEGER PRIMARY KEY A
UTOINCREMENT NOT NULL,\r\n    LastName NVARCHAR(20)  NOT NULL,\r\n    FirstN
ame NVARCHAR(20)  NOT NULL,\r\n    Title NVARCHAR(30),\r\n    ReportsTo INTE
GER,\r\n    BirthDate DATETIME,\r\n    HireDate DATETIME,\r\n    Address NVA
RCHAR(70),\r\n    City NVARCHAR(40),\r\n    State NVARCHAR(40),\r\n    Count
ry NVARCHAR(40),\r\n    PostalCode NVARCHAR(10),\r\n    Phone NVARCHAR(2
4),\r\n    Fax NVARCHAR(24),\r\n    Email NVARCHAR(60),\r\n    FOREIGN KEY
(ReportsTo) REFERENCES "employees" (EmployeeId) \r\n\t\tON DELETE NO ACTION
ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE "playlist_track"\r\n(\r\n    Playli
stId INTEGER  NOT NULL,\r\n    TrackId INTEGER  NOT NULL,\r\n    CONSTRAINT
PK_PlaylistTrack PRIMARY KEY  (PlaylistId, TrackId),\r\n    FOREIGN KEY (Pla
ylistId) REFERENCES "playlists" (PlaylistId) \r\n\t\tON DELETE NO ACTION ON
UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackId) REFERENCES "tracks" (TrackId)
\r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE "album
s"\r\n(\r\n    AlbumId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Ti
tle NVARCHAR(160)  NOT NULL,\r\n    ArtistId INTEGER  NOT NULL,\r\n    FOREI
GN KEY (ArtistId) REFERENCES "artists" (ArtistId) \r\n\t\tON DELETE NO ACTIO
N ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_CustomerSupportRepId ON "cust
omers" (SupportRepId)\n\nCREATE TABLE "playlists"\r\n(\r\n    PlaylistId INT
EGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(120)\r\n)\n\nC
REATE TABLE "tracks"\r\n(\r\n    TrackId INTEGER PRIMARY KEY AUTOINCREMENT N
OT NULL,\r\n    Name NVARCHAR(200)  NOT NULL,\r\n    AlbumId INTEGER,\r\n
MediaTypeId INTEGER  NOT NULL,\r\n    GenreId INTEGER,\r\n    Composer NVARC
HAR(220),\r\n    Milliseconds INTEGER  NOT NULL,\r\n    Bytes INTEGER,\r\n
UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    FOREIGN KEY (AlbumId) REFERENCES
"albums" (AlbumId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    F
OREIGN KEY (GenreId) REFERENCES "genres" (GenreId) \r\n\t\tON DELETE NO ACTI
ON ON UPDATE NO ACTION,\r\n    FOREIGN KEY (MediaTypeId) REFERENCES "media_t
ypes" (MediaTypeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\n
CREATE INDEX IFK_InvoiceCustomerId ON "invoices" (CustomerId)\n\n\n===Additi
onal Context \n\nIn the chinook database invoice means order\n\n===Response
Guidelines \n1. If the provided context is sufficient, please generate a val
```

```
id SQL query without any explanations for the question. \n2. If the provided
context is almost sufficient but requires knowledge of a specific string in
a particular column, please generate an intermediate SQL query to find the d
istinct strings in that column. Prepend the query with a comment saying inte
rmediate_sql \n3. If the provided context is insufficient, please explain wh
y it can\'t be generated. \n4. Please use the most relevant table(s). \n5. I
f the question has been asked and answered before, please repeat the answer
exactly as it was given before. \n'}, {'role': 'user', 'content': "which tab
le stores customer's orders"}, {'role': 'assistant', 'content': 'SELECT * FR
OM invoices'}, {'role': 'user', 'content': '  \n    Get the total number of
invoices for each customer\n'}, {'role': 'assistant', 'content': 'SELECT Cus
tomerId, COUNT(*) AS NumberOfInvoices\nFROM invoices\nGROUP BY CustomerId'},
{'role': 'user', 'content': 'How many customers are there'}, {'role': 'assis
tant', 'content': 'SELECT COUNT(*) AS NumberOfCustomers FROM customers'},
{'role': 'user', 'content': '  \n    List all invoices with a total exceedin
g $10:\n'}, {'role': 'assistant', 'content': 'SELECT * FROM invoices WHERE T
otal > 10.0'}, {'role': 'user', 'content': '  \n    Get the average invoice
total for each customer:\n'}, {'role': 'assistant', 'content': 'SELECT Custo
merId, AVG(Total) AS AverageInvoiceTotal\nFROM invoices\nGROUP BY CustomerI
d'}, {'role': 'user', 'content': '  \n    Find all invoices since 2010 and t
he total amount invoiced:\n'}, {'role': 'assistant', 'content': "SELECT Invo
iceId, Total\nFROM invoices\nWHERE strftime('%Y', InvoiceDate) >= '2010'"},
{'role': 'user', 'content': "  \n    List all employees and their reporting
manager's name (if any):\n"}, {'role': 'assistant', 'content': 'SELECT e1.Fi
rstName AS EmployeeFirstname, e2.FirstName AS ManagerName\nFROM employees e1
\nLEFT JOIN employees e2 ON e1.ReportsTo = e2.EmployeeId'}, {'role': 'user',
'content': '  \n    List all albums and their corresponding artist names
\n'}, {'role': 'assistant', 'content': 'SELECT a.Title, ar.Name AS ArtistNam
e\nFROM albums a\nJOIN artists ar ON a.ArtistId = ar.ArtistId'}, {'role': 'u
ser', 'content': '  \n    Find the top 5 most expensive tracks (based on uni
t price):\n'}, {'role': 'assistant', 'content': 'SELECT TrackId, Name, UnitP
rice FROM tracks ORDER BY UnitPrice DESC LIMIT 5'}, {'role': 'user', 'conten
t': '  \n    List all genres and the number of tracks in each genre:\n'},
{'role': 'assistant', 'content': 'SELECT g.GenreId, g.Name, COUNT(t.TrackId)
AS NumberOfTracks\nFROM genres g\nLEFT JOIN tracks t ON g.GenreId = t.GenreI
d\nGROUP BY g.GenreId'}, {'role': 'user', 'content': '  \n    List all custo
mers who have not placed any orders:\n'}]
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "You are a SQLite expert. Please help to gene
rate a SQL query to answer the question. Your response should ONLY be based
on the given context and follow the response guidelines and format instructi
ons. \n===Tables \nCREATE TABLE \"invoices\"\r\n(\r\n    InvoiceId INTEGER P
RIMARY KEY AUTOINCREMENT NOT NULL,\r\n    CustomerId INTEGER  NOT NULL,\r\n
InvoiceDate DATETIME  NOT NULL,\r\n    BillingAddress NVARCHAR(70),\r\n    B
illingCity NVARCHAR(40),\r\n    BillingState NVARCHAR(40),\r\n    BillingCou
ntry NVARCHAR(40),\r\n    BillingPostalCode NVARCHAR(10),\r\n    Total NUMER
IC(10,2)  NOT NULL,\r\n    FOREIGN KEY (CustomerId) REFERENCES \"customers\"
(CustomerId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE
TABLE \"customers\"\r\n(\r\n    CustomerId INTEGER PRIMARY KEY AUTOINCREMENT
NOT NULL,\r\n    FirstName NVARCHAR(40)  NOT NULL,\r\n    LastName NVARCHAR
(20)  NOT NULL,\r\n    Company NVARCHAR(80),\r\n    Address NVARCHAR(70),\r
\n    City NVARCHAR(40),\r\n    State NVARCHAR(40),\r\n    Country NVARCHAR
(40),\r\n    PostalCode NVARCHAR(10),\r\n    Phone NVARCHAR(24),\r\n    Fax
NVARCHAR(24),\r\n    Email NVARCHAR(60)  NOT NULL,\r\n    SupportRepId INTEG
ER,\r\n    FOREIGN KEY (SupportRepId) REFERENCES \"employees\" (EmployeeId)
\r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE \"invoi
ce_items\"\r\n(\r\n    InvoiceLineId INTEGER PRIMARY KEY AUTOINCREMENT NOT N
ULL,\r\n    InvoiceId INTEGER  NOT NULL,\r\n    TrackId INTEGER  NOT NULL,\r
\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    Quantity INTEGER  NOT NUL
L,\r\n    FOREIGN KEY (InvoiceId) REFERENCES \"invoices\" (InvoiceId) \r\n\t
\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackId) REFE
RENCES \"tracks\" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION
\r\n)\n\nCREATE TABLE \"employees\"\r\n(\r\n    EmployeeId INTEGER PRIMARY K
EY AUTOINCREMENT NOT NULL,\r\n    LastName NVARCHAR(20)  NOT NULL,\r\n    Fi
rstName NVARCHAR(20)  NOT NULL,\r\n    Title NVARCHAR(30),\r\n    ReportsTo
INTEGER,\r\n    BirthDate DATETIME,\r\n    HireDate DATETIME,\r\n    Address
NVARCHAR(70),\r\n    City NVARCHAR(40),\r\n    State NVARCHAR(40),\r\n    Co
untry NVARCHAR(40),\r\n    PostalCode NVARCHAR(10),\r\n    Phone NVARCHAR(2
4),\r\n    Fax NVARCHAR(24),\r\n    Email NVARCHAR(60),\r\n    FOREIGN KEY
(ReportsTo) REFERENCES \"employees\" (EmployeeId) \r\n\t\tON DELETE NO ACTIO
N ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE \"playlist_track\"\r\n(\r\n    Pl
aylistId INTEGER  NOT NULL,\r\n    TrackId INTEGER  NOT NULL,\r\n    CONSTRA
INT PK_PlaylistTrack PRIMARY KEY  (PlaylistId, TrackId),\r\n    FOREIGN KEY
(PlaylistId) REFERENCES \"playlists\" (PlaylistId) \r\n\t\tON DELETE NO ACTI
ON ON UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackId) REFERENCES \"tracks\"
(TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TAB
LE \"albums\"\r\n(\r\n    AlbumId INTEGER PRIMARY KEY AUTOINCREMENT NOT NUL
L,\r\n    Title NVARCHAR(160)  NOT NULL,\r\n    ArtistId INTEGER  NOT NUL
L,\r\n    FOREIGN KEY (ArtistId) REFERENCES \"artists\" (ArtistId) \r\n\t\tO
N DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_CustomerSupp
ortRepId ON \"customers\" (SupportRepId)\n\nCREATE TABLE \"playlists\"\r\n
(\r\n    PlaylistId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name
NVARCHAR(120)\r\n)\n\nCREATE TABLE \"tracks\"\r\n(\r\n    TrackId INTEGER PR
IMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(200)  NOT NULL,\r\n
AlbumId INTEGER,\r\n    MediaTypeId INTEGER  NOT NULL,\r\n    GenreId INTEGE
R,\r\n    Composer NVARCHAR(220),\r\n    Milliseconds INTEGER  NOT NULL,\r\n
Bytes INTEGER,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    FOREIGN KEY
(AlbumId) REFERENCES \"albums\" (AlbumId) \r\n\t\tON DELETE NO ACTION ON UPD
ATE NO ACTION,\r\n    FOREIGN KEY (GenreId) REFERENCES \"genres\" (GenreId)
\r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (MediaTy
peId) REFERENCES \"media_types\" (MediaTypeId) \r\n\t\tON DELETE NO ACTION O
N UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_InvoiceCustomerId ON \"invoices
\" (CustomerId)\n\n\n===Additional Context \n\nIn the chinook database invoi
ce means order\n\n===Response Guidelines \n1. If the provided context is suf
ficient, please generate a valid SQL query without any explanations for the
question. \n2. If the provided context is almost sufficient but requires kno
wledge of a specific string in a particular column, please generate an inter
mediate SQL query to find the distinct strings in that column. Prepend the q
uery with a comment saying intermediate_sql \n3. If the provided context is
insufficient, please explain why it can't be generated. \n4. Please use the
most relevant table(s). \n5. If the question has been asked and answered bef
ore, please repeat the answer exactly as it was given before. \n"}, {"role":
"user", "content": "which table stores customer's orders"}, {"role": "assist
ant", "content": "SELECT * FROM invoices"}, {"role": "user", "content": "
\n    Get the total number of invoices for each customer\n"}, {"role": "assi
stant", "content": "SELECT CustomerId, COUNT(*) AS NumberOfInvoices\nFROM in
voices\nGROUP BY CustomerId"}, {"role": "user", "content": "How many custome
rs are there"}, {"role": "assistant", "content": "SELECT COUNT(*) AS NumberO
```

```
fCustomers FROM customers"}, {"role": "user", "content": "  \n    List all i
nvoices with a total exceeding $10:\n"}, {"role": "assistant", "content": "S
ELECT * FROM invoices WHERE Total > 10.0"}, {"role": "user", "content": "  
\n    Get the average invoice total for each customer:\n"}, {"role": "assist
ant", "content": "SELECT CustomerId, AVG(Total) AS AverageInvoiceTotal\nFROM
invoices\nGROUP BY CustomerId"}, {"role": "user", "content": "  \n    Find a
ll invoices since 2010 and the total amount invoiced:\n"}, {"role": "assista
nt", "content": "SELECT InvoiceId, Total\nFROM invoices\nWHERE strftime('%
Y', InvoiceDate) >= '2010'"}, {"role": "user", "content": "  \n    List all
employees and their reporting manager's name (if any):\n"}, {"role": "assist
ant", "content": "SELECT e1.FirstName AS EmployeeFirstname, e2.FirstName AS
ManagerName\nFROM employees e1\nLEFT JOIN employees e2 ON e1.ReportsTo = e2.
EmployeeId"}, {"role": "user", "content": "  \n    List all albums and their
corresponding artist names  \n"}, {"role": "assistant", "content": "SELECT
a.Title, ar.Name AS ArtistName\nFROM albums a\nJOIN artists ar ON a.ArtistId
= ar.ArtistId"}, {"role": "user", "content": "  \n    Find the top 5 most ex
pensive tracks (based on unit price):\n"}, {"role": "assistant", "content":
"SELECT TrackId, Name, UnitPrice FROM tracks ORDER BY UnitPrice DESC LIMIT
5"}, {"role": "user", "content": "  \n    List all genres and the number of
tracks in each genre:\n"}, {"role": "assistant", "content": "SELECT g.GenreI
d, g.Name, COUNT(t.TrackId) AS NumberOfTracks\nFROM genres g\nLEFT JOIN trac
ks t ON g.GenreId = t.GenreId\nGROUP BY g.GenreId"}, {"role": "user", "conte
nt": "  \n    List all customers who have not placed any orders:\n"}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T22:07:58.416400197Z', 'messa
ge': {'role': 'assistant', 'content': 'SELECT c.* FROM customers c\nWHERE NO
T EXISTS (SELECT 1 FROM invoices i WHERE i.CustomerId = c.CustomerId)'}, 'do
ne_reason': 'stop', 'done': True, 'total_duration': 139988644325, 'load_dura
tion': 3183472, 'prompt_eval_count': 2023, 'prompt_eval_duration': 129636577
000, 'eval_count': 31, 'eval_duration': 9308785000}
LLM Response: SELECT c.* FROM customers c
WHERE NOT EXISTS (SELECT 1 FROM invoices i WHERE i.CustomerId = c.CustomerI
d)
SELECT c.* FROM customers c
WHERE NOT EXISTS (SELECT 1 FROM invoices i WHERE i.CustomerId = c.CustomerI
d)
Empty DataFrame
Columns: [CustomerId, FirstName, LastName, Company, Address, City, State, Co
untry, PostalCode, Phone, Fax, Email, SupportRepId]
Index: []
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "The following is a pandas DataFrame that con
tains the results of the query that answers the question the user asked: '  
\n    List all customers who have not placed any orders:\n'\n\nThe DataFrame
was produced using this query: SELECT c.* FROM customers c\nWHERE NOT EXISTS
(SELECT 1 FROM invoices i WHERE i.CustomerId = c.CustomerId)\n\nThe followin
g is information about the resulting pandas DataFrame 'df': \nRunning df.dty
pes gives:\n CustomerId      object\nFirstName       object\nLastName      
object\nCompany         object\nAddress         object\nCity            obje
ct\nState           object\nCountry         object\nPostalCode      object\n
Phone           object\nFax             object\nEmail           object\nSupp
ortRepId    object\ndtype: object"}, {"role": "user", "content": "Can you ge
```

```
nerate the Python plotly code to chart the results of the dataframe? Assume
the data is in a pandas dataframe called 'df'. If there is only one value in
the dataframe, use an Indicator. Respond with only Python code. Do not answe
r with any explanations -- just the code."}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T22:09:08.30943483Z', 'messag
e': {'role': 'assistant', 'content': "```python\n# First import necessary li
braries\nimport plotly.express as px\nfrom plotly.subplots import make_subpl
ots\n\ndef chart_customer(df):\n    # Check if dataframe has more than one r
ow, if so create a scatter plot\n    if len(df) > 1:\n        fig = px.scatt
er(df, x='CustomerId', y='FirstName')\n    else:\n        fig = go.Figure()
\n    \n    # Add indicators to the figure for each customer row\n    for in
dex, row in df.iterrows():\n        fig.add_trace(go.Indicator())\n
\n    # Create subplot with 1 column and number of rows equal to dataframes
length\n    fig = make_cuptles((len(df), 1))[0]\n    \n    return fig\n``
`"}, 'done_reason': 'stop', 'done': True, 'total_duration': 69890457705, 'lo
ad_duration': 44051661, 'prompt_eval_count': 244, 'prompt_eval_duration': 14
548848000, 'eval_count': 190, 'eval_duration': 55247780000}
Couldn't run plotly code:  'NoneType' object has no attribute 'show'
```

```
Traceback (most recent call last):
  File "/home/gongai/anaconda3/envs/vanna/lib/python3.11/site-packages/vann
a/base/base.py", line 1684, in ask
    img_bytes = fig.to_image(format="png", scale=2)
                ^^^^^^^^^^^^
AttributeError: 'NoneType' object has no attribute 'to_image'

During handling of the above exception, another exception occurred:

Traceback (most recent call last):
  File "/home/gongai/anaconda3/envs/vanna/lib/python3.11/site-packages/vann
a/base/base.py", line 1687, in ask
    fig.show()
    ^^^^^^^^
AttributeError: 'NoneType' object has no attribute 'show'
```

In [32]:

```
question = """
    There are 3 tables: artists, albums and tracks, where albums and artists
    Can you find the top 10 most popular artists based on the number of trac
"""

vn.ask(question=question)
```

```
Number of requested results 10 is greater than number of elements in index
1, updating n_results = 1
```

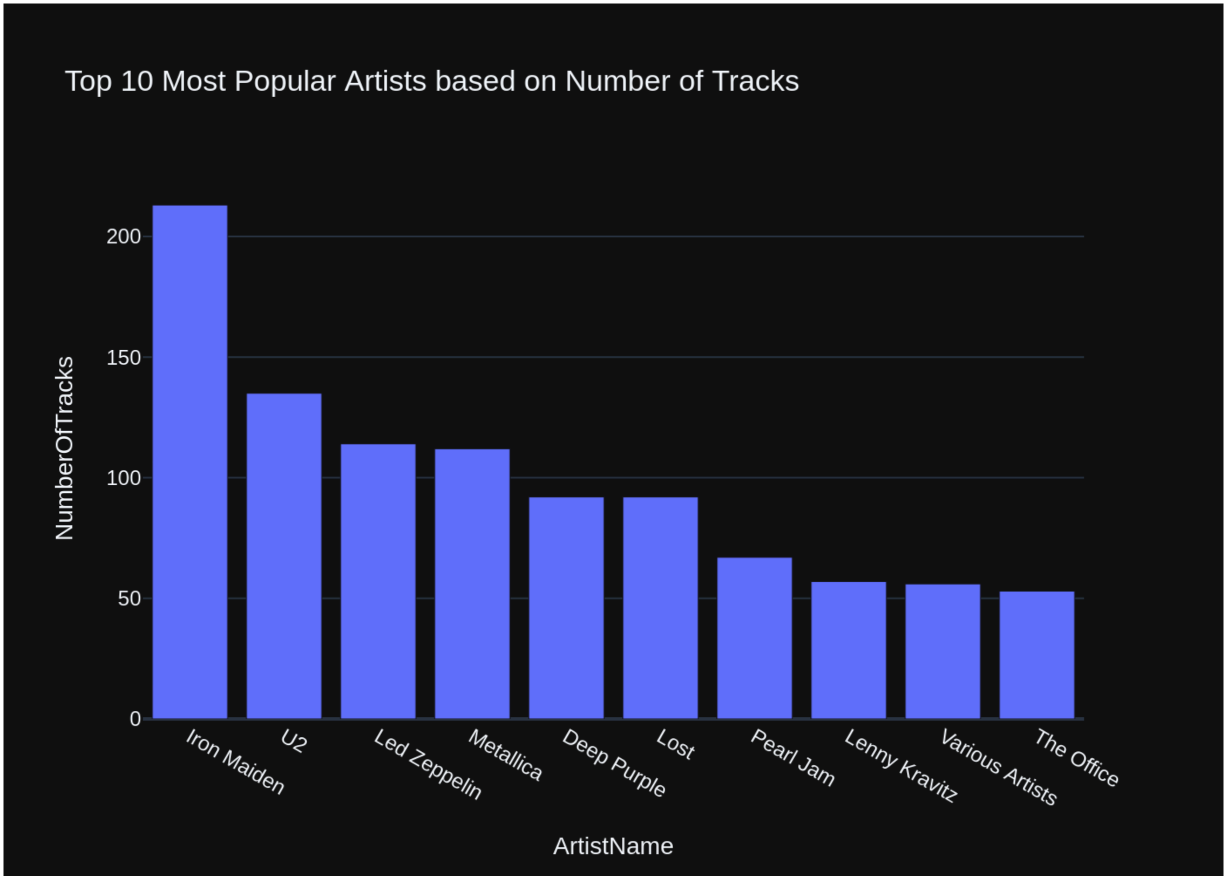
```
SQL Prompt: [{'role': 'system', 'content': 'You are a SQLite expert. Please
help to generate a SQL query to answer the question. Your response should ON
LY be based on the given context and follow the response guidelines and form
at instructions. \n===Tables \nCREATE TABLE "tracks"\r\n(\r\n    TrackId INT
EGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(200)  NOT NUL
L,\r\n    AlbumId INTEGER,\r\n    MediaTypeId INTEGER  NOT NULL,\r\n    Genr
eId INTEGER,\r\n    Composer NVARCHAR(220),\r\n    Milliseconds INTEGER  NOT
NULL,\r\n    Bytes INTEGER,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n
FOREIGN KEY (AlbumId) REFERENCES "albums" (AlbumId) \r\n\t\tON DELETE NO ACT
ION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (GenreId) REFERENCES "genres" (G
enreId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY
(MediaTypeId) REFERENCES "media_types" (MediaTypeId) \r\n\t\tON DELETE NO AC
TION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE "albums"\r\n(\r\n    AlbumId I
NTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Title NVARCHAR(160)  NOT N
ULL,\r\n    ArtistId INTEGER  NOT NULL,\r\n    FOREIGN KEY (ArtistId) REFERE
NCES "artists" (ArtistId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r
\n)\n\nCREATE TABLE "artists"\r\n(\r\n    ArtistId INTEGER PRIMARY KEY AUTOI
NCREMENT NOT NULL,\r\n    Name NVARCHAR(120)\r\n)\n\nCREATE INDEX IFK_AlbumA
rtistId ON "albums" (ArtistId)\n\nCREATE INDEX IFK_TrackAlbumId ON "tracks"
(AlbumId)\n\nCREATE TABLE "playlists"\r\n(\r\n    PlaylistId INTEGER PRIMARY
KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(120)\r\n)\n\nCREATE TABLE
"genres"\r\n(\r\n    GenreId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n
Name NVARCHAR(120)\r\n)\n\nCREATE TABLE "playlist_track"\r\n(\r\n    Playlis
tId INTEGER  NOT NULL,\r\n    TrackId INTEGER  NOT NULL,\r\n    CONSTRAINT P
K_PlaylistTrack PRIMARY KEY  (PlaylistId, TrackId),\r\n    FOREIGN KEY (Play
listId) REFERENCES "playlists" (PlaylistId) \r\n\t\tON DELETE NO ACTION ON U
PDATE NO ACTION,\r\n    FOREIGN KEY (TrackId) REFERENCES "tracks" (TrackId)
\r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_Tra
ckGenreId ON "tracks" (GenreId)\n\nCREATE INDEX IFK_PlaylistTrackTrackId ON
"playlist_track" (TrackId)\n\n\n===Additional Context \n\nIn the chinook dat
abase invoice means order\n\n===Response Guidelines \n1. If the provided con
text is sufficient, please generate a valid SQL query without any explanatio
ns for the question. \n2. If the provided context is almost sufficient but r
equires knowledge of a specific string in a particular column, please genera
te an intermediate SQL query to find the distinct strings in that column. Pr
epend the query with a comment saying intermediate_sql \n3. If the provided
context is insufficient, please explain why it can\'t be generated. \n4. Ple
ase use the most relevant table(s). \n5. If the question has been asked and
answered before, please repeat the answer exactly as it was given before.
\n'}, {'role': 'user', 'content': '  \n    List all genres and the number of
tracks in each genre:\n'}, {'role': 'assistant', 'content': 'SELECT g.GenreI
d, g.Name, COUNT(t.TrackId) AS NumberOfTracks\nFROM genres g\nLEFT JOIN trac
ks t ON g.GenreId = t.GenreId\nGROUP BY g.GenreId'}, {'role': 'user', 'conte
nt': '  \n    List all albums and their corresponding artist names  \n'},
{'role': 'assistant', 'content': 'SELECT a.Title, ar.Name AS ArtistName\nFRO
M albums a\nJOIN artists ar ON a.ArtistId = ar.ArtistId'}, {'role': 'user',
'content': '  \n    Find the top 5 most expensive tracks (based on unit pric
e):\n'}, {'role': 'assistant', 'content': 'SELECT TrackId, Name, UnitPrice F
ROM tracks ORDER BY UnitPrice DESC LIMIT 5'}, {'role': 'user', 'content': 'C
an you list all tables in the SQLite database catalog?'}, {'role': 'assistan
t', 'content': "SELECT name FROM sqlite_master WHERE type = 'table'"}, {'rol
e': 'user', 'content': 'How many customers are there'}, {'role': 'assistan
t', 'content': 'SELECT COUNT(*) AS NumberOfCustomers FROM customers'}, {'rol
e': 'user', 'content': '  \n    Get the total number of invoices for each cu
stomer\n'}, {'role': 'assistant', 'content': 'SELECT CustomerId, COUNT(*) AS
NumberOfInvoices\nFROM invoices\nGROUP BY CustomerId'}, {'role': 'user', 'co
```

```
ntent': '  \n    Find all tracks with a name containing "What" (case-insensi
tive)\n'}, {'role': 'assistant', 'content': "SELECT DISTINCT LOWER(name) AS
Pattern FROM tracks WHERE lower(Name) LIKE '%what%'"}, {'role': 'user', 'con
tent': '  \n    Get the average invoice total for each customer:\n'}, {'rol
e': 'assistant', 'content': 'SELECT CustomerId, AVG(Total) AS AverageInvoice
Total\nFROM invoices\nGROUP BY CustomerId'}, {'role': 'user', 'content': '
\n    List all invoices with a total exceeding $10:\n'}, {'role': 'assistan
t', 'content': 'SELECT * FROM invoices WHERE Total > 10.0'}, {'role': 'use
r', 'content': "  \n    List all employees and their reporting manager's nam
e (if any):\n"}, {'role': 'assistant', 'content': 'SELECT e1.FirstName AS Em
ployeeFirstname, e2.FirstName AS ManagerName\nFROM employees e1\nLEFT JOIN e
mployees e2 ON e1.ReportsTo = e2.EmployeeId'}, {'role': 'user', 'content': '
\n    There are 3 tables: artists, albums and tracks, where albums and artis
ts are linked by ArtistId, albums and tracks are linked by AlbumId,\n    Can
you find the top 10 most popular artists based on the number of tracks\n'}]
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "You are a SQLite expert. Please help to gene
rate a SQL query to answer the question. Your response should ONLY be based
on the given context and follow the response guidelines and format instructi
ons. \n===Tables \nCREATE TABLE \"tracks\"\r\n(\r\n    TrackId INTEGER PRIMA
RY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(200)  NOT NULL,\r\n    A
lbumId INTEGER,\r\n    MediaTypeId INTEGER  NOT NULL,\r\n    GenreId INTEGE
R,\r\n    Composer NVARCHAR(220),\r\n    Milliseconds INTEGER  NOT NULL,\r\n
Bytes INTEGER,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    FOREIGN KEY
(AlbumId) REFERENCES \"albums\" (AlbumId) \r\n\t\tON DELETE NO ACTION ON UPD
ATE NO ACTION,\r\n    FOREIGN KEY (GenreId) REFERENCES \"genres\" (GenreId)
\r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (MediaTy
peId) REFERENCES \"media_types\" (MediaTypeId) \r\n\t\tON DELETE NO ACTION O
N UPDATE NO ACTION\r\n)\n\nCREATE TABLE \"albums\"\r\n(\r\n    AlbumId INTEG
ER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Title NVARCHAR(160)  NOT NUL
L,\r\n    ArtistId INTEGER  NOT NULL,\r\n    FOREIGN KEY (ArtistId) REFERENC
ES \"artists\" (ArtistId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r
\n)\n\nCREATE TABLE \"artists\"\r\n(\r\n    ArtistId INTEGER PRIMARY KEY AUT
OINCREMENT NOT NULL,\r\n    Name NVARCHAR(120)\r\n)\n\nCREATE INDEX IFK_Albu
mArtistId ON \"albums\" (ArtistId)\n\nCREATE INDEX IFK_TrackAlbumId ON \"tra
cks\" (AlbumId)\n\nCREATE TABLE \"playlists\"\r\n(\r\n    PlaylistId INTEGER
PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(120)\r\n)\n\nCREATE
TABLE \"genres\"\r\n(\r\n    GenreId INTEGER PRIMARY KEY AUTOINCREMENT NOT N
ULL,\r\n    Name NVARCHAR(120)\r\n)\n\nCREATE TABLE \"playlist_track\"\r\n
(\r\n    PlaylistId INTEGER  NOT NULL,\r\n    TrackId INTEGER  NOT NULL,\r\n
CONSTRAINT PK_PlaylistTrack PRIMARY KEY  (PlaylistId, TrackId),\r\n    FOREI
GN KEY (PlaylistId) REFERENCES \"playlists\" (PlaylistId) \r\n\t\tON DELETE
NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackId) REFERENCES \"tra
cks\" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREA
TE INDEX IFK_TrackGenreId ON \"tracks\" (GenreId)\n\nCREATE INDEX IFK_Playli
stTrackTrackId ON \"playlist_track\" (TrackId)\n\n\n===Additional Context \n
\nIn the chinook database invoice means order\n\n===Response Guidelines \n1.
If the provided context is sufficient, please generate a valid SQL query wit
hout any explanations for the question. \n2. If the provided context is almo
st sufficient but requires knowledge of a specific string in a particular co
lumn, please generate an intermediate SQL query to find the distinct strings
in that column. Prepend the query with a comment saying intermediate_sql \n
```

```
3. If the provided context is insufficient, please explain why it can't be g
enerated. \n4. Please use the most relevant table(s). \n5. If the question h
as been asked and answered before, please repeat the answer exactly as it wa
s given before. \n"}, {"role": "user", "content": "  \n    List all genres a
nd the number of tracks in each genre:\n"}, {"role": "assistant", "content":
"SELECT g.GenreId, g.Name, COUNT(t.TrackId) AS NumberOfTracks\nFROM genres g
\nLEFT JOIN tracks t ON g.GenreId = t.GenreId\nGROUP BY g.GenreId"}, {"rol
e": "user", "content": "  \n    List all albums and their corresponding arti
st names  \n"}, {"role": "assistant", "content": "SELECT a.Title, ar.Name AS
ArtistName\nFROM albums a\nJOIN artists ar ON a.ArtistId = ar.ArtistId"},
{"role": "user", "content": "  \n    Find the top 5 most expensive tracks (b
ased on unit price):\n"}, {"role": "assistant", "content": "SELECT TrackId,
Name, UnitPrice FROM tracks ORDER BY UnitPrice DESC LIMIT 5"}, {"role": "use
r", "content": "Can you list all tables in the SQLite database catalog?"},
{"role": "assistant", "content": "SELECT name FROM sqlite_master WHERE type
= 'table'"}, {"role": "user", "content": "How many customers are there"},
{"role": "assistant", "content": "SELECT COUNT(*) AS NumberOfCustomers FROM
customers"}, {"role": "user", "content": "  \n    Get the total number of in
voices for each customer\n"}, {"role": "assistant", "content": "SELECT Custo
merId, COUNT(*) AS NumberOfInvoices\nFROM invoices\nGROUP BY CustomerId"},
{"role": "user", "content": "  \n    Find all tracks with a name containing
\"What\" (case-insensitive)\n"}, {"role": "assistant", "content": "SELECT DI
STINCT LOWER(name) AS Pattern FROM tracks WHERE lower(Name) LIKE '%what%'"},
{"role": "user", "content": "  \n    Get the average invoice total for each
customer:\n"}, {"role": "assistant", "content": "SELECT CustomerId, AVG(Tota
l) AS AverageInvoiceTotal\nFROM invoices\nGROUP BY CustomerId"}, {"role": "u
ser", "content": "  \n    List all invoices with a total exceeding $10:\n"},
{"role": "assistant", "content": "SELECT * FROM invoices WHERE Total > 10.
0"}, {"role": "user", "content": "  \n    List all employees and their repor
ting manager's name (if any):\n"}, {"role": "assistant", "content": "SELECT
e1.FirstName AS EmployeeFirstname, e2.FirstName AS ManagerName\nFROM employe
es e1\nLEFT JOIN employees e2 ON e1.ReportsTo = e2.EmployeeId"}, {"role": "u
ser", "content": " \n    There are 3 tables: artists, albums and tracks, whe
re albums and artists are linked by ArtistId, albums and tracks are linked b
y AlbumId,\n    Can you find the top 10 most popular artists based on the nu
mber of tracks\n"}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T22:11:08.023532603Z', 'messa
ge': {'role': 'assistant', 'content': 'SELECT ar.Name AS ArtistName, COUNT
(t.TrackId) as NumberOfTracks\nFROM artists ar\nJOIN albums al ON ar.ArtistI
d = al.ArtistId\nJOIN tracks t ON al.AlbumId = t.AlbumId\nGROUP BY ar.Artist
Id\nORDER BY NumberOfTracks DESC\nLIMIT \n  10;'}, 'done_reason': 'stop', 'd
one': True, 'total_duration': 119676729734, 'load_duration': 3286248, 'promp
t_eval_count': 1478, 'prompt_eval_duration': 92941585000, 'eval_count': 85,
'eval_duration': 25506525000}
LLM Response: SELECT ar.Name AS ArtistName, COUNT(t.TrackId) as NumberOfTrac
ks
FROM artists ar
JOIN albums al ON ar.ArtistId = al.ArtistId
JOIN tracks t ON al.AlbumId = t.AlbumId
GROUP BY ar.ArtistId
ORDER BY NumberOfTracks DESC
LIMIT
  10;
Info: Output from LLM: SELECT ar.Name AS ArtistName, COUNT(t.TrackId) as Num
berOfTracks
```

```
FROM artists ar
JOIN albums al ON ar.ArtistId = al.ArtistId
JOIN tracks t ON al.AlbumId = t.AlbumId
GROUP BY ar.ArtistId
ORDER BY NumberOfTracks DESC
LIMIT
  10;
Extracted SQL: SELECT ar.Name AS ArtistName, COUNT(t.TrackId) as NumberOfTra
cks
FROM artists ar
JOIN albums al ON ar.ArtistId = al.ArtistId
JOIN tracks t ON al.AlbumId = t.AlbumId
GROUP BY ar.ArtistId
ORDER BY NumberOfTracks DESC
LIMIT
  10
SELECT ar.Name AS ArtistName, COUNT(t.TrackId) as NumberOfTracks
FROM artists ar
JOIN albums al ON ar.ArtistId = al.ArtistId
JOIN tracks t ON al.AlbumId = t.AlbumId
GROUP BY ar.ArtistId
ORDER BY NumberOfTracks DESC
LIMIT
  10
        ArtistName  NumberOfTracks
0      Iron Maiden             213
1               U2             135
2     Led Zeppelin             114
3        Metallica             112
4      Deep Purple              92
5             Lost              92
6        Pearl Jam              67
7    Lenny Kravitz              57
8  Various Artists              56
9       The Office              53
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "The following is a pandas DataFrame that con
tains the results of the query that answers the question the user asked: '
\n    There are 3 tables: artists, albums and tracks, where albums and artis
ts are linked by ArtistId, albums and tracks are linked by AlbumId,\n    Can
you find the top 10 most popular artists based on the number of tracks\n'\n
\nThe DataFrame was produced using this query: SELECT ar.Name AS ArtistName,
COUNT(t.TrackId) as NumberOfTracks\nFROM artists ar\nJOIN albums al ON ar.Ar
tistId = al.ArtistId\nJOIN tracks t ON al.AlbumId = t.AlbumId\nGROUP BY ar.A
rtistId\nORDER BY NumberOfTracks DESC\nLIMIT \n  10\n\nThe following is info
rmation about the resulting pandas DataFrame 'df': \nRunning df.dtypes give
s:\n ArtistName        object\nNumberOfTracks     int64\ndtype: object"},
{"role": "user", "content": "Can you generate the Python plotly code to char
t the results of the dataframe? Assume the data is in a pandas dataframe cal
led 'df'. If there is only one value in the dataframe, use an Indicator. Res
pond with only Python code. Do not answer with any explanations -- just the
code."}]
```

```
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T22:11:59.115007638Z', 'messa
ge': {'role': 'assistant', 'content': '```python\nimport plotly.express as p
x\n\n# Assuming df is your DataFrame\nif len(df) > 1:\n    fig = px.bar(df,
x=\'ArtistName\', y="NumberOfTracks", title=\'Top 10 Most Popular Artists\')
\n    fig.update_layout(title=\'Top 10 Most Popular Artists based on Number
of Tracks\')\n    fig.show()\nelse:\n    indicator = \'Only one artist in th
e DataFrame.\'\n```'}, 'done_reason': 'stop', 'done': True, 'total_duratio
n': 51066967969, 'load_duration': 50840123, 'prompt_eval_count': 294, 'promp
t_eval_duration': 17661780000, 'eval_count': 115, 'eval_duration': 333042520
00}
```



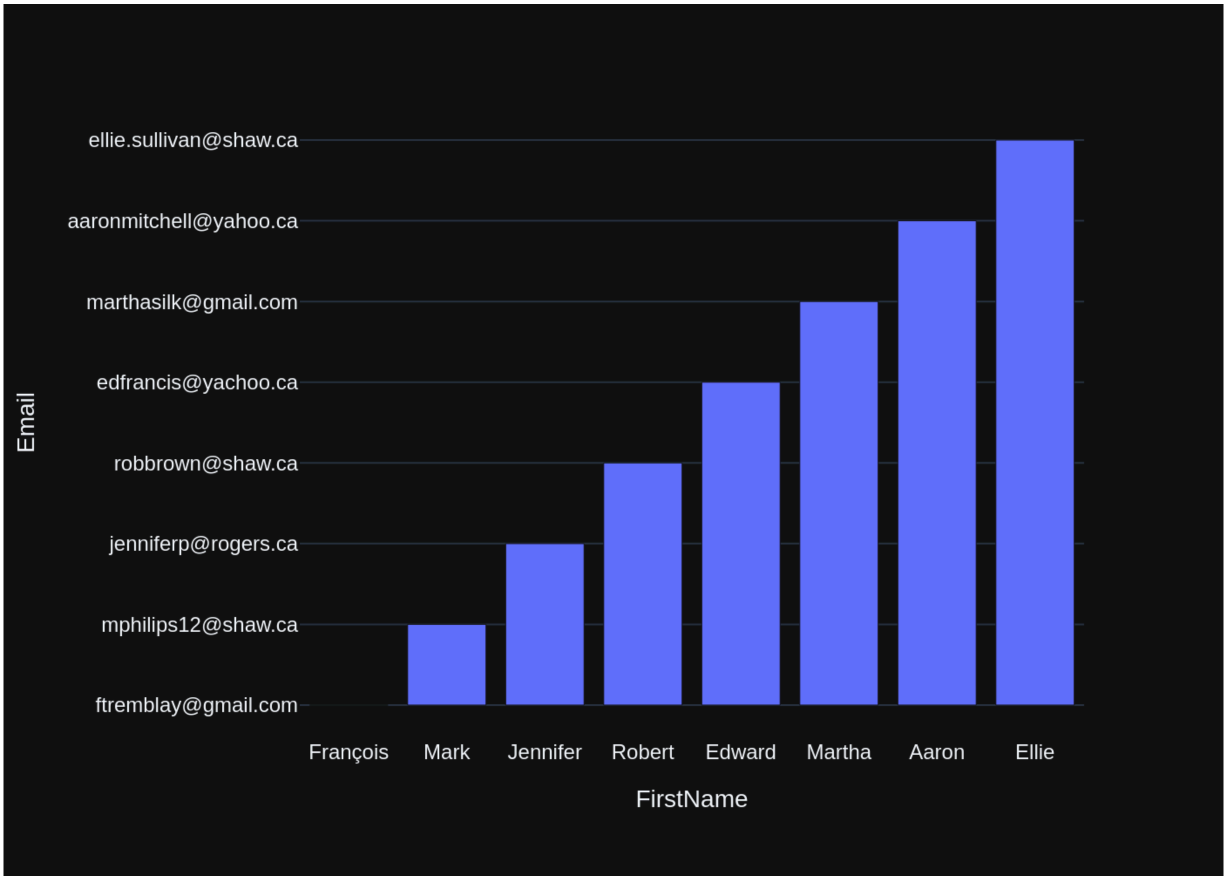
```
Out[32]: ('SELECT ar.Name AS ArtistName, COUNT(t.TrackId) as NumberOfTracks\nFROM ar
         tists ar\nJOIN albums al ON ar.ArtistId = al.ArtistId\nJOIN tracks t ON al.
         AlbumId = t.AlbumId\nGROUP BY ar.ArtistId\nORDER BY NumberOfTracks DESC\nLI
         MIT \n  10',
                  ArtistName  NumberOfTracks
          0      Iron Maiden             213
          1               U2             135
          2     Led Zeppelin             114
          3        Metallica             112
          4      Deep Purple              92
          5             Lost              92
          6        Pearl Jam              67
          7    Lenny Kravitz              57
          8  Various Artists              56
          9       The Office              53,
          Figure({
              'data': [{'alignmentgroup': 'True',
                        'hovertemplate': 'ArtistName=%{x}<br>NumberOfTracks=%{y}<ext
         ra></extra>',
                        'legendgroup': '',
                        'marker': {'color': '#636efa', 'pattern': {'shape': ''}},
                        'name': '',
                        'offsetgroup': '',
                        'orientation': 'v',
                        'showlegend': False,
                        'textposition': 'auto',
                        'type': 'bar',
                        'x': array(['Iron Maiden', 'U2', 'Led Zeppelin', 'Metallic
         a', 'Deep Purple', 'Lost',
                                    'Pearl Jam', 'Lenny Kravitz', 'Various Artists',
         'The Office'],
                                   dtype=object),
                        'xaxis': 'x',
                        'y': array([213, 135, 114, 112,  92,  92,  67,  57,  56,  5
         3]),
                        'yaxis': 'y'}],
              'layout': {'barmode': 'relative',
                         'legend': {'tracegroupgap': 0},
                         'template': '...',
                         'title': {'text': 'Top 10 Most Popular Artists based on Num
         ber of Tracks'},
                         'xaxis': {'anchor': 'y', 'domain': [0.0, 1.0], 'title': {'t
         ext': 'ArtistName'}},
                         'yaxis': {'anchor': 'x', 'domain': [0.0, 1.0], 'title': {'t
         ext': 'NumberOfTracks'}}}
          }))
```

```
In [33]: question = """
              List all customers from Canada and their email addresses:
         """

         vn.ask(question=question)
```

```
Number of requested results 10 is greater than number of elements in index
1, updating n_results = 1
```

```
SQL Prompt: [{'role': 'system', 'content': 'You are a SQLite expert. Please
help to generate a SQL query to answer the question. Your response should ON
LY be based on the given context and follow the response guidelines and form
at instructions. \n===Tables \nCREATE INDEX IFK_CustomerSupportRepId ON "cus
tomers" (SupportRepId)\n\nCREATE TABLE "customers"\r\n(\r\n    CustomerId IN
TEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    FirstName NVARCHAR(40)  NOT
NULL,\r\n    LastName NVARCHAR(20)  NOT NULL,\r\n    Company NVARCHAR(80),\r
\n    Address NVARCHAR(70),\r\n    City NVARCHAR(40),\r\n    State NVARCHAR
(40),\r\n    Country NVARCHAR(40),\r\n    PostalCode NVARCHAR(10),\r\n    Ph
one NVARCHAR(24),\r\n    Fax NVARCHAR(24),\r\n    Email NVARCHAR(60)  NOT NU
LL,\r\n    SupportRepId INTEGER,\r\n    FOREIGN KEY (SupportRepId) REFERENCE
S "employees" (EmployeeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r
\n)\n\nCREATE TABLE "invoices"\r\n(\r\n    InvoiceId INTEGER PRIMARY KEY AUT
OINCREMENT NOT NULL,\r\n    CustomerId INTEGER  NOT NULL,\r\n    InvoiceDate
DATETIME  NOT NULL,\r\n    BillingAddress NVARCHAR(70),\r\n    BillingCity N
VARCHAR(40),\r\n    BillingState NVARCHAR(40),\r\n    BillingCountry NVARCHA
R(40),\r\n    BillingPostalCode NVARCHAR(10),\r\n    Total NUMERIC(10,2)  NO
T NULL,\r\n    FOREIGN KEY (CustomerId) REFERENCES "customers" (CustomerId)
\r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_Inv
oiceCustomerId ON "invoices" (CustomerId)\n\nCREATE TABLE "employees"\r\n(\r
\n    EmployeeId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    LastName
NVARCHAR(20)  NOT NULL,\r\n    FirstName NVARCHAR(20)  NOT NULL,\r\n    Titl
e NVARCHAR(30),\r\n    ReportsTo INTEGER,\r\n    BirthDate DATETIME,\r\n
HireDate DATETIME,\r\n    Address NVARCHAR(70),\r\n    City NVARCHAR(40),\r
\n    State NVARCHAR(40),\r\n    Country NVARCHAR(40),\r\n    PostalCode NVA
RCHAR(10),\r\n    Phone NVARCHAR(24),\r\n    Fax NVARCHAR(24),\r\n    Email
NVARCHAR(60),\r\n    FOREIGN KEY (ReportsTo) REFERENCES "employees" (Employe
eId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE "i
nvoice_items"\r\n(\r\n    InvoiceLineId INTEGER PRIMARY KEY AUTOINCREMENT NO
T NULL,\r\n    InvoiceId INTEGER  NOT NULL,\r\n    TrackId INTEGER  NOT NUL
L,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    Quantity INTEGER  NOT NU
LL,\r\n    FOREIGN KEY (InvoiceId) REFERENCES "invoices" (InvoiceId) \r\n\t
\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackId) REFE
RENCES "tracks" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r
\n)\n\nCREATE TABLE sqlite_sequence(name,seq)\n\nCREATE TABLE "playlist_trac
k"\r\n(\r\n    PlaylistId INTEGER  NOT NULL,\r\n    TrackId INTEGER  NOT NUL
L,\r\n    CONSTRAINT PK_PlaylistTrack PRIMARY KEY  (PlaylistId, TrackId),\r
\n    FOREIGN KEY (PlaylistId) REFERENCES "playlists" (PlaylistId) \r\n\t\tO
N DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackId) REFEREN
CES "tracks" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)
\n\nCREATE INDEX IFK_EmployeeReportsTo ON "employees" (ReportsTo)\n\nCREATE
TABLE "albums"\r\n(\r\n    AlbumId INTEGER PRIMARY KEY AUTOINCREMENT NOT NUL
L,\r\n    Title NVARCHAR(160)  NOT NULL,\r\n    ArtistId INTEGER  NOT NUL
L,\r\n    FOREIGN KEY (ArtistId) REFERENCES "artists" (ArtistId) \r\n\t\tON
DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\n\n===Additional Context \n\nIn
the chinook database invoice means order\n\n===Response Guidelines \n1. If t
he provided context is sufficient, please generate a valid SQL query without
any explanations for the question. \n2. If the provided context is almost su
fficient but requires knowledge of a specific string in a particular column,
please generate an intermediate SQL query to find the distinct strings in th
at column. Prepend the query with a comment saying intermediate_sql \n3. If
the provided context is insufficient, please explain why it can\'t be genera
ted. \n4. Please use the most relevant table(s). \n5. If the question has be
en asked and answered before, please repeat the answer exactly as it was giv
en before. \n'}, {'role': 'user', 'content': 'How many customers are ther
e'}, {'role': 'assistant', 'content': 'SELECT COUNT(*) AS NumberOfCustomers
```

```
FROM customers'}, {'role': 'user', 'content': "which table stores customer's
orders"}, {'role': 'assistant', 'content': 'SELECT * FROM invoices'}, {'rol
e': 'user', 'content': '  \n    Get the total number of invoices for each cu
stomer\n'}, {'role': 'assistant', 'content': 'SELECT CustomerId, COUNT(*) AS
NumberOfInvoices\nFROM invoices\nGROUP BY CustomerId'}, {'role': 'user', 'co
ntent': "  \n    List all employees and their reporting manager's name (if a
ny):\n"}, {'role': 'assistant', 'content': 'SELECT e1.FirstName AS EmployeeF
irstname, e2.FirstName AS ManagerName\nFROM employees e1\nLEFT JOIN employee
s e2 ON e1.ReportsTo = e2.EmployeeId'}, {'role': 'user', 'content': '  \n
Get the average invoice total for each customer:\n'}, {'role': 'assistant',
'content': 'SELECT CustomerId, AVG(Total) AS AverageInvoiceTotal\nFROM invoi
ces\nGROUP BY CustomerId'}, {'role': 'user', 'content': '  \n    List all in
voices with a total exceeding $10:\n'}, {'role': 'assistant', 'content': 'SE
LECT * FROM invoices WHERE Total > 10.0'}, {'role': 'user', 'content': '  \n
Find all invoices since 2010 and the total amount invoiced:\n'}, {'role': 'a
ssistant', 'content': "SELECT InvoiceId, Total\nFROM invoices\nWHERE strftim
e('%Y', InvoiceDate) >= '2010'"}, {'role': 'user', 'content': 'Can you list
all tables in the SQLite database catalog?'}, {'role': 'assistant', 'conten
t': "SELECT name FROM sqlite_master WHERE type = 'table'"}, {'role': 'user',
'content': '  \n    Find the top 5 most expensive tracks (based on unit pric
e):\n'}, {'role': 'assistant', 'content': 'SELECT TrackId, Name, UnitPrice F
ROM tracks ORDER BY UnitPrice DESC LIMIT 5'}, {'role': 'user', 'content': '
\n    List all albums and their corresponding artist names  \n'}, {'role':
'assistant', 'content': 'SELECT a.Title, ar.Name AS ArtistName\nFROM albums
a\nJOIN artists ar ON a.ArtistId = ar.ArtistId'}, {'role': 'user', 'conten
t': '  \n     List all customers from Canada and their email addresses:\n'}]
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "You are a SQLite expert. Please help to gene
rate a SQL query to answer the question. Your response should ONLY be based
on the given context and follow the response guidelines and format instructi
ons. \n===Tables \nCREATE INDEX IFK_CustomerSupportRepId ON \"customers\" (S
upportRepId)\n\nCREATE TABLE \"customers\"\r\n(\r\n    CustomerId INTEGER PR
IMARY KEY AUTOINCREMENT NOT NULL,\r\n    FirstName NVARCHAR(40)  NOT NULL,\r
\n    LastName NVARCHAR(20)  NOT NULL,\r\n    Company NVARCHAR(80),\r\n    A
ddress NVARCHAR(70),\r\n    City NVARCHAR(40),\r\n    State NVARCHAR(40),\r
\n    Country NVARCHAR(40),\r\n    PostalCode NVARCHAR(10),\r\n    Phone NVA
RCHAR(24),\r\n    Fax NVARCHAR(24),\r\n    Email NVARCHAR(60)  NOT NULL,\r\n
SupportRepId INTEGER,\r\n    FOREIGN KEY (SupportRepId) REFERENCES \"employe
es\" (EmployeeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCR
EATE TABLE \"invoices\"\r\n(\r\n    InvoiceId INTEGER PRIMARY KEY AUTOINCREM
ENT NOT NULL,\r\n    CustomerId INTEGER  NOT NULL,\r\n    InvoiceDate DATETI
ME  NOT NULL,\r\n    BillingAddress NVARCHAR(70),\r\n    BillingCity NVARCHA
R(40),\r\n    BillingState NVARCHAR(40),\r\n    BillingCountry NVARCHAR(4
0),\r\n    BillingPostalCode NVARCHAR(10),\r\n    Total NUMERIC(10,2)  NOT N
ULL,\r\n    FOREIGN KEY (CustomerId) REFERENCES \"customers\" (CustomerId)
\r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_Inv
oiceCustomerId ON \"invoices\" (CustomerId)\n\nCREATE TABLE \"employees\"\r
\n(\r\n    EmployeeId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Las
tName NVARCHAR(20)  NOT NULL,\r\n    FirstName NVARCHAR(20)  NOT NULL,\r\n
Title NVARCHAR(30),\r\n    ReportsTo INTEGER,\r\n    BirthDate DATETIME,\r\n
HireDate DATETIME,\r\n    Address NVARCHAR(70),\r\n    City NVARCHAR(40),\r
\n    State NVARCHAR(40),\r\n    Country NVARCHAR(40),\r\n    PostalCode NVA
RCHAR(10),\r\n    Phone NVARCHAR(24),\r\n    Fax NVARCHAR(24),\r\n    Email 
NVARCHAR(60),\r\n    FOREIGN KEY (ReportsTo) REFERENCES \"employees\" (Emplo
yeeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE
\"invoice_items\"\r\n(\r\n    InvoiceLineId INTEGER PRIMARY KEY AUTOINCREMEN
T NOT NULL,\r\n    InvoiceId INTEGER  NOT NULL,\r\n    TrackId INTEGER  NOT
NULL,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    Quantity INTEGER  NOT
NULL,\r\n    FOREIGN KEY (InvoiceId) REFERENCES \"invoices\" (InvoiceId) \r
\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackId)
REFERENCES \"tracks\" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACT
ION\r\n)\n\nCREATE TABLE sqlite_sequence(name,seq)\n\nCREATE TABLE \"playlis
t_track\"\r\n(\r\n    PlaylistId INTEGER  NOT NULL,\r\n    TrackId INTEGER
NOT NULL,\r\n    CONSTRAINT PK_PlaylistTrack PRIMARY KEY  (PlaylistId, Track
Id),\r\n    FOREIGN KEY (PlaylistId) REFERENCES \"playlists\" (PlaylistId)
\r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackI
d) REFERENCES \"tracks\" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO
ACTION\r\n)\n\nCREATE INDEX IFK_EmployeeReportsTo ON \"employees\" (ReportsT
o)\n\nCREATE TABLE \"albums\"\r\n(\r\n    AlbumId INTEGER PRIMARY KEY AUTOIN
CREMENT NOT NULL,\r\n    Title NVARCHAR(160)  NOT NULL,\r\n    ArtistId INTE
GER  NOT NULL,\r\n    FOREIGN KEY (ArtistId) REFERENCES \"artists\" (ArtistI
d) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\n\n===Additional C
ontext \n\nIn the chinook database invoice means order\n\n===Response Guidel
ines \n1. If the provided context is sufficient, please generate a valid SQL
query without any explanations for the question. \n2. If the provided contex
t is almost sufficient but requires knowledge of a specific string in a part
icular column, please generate an intermediate SQL query to find the distinc
t strings in that column. Prepend the query with a comment saying intermedia
te_sql \n3. If the provided context is insufficient, please explain why it c
an't be generated. \n4. Please use the most relevant table(s). \n5. If the q
uestion has been asked and answered before, please repeat the answer exactly
as it was given before. \n"}, {"role": "user", "content": "How many customer
s are there"}, {"role": "assistant", "content": "SELECT COUNT(*) AS NumberOf
Customers FROM customers"}, {"role": "user", "content": "which table stores
customer's orders"}, {"role": "assistant", "content": "SELECT * FROM invoice
s"}, {"role": "user", "content": "  \n    Get the total number of invoices f
or each customer\n"}, {"role": "assistant", "content": "SELECT CustomerId, C
OUNT(*) AS NumberOfInvoices\nFROM invoices\nGROUP BY CustomerId"}, {"role":
"user", "content": "  \n    List all employees and their reporting manager's
name (if any):\n"}, {"role": "assistant", "content": "SELECT e1.FirstName AS
EmployeeFirstname, e2.FirstName AS ManagerName\nFROM employees e1\nLEFT JOIN
employees e2 ON e1.ReportsTo = e2.EmployeeId"}, {"role": "user", "content":
"  \n    Get the average invoice total for each customer:\n"}, {"role": "ass
istant", "content": "SELECT CustomerId, AVG(Total) AS AverageInvoiceTotal\nF
ROM invoices\nGROUP BY CustomerId"}, {"role": "user", "content": "  \n    Li
st all invoices with a total exceeding $10:\n"}, {"role": "assistant", "cont
ent": "SELECT * FROM invoices WHERE Total > 10.0"}, {"role": "user", "conten
t": "  \n    Find all invoices since 2010 and the total amount invoice
d:\n"}, {"role": "assistant", "content": "SELECT InvoiceId, Total\nFROM invo
ices\nWHERE strftime('%Y', InvoiceDate) >= '2010'"}, {"role": "user", "conte
nt": "Can you list all tables in the SQLite database catalog?"}, {"role": "a
ssistant", "content": "SELECT name FROM sqlite_master WHERE type = 'tabl
e'"}, {"role": "user", "content": "  \n    Find the top 5 most expensive tra
cks (based on unit price):\n"}, {"role": "assistant", "content": "SELECT Tra
ckId, Name, UnitPrice FROM tracks ORDER BY UnitPrice DESC LIMIT 5"}, {"rol
e": "user", "content": "  \n    List all albums and their corresponding arti
st names  \n"}, {"role": "assistant", "content": "SELECT a.Title, ar.Name AS
ArtistName\nFROM albums a\nJOIN artists ar ON a.ArtistId = ar.ArtistId"},
{"role": "user", "content": "  \n     List all customers from Canada and the
ir email addresses:\n"}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T22:14:00.066085881Z', 'messa
ge': {'role': 'assistant', 'content': "SELECT FirstName, LastName, Email\nFR
OM customers\nWHERE Country = 'Canada'"}, 'done_reason': 'stop', 'done': Tru
e, 'total_duration': 120857716853, 'load_duration': 3463799, 'prompt_eval_co
unt': 1794, 'prompt_eval_duration': 113861009000, 'eval_count': 20, 'eval_du
ration': 5757836000}
LLM Response: SELECT FirstName, LastName, Email
FROM customers
WHERE Country = 'Canada'
SELECT FirstName, LastName, Email
FROM customers
WHERE Country = 'Canada'
  FirstName  LastName                   Email
0  François  Tremblay     ftremblay@gmail.com
1      Mark   Philips      mphilips12@shaw.ca
2  Jennifer  Peterson     jenniferp@rogers.ca
3    Robert     Brown        robbrown@shaw.ca
4    Edward   Francis     edfrancis@yachoo.ca
5    Martha      Silk    marthasilk@gmail.com
6     Aaron  Mitchell  aaronmitchell@yahoo.ca
7     Ellie  Sullivan  ellie.sullivan@shaw.ca
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "The following is a pandas DataFrame that con
tains the results of the query that answers the question the user asked: '
\n     List all customers from Canada and their email addresses:\n'\n\nThe D
ataFrame was produced using this query: SELECT FirstName, LastName, Email\nF
ROM customers\nWHERE Country = 'Canada'\n\nThe following is information abou
t the resulting pandas DataFrame 'df': \nRunning df.dtypes gives:\n FirstNam
e    object\nLastName     object\nEmail        object\ndtype: object"}, {"ro
le": "user", "content": "Can you generate the Python plotly code to chart th
e results of the dataframe? Assume the data is in a pandas dataframe called
'df'. If there is only one value in the dataframe, use an Indicator. Respond
with only Python code. Do not answer with any explanations -- just the cod
e."}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T22:14:42.21502909Z', 'messag
e': {'role': 'assistant', 'content': "```python\nimport plotly.express as px
\n\n# Check if DataFrame has more than one row and contains 'FirstName', 'La
stName' columns\nif len(df) > 1:\n    fig = px.bar(df, x='FirstName', y='Ema
il')\nelse:\n    # If only one value in the dataframe, use an Indicator\n
fig = px.indicator(values=[len(df)])\n    \nfig.show()\n```"}, 'done_reaso
n': 'stop', 'done': True, 'total_duration': 42124281203, 'load_duration': 44
067868, 'prompt_eval_count': 188, 'prompt_eval_duration': 11148982000, 'eval
_count': 107, 'eval_duration': 30881495000}
```


ellie.sullivan@shaw.ca
aaronmitchell@yahoo.ca
marthasilk@gmail.com
edfrancis@yachoo.ca
robbrown@shaw.ca
jenniferp@rogers.ca
mphilips12@shaw.ca
ftremblay@gmail.com
Email
François
Mark
Jennifer
Robert
Edward
Martha
Aaron
Ellie
FirstName

Out[33]:

```
("SELECT FirstName, LastName, Email\nFROM customers\nWHERE Country = 'Canad
a'",
   FirstName  LastName                    Email
 0  François  Tremblay     ftremblay@gmail.com
 1      Mark   Philips      mphilips12@shaw.ca
 2  Jennifer  Peterson     jenniferp@rogers.ca
 3    Robert     Brown        robbrown@shaw.ca
 4    Edward   Francis     edfrancis@yachoo.ca
 5    Martha      Silk    marthasilk@gmail.com
 6     Aaron  Mitchell  aaronmitchell@yahoo.ca
 7     Ellie  Sullivan  ellie.sullivan@shaw.ca,
 Figure({
     'data': [{'alignmentgroup': 'True',
               'hovertemplate': 'FirstName=%{x}<br>Email=%{y}<extra></extra
>',
               'legendgroup': '',
               'marker': {'color': '#636efa', 'pattern': {'shape': ''}},
               'name': '',
               'offsetgroup': '',
               'orientation': 'v',
               'showlegend': False,
               'textposition': 'auto',
               'type': 'bar',
               'x': array(['François', 'Mark', 'Jennifer', 'Robert', 'Edwar
d', 'Martha', 'Aaron',
                           'Ellie'], dtype=object),
               'xaxis': 'x',
               'y': array(['ftremblay@gmail.com', 'mphilips12@shaw.ca', 'je
nniferp@rogers.ca',
                           'robbrown@shaw.ca', 'edfrancis@yachoo.ca', 'mart
hasilk@gmail.com',
                           'aaronmitchell@yahoo.ca', 'ellie.sullivan@shaw.c
a'], dtype=object),
               'yaxis': 'y'}],
     'layout': {'barmode': 'relative',
                'legend': {'tracegroupgap': 0},
                'margin': {'t': 60},
                'template': '...',
                'xaxis': {'anchor': 'y', 'domain': [0.0, 1.0], 'title': {'t
ext': 'FirstName'}},
                'yaxis': {'anchor': 'x', 'domain': [0.0, 1.0], 'title': {'t
ext': 'Email'}}}
 }))
```

In [34]:

```
question = """
     Find the customer with the most invoices
"""

vn.ask(question=question)
```

```
Number of requested results 10 is greater than number of elements in index
1, updating n_results = 1
```

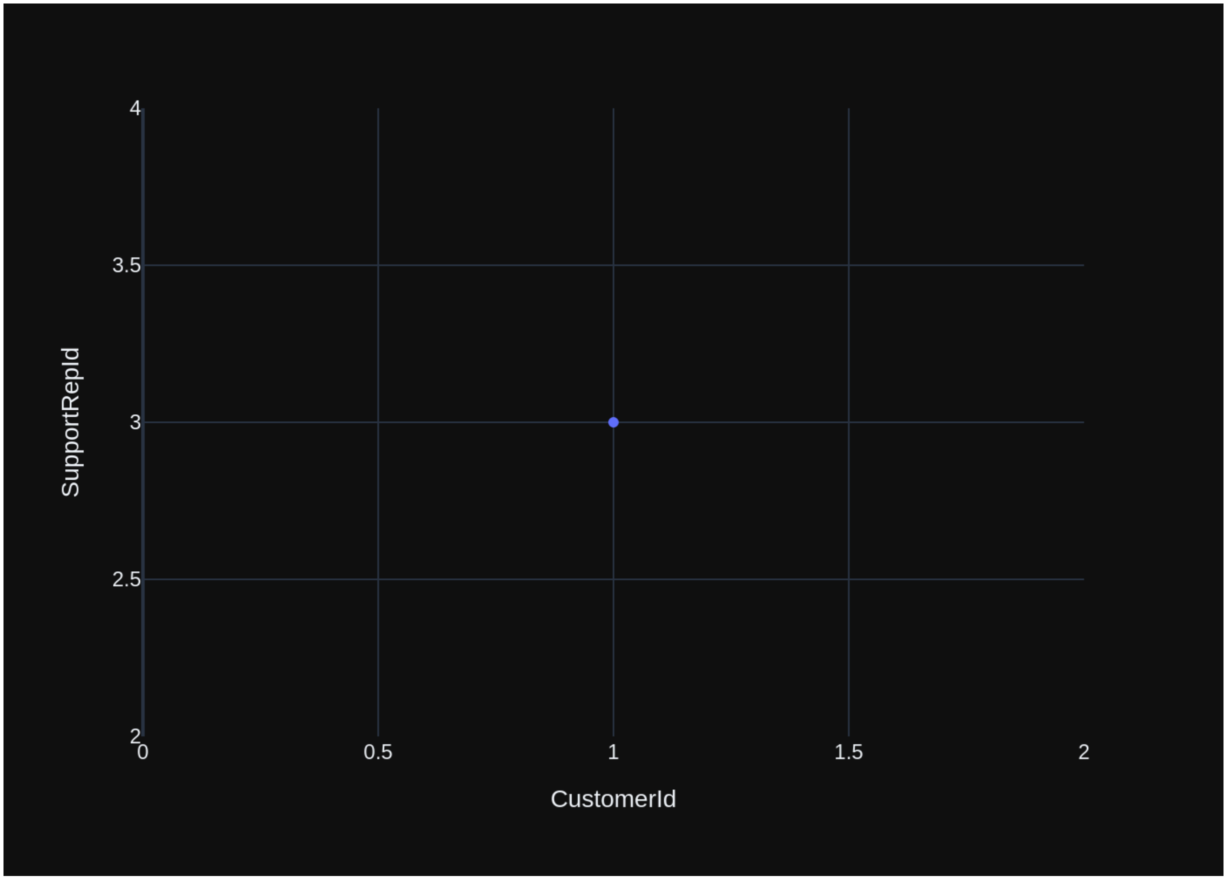
```
SQL Prompt: [{'role': 'system', 'content': 'You are a SQLite expert. Please
help to generate a SQL query to answer the question. Your response should ON
LY be based on the given context and follow the response guidelines and form
at instructions. \n===Tables \nCREATE TABLE "invoices"\r\n(\r\n    InvoiceId
INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    CustomerId INTEGER  NOT N
ULL,\r\n    InvoiceDate DATETIME  NOT NULL,\r\n    BillingAddress NVARCHAR(7
0),\r\n    BillingCity NVARCHAR(40),\r\n    BillingState NVARCHAR(40),\r\n
BillingCountry NVARCHAR(40),\r\n    BillingPostalCode NVARCHAR(10),\r\n    T
otal NUMERIC(10,2)  NOT NULL,\r\n    FOREIGN KEY (CustomerId) REFERENCES "cu
stomers" (CustomerId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n
\nCREATE INDEX IFK_InvoiceCustomerId ON "invoices" (CustomerId)\n\nCREATE IN
DEX IFK_InvoiceLineInvoiceId ON "invoice_items" (InvoiceId)\n\nCREATE TABLE
"invoice_items"\r\n(\r\n    InvoiceLineId INTEGER PRIMARY KEY AUTOINCREMENT
NOT NULL,\r\n    InvoiceId INTEGER  NOT NULL,\r\n    TrackId INTEGER  NOT NU
LL,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    Quantity INTEGER  NOT N
ULL,\r\n    FOREIGN KEY (InvoiceId) REFERENCES "invoices" (InvoiceId) \r\n\t
\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackId) REFE
RENCES "tracks" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r
\n)\n\nCREATE INDEX IFK_InvoiceLineTrackId ON "invoice_items" (TrackId)\n\nC
REATE TABLE "customers"\r\n(\r\n    CustomerId INTEGER PRIMARY KEY AUTOINCRE
MENT NOT NULL,\r\n    FirstName NVARCHAR(40)  NOT NULL,\r\n    LastName NVAR
CHAR(20)  NOT NULL,\r\n    Company NVARCHAR(80),\r\n    Address NVARCHAR(7
0),\r\n    City NVARCHAR(40),\r\n    State NVARCHAR(40),\r\n    Country NVAR
CHAR(40),\r\n    PostalCode NVARCHAR(10),\r\n    Phone NVARCHAR(24),\r\n
Fax NVARCHAR(24),\r\n    Email NVARCHAR(60)  NOT NULL,\r\n    SupportRepId I
NTEGER,\r\n    FOREIGN KEY (SupportRepId) REFERENCES "employees" (EmployeeI
d) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_
CustomerSupportRepId ON "customers" (SupportRepId)\n\nCREATE TABLE "employee
s"\r\n(\r\n    EmployeeId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n
LastName NVARCHAR(20)  NOT NULL,\r\n    FirstName NVARCHAR(20)  NOT NULL,\r
\n    Title NVARCHAR(30),\r\n    ReportsTo INTEGER,\r\n    BirthDate DATETIM
E,\r\n    HireDate DATETIME,\r\n    Address NVARCHAR(70),\r\n    City NVARCH
AR(40),\r\n    State NVARCHAR(40),\r\n    Country NVARCHAR(40),\r\n    Posta
lCode NVARCHAR(10),\r\n    Phone NVARCHAR(24),\r\n    Fax NVARCHAR(24),\r\n
Email NVARCHAR(60),\r\n    FOREIGN KEY (ReportsTo) REFERENCES "employees" (E
mployeeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE IN
DEX IFK_EmployeeReportsTo ON "employees" (ReportsTo)\n\nCREATE TABLE "track
s"\r\n(\r\n    TrackId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Na
me NVARCHAR(200)  NOT NULL,\r\n    AlbumId INTEGER,\r\n    MediaTypeId INTEG
ER  NOT NULL,\r\n    GenreId INTEGER,\r\n    Composer NVARCHAR(220),\r\n
Milliseconds INTEGER  NOT NULL,\r\n    Bytes INTEGER,\r\n    UnitPrice NUMER
IC(10,2)  NOT NULL,\r\n    FOREIGN KEY (AlbumId) REFERENCES "albums" (AlbumI
d) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (Genr
eId) REFERENCES "genres" (GenreId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO
ACTION,\r\n    FOREIGN KEY (MediaTypeId) REFERENCES "media_types" (MediaType
Id) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\n\n===Additional
Context \n\nIn the chinook database invoice means order\n\n===Response Guide
lines \n1. If the provided context is sufficient, please generate a valid SQ
L query without any explanations for the question. \n2. If the provided cont
ext is almost sufficient but requires knowledge of a specific string in a pa
rticular column, please generate an intermediate SQL query to find the disti
nct strings in that column. Prepend the query with a comment saying intermed
iate_sql \n3. If the provided context is insufficient, please explain why it
can\'t be generated. \n4. Please use the most relevant table(s). \n5. If the
question has been asked and answered before, please repeat the answer exactl
y as it was given before. \n'}, {'role': 'user', 'content': '  \n    Get the
```

```
total number of invoices for each customer\n'}, {'role': 'assistant', 'conte
nt': 'SELECT CustomerId, COUNT(*) AS NumberOfInvoices\nFROM invoices\nGROUP
BY CustomerId'}, {'role': 'user', 'content': "which table stores customer's
orders"}, {'role': 'assistant', 'content': 'SELECT * FROM invoices'}, {'rol
e': 'user', 'content': '  \n    List all invoices with a total exceeding $1
0:\n'}, {'role': 'assistant', 'content': 'SELECT * FROM invoices WHERE Total
> 10.0'}, {'role': 'user', 'content': '  \n    Get the average invoice total
for each customer:\n'}, {'role': 'assistant', 'content': 'SELECT CustomerId,
AVG(Total) AS AverageInvoiceTotal\nFROM invoices\nGROUP BY CustomerId'}, {'r
ole': 'user', 'content': '  \n    Find all invoices since 2010 and the total
amount invoiced:\n'}, {'role': 'assistant', 'content': "SELECT InvoiceId, To
tal\nFROM invoices\nWHERE strftime('%Y', InvoiceDate) >= '2010'"}, {'role':
'user', 'content': '  \n    Find the top 5 most expensive tracks (based on u
nit price):\n'}, {'role': 'assistant', 'content': 'SELECT TrackId, Name, Uni
tPrice FROM tracks ORDER BY UnitPrice DESC LIMIT 5'}, {'role': 'user', 'cont
ent': 'How many customers are there'}, {'role': 'assistant', 'content': 'SEL
ECT COUNT(*) AS NumberOfCustomers FROM customers'}, {'role': 'user', 'conten
t': '  \n     List all customers from Canada and their email addresses:\n'},
{'role': 'assistant', 'content': "SELECT FirstName, LastName, Email\nFROM cu
stomers\nWHERE Country = 'Canada'"}, {'role': 'user', 'content': "  \n    Li
st all employees and their reporting manager's name (if any):\n"}, {'role':
'assistant', 'content': 'SELECT e1.FirstName AS EmployeeFirstname, e2.FirstN
ame AS ManagerName\nFROM employees e1\nLEFT JOIN employees e2 ON e1.ReportsT
o = e2.EmployeeId'}, {'role': 'user', 'content': ' \n    There are 3 tables:
artists, albums and tracks, where albums and artists are linked by ArtistId,
albums and tracks are linked by AlbumId,\n    Can you find the top 10 most p
opular artists based on the number of tracks\n'}, {'role': 'assistant', 'con
tent': 'SELECT ar.Name AS ArtistName, COUNT(t.TrackId) as NumberOfTracks\nFR
OM artists ar\nJOIN albums al ON ar.ArtistId = al.ArtistId\nJOIN tracks t ON
al.AlbumId = t.AlbumId\nGROUP BY ar.ArtistId\nORDER BY NumberOfTracks DESC\n
LIMIT \n  10'}, {'role': 'user', 'content': '  \n     Find the customer with
the most invoices \n'}]
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "You are a SQLite expert. Please help to gene
rate a SQL query to answer the question. Your response should ONLY be based
on the given context and follow the response guidelines and format instructi
ons. \n===Tables \nCREATE TABLE \"invoices\"\r\n(\r\n    InvoiceId INTEGER P
RIMARY KEY AUTOINCREMENT NOT NULL,\r\n    CustomerId INTEGER  NOT NULL,\r\n
InvoiceDate DATETIME  NOT NULL,\r\n    BillingAddress NVARCHAR(70),\r\n    B
illingCity NVARCHAR(40),\r\n    BillingState NVARCHAR(40),\r\n    BillingCou
ntry NVARCHAR(40),\r\n    BillingPostalCode NVARCHAR(10),\r\n    Total NUMER
IC(10,2)  NOT NULL,\r\n    FOREIGN KEY (CustomerId) REFERENCES \"customers\"
(CustomerId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE
INDEX IFK_InvoiceCustomerId ON \"invoices\" (CustomerId)\n\nCREATE INDEX IFK
_InvoiceLineInvoiceId ON \"invoice_items\" (InvoiceId)\n\nCREATE TABLE \"inv
oice_items\"\r\n(\r\n    InvoiceLineId INTEGER PRIMARY KEY AUTOINCREMENT NOT
NULL,\r\n    InvoiceId INTEGER  NOT NULL,\r\n    TrackId INTEGER  NOT NUL
L,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    Quantity INTEGER  NOT NU
LL,\r\n    FOREIGN KEY (InvoiceId) REFERENCES \"invoices\" (InvoiceId) \r\n
\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackId) RE
FERENCES \"tracks\" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTIO
N\r\n)\n\nCREATE INDEX IFK_InvoiceLineTrackId ON \"invoice_items\" (TrackId)
```

```
\n\nCREATE TABLE \"customers\"\r\n(\r\n    CustomerId INTEGER PRIMARY KEY AU
TOINCREMENT NOT NULL,\r\n    FirstName NVARCHAR(40)  NOT NULL,\r\n    LastNa
me NVARCHAR(20)  NOT NULL,\r\n    Company NVARCHAR(80),\r\n    Address NVARC
HAR(70),\r\n    City NVARCHAR(40),\r\n    State NVARCHAR(40),\r\n    Country
NVARCHAR(40),\r\n    PostalCode NVARCHAR(10),\r\n    Phone NVARCHAR(24),\r\n
Fax NVARCHAR(24),\r\n    Email NVARCHAR(60)  NOT NULL,\r\n    SupportRepId I
NTEGER,\r\n    FOREIGN KEY (SupportRepId) REFERENCES \"employees\" (Employee
Id) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK
_CustomerSupportRepId ON \"customers\" (SupportRepId)\n\nCREATE TABLE \"empl
oyees\"\r\n(\r\n    EmployeeId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r
\n    LastName NVARCHAR(20)  NOT NULL,\r\n    FirstName NVARCHAR(20)  NOT NU
LL,\r\n    Title NVARCHAR(30),\r\n    ReportsTo INTEGER,\r\n    BirthDate DA
TETIME,\r\n    HireDate DATETIME,\r\n    Address NVARCHAR(70),\r\n    City N
VARCHAR(40),\r\n    State NVARCHAR(40),\r\n    Country NVARCHAR(40),\r\n
PostalCode NVARCHAR(10),\r\n    Phone NVARCHAR(24),\r\n    Fax NVARCHAR(2
4),\r\n    Email NVARCHAR(60),\r\n    FOREIGN KEY (ReportsTo) REFERENCES \"e
mployees\" (EmployeeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)
\n\nCREATE INDEX IFK_EmployeeReportsTo ON \"employees\" (ReportsTo)\n\nCREAT
E TABLE \"tracks\"\r\n(\r\n    TrackId INTEGER PRIMARY KEY AUTOINCREMENT NOT
NULL,\r\n    Name NVARCHAR(200)  NOT NULL,\r\n    AlbumId INTEGER,\r\n    Me
diaTypeId INTEGER  NOT NULL,\r\n    GenreId INTEGER,\r\n    Composer NVARCHA
R(220),\r\n    Milliseconds INTEGER  NOT NULL,\r\n    Bytes INTEGER,\r\n
UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    FOREIGN KEY (AlbumId) REFERENCES
\"albums\" (AlbumId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n
FOREIGN KEY (GenreId) REFERENCES \"genres\" (GenreId) \r\n\t\tON DELETE NO A
CTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (MediaTypeId) REFERENCES \"med
ia_types\" (MediaTypeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r
\n)\n\n\n===Additional Context \n\nIn the chinook database invoice means ord
er\n\n===Response Guidelines \n1. If the provided context is sufficient, ple
ase generate a valid SQL query without any explanations for the question. \n
2. If the provided context is almost sufficient but requires knowledge of a
specific string in a particular column, please generate an intermediate SQL
query to find the distinct strings in that column. Prepend the query with a
comment saying intermediate_sql \n3. If the provided context is insufficien
t, please explain why it can't be generated. \n4. Please use the most releva
nt table(s). \n5. If the question has been asked and answered before, please
repeat the answer exactly as it was given before. \n"}, {"role": "user", "co
ntent": "  \n    Get the total number of invoices for each customer\n"}, {"r
ole": "assistant", "content": "SELECT CustomerId, COUNT(*) AS NumberOfInvoic
es\nFROM invoices\nGROUP BY CustomerId"}, {"role": "user", "content": "which
table stores customer's orders"}, {"role": "assistant", "content": "SELECT *
FROM invoices"}, {"role": "user", "content": "  \n    List all invoices with
a total exceeding $10:\n"}, {"role": "assistant", "content": "SELECT * FROM
invoices WHERE Total > 10.0"}, {"role": "user", "content": "  \n    Get the
average invoice total for each customer:\n"}, {"role": "assistant", "conten
t": "SELECT CustomerId, AVG(Total) AS AverageInvoiceTotal\nFROM invoices\nGR
OUP BY CustomerId"}, {"role": "user", "content": "  \n    Find all invoices
since 2010 and the total amount invoiced:\n"}, {"role": "assistant", "conten
t": "SELECT InvoiceId, Total\nFROM invoices\nWHERE strftime('%Y', InvoiceDat
e) >= '2010'"}, {"role": "user", "content": "  \n    Find the top 5 most exp
ensive tracks (based on unit price):\n"}, {"role": "assistant", "content":
"SELECT TrackId, Name, UnitPrice FROM tracks ORDER BY UnitPrice DESC LIMIT
5"}, {"role": "user", "content": "How many customers are there"}, {"role":
"assistant", "content": "SELECT COUNT(*) AS NumberOfCustomers FROM customer
s"}, {"role": "user", "content": "  \n     List all customers from Canada an
d their email addresses:\n"}, {"role": "assistant", "content": "SELECT First
```

```
Name, LastName, Email\nFROM customers\nWHERE Country = 'Canada'"}, {"role":
"user", "content": "  \n    List all employees and their reporting manager's
name (if any):\n"}, {"role": "assistant", "content": "SELECT e1.FirstName AS
EmployeeFirstname, e2.FirstName AS ManagerName\nFROM employees e1\nLEFT JOIN
employees e2 ON e1.ReportsTo = e2.EmployeeId"}, {"role": "user", "content":
" \n    There are 3 tables: artists, albums and tracks, where albums and art
ists are linked by ArtistId, albums and tracks are linked by AlbumId,\n    C
an you find the top 10 most popular artists based on the number of tracks
\n"}, {"role": "assistant", "content": "SELECT ar.Name AS ArtistName, COUNT
(t.TrackId) as NumberOfTracks\nFROM artists ar\nJOIN albums al ON ar.ArtistI
d = al.ArtistId\nJOIN tracks t ON al.AlbumId = t.AlbumId\nGROUP BY ar.Artist
Id\nORDER BY NumberOfTracks DESC\nLIMIT \n  10"}, {"role": "user", "conten
t": "  \n     Find the customer with the most invoices \n"}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T22:17:03.983993835Z', 'messa
ge': {'role': 'assistant', 'content': 'SELECT c.*, COUNT(i.InvoiceId) AS Inv
oiceCount\nFROM customers c JOIN invoices i ON c.CustomerId = i.CustomerId\n
GROUP BY c.CustomerId\nORDER BY InvoiceCount DESC\nLIMIT 1'}, 'done_reason':
'stop', 'done': True, 'total_duration': 141671171238, 'load_duration': 36280
95, 'prompt_eval_count': 1942, 'prompt_eval_duration': 123947471000, 'eval_c
ount': 55, 'eval_duration': 16492047000}
LLM Response: SELECT c.*, COUNT(i.InvoiceId) AS InvoiceCount
FROM customers c JOIN invoices i ON c.CustomerId = i.CustomerId
GROUP BY c.CustomerId
ORDER BY InvoiceCount DESC
LIMIT 1
SELECT c.*, COUNT(i.InvoiceId) AS InvoiceCount
FROM customers c JOIN invoices i ON c.CustomerId = i.CustomerId
GROUP BY c.CustomerId
ORDER BY InvoiceCount DESC
LIMIT 1
   CustomerId FirstName   LastName  \
0           1      Luís  Gonçalves

                                            Company  \
0  Embraer - Empresa Brasileira de Aeronáutica S.A.

                           Address                 City State Country  \
0  Av. Brigadeiro Faria Lima, 2170  São José dos Campos    SP  Brazil

  PostalCode               Phone                 Fax                 Email
\
0  12227-000  +55 (12) 3923-5555  +55 (12) 3923-5566  luisg@embraer.com.br

   SupportRepId  InvoiceCount
0             3             7
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "The following is a pandas DataFrame that con
tains the results of the query that answers the question the user asked: '
\n     Find the customer with the most invoices \n'\n\nThe DataFrame was pro
duced using this query: SELECT c.*, COUNT(i.InvoiceId) AS InvoiceCount\nFROM
customers c JOIN invoices i ON c.CustomerId = i.CustomerId\nGROUP BY c.Custo
```

```
merId\nORDER BY InvoiceCount DESC\nLIMIT 1\n\nThe following is information a
bout the resulting pandas DataFrame 'df': \nRunning df.dtypes gives:\n Custo
merId       int64\nFirstName        object\nLastName        object\nCompany  
object\nAddress        object\nCity            object\nState           obje
ct\nCountry         object\nPostalCode      object\nPhone           object\n
Fax             object\nEmail           object\nSupportRepId     int64\nInvo
iceCount     int64\ndtype: object"}, {"role": "user", "content": "Can you ge
nerate the Python plotly code to chart the results of the dataframe? Assume
the data is in a pandas dataframe called 'df'. If there is only one value in
the dataframe, use an Indicator. Respond with only Python code. Do not answe
r with any explanations -- just the code."}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T22:17:57.359145473Z', 'messa
ge': {'role': 'assistant', 'content': "```python\nimport plotly.express as p
x\n\n# Check if df has only one row\nif len(df) == 1:\n    fig = px.scatter_
pie(df, labels={'CustomerId': 'Customer ID', 'InvoiceCount': 'Number of Invo
ices'}, title='Single Customer with the Most Invoices')\nelse:\n    fig = p
x.bar(df, x=['InvoiceCount'], y=['CustomerId'], title='Top Customers by Numb
er of Invoices')\nfig.show()\n```"}, 'done_reason': 'stop', 'done': True, 't
otal_duration': 53348952462, 'load_duration': 48529520, 'prompt_eval_count':
280, 'prompt_eval_duration': 16887891000, 'eval_count': 125, 'eval_duratio
n': 36362079000}
```



Out[34]:

```
('SELECT c.*, COUNT(i.InvoiceId) AS InvoiceCount\nFROM customers c JOIN inv
oices i ON c.CustomerId = i.CustomerId\nGROUP BY c.CustomerId\nORDER BY Inv
oiceCount DESC\nLIMIT 1',
    CustomerId FirstName   LastName  \
 0           1      Luís  Gonçalves

                                             Company  \
 0  Embraer - Empresa Brasileira de Aeronáutica S.A.

                            Address                 City State Country  \
 0  Av. Brigadeiro Faria Lima, 2170  São José dos Campos    SP  Brazil

   PostalCode               Phone                 Fax                 Email
\
 0  12227-000  +55 (12) 3923-5555  +55 (12) 3923-5566  luisg@embraer.com.br

    SupportRepId  InvoiceCount
 0             3             7  ,
 Figure({
     'data': [{'hovertemplate': 'CustomerId=%{x}<br>SupportRepId=%{y}<extra
></extra>',
               'legendgroup': '',
               'marker': {'color': '#636efa', 'symbol': 'circle'},
               'mode': 'markers',
               'name': '',
               'orientation': 'v',
               'showlegend': False,
               'type': 'scatter',
               'x': array([1]),
               'xaxis': 'x',
               'y': array([3]),
               'yaxis': 'y'}],
     'layout': {'legend': {'tracegroupgap': 0},
                'margin': {'t': 60},
                'template': '...',
                'xaxis': {'anchor': 'y', 'domain': [0.0, 1.0], 'title': {'t
ext': 'CustomerId'}},
                'yaxis': {'anchor': 'x', 'domain': [0.0, 1.0], 'title': {'t
ext': 'SupportRepId'}}}
 }))
```

In [ ]:

Advanced SQL questions

In [35]:

```
question = """
     Find the customer who bought the most albums in total quantity (across
"""

vn.ask(question=question)
```

```
Number of requested results 10 is greater than number of elements in index
1, updating n_results = 1
```

```
SQL Prompt: [{'role': 'system', 'content': 'You are a SQLite expert. Please
help to generate a SQL query to answer the question. Your response should ON
LY be based on the given context and follow the response guidelines and form
at instructions. \n===Tables \nCREATE TABLE "tracks"\r\n(\r\n    TrackId INT
EGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(200)  NOT NUL
L,\r\n    AlbumId INTEGER,\r\n    MediaTypeId INTEGER  NOT NULL,\r\n    Genr
eId INTEGER,\r\n    Composer NVARCHAR(220),\r\n    Milliseconds INTEGER  NOT
NULL,\r\n    Bytes INTEGER,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n
FOREIGN KEY (AlbumId) REFERENCES "albums" (AlbumId) \r\n\t\tON DELETE NO ACT
ION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (GenreId) REFERENCES "genres" (G
enreId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY
(MediaTypeId) REFERENCES "media_types" (MediaTypeId) \r\n\t\tON DELETE NO AC
TION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE "invoice_items"\r\n(\r\n    In
voiceLineId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    InvoiceId INT
EGER  NOT NULL,\r\n    TrackId INTEGER  NOT NULL,\r\n    UnitPrice NUMERIC(1
0,2)  NOT NULL,\r\n    Quantity INTEGER  NOT NULL,\r\n    FOREIGN KEY (Invoi
ceId) REFERENCES "invoices" (InvoiceId) \r\n\t\tON DELETE NO ACTION ON UPDAT
E NO ACTION,\r\n    FOREIGN KEY (TrackId) REFERENCES "tracks" (TrackId) \r\n
\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE "albums"\r
\n(\r\n    AlbumId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Title
NVARCHAR(160)  NOT NULL,\r\n    ArtistId INTEGER  NOT NULL,\r\n    FOREIGN K
EY (ArtistId) REFERENCES "artists" (ArtistId) \r\n\t\tON DELETE NO ACTION ON
UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_AlbumArtistId ON "albums" (ArtistI
d)\n\nCREATE TABLE "invoices"\r\n(\r\n    InvoiceId INTEGER PRIMARY KEY AUTO
INCREMENT NOT NULL,\r\n    CustomerId INTEGER  NOT NULL,\r\n    InvoiceDate
DATETIME  NOT NULL,\r\n    BillingAddress NVARCHAR(70),\r\n    BillingCity N
VARCHAR(40),\r\n    BillingState NVARCHAR(40),\r\n    BillingCountry NVARCHA
R(40),\r\n    BillingPostalCode NVARCHAR(10),\r\n    Total NUMERIC(10,2)  NO
T NULL,\r\n    FOREIGN KEY (CustomerId) REFERENCES "customers" (CustomerId)
\r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_Inv
oiceLineTrackId ON "invoice_items" (TrackId)\n\nCREATE INDEX IFK_InvoiceLine
InvoiceId ON "invoice_items" (InvoiceId)\n\nCREATE INDEX IFK_InvoiceCustomer
Id ON "invoices" (CustomerId)\n\nCREATE INDEX IFK_TrackAlbumId ON "tracks"
(AlbumId)\n\nCREATE TABLE "artists"\r\n(\r\n    ArtistId INTEGER PRIMARY KEY
AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(120)\r\n)\n\n\n===Additional Co
ntext \n\nIn the chinook database invoice means order\n\n===Response Guideli
nes \n1. If the provided context is sufficient, please generate a valid SQL
query without any explanations for the question. \n2. If the provided contex
t is almost sufficient but requires knowledge of a specific string in a part
icular column, please generate an intermediate SQL query to find the distinc
t strings in that column. Prepend the query with a comment saying intermedia
te_sql \n3. If the provided context is insufficient, please explain why it c
an\'t be generated. \n4. Please use the most relevant table(s). \n5. If the
question has been asked and answered before, please repeat the answer exactl
y as it was given before. \n'}, {'role': 'user', 'content': '  \n     Find t
he customer with the most invoices \n'}, {'role': 'assistant', 'content': 'S
ELECT c.*, COUNT(i.InvoiceId) AS InvoiceCount\nFROM customers c JOIN invoice
s i ON c.CustomerId = i.CustomerId\nGROUP BY c.CustomerId\nORDER BY InvoiceC
ount DESC\nLIMIT 1'}, {'role': 'user', 'content': ' \n    There are 3 table
s: artists, albums and tracks, where albums and artists are linked by Artist
Id, albums and tracks are linked by AlbumId,\n    Can you find the top 10 mo
st popular artists based on the number of tracks\n'}, {'role': 'assistant',
'content': 'SELECT ar.Name AS ArtistName, COUNT(t.TrackId) as NumberOfTracks
\nFROM artists ar\nJOIN albums al ON ar.ArtistId = al.ArtistId\nJOIN tracks
t ON al.AlbumId = t.AlbumId\nGROUP BY ar.ArtistId\nORDER BY NumberOfTracks D
ESC\nLIMIT \n  10'}, {'role': 'user', 'content': '  \n    Get the total numb
```

```
er of invoices for each customer\n'}, {'role': 'assistant', 'content': 'SELE
CT CustomerId, COUNT(*) AS NumberOfInvoices\nFROM invoices\nGROUP BY Custome
rId'}, {'role': 'user', 'content': '  \n    Find the top 5 most expensive tr
acks (based on unit price):\n'}, {'role': 'assistant', 'content': 'SELECT Tr
ackId, Name, UnitPrice FROM tracks ORDER BY UnitPrice DESC LIMIT 5'}, {'rol
e': 'user', 'content': '  \n    List all invoices with a total exceeding $1
0:\n'}, {'role': 'assistant', 'content': 'SELECT * FROM invoices WHERE Total
> 10.0'}, {'role': 'user', 'content': '  \n    Get the average invoice total
for each customer:\n'}, {'role': 'assistant', 'content': 'SELECT CustomerId,
AVG(Total) AS AverageInvoiceTotal\nFROM invoices\nGROUP BY CustomerId'}, {'r
ole': 'user', 'content': '  \n    Find all invoices since 2010 and the total
amount invoiced:\n'}, {'role': 'assistant', 'content': "SELECT InvoiceId, To
tal\nFROM invoices\nWHERE strftime('%Y', InvoiceDate) >= '2010'"}, {'role':
'user', 'content': '  \n    List all albums and their corresponding artist n
ames  \n'}, {'role': 'assistant', 'content': 'SELECT a.Title, ar.Name AS Art
istName\nFROM albums a\nJOIN artists ar ON a.ArtistId = ar.ArtistId'}, {'rol
e': 'user', 'content': "which table stores customer's orders"}, {'role': 'as
sistant', 'content': 'SELECT * FROM invoices'}, {'role': 'user', 'content':
'  \n    List all genres and the number of tracks in each genre:\n'}, {'rol
e': 'assistant', 'content': 'SELECT g.GenreId, g.Name, COUNT(t.TrackId) AS N
umberOfTracks\nFROM genres g\nLEFT JOIN tracks t ON g.GenreId = t.GenreId\nG
ROUP BY g.GenreId'}, {'role': 'user', 'content': '  \n     Find the customer
who bought the most albums in total quantity (across all invoices): \n'}]
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "You are a SQLite expert. Please help to gene
rate a SQL query to answer the question. Your response should ONLY be based
on the given context and follow the response guidelines and format instructi
ons. \n===Tables \nCREATE TABLE \"tracks\"\r\n(\r\n    TrackId INTEGER PRIMA
RY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(200)  NOT NULL,\r\n    A
lbumId INTEGER,\r\n    MediaTypeId INTEGER  NOT NULL,\r\n    GenreId INTEGE
R,\r\n    Composer NVARCHAR(220),\r\n    Milliseconds INTEGER  NOT NULL,\r\n
Bytes INTEGER,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    FOREIGN KEY
(AlbumId) REFERENCES \"albums\" (AlbumId) \r\n\t\tON DELETE NO ACTION ON UPD
ATE NO ACTION,\r\n    FOREIGN KEY (GenreId) REFERENCES \"genres\" (GenreId)
\r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (MediaTy
peId) REFERENCES \"media_types\" (MediaTypeId) \r\n\t\tON DELETE NO ACTION O
N UPDATE NO ACTION\r\n)\n\nCREATE TABLE \"invoice_items\"\r\n(\r\n    Invoic
eLineId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    InvoiceId INTEGER
NOT NULL,\r\n    TrackId INTEGER  NOT NULL,\r\n    UnitPrice NUMERIC(10,2)
NOT NULL,\r\n    Quantity INTEGER  NOT NULL,\r\n    FOREIGN KEY (InvoiceId)
REFERENCES \"invoices\" (InvoiceId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO
ACTION,\r\n    FOREIGN KEY (TrackId) REFERENCES \"tracks\" (TrackId) \r\n\t
\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE \"albums\"\r
\n(\r\n    AlbumId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Title
NVARCHAR(160)  NOT NULL,\r\n    ArtistId INTEGER  NOT NULL,\r\n    FOREIGN K
EY (ArtistId) REFERENCES \"artists\" (ArtistId) \r\n\t\tON DELETE NO ACTION
ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_AlbumArtistId ON \"albums\" (Ar
tistId)\n\nCREATE TABLE \"invoices\"\r\n(\r\n    InvoiceId INTEGER PRIMARY K
EY AUTOINCREMENT NOT NULL,\r\n    CustomerId INTEGER  NOT NULL,\r\n    Invoi
ceDate DATETIME  NOT NULL,\r\n    BillingAddress NVARCHAR(70),\r\n    Billin
gCity NVARCHAR(40),\r\n    BillingState NVARCHAR(40),\r\n    BillingCountry
NVARCHAR(40),\r\n    BillingPostalCode NVARCHAR(10),\r\n    Total NUMERIC(1
0,2)  NOT NULL,\r\n    FOREIGN KEY (CustomerId) REFERENCES \"customers\" (Cu
stomerId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE IND
EX IFK_InvoiceLineTrackId ON \"invoice_items\" (TrackId)\n\nCREATE INDEX IFK
_InvoiceLineInvoiceId ON \"invoice_items\" (InvoiceId)\n\nCREATE INDEX IFK_I
nvoiceCustomerId ON \"invoices\" (CustomerId)\n\nCREATE INDEX IFK_TrackAlbum
Id ON \"tracks\" (AlbumId)\n\nCREATE TABLE \"artists\"\r\n(\r\n    ArtistId 
INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(120)\r\n)\n
\n\n===Additional Context \n\nIn the chinook database invoice means order\n
\n===Response Guidelines \n1. If the provided context is sufficient, please 
generate a valid SQL query without any explanations for the question. \n2. I
f the provided context is almost sufficient but requires knowledge of a spec
ific string in a particular column, please generate an intermediate SQL quer
y to find the distinct strings in that column. Prepend the query with a comm
ent saying intermediate_sql \n3. If the provided context is insufficient, pl
ease explain why it can't be generated. \n4. Please use the most relevant ta
ble(s). \n5. If the question has been asked and answered before, please repe
at the answer exactly as it was given before. \n"}, {"role": "user", "conten
t": "  \n     Find the customer with the most invoices \n"}, {"role": "assis
tant", "content": "SELECT c.*, COUNT(i.InvoiceId) AS InvoiceCount\nFROM cust
omers c JOIN invoices i ON c.CustomerId = i.CustomerId\nGROUP BY c.CustomerI
d\nORDER BY InvoiceCount DESC\nLIMIT 1"}, {"role": "user", "content": " \n 
There are 3 tables: artists, albums and tracks, where albums and artists are
linked by ArtistId, albums and tracks are linked by AlbumId,\n    Can you fi
nd the top 10 most popular artists based on the number of tracks\n"}, {"rol
e": "assistant", "content": "SELECT ar.Name AS ArtistName, COUNT(t.TrackId) 
as NumberOfTracks\nFROM artists ar\nJOIN albums al ON ar.ArtistId = al.Artis
tId\nJOIN tracks t ON al.AlbumId = t.AlbumId\nGROUP BY ar.ArtistId\nORDER BY
NumberOfTracks DESC\nLIMIT \n  10"}, {"role": "user", "content": "  \n    Ge
t the total number of invoices for each customer\n"}, {"role": "assistant", 
"content": "SELECT CustomerId, COUNT(*) AS NumberOfInvoices\nFROM invoices\n
GROUP BY CustomerId"}, {"role": "user", "content": "  \n    Find the top 5 m
ost expensive tracks (based on unit price):\n"}, {"role": "assistant", "cont
ent": "SELECT TrackId, Name, UnitPrice FROM tracks ORDER BY UnitPrice DESC L
IMIT 5"}, {"role": "user", "content": "  \n    List all invoices with a tota
l exceeding $10:\n"}, {"role": "assistant", "content": "SELECT * FROM invoic
es WHERE Total > 10.0"}, {"role": "user", "content": "  \n    Get the averag
e invoice total for each customer:\n"}, {"role": "assistant", "content": "SE
LECT CustomerId, AVG(Total) AS AverageInvoiceTotal\nFROM invoices\nGROUP BY 
CustomerId"}, {"role": "user", "content": "  \n    Find all invoices since 2
010 and the total amount invoiced:\n"}, {"role": "assistant", "content": "SE
LECT InvoiceId, Total\nFROM invoices\nWHERE strftime('%Y', InvoiceDate) >= 
'2010'"}, {"role": "user", "content": "  \n    List all albums and their cor
responding artist names  \n"}, {"role": "assistant", "content": "SELECT a.Ti
tle, ar.Name AS ArtistName\nFROM albums a\nJOIN artists ar ON a.ArtistId = a
r.ArtistId"}, {"role": "user", "content": "which table stores customer's ord
ers"}, {"role": "assistant", "content": "SELECT * FROM invoices"}, {"role": 
"user", "content": "  \n    List all genres and the number of tracks in each
genre:\n"}, {"role": "assistant", "content": "SELECT g.GenreId, g.Name, COUN
T(t.TrackId) AS NumberOfTracks\nFROM genres g\nLEFT JOIN tracks t ON g.Genre
Id = t.GenreId\nGROUP BY g.GenreId"}, {"role": "user", "content": "  \n    
Find the customer who bought the most albums in total quantity (across all i
nvoices): \n"}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T22:20:09.944554945Z', 'messa
ge': {'role': 'assistant', 'content': 'SELECT CustomerId, SUM(i.Quantity) AS
TotalAlbumsPurchased\nFROM invoice_items i\nJOIN tracks t ON i.TrackId = t.T
```

```
rackId\nGROUP BY CustomerId\nORDER BY TotalAlbumsPurchased DESC\nLIMIT 1'},
'done_reason': 'stop', 'done': True, 'total_duration': 132499127797, 'load_d
uration': 3857470, 'prompt_eval_count': 1723, 'prompt_eval_duration': 111583
030000, 'eval_count': 63, 'eval_duration': 19686276000}
LLM Response: SELECT CustomerId, SUM(i.Quantity) AS TotalAlbumsPurchased
FROM invoice_items i
JOIN tracks t ON i.TrackId = t.TrackId
GROUP BY CustomerId
ORDER BY TotalAlbumsPurchased DESC
LIMIT 1
SELECT CustomerId, SUM(i.Quantity) AS TotalAlbumsPurchased
FROM invoice_items i
JOIN tracks t ON i.TrackId = t.TrackId
GROUP BY CustomerId
ORDER BY TotalAlbumsPurchased DESC
LIMIT 1
Couldn't run sql:  Execution failed on sql 'SELECT CustomerId, SUM(i.Quantit
y) AS TotalAlbumsPurchased
FROM invoice_items i
JOIN tracks t ON i.TrackId = t.TrackId
GROUP BY CustomerId
ORDER BY TotalAlbumsPurchased DESC
LIMIT 1': no such column: CustomerId
```

In [36]:

```python
question = """
    Hint: album quantity is found in invoice_items,

    Find the top 5 customers who bought the most albums in total quantity (a
"""

vn.ask(question=question)
```

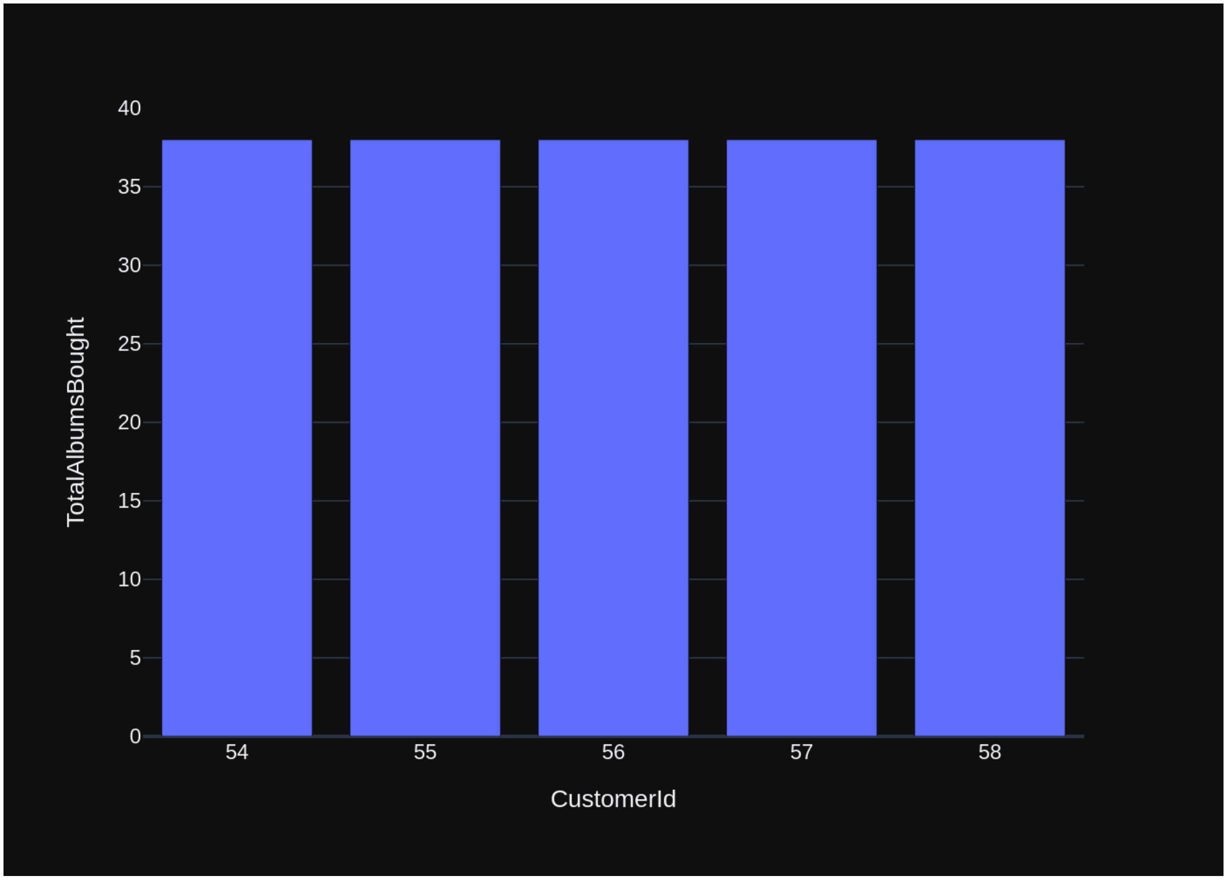
```
Number of requested results 10 is greater than number of elements in index
1, updating n_results = 1
```

```
SQL Prompt: [{'role': 'system', 'content': 'You are a SQLite expert. Please
help to generate a SQL query to answer the question. Your response should ON
LY be based on the given context and follow the response guidelines and form
at instructions. \n===Tables \nCREATE TABLE "invoice_items"\r\n(\r\n    Invo
iceLineId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    InvoiceId INTEG
ER  NOT NULL,\r\n    TrackId INTEGER  NOT NULL,\r\n    UnitPrice NUMERIC(10,
2)  NOT NULL,\r\n    Quantity INTEGER  NOT NULL,\r\n    FOREIGN KEY (Invoice
Id) REFERENCES "invoices" (InvoiceId) \r\n\t\tON DELETE NO ACTION ON UPDATE
NO ACTION,\r\n    FOREIGN KEY (TrackId) REFERENCES "tracks" (TrackId) \r\n\t
\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE "tracks"\r\n
(\r\n    TrackId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVA
RCHAR(200)  NOT NULL,\r\n    AlbumId INTEGER,\r\n    MediaTypeId INTEGER  NO
T NULL,\r\n    GenreId INTEGER,\r\n    Composer NVARCHAR(220),\r\n    Millis
econds INTEGER  NOT NULL,\r\n    Bytes INTEGER,\r\n    UnitPrice NUMERIC(10,
2)  NOT NULL,\r\n    FOREIGN KEY (AlbumId) REFERENCES "albums" (AlbumId) \r
\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (GenreId)
REFERENCES "genres" (GenreId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTIO
N,\r\n    FOREIGN KEY (MediaTypeId) REFERENCES "media_types" (MediaTypeId)
\r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE "album
s"\r\n(\r\n    AlbumId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Ti
tle NVARCHAR(160)  NOT NULL,\r\n    ArtistId INTEGER  NOT NULL,\r\n    FOREI
GN KEY (ArtistId) REFERENCES "artists" (ArtistId) \r\n\t\tON DELETE NO ACTIO
N ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_AlbumArtistId ON "albums" (Ar
tistId)\n\nCREATE INDEX IFK_InvoiceLineInvoiceId ON "invoice_items" (Invoice
Id)\n\nCREATE INDEX IFK_InvoiceLineTrackId ON "invoice_items" (TrackId)\n\nC
REATE TABLE "invoices"\r\n(\r\n    InvoiceId INTEGER PRIMARY KEY AUTOINCREME
NT NOT NULL,\r\n    CustomerId INTEGER  NOT NULL,\r\n    InvoiceDate DATETIM
E  NOT NULL,\r\n    BillingAddress NVARCHAR(70),\r\n    BillingCity NVARCHAR
(40),\r\n    BillingState NVARCHAR(40),\r\n    BillingCountry NVARCHAR(4
0),\r\n    BillingPostalCode NVARCHAR(10),\r\n    Total NUMERIC(10,2)  NOT N
ULL,\r\n    FOREIGN KEY (CustomerId) REFERENCES "customers" (CustomerId) \r
\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_Invoi
ceCustomerId ON "invoices" (CustomerId)\n\nCREATE INDEX IFK_TrackAlbumId ON
"tracks" (AlbumId)\n\nCREATE TABLE "artists"\r\n(\r\n    ArtistId INTEGER PR
IMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(120)\r\n)\n\n\n===Add
itional Context \n\nIn the chinook database invoice means order\n\n===Respon
se Guidelines \n1. If the provided context is sufficient, please generate a
valid SQL query without any explanations for the question. \n2. If the provi
ded context is almost sufficient but requires knowledge of a specific string
in a particular column, please generate an intermediate SQL query to find th
e distinct strings in that column. Prepend the query with a comment saying i
ntermediate_sql \n3. If the provided context is insufficient, please explain
why it can\'t be generated. \n4. Please use the most relevant table(s). \n5.
If the question has been asked and answered before, please repeat the answer
exactly as it was given before. \n'}, {'role': 'user', 'content': '  \n
Find the customer with the most invoices \n'}, {'role': 'assistant', 'conten
t': 'SELECT c.*, COUNT(i.InvoiceId) AS InvoiceCount\nFROM customers c JOIN i
nvoices i ON c.CustomerId = i.CustomerId\nGROUP BY c.CustomerId\nORDER BY In
voiceCount DESC\nLIMIT 1'}, {'role': 'user', 'content': ' \n    There are 3
tables: artists, albums and tracks, where albums and artists are linked by A
rtistId, albums and tracks are linked by AlbumId,\n    Can you find the top
10 most popular artists based on the number of tracks\n'}, {'role': 'assista
nt', 'content': 'SELECT ar.Name AS ArtistName, COUNT(t.TrackId) as NumberOfT
racks\nFROM artists ar\nJOIN albums al ON ar.ArtistId = al.ArtistId\nJOIN tr
acks t ON al.AlbumId = t.AlbumId\nGROUP BY ar.ArtistId\nORDER BY NumberOfTra
cks DESC\nLIMIT \n  10'}, {'role': 'user', 'content': '  \n    Find the top
```

```
5 most expensive tracks (based on unit price):\n'}, {'role': 'assistant', 'c
ontent': 'SELECT TrackId, Name, UnitPrice FROM tracks ORDER BY UnitPrice DES
C LIMIT 5'}, {'role': 'user', 'content': '  \n    List all invoices with a t
otal exceeding $10:\n'}, {'role': 'assistant', 'content': 'SELECT * FROM inv
oices WHERE Total > 10.0'}, {'role': 'user', 'content': '  \n    Get the tot
al number of invoices for each customer\n'}, {'role': 'assistant', 'conten
t': 'SELECT CustomerId, COUNT(*) AS NumberOfInvoices\nFROM invoices\nGROUP B
Y CustomerId'}, {'role': 'user', 'content': '  \n    Get the average invoice
total for each customer:\n'}, {'role': 'assistant', 'content': 'SELECT Custo
merId, AVG(Total) AS AverageInvoiceTotal\nFROM invoices\nGROUP BY CustomerI
d'}, {'role': 'user', 'content': '  \n    List all albums and their correspo
nding artist names  \n'}, {'role': 'assistant', 'content': 'SELECT a.Title,
ar.Name AS ArtistName\nFROM albums a\nJOIN artists ar ON a.ArtistId = ar.Art
istId'}, {'role': 'user', 'content': '  \n    Find all invoices since 2010 a
nd the total amount invoiced:\n'}, {'role': 'assistant', 'content': "SELECT
InvoiceId, Total\nFROM invoices\nWHERE strftime('%Y', InvoiceDate) >= '201
0'"}, {'role': 'user', 'content': "which table stores customer's orders"},
{'role': 'assistant', 'content': 'SELECT * FROM invoices'}, {'role': 'user',
'content': '  \n    List all genres and the number of tracks in each genr
e:\n'}, {'role': 'assistant', 'content': 'SELECT g.GenreId, g.Name, COUNT(t.
TrackId) AS NumberOfTracks\nFROM genres g\nLEFT JOIN tracks t ON g.GenreId =
t.GenreId\nGROUP BY g.GenreId'}, {'role': 'user', 'content': '  \n    Hint:
album quantity is found in invoice_items, \n    \n    Find the top 5 custome
rs who bought the most albums in total quantity (across all invoices):\n'}]
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "You are a SQLite expert. Please help to gene
rate a SQL query to answer the question. Your response should ONLY be based
on the given context and follow the response guidelines and format instructi
ons. \n===Tables \nCREATE TABLE \"invoice_items\"\r\n(\r\n    InvoiceLineId
INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    InvoiceId INTEGER  NOT NU
LL,\r\n    TrackId INTEGER  NOT NULL,\r\n    UnitPrice NUMERIC(10,2)  NOT NU
LL,\r\n    Quantity INTEGER  NOT NULL,\r\n    FOREIGN KEY (InvoiceId) REFERE
NCES \"invoices\" (InvoiceId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTIO
N,\r\n    FOREIGN KEY (TrackId) REFERENCES \"tracks\" (TrackId) \r\n\t\tON D
ELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE \"tracks\"\r\n(\r\n
TrackId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(20
0)  NOT NULL,\r\n    AlbumId INTEGER,\r\n    MediaTypeId INTEGER  NOT NUL
L,\r\n    GenreId INTEGER,\r\n    Composer NVARCHAR(220),\r\n    Millisecond
s INTEGER  NOT NULL,\r\n    Bytes INTEGER,\r\n    UnitPrice NUMERIC(10,2)  N
OT NULL,\r\n    FOREIGN KEY (AlbumId) REFERENCES \"albums\" (AlbumId) \r\n\t
\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (GenreId) REFE
RENCES \"genres\" (GenreId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTIO
N,\r\n    FOREIGN KEY (MediaTypeId) REFERENCES \"media_types\" (MediaTypeId)
\r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE \"album
s\"\r\n(\r\n    AlbumId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    T
itle NVARCHAR(160)  NOT NULL,\r\n    ArtistId INTEGER  NOT NULL,\r\n    FORE
IGN KEY (ArtistId) REFERENCES \"artists\" (ArtistId) \r\n\t\tON DELETE NO AC
TION ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_AlbumArtistId ON \"albums
\" (ArtistId)\n\nCREATE INDEX IFK_InvoiceLineInvoiceId ON \"invoice_items\"
(InvoiceId)\n\nCREATE INDEX IFK_InvoiceLineTrackId ON \"invoice_items\" (Tra
ckId)\n\nCREATE TABLE \"invoices\"\r\n(\r\n    InvoiceId INTEGER PRIMARY KEY
AUTOINCREMENT NOT NULL,\r\n    CustomerId INTEGER  NOT NULL,\r\n    InvoiceD
ate DATETIME  NOT NULL,\r\n    BillingAddress NVARCHAR(70),\r\n    BillingCi
ty NVARCHAR(40),\r\n    BillingState NVARCHAR(40),\r\n    BillingCountry NVA
RCHAR(40),\r\n    BillingPostalCode NVARCHAR(10),\r\n    Total NUMERIC(10,2)
NOT NULL,\r\n    FOREIGN KEY (CustomerId) REFERENCES \"customers\" (Customer
Id) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK
_InvoiceCustomerId ON \"invoices\" (CustomerId)\n\nCREATE INDEX IFK_TrackAlb
umId ON \"tracks\" (AlbumId)\n\nCREATE TABLE \"artists\"\r\n(\r\n    ArtistI
d INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(120)\r\n)
\n\n\n===Additional Context \n\nIn the chinook database invoice means order 
\n\n===Response Guidelines \n1. If the provided context is sufficient, pleas
e generate a valid SQL query without any explanations for the question. \n2.
If the provided context is almost sufficient but requires knowledge of a spe
cific string in a particular column, please generate an intermediate SQL que
ry to find the distinct strings in that column. Prepend the query with a com
ment saying intermediate_sql \n3. If the provided context is insufficient, p
lease explain why it can't be generated. \n4. Please use the most relevant t
able(s). \n5. If the question has been asked and answered before, please rep
eat the answer exactly as it was given before. \n"}, {"role": "user", "conte
nt": "  \n     Find the customer with the most invoices \n"}, {"role": "assi
stant", "content": "SELECT c.*, COUNT(i.InvoiceId) AS InvoiceCount\nFROM cus
tomers c JOIN invoices i ON c.CustomerId = i.CustomerId\nGROUP BY c.Customer
Id\nORDER BY InvoiceCount DESC\nLIMIT 1"}, {"role": "user", "content": " \n 
There are 3 tables: artists, albums and tracks, where albums and artists are
linked by ArtistId, albums and tracks are linked by AlbumId,\n    Can you fi
nd the top 10 most popular artists based on the number of tracks\n"}, {"rol
e": "assistant", "content": "SELECT ar.Name AS ArtistName, COUNT(t.TrackId) 
as NumberOfTracks\nFROM artists ar\nJOIN albums al ON ar.ArtistId = al.Artis
tId\nJOIN tracks t ON al.AlbumId = t.AlbumId\nGROUP BY ar.ArtistId\nORDER BY
NumberOfTracks DESC\nLIMIT \n  10"}, {"role": "user", "content": "  \n    Fi
nd the top 5 most expensive tracks (based on unit price):\n"}, {"role": "ass
istant", "content": "SELECT TrackId, Name, UnitPrice FROM tracks ORDER BY Un
itPrice DESC LIMIT 5"}, {"role": "user", "content": "  \n    List all invoic
es with a total exceeding $10:\n"}, {"role": "assistant", "content": "SELECT
* FROM invoices WHERE Total > 10.0"}, {"role": "user", "content": "  \n    G
et the total number of invoices for each customer\n"}, {"role": "assistant",
"content": "SELECT CustomerId, COUNT(*) AS NumberOfInvoices\nFROM invoices\n
GROUP BY CustomerId"}, {"role": "user", "content": "  \n    Get the average 
invoice total for each customer:\n"}, {"role": "assistant", "content": "SELE
CT CustomerId, AVG(Total) AS AverageInvoiceTotal\nFROM invoices\nGROUP BY Cu
stomerId"}, {"role": "user", "content": "  \n    List all albums and their c
orresponding artist names  \n"}, {"role": "assistant", "content": "SELECT a.
Title, ar.Name AS ArtistName\nFROM albums a\nJOIN artists ar ON a.ArtistId =
ar.ArtistId"}, {"role": "user", "content": "  \n    Find all invoices since 
2010 and the total amount invoiced:\n"}, {"role": "assistant", "content": "S
ELECT InvoiceId, Total\nFROM invoices\nWHERE strftime('%Y', InvoiceDate) >= 
'2010'"}, {"role": "user", "content": "which table stores customer's order
s"}, {"role": "assistant", "content": "SELECT * FROM invoices"}, {"role": "u
ser", "content": "  \n    List all genres and the number of tracks in each g
enre:\n"}, {"role": "assistant", "content": "SELECT g.GenreId, g.Name, COUNT
(t.TrackId) AS NumberOfTracks\nFROM genres g\nLEFT JOIN tracks t ON g.GenreI
d = t.GenreId\nGROUP BY g.GenreId"}, {"role": "user", "content": "  \n    Hi
nt: album quantity is found in invoice_items, \n    \n    Find the top 5 cus
tomers who bought the most albums in total quantity (across all invoice
s):\n"}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T22:22:34.972710425Z', 'messa
```

```
ge': {'role': 'assistant', 'content': 'SELECT i.CustomerId, SUM(ii.Quantity)
AS TotalAlbumsBought\nFROM invoices i\nJOIN invoice_items ii ON i.InvoiceId
= ii.InvoiceId\nJOIN albums al ON ii.TrackId IN (SELECT TrackId FROM tracks
WHERE AlbumId = al.AlbumId)\nGROUP BY i.CustomerId\nORDER BY TotalAlbumsBoug
ht DESC \nLIMIT 5'}, 'done_reason': 'stop', 'done': True, 'total_duration':
144982852806, 'load_duration': 5063113, 'prompt_eval_count': 1743, 'prompt_e
val_duration': 113661638000, 'eval_count': 97, 'eval_duration': 30091693000}
LLM Response: SELECT i.CustomerId, SUM(ii.Quantity) AS TotalAlbumsBought
FROM invoices i
JOIN invoice_items ii ON i.InvoiceId = ii.InvoiceId
JOIN albums al ON ii.TrackId IN (SELECT TrackId FROM tracks WHERE AlbumId =
al.AlbumId)
GROUP BY i.CustomerId
ORDER BY TotalAlbumsBought DESC
LIMIT 5
SELECT i.CustomerId, SUM(ii.Quantity) AS TotalAlbumsBought
FROM invoices i
JOIN invoice_items ii ON i.InvoiceId = ii.InvoiceId
JOIN albums al ON ii.TrackId IN (SELECT TrackId FROM tracks WHERE AlbumId =
al.AlbumId)
GROUP BY i.CustomerId
ORDER BY TotalAlbumsBought DESC
LIMIT 5
   CustomerId  TotalAlbumsBought
0          58                 38
1          57                 38
2          56                 38
3          55                 38
4          54                 38
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "The following is a pandas DataFrame that con
tains the results of the query that answers the question the user asked: '
\n    Hint: album quantity is found in invoice_items, \n    \n    Find the t
op 5 customers who bought the most albums in total quantity (across all invo
ices):\n'\n\nThe DataFrame was produced using this query: SELECT i.CustomerI
d, SUM(ii.Quantity) AS TotalAlbumsBought\nFROM invoices i\nJOIN invoice_item
s ii ON i.InvoiceId = ii.InvoiceId\nJOIN albums al ON ii.TrackId IN (SELECT
TrackId FROM tracks WHERE AlbumId = al.AlbumId)\nGROUP BY i.CustomerId\nORDE
R BY TotalAlbumsBought DESC \nLIMIT 5\n\nThe following is information about
the resulting pandas DataFrame 'df': \nRunning df.dtypes gives:\n CustomerId
int64\nTotalAlbumsBought    int64\ndtype: object"}, {"role": "user", "conten
t": "Can you generate the Python plotly code to chart the results of the dat
aframe? Assume the data is in a pandas dataframe called 'df'. If there is on
ly one value in the dataframe, use an Indicator. Respond with only Python co
de. Do not answer with any explanations -- just the code."}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T22:23:31.114344914Z', 'messa
ge': {'role': 'assistant', 'content': '```python\nimport plotly.express as p
x\n\n# Assuming \'CustomerId\' and \'TotalAlbumsBought\' are the columns in
df DataFrame.\nif len(df) > 1:\n    fig = px.bar(df, x=\'CustomerId\', y=\'T
otalAlbumsBought\')\nelse:\n    # If there is only one value in the datafram
e use an Indicator\n    fig = px.indicator(data_frame=df[[\'CustomerId\']],
```

```
title="Top 5 customers")\nfig.show()\n```'}, 'done_reason': 'stop', 'done':
True, 'total_duration': 56112913973, 'load_duration': 45353546, 'prompt_eval
_count': 299, 'prompt_eval_duration': 18247097000, 'eval_count': 126, 'eval_
duration': 37766507000}
```



```
Out[36]: ('SELECT i.CustomerId, SUM(ii.Quantity) AS TotalAlbumsBought\nFROM invoices
         i\nJOIN invoice_items ii ON i.InvoiceId = ii.InvoiceId\nJOIN albums al ON i
         i.TrackId IN (SELECT TrackId FROM tracks WHERE AlbumId = al.AlbumId)\nGROUP
         BY i.CustomerId\nORDER BY TotalAlbumsBought DESC \nLIMIT 5',
             CustomerId  TotalAlbumsBought
          0          58                 38
          1          57                 38
          2          56                 38
          3          55                 38
          4          54                 38,
          Figure({
              'data': [{'alignmentgroup': 'True',
                        'hovertemplate': 'CustomerId=%{x}<br>TotalAlbumsBought=%{y}<
         extra></extra>',
                        'legendgroup': '',
                        'marker': {'color': '#636efa', 'pattern': {'shape': ''}},
                        'name': '',
                        'offsetgroup': '',
                        'orientation': 'v',
                        'showlegend': False,
                        'textposition': 'auto',
                        'type': 'bar',
                        'x': array([58, 57, 56, 55, 54]),
                        'xaxis': 'x',
                        'y': array([38, 38, 38, 38, 38]),
                        'yaxis': 'y'}],
              'layout': {'barmode': 'relative',
                         'legend': {'tracegroupgap': 0},
                         'margin': {'t': 60},
                         'template': '...',
                         'xaxis': {'anchor': 'y', 'domain': [0.0, 1.0], 'title': {'t
         ext': 'CustomerId'}},
                         'yaxis': {'anchor': 'x', 'domain': [0.0, 1.0], 'title': {'t
         ext': 'TotalAlbumsBought'}}}
          }))
```

```sql
SELECT c.CustomerId, SUM(il.Quantity) AS TotalAlbums
FROM Customers c
JOIN invoices i  ON c.CustomerId = i.CustomerId
JOIN invoice_items il ON i.InvoiceId = il.InvoiceId
GROUP BY c.CustomerId
ORDER BY TotalAlbums DESC
LIMIT 5
```

```
In [37]: question = """
              Find the top 5 customers who spent the most money overall,

              Hint: order total can be found on invoices table, calculation using inv
         """

         vn.ask(question=question)
```

```
Number of requested results 10 is greater than number of elements in index
1, updating n_results = 1
```

```
SQL Prompt: [{'role': 'system', 'content': 'You are a SQLite expert. Please
help to generate a SQL query to answer the question. Your response should ON
LY be based on the given context and follow the response guidelines and form
at instructions. \n===Tables \nCREATE TABLE "invoices"\r\n(\r\n    InvoiceId
INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    CustomerId INTEGER  NOT N
ULL,\r\n    InvoiceDate DATETIME  NOT NULL,\r\n    BillingAddress NVARCHAR(7
0),\r\n    BillingCity NVARCHAR(40),\r\n    BillingState NVARCHAR(40),\r\n
BillingCountry NVARCHAR(40),\r\n    BillingPostalCode NVARCHAR(10),\r\n    T
otal NUMERIC(10,2)  NOT NULL,\r\n    FOREIGN KEY (CustomerId) REFERENCES "cu
stomers" (CustomerId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n
\nCREATE TABLE "invoice_items"\r\n(\r\n    InvoiceLineId INTEGER PRIMARY KEY
AUTOINCREMENT NOT NULL,\r\n    InvoiceId INTEGER  NOT NULL,\r\n    TrackId I
NTEGER  NOT NULL,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    Quantity
INTEGER  NOT NULL,\r\n    FOREIGN KEY (InvoiceId) REFERENCES "invoices" (Inv
oiceId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY
(TrackId) REFERENCES "tracks" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDAT
E NO ACTION\r\n)\n\nCREATE INDEX IFK_InvoiceLineInvoiceId ON "invoice_items"
(InvoiceId)\n\nCREATE INDEX IFK_InvoiceCustomerId ON "invoices" (CustomerId)
\n\nCREATE INDEX IFK_InvoiceLineTrackId ON "invoice_items" (TrackId)\n\nCREA
TE TABLE "customers"\r\n(\r\n    CustomerId INTEGER PRIMARY KEY AUTOINCREMEN
T NOT NULL,\r\n    FirstName NVARCHAR(40)  NOT NULL,\r\n    LastName NVARCHA
R(20)  NOT NULL,\r\n    Company NVARCHAR(80),\r\n    Address NVARCHAR(70),\r
\n    City NVARCHAR(40),\r\n    State NVARCHAR(40),\r\n    Country NVARCHAR
(40),\r\n    PostalCode NVARCHAR(10),\r\n    Phone NVARCHAR(24),\r\n    Fax
NVARCHAR(24),\r\n    Email NVARCHAR(60)  NOT NULL,\r\n    SupportRepId INTEG
ER,\r\n    FOREIGN KEY (SupportRepId) REFERENCES "employees" (EmployeeId) \r
\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE "employee
s"\r\n(\r\n    EmployeeId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n
LastName NVARCHAR(20)  NOT NULL,\r\n    FirstName NVARCHAR(20)  NOT NULL,\r
\n    Title NVARCHAR(30),\r\n    ReportsTo INTEGER,\r\n    BirthDate DATETIM
E,\r\n    HireDate DATETIME,\r\n    Address NVARCHAR(70),\r\n    City NVARCH
AR(40),\r\n    State NVARCHAR(40),\r\n    Country NVARCHAR(40),\r\n    Posta
lCode NVARCHAR(10),\r\n    Phone NVARCHAR(24),\r\n    Fax NVARCHAR(24),\r\n
Email NVARCHAR(60),\r\n    FOREIGN KEY (ReportsTo) REFERENCES "employees" (E
mployeeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TA
BLE "tracks"\r\n(\r\n    TrackId INTEGER PRIMARY KEY AUTOINCREMENT NOT NUL
L,\r\n    Name NVARCHAR(200)  NOT NULL,\r\n    AlbumId INTEGER,\r\n    Media
TypeId INTEGER  NOT NULL,\r\n    GenreId INTEGER,\r\n    Composer NVARCHAR(2
20),\r\n    Milliseconds INTEGER  NOT NULL,\r\n    Bytes INTEGER,\r\n    Uni
tPrice NUMERIC(10,2)  NOT NULL,\r\n    FOREIGN KEY (AlbumId) REFERENCES "alb
ums" (AlbumId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREI
GN KEY (GenreId) REFERENCES "genres" (GenreId) \r\n\t\tON DELETE NO ACTION O
N UPDATE NO ACTION,\r\n    FOREIGN KEY (MediaTypeId) REFERENCES "media_type
s" (MediaTypeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCRE
ATE TABLE "playlist_track"\r\n(\r\n    PlaylistId INTEGER  NOT NULL,\r\n
TrackId INTEGER  NOT NULL,\r\n    CONSTRAINT PK_PlaylistTrack PRIMARY KEY
(PlaylistId, TrackId),\r\n    FOREIGN KEY (PlaylistId) REFERENCES "playlist
s" (PlaylistId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FORE
IGN KEY (TrackId) REFERENCES "tracks" (TrackId) \r\n\t\tON DELETE NO ACTION
ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_EmployeeReportsTo ON "employee
s" (ReportsTo)\n\n\n===Additional Context \n\nIn the chinook database invoic
e means order\n\n===Response Guidelines \n1. If the provided context is suff
icient, please generate a valid SQL query without any explanations for the q
uestion. \n2. If the provided context is almost sufficient but requires know
ledge of a specific string in a particular column, please generate an interm
ediate SQL query to find the distinct strings in that column. Prepend the qu
```

```
ery with a comment saying intermediate_sql \n3. If the provided context is i
nsufficient, please explain why it can\'t be generated. \n4. Please use the
most relevant table(s). \n5. If the question has been asked and answered bef
ore, please repeat the answer exactly as it was given before. \n'}, {'role':
'user', 'content': '  \n    Hint: album quantity is found in invoice_items,
\n    \n    Find the top 5 customers who bought the most albums in total qua
ntity (across all invoices):\n'}, {'role': 'assistant', 'content': 'SELECT
i.CustomerId, SUM(ii.Quantity) AS TotalAlbumsBought\nFROM invoices i\nJOIN i
nvoice_items ii ON i.InvoiceId = ii.InvoiceId\nJOIN albums al ON ii.TrackId
IN (SELECT TrackId FROM tracks WHERE AlbumId = al.AlbumId)\nGROUP BY i.Custo
merId\nORDER BY TotalAlbumsBought DESC \nLIMIT 5'}, {'role': 'user', 'conten
t': '  \n     Find the customer with the most invoices \n'}, {'role': 'assis
tant', 'content': 'SELECT c.*, COUNT(i.InvoiceId) AS InvoiceCount\nFROM cust
omers c JOIN invoices i ON c.CustomerId = i.CustomerId\nGROUP BY c.CustomerI
d\nORDER BY InvoiceCount DESC\nLIMIT 1'}, {'role': 'user', 'content': '  \n
Get the average invoice total for each customer:\n'}, {'role': 'assistant',
'content': 'SELECT CustomerId, AVG(Total) AS AverageInvoiceTotal\nFROM invoi
ces\nGROUP BY CustomerId'}, {'role': 'user', 'content': '  \n    Find the to
p 5 most expensive tracks (based on unit price):\n'}, {'role': 'assistant',
'content': 'SELECT TrackId, Name, UnitPrice FROM tracks ORDER BY UnitPrice D
ESC LIMIT 5'}, {'role': 'user', 'content': '  \n    Get the total number of
invoices for each customer\n'}, {'role': 'assistant', 'content': 'SELECT Cus
tomerId, COUNT(*) AS NumberOfInvoices\nFROM invoices\nGROUP BY CustomerId'},
{'role': 'user', 'content': '  \n    List all invoices with a total exceedin
g $10:\n'}, {'role': 'assistant', 'content': 'SELECT * FROM invoices WHERE T
otal > 10.0'}, {'role': 'user', 'content': '  \n    Find all invoices since
2010 and the total amount invoiced:\n'}, {'role': 'assistant', 'content': "S
ELECT InvoiceId, Total\nFROM invoices\nWHERE strftime('%Y', InvoiceDate) >=
'2010'"}, {'role': 'user', 'content': ' \n    There are 3 tables: artists, a
lbums and tracks, where albums and artists are linked by ArtistId, albums an
d tracks are linked by AlbumId,\n    Can you find the top 10 most popular ar
tists based on the number of tracks\n'}, {'role': 'assistant', 'content': 'S
ELECT ar.Name AS ArtistName, COUNT(t.TrackId) as NumberOfTracks\nFROM artist
s ar\nJOIN albums al ON ar.ArtistId = al.ArtistId\nJOIN tracks t ON al.Album
Id = t.AlbumId\nGROUP BY ar.ArtistId\nORDER BY NumberOfTracks DESC\nLIMIT \n
10'}, {'role': 'user', 'content': "which table stores customer's orders"},
{'role': 'assistant', 'content': 'SELECT * FROM invoices'}, {'role': 'user',
'content': 'How many customers are there'}, {'role': 'assistant', 'content':
'SELECT COUNT(*) AS NumberOfCustomers FROM customers'}, {'role': 'user', 'co
ntent': '  \n     Find the top 5 customers who spent the most money overall,
\n     \n     Hint: order total can be found on invoices table, calculation
using invoice_items detail table is unnecessary \n'}]
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "You are a SQLite expert. Please help to gene
rate a SQL query to answer the question. Your response should ONLY be based
on the given context and follow the response guidelines and format instructi
ons. \n===Tables \nCREATE TABLE \"invoices\"\r\n(\r\n    InvoiceId INTEGER P
RIMARY KEY AUTOINCREMENT NOT NULL,\r\n    CustomerId INTEGER  NOT NULL,\r\n
InvoiceDate DATETIME  NOT NULL,\r\n    BillingAddress NVARCHAR(70),\r\n    B
illingCity NVARCHAR(40),\r\n    BillingState NVARCHAR(40),\r\n    BillingCou
ntry NVARCHAR(40),\r\n    BillingPostalCode NVARCHAR(10),\r\n    Total NUMER
IC(10,2)  NOT NULL,\r\n    FOREIGN KEY (CustomerId) REFERENCES \"customers\"
(CustomerId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE
TABLE \"invoice_items\"\r\n(\r\n    InvoiceLineId INTEGER PRIMARY KEY AUTOIN
CREMENT NOT NULL,\r\n    InvoiceId INTEGER  NOT NULL,\r\n    TrackId INTEGER
NOT NULL,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    Quantity INTEGER
NOT NULL,\r\n    FOREIGN KEY (InvoiceId) REFERENCES \"invoices\" (InvoiceId)
\r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackI
d) REFERENCES \"tracks\" (TrackId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO
ACTION\r\n)\n\nCREATE INDEX IFK_InvoiceLineInvoiceId ON \"invoice_items\" (I
nvoiceId)\n\nCREATE INDEX IFK_InvoiceCustomerId ON \"invoices\" (CustomerId)
\n\nCREATE INDEX IFK_InvoiceLineTrackId ON \"invoice_items\" (TrackId)\n\nCR
EATE TABLE \"customers\"\r\n(\r\n    CustomerId INTEGER PRIMARY KEY AUTOINCR
EMENT NOT NULL,\r\n    FirstName NVARCHAR(40)  NOT NULL,\r\n    LastName NVA
RCHAR(20)  NOT NULL,\r\n    Company NVARCHAR(80),\r\n    Address NVARCHAR(7
0),\r\n    City NVARCHAR(40),\r\n    State NVARCHAR(40),\r\n    Country NVAR
CHAR(40),\r\n    PostalCode NVARCHAR(10),\r\n    Phone NVARCHAR(24),\r\n
Fax NVARCHAR(24),\r\n    Email NVARCHAR(60)  NOT NULL,\r\n    SupportRepId I
NTEGER,\r\n    FOREIGN KEY (SupportRepId) REFERENCES \"employees\" (Employee
Id) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE \"e
mployees\"\r\n(\r\n    EmployeeId INTEGER PRIMARY KEY AUTOINCREMENT NOT NUL
L,\r\n    LastName NVARCHAR(20)  NOT NULL,\r\n    FirstName NVARCHAR(20)  NO
T NULL,\r\n    Title NVARCHAR(30),\r\n    ReportsTo INTEGER,\r\n    BirthDat
e DATETIME,\r\n    HireDate DATETIME,\r\n    Address NVARCHAR(70),\r\n    Ci
ty NVARCHAR(40),\r\n    State NVARCHAR(40),\r\n    Country NVARCHAR(40),\r\n
PostalCode NVARCHAR(10),\r\n    Phone NVARCHAR(24),\r\n    Fax NVARCHAR(2
4),\r\n    Email NVARCHAR(60),\r\n    FOREIGN KEY (ReportsTo) REFERENCES \"e
mployees\" (EmployeeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)
\n\nCREATE TABLE \"tracks\"\r\n(\r\n    TrackId INTEGER PRIMARY KEY AUTOINCR
EMENT NOT NULL,\r\n    Name NVARCHAR(200)  NOT NULL,\r\n    AlbumId INTEGE
R,\r\n    MediaTypeId INTEGER  NOT NULL,\r\n    GenreId INTEGER,\r\n    Comp
oser NVARCHAR(220),\r\n    Milliseconds INTEGER  NOT NULL,\r\n    Bytes INTE
GER,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    FOREIGN KEY (AlbumId)
REFERENCES \"albums\" (AlbumId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACT
ION,\r\n    FOREIGN KEY (GenreId) REFERENCES \"genres\" (GenreId) \r\n\t\tON
DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (MediaTypeId) REFER
ENCES \"media_types\" (MediaTypeId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO
ACTION\r\n)\n\nCREATE TABLE \"playlist_track\"\r\n(\r\n    PlaylistId INTEGE
R  NOT NULL,\r\n    TrackId INTEGER  NOT NULL,\r\n    CONSTRAINT PK_Playlist
Track PRIMARY KEY  (PlaylistId, TrackId),\r\n    FOREIGN KEY (PlaylistId) RE
FERENCES \"playlists\" (PlaylistId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO
ACTION,\r\n    FOREIGN KEY (TrackId) REFERENCES \"tracks\" (TrackId) \r\n\t
\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_EmployeeR
eportsTo ON \"employees\" (ReportsTo)\n\n\n===Additional Context \n\nIn the
chinook database invoice means order\n\n===Response Guidelines \n1. If the p
rovided context is sufficient, please generate a valid SQL query without any
explanations for the question. \n2. If the provided context is almost suffic
ient but requires knowledge of a specific string in a particular column, ple
ase generate an intermediate SQL query to find the distinct strings in that
column. Prepend the query with a comment saying intermediate_sql \n3. If the
provided context is insufficient, please explain why it can't be generated.
\n4. Please use the most relevant table(s). \n5. If the question has been as
ked and answered before, please repeat the answer exactly as it was given be
fore. \n"}, {"role": "user", "content": "  \n    Hint: album quantity is fou
nd in invoice_items, \n    \n    Find the top 5 customers who bought the mos
t albums in total quantity (across all invoices):\n"}, {"role": "assistant",
"content": "SELECT i.CustomerId, SUM(ii.Quantity) AS TotalAlbumsBought\nFROM
invoices i\nJOIN invoice_items ii ON i.InvoiceId = ii.InvoiceId\nJOIN albums
```

```
al ON ii.TrackId IN (SELECT TrackId FROM tracks WHERE AlbumId = al.AlbumId)
\nGROUP BY i.CustomerId\nORDER BY TotalAlbumsBought DESC \nLIMIT 5"}, {"rol
e": "user", "content": "  \n     Find the customer with the most invoices
\n"}, {"role": "assistant", "content": "SELECT c.*, COUNT(i.InvoiceId) AS In
voiceCount\nFROM customers c JOIN invoices i ON c.CustomerId = i.CustomerId
\nGROUP BY c.CustomerId\nORDER BY InvoiceCount DESC\nLIMIT 1"}, {"role": "us
er", "content": "  \n    Get the average invoice total for each custome
r:\n"}, {"role": "assistant", "content": "SELECT CustomerId, AVG(Total) AS A
verageInvoiceTotal\nFROM invoices\nGROUP BY CustomerId"}, {"role": "user",
"content": "  \n    Find the top 5 most expensive tracks (based on unit pric
e):\n"}, {"role": "assistant", "content": "SELECT TrackId, Name, UnitPrice F
ROM tracks ORDER BY UnitPrice DESC LIMIT 5"}, {"role": "user", "content": "
\n    Get the total number of invoices for each customer\n"}, {"role": "assi
stant", "content": "SELECT CustomerId, COUNT(*) AS NumberOfInvoices\nFROM in
voices\nGROUP BY CustomerId"}, {"role": "user", "content": "  \n    List all
invoices with a total exceeding $10:\n"}, {"role": "assistant", "content":
"SELECT * FROM invoices WHERE Total > 10.0"}, {"role": "user", "content": "
\n    Find all invoices since 2010 and the total amount invoiced:\n"}, {"rol
e": "assistant", "content": "SELECT InvoiceId, Total\nFROM invoices\nWHERE s
trftime('%Y', InvoiceDate) >= '2010'"}, {"role": "user", "content": " \n
There are 3 tables: artists, albums and tracks, where albums and artists are
linked by ArtistId, albums and tracks are linked by AlbumId,\n    Can you fi
nd the top 10 most popular artists based on the number of tracks\n"}, {"rol
e": "assistant", "content": "SELECT ar.Name AS ArtistName, COUNT(t.TrackId)
as NumberOfTracks\nFROM artists ar\nJOIN albums al ON ar.ArtistId = al.Artis
tId\nJOIN tracks t ON al.AlbumId = t.AlbumId\nGROUP BY ar.ArtistId\nORDER BY
NumberOfTracks DESC\nLIMIT \n  10"}, {"role": "user", "content": "which tabl
e stores customer's orders"}, {"role": "assistant", "content": "SELECT * FRO
M invoices"}, {"role": "user", "content": "How many customers are there"},
{"role": "assistant", "content": "SELECT COUNT(*) AS NumberOfCustomers FROM
customers"}, {"role": "user", "content": "  \n     Find the top 5 customers
who spent the most money overall, \n     \n     Hint: order total can be fou
nd on invoices table, calculation using invoice_items detail table is unnece
ssary \n"}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T22:26:07.998978405Z', 'messa
ge': {'role': 'assistant', 'content': 'SELECT c.CustomerId, c.CompanyName AS
CustomerName, SUM(i.Total) as TotalSpent\nFROM customers c JOIN invoices i O
N c.CustomerId = i.CustomerId\nGROUP BY c.CustomerId\nORDER BY Total DESC\nL
IMIT 5'}, 'done_reason': 'stop', 'done': True, 'total_duration': 15675453423
4, 'load_duration': 4220845, 'prompt_eval_count': 2040, 'prompt_eval_duratio
n': 136694064000, 'eval_count': 61, 'eval_duration': 18987358000}
LLM Response: SELECT c.CustomerId, c.CompanyName AS CustomerName, SUM(i.Tota
l) as TotalSpent
FROM customers c JOIN invoices i ON c.CustomerId = i.CustomerId
GROUP BY c.CustomerId
ORDER BY Total DESC
LIMIT 5
SELECT c.CustomerId, c.CompanyName AS CustomerName, SUM(i.Total) as TotalSpe
nt
FROM customers c JOIN invoices i ON c.CustomerId = i.CustomerId
GROUP BY c.CustomerId
ORDER BY Total DESC
LIMIT 5
Couldn't run sql:  Execution failed on sql 'SELECT c.CustomerId, c.CompanyNa
me AS CustomerName, SUM(i.Total) as TotalSpent
```

```
FROM customers c JOIN invoices i ON c.CustomerId = i.CustomerId
GROUP BY c.CustomerId
ORDER BY Total DESC
LIMIT 5': no such column: c.CompanyName
```

In [38]:

```
question = """
     Get all playlists containing at least 10 tracks and the total duration
"""

vn.ask(question=question)
```

```
Number of requested results 10 is greater than number of elements in index
1, updating n_results = 1
```

```
SQL Prompt: [{'role': 'system', 'content': 'You are a SQLite expert. Please
help to generate a SQL query to answer the question. Your response should ON
LY be based on the given context and follow the response guidelines and form
at instructions. \n===Tables \nCREATE INDEX IFK_PlaylistTrackTrackId ON "pla
ylist_track" (TrackId)\n\nCREATE TABLE "playlists"\r\n(\r\n    PlaylistId IN
TEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(120)\r\n)\n\n
CREATE TABLE "playlist_track"\r\n(\r\n    PlaylistId INTEGER  NOT NULL,\r\n
TrackId INTEGER  NOT NULL,\r\n    CONSTRAINT PK_PlaylistTrack PRIMARY KEY
(PlaylistId, TrackId),\r\n    FOREIGN KEY (PlaylistId) REFERENCES "playlist
s" (PlaylistId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FORE
IGN KEY (TrackId) REFERENCES "tracks" (TrackId) \r\n\t\tON DELETE NO ACTION
ON UPDATE NO ACTION\r\n)\n\nCREATE TABLE "tracks"\r\n(\r\n    TrackId INTEGE
R PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(200)  NOT NULL,\r
\n    AlbumId INTEGER,\r\n    MediaTypeId INTEGER  NOT NULL,\r\n    GenreId
INTEGER,\r\n    Composer NVARCHAR(220),\r\n    Milliseconds INTEGER  NOT NUL
L,\r\n    Bytes INTEGER,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    FO
REIGN KEY (AlbumId) REFERENCES "albums" (AlbumId) \r\n\t\tON DELETE NO ACTIO
N ON UPDATE NO ACTION,\r\n    FOREIGN KEY (GenreId) REFERENCES "genres" (Gen
reId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (M
ediaTypeId) REFERENCES "media_types" (MediaTypeId) \r\n\t\tON DELETE NO ACTI
ON ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_TrackGenreId ON "tracks" (Ge
nreId)\n\nCREATE INDEX IFK_TrackAlbumId ON "tracks" (AlbumId)\n\nCREATE INDE
X IFK_TrackMediaTypeId ON "tracks" (MediaTypeId)\n\nCREATE INDEX IFK_AlbumAr
tistId ON "albums" (ArtistId)\n\nCREATE TABLE "albums"\r\n(\r\n    AlbumId I
NTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Title NVARCHAR(160)  NOT N
ULL,\r\n    ArtistId INTEGER  NOT NULL,\r\n    FOREIGN KEY (ArtistId) REFERE
NCES "artists" (ArtistId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r
\n)\n\nCREATE TABLE "genres"\r\n(\r\n    GenreId INTEGER PRIMARY KEY AUTOINC
REMENT NOT NULL,\r\n    Name NVARCHAR(120)\r\n)\n\n\n===Additional Context
\n\nIn the chinook database invoice means order\n\n===Response Guidelines \n
1. If the provided context is sufficient, please generate a valid SQL query
without any explanations for the question. \n2. If the provided context is a
lmost sufficient but requires knowledge of a specific string in a particular
column, please generate an intermediate SQL query to find the distinct strin
gs in that column. Prepend the query with a comment saying intermediate_sql
\n3. If the provided context is insufficient, please explain why it can\'t b
e generated. \n4. Please use the most relevant table(s). \n5. If the questio
n has been asked and answered before, please repeat the answer exactly as it
was given before. \n'}, {'role': 'user', 'content': '  \n    List all genres
and the number of tracks in each genre:\n'}, {'role': 'assistant', 'conten
t': 'SELECT g.GenreId, g.Name, COUNT(t.TrackId) AS NumberOfTracks\nFROM genr
es g\nLEFT JOIN tracks t ON g.GenreId = t.GenreId\nGROUP BY g.GenreId'}, {'r
ole': 'user', 'content': ' \n    There are 3 tables: artists, albums and tra
cks, where albums and artists are linked by ArtistId, albums and tracks are
linked by AlbumId,\n    Can you find the top 10 most popular artists based o
n the number of tracks\n'}, {'role': 'assistant', 'content': 'SELECT ar.Name
AS ArtistName, COUNT(t.TrackId) as NumberOfTracks\nFROM artists ar\nJOIN alb
ums al ON ar.ArtistId = al.ArtistId\nJOIN tracks t ON al.AlbumId = t.AlbumId
\nGROUP BY ar.ArtistId\nORDER BY NumberOfTracks DESC\nLIMIT \n  10'}, {'rol
e': 'user', 'content': '  \n    Hint: album quantity is found in invoice_ite
ms, \n    \n    Find the top 5 customers who bought the most albums in total
quantity (across all invoices):\n'}, {'role': 'assistant', 'content': 'SELEC
T i.CustomerId, SUM(ii.Quantity) AS TotalAlbumsBought\nFROM invoices i\nJOIN
invoice_items ii ON i.InvoiceId = ii.InvoiceId\nJOIN albums al ON ii.TrackId
IN (SELECT TrackId FROM tracks WHERE AlbumId = al.AlbumId)\nGROUP BY i.Custo
merId\nORDER BY TotalAlbumsBought DESC \nLIMIT 5'}, {'role': 'user', 'conten
```

```
t': '  \n    List all albums and their corresponding artist names  \n'}, {'r
ole': 'assistant', 'content': 'SELECT a.Title, ar.Name AS ArtistName\nFROM a
lbums a\nJOIN artists ar ON a.ArtistId = ar.ArtistId'}, {'role': 'user', 'co
ntent': '  \n    Find the top 5 most expensive tracks (based on unit pric
e):\n'}, {'role': 'assistant', 'content': 'SELECT TrackId, Name, UnitPrice F
ROM tracks ORDER BY UnitPrice DESC LIMIT 5'}, {'role': 'user', 'content': '
\n    Find all invoices since 2010 and the total amount invoiced:\n'}, {'rol
e': 'assistant', 'content': "SELECT InvoiceId, Total\nFROM invoices\nWHERE s
trftime('%Y', InvoiceDate) >= '2010'"}, {'role': 'user', 'content': '  \n
Find all tracks with a name containing "What" (case-insensitive)\n'}, {'rol
e': 'assistant', 'content': "SELECT DISTINCT LOWER(name) AS Pattern FROM tra
cks WHERE lower(Name) LIKE '%what%'"}, {'role': 'user', 'content': 'Can you
list all tables in the SQLite database catalog?'}, {'role': 'assistant', 'co
ntent': "SELECT name FROM sqlite_master WHERE type = 'table'"}, {'role': 'us
er', 'content': '  \n    List all invoices with a total exceeding $10:\n'},
{'role': 'assistant', 'content': 'SELECT * FROM invoices WHERE Total > 10.
0'}, {'role': 'user', 'content': '  \n    Get the total number of invoices f
or each customer\n'}, {'role': 'assistant', 'content': 'SELECT CustomerId, C
OUNT(*) AS NumberOfInvoices\nFROM invoices\nGROUP BY CustomerId'}, {'role':
'user', 'content': '  \n     Get all playlists containing at least 10 tracks
and the total duration of those tracks:\n'}]
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "You are a SQLite expert. Please help to gene
rate a SQL query to answer the question. Your response should ONLY be based
on the given context and follow the response guidelines and format instructi
ons. \n===Tables \nCREATE INDEX IFK_PlaylistTrackTrackId ON \"playlist_track
\" (TrackId)\n\nCREATE TABLE \"playlists\"\r\n(\r\n    PlaylistId INTEGER PR
IMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(120)\r\n)\n\nCREATE T
ABLE \"playlist_track\"\r\n(\r\n    PlaylistId INTEGER  NOT NULL,\r\n    Tra
ckId INTEGER  NOT NULL,\r\n    CONSTRAINT PK_PlaylistTrack PRIMARY KEY  (Pla
ylistId, TrackId),\r\n    FOREIGN KEY (PlaylistId) REFERENCES \"playlists\"
(PlaylistId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN
KEY (TrackId) REFERENCES \"tracks\" (TrackId) \r\n\t\tON DELETE NO ACTION ON
UPDATE NO ACTION\r\n)\n\nCREATE TABLE \"tracks\"\r\n(\r\n    TrackId INTEGER
PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(200)  NOT NULL,\r\n
AlbumId INTEGER,\r\n    MediaTypeId INTEGER  NOT NULL,\r\n    GenreId INTEGE
R,\r\n    Composer NVARCHAR(220),\r\n    Milliseconds INTEGER  NOT NULL,\r\n
Bytes INTEGER,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    FOREIGN KEY
(AlbumId) REFERENCES \"albums\" (AlbumId) \r\n\t\tON DELETE NO ACTION ON UPD
ATE NO ACTION,\r\n    FOREIGN KEY (GenreId) REFERENCES \"genres\" (GenreId)
\r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (MediaTy
peId) REFERENCES \"media_types\" (MediaTypeId) \r\n\t\tON DELETE NO ACTION O
N UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_TrackGenreId ON \"tracks\" (Genr
eId)\n\nCREATE INDEX IFK_TrackAlbumId ON \"tracks\" (AlbumId)\n\nCREATE INDE
X IFK_TrackMediaTypeId ON \"tracks\" (MediaTypeId)\n\nCREATE INDEX IFK_Album
ArtistId ON \"albums\" (ArtistId)\n\nCREATE TABLE \"albums\"\r\n(\r\n    Alb
umId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Title NVARCHAR(160)
NOT NULL,\r\n    ArtistId INTEGER  NOT NULL,\r\n    FOREIGN KEY (ArtistId) R
EFERENCES \"artists\" (ArtistId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO AC
TION\r\n)\n\nCREATE TABLE \"genres\"\r\n(\r\n    GenreId INTEGER PRIMARY KEY
AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(120)\r\n)\n\n\n===Additional Co
ntext \n\nIn the chinook database invoice means order\n\n===Response Guideli
```

```
nes \n1. If the provided context is sufficient, please generate a valid SQL
query without any explanations for the question. \n2. If the provided contex
t is almost sufficient but requires knowledge of a specific string in a part
icular column, please generate an intermediate SQL query to find the distinc
t strings in that column. Prepend the query with a comment saying intermedia
te_sql \n3. If the provided context is insufficient, please explain why it c
an't be generated. \n4. Please use the most relevant table(s). \n5. If the q
uestion has been asked and answered before, please repeat the answer exactly
as it was given before. \n"}, {"role": "user", "content": "  \n    List all
genres and the number of tracks in each genre:\n"}, {"role": "assistant", "c
ontent": "SELECT g.GenreId, g.Name, COUNT(t.TrackId) AS NumberOfTracks\nFROM
genres g\nLEFT JOIN tracks t ON g.GenreId = t.GenreId\nGROUP BY g.GenreId"},
{"role": "user", "content": " \n    There are 3 tables: artists, albums and
tracks, where albums and artists are linked by ArtistId, albums and tracks a
re linked by AlbumId,\n    Can you find the top 10 most popular artists base
d on the number of tracks\n"}, {"role": "assistant", "content": "SELECT ar.N
ame AS ArtistName, COUNT(t.TrackId) as NumberOfTracks\nFROM artists ar\nJOIN
albums al ON ar.ArtistId = al.ArtistId\nJOIN tracks t ON al.AlbumId = t.Albu
mId\nGROUP BY ar.ArtistId\nORDER BY NumberOfTracks DESC\nLIMIT \n  10"}, {"r
ole": "user", "content": "  \n    Hint: album quantity is found in invoice_i
tems, \n    \n    Find the top 5 customers who bought the most albums in tot
al quantity (across all invoices):\n"}, {"role": "assistant", "content": "SE
LECT i.CustomerId, SUM(ii.Quantity) AS TotalAlbumsBought\nFROM invoices i\nJ
OIN invoice_items ii ON i.InvoiceId = ii.InvoiceId\nJOIN albums al ON ii.Tra
ckId IN (SELECT TrackId FROM tracks WHERE AlbumId = al.AlbumId)\nGROUP BY i.
CustomerId\nORDER BY TotalAlbumsBought DESC \nLIMIT 5"}, {"role": "user", "c
ontent": "  \n    List all albums and their corresponding artist names
\n"}, {"role": "assistant", "content": "SELECT a.Title, ar.Name AS ArtistNam
e\nFROM albums a\nJOIN artists ar ON a.ArtistId = ar.ArtistId"}, {"role": "u
ser", "content": "  \n    Find the top 5 most expensive tracks (based on uni
t price):\n"}, {"role": "assistant", "content": "SELECT TrackId, Name, UnitP
rice FROM tracks ORDER BY UnitPrice DESC LIMIT 5"}, {"role": "user", "conten
t": "  \n    Find all invoices since 2010 and the total amount invoice
d:\n"}, {"role": "assistant", "content": "SELECT InvoiceId, Total\nFROM invo
ices\nWHERE strftime('%Y', InvoiceDate) >= '2010'"}, {"role": "user", "conte
nt": "  \n    Find all tracks with a name containing \"What\" (case-insensit
ive)\n"}, {"role": "assistant", "content": "SELECT DISTINCT LOWER(name) AS P
attern FROM tracks WHERE lower(Name) LIKE '%what%'"}, {"role": "user", "cont
ent": "Can you list all tables in the SQLite database catalog?"}, {"role":
"assistant", "content": "SELECT name FROM sqlite_master WHERE type = 'tabl
e'"}, {"role": "user", "content": "  \n    List all invoices with a total ex
ceeding $10:\n"}, {"role": "assistant", "content": "SELECT * FROM invoices W
HERE Total > 10.0"}, {"role": "user", "content": "  \n    Get the total numb
er of invoices for each customer\n"}, {"role": "assistant", "content": "SELE
CT CustomerId, COUNT(*) AS NumberOfInvoices\nFROM invoices\nGROUP BY Custome
rId"}, {"role": "user", "content": "  \n     Get all playlists containing at
least 10 tracks and the total duration of those tracks:\n"}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T22:28:24.761455664Z', 'messa
ge': {'role': 'assistant', 'content': 'SELECT p.PlaylistId, ar.Name as Playl
istName, SUM(t.Milliseconds) AS TotalDuration\nFROM playlist_track pt\nJOIN
playlists p ON pt.PlaylistId = p.PlaylistId\nJOIN tracks t ON pt.TrackId =
t.TrackId\nGROUP BY p.PlaylistId\nHAVING COUNT(pt.TrackId) >= 10'}, 'done_re
ason': 'stop', 'done': True, 'total_duration': 136722648465, 'load_duratio
n': 4080576, 'prompt_eval_count': 1646, 'prompt_eval_duration': 10797840300
0, 'eval_count': 89, 'eval_duration': 27386827000}
```

```
LLM Response: SELECT p.PlaylistId, ar.Name as PlaylistName, SUM(t.Millisecon
ds) AS TotalDuration
FROM playlist_track pt
JOIN playlists p ON pt.PlaylistId = p.PlaylistId
JOIN tracks t ON pt.TrackId = t.TrackId
GROUP BY p.PlaylistId
HAVING COUNT(pt.TrackId) >= 10
SELECT p.PlaylistId, ar.Name as PlaylistName, SUM(t.Milliseconds) AS TotalDu
ration
FROM playlist_track pt
JOIN playlists p ON pt.PlaylistId = p.PlaylistId
JOIN tracks t ON pt.TrackId = t.TrackId
GROUP BY p.PlaylistId
HAVING COUNT(pt.TrackId) >= 10
Couldn't run sql:  Execution failed on sql 'SELECT p.PlaylistId, ar.Name as
PlaylistName, SUM(t.Milliseconds) AS TotalDuration
FROM playlist_track pt
JOIN playlists p ON pt.PlaylistId = p.PlaylistId
JOIN tracks t ON pt.TrackId = t.TrackId
GROUP BY p.PlaylistId
HAVING COUNT(pt.TrackId) >= 10': no such column: ar.Name
```

In [39]:

```
question = """
     Identify artists who have albums with tracks appearing in multiple genr


"""

vn.ask(question=question)
```

```
Number of requested results 10 is greater than number of elements in index
1, updating n_results = 1
```

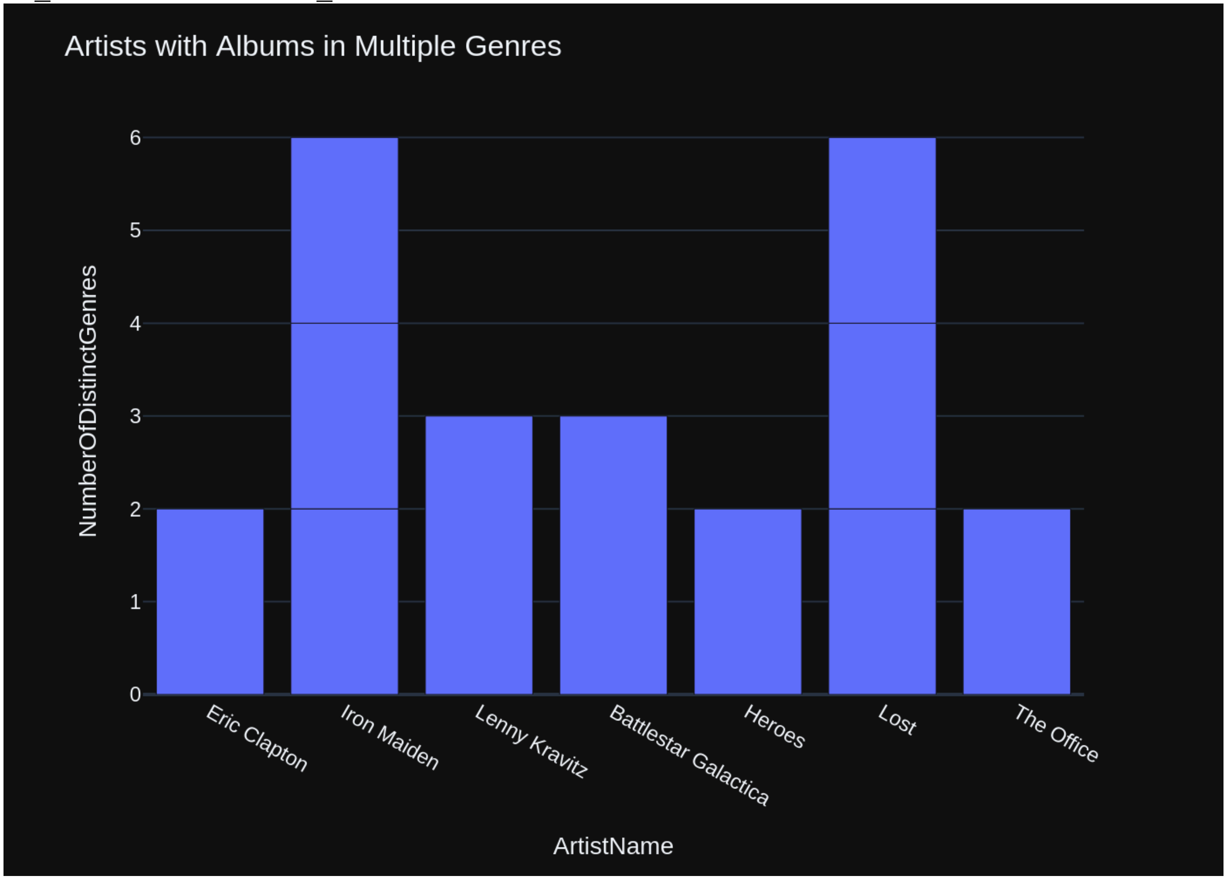
```
SQL Prompt: [{'role': 'system', 'content': 'You are a SQLite expert. Please
help to generate a SQL query to answer the question. Your response should ON
LY be based on the given context and follow the response guidelines and form
at instructions. \n===Tables \nCREATE TABLE "tracks"\r\n(\r\n    TrackId INT
EGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(200)  NOT NUL
L,\r\n    AlbumId INTEGER,\r\n    MediaTypeId INTEGER  NOT NULL,\r\n    Genr
eId INTEGER,\r\n    Composer NVARCHAR(220),\r\n    Milliseconds INTEGER  NOT
NULL,\r\n    Bytes INTEGER,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n
FOREIGN KEY (AlbumId) REFERENCES "albums" (AlbumId) \r\n\t\tON DELETE NO ACT
ION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (GenreId) REFERENCES "genres" (G
enreId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY
(MediaTypeId) REFERENCES "media_types" (MediaTypeId) \r\n\t\tON DELETE NO AC
TION ON UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_AlbumArtistId ON "albums"
(ArtistId)\n\nCREATE INDEX IFK_TrackGenreId ON "tracks" (GenreId)\n\nCREATE
INDEX IFK_TrackAlbumId ON "tracks" (AlbumId)\n\nCREATE TABLE "albums"\r\n(\r
\n    AlbumId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Title NVARC
HAR(160)  NOT NULL,\r\n    ArtistId INTEGER  NOT NULL,\r\n    FOREIGN KEY (A
rtistId) REFERENCES "artists" (ArtistId) \r\n\t\tON DELETE NO ACTION ON UPDA
TE NO ACTION\r\n)\n\nCREATE INDEX IFK_TrackMediaTypeId ON "tracks" (MediaTyp
eId)\n\nCREATE TABLE "genres"\r\n(\r\n    GenreId INTEGER PRIMARY KEY AUTOIN
CREMENT NOT NULL,\r\n    Name NVARCHAR(120)\r\n)\n\nCREATE INDEX IFK_Playlis
tTrackTrackId ON "playlist_track" (TrackId)\n\nCREATE TABLE "artists"\r\n(\r
\n    ArtistId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARC
HAR(120)\r\n)\n\nCREATE TABLE "playlist_track"\r\n(\r\n    PlaylistId INTEGE
R  NOT NULL,\r\n    TrackId INTEGER  NOT NULL,\r\n    CONSTRAINT PK_Playlist
Track PRIMARY KEY  (PlaylistId, TrackId),\r\n    FOREIGN KEY (PlaylistId) RE
FERENCES "playlists" (PlaylistId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO A
CTION,\r\n    FOREIGN KEY (TrackId) REFERENCES "tracks" (TrackId) \r\n\t\tON
DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\n\n===Additional Context \n\nIn
the chinook database invoice means order\n\n===Response Guidelines \n1. If t
he provided context is sufficient, please generate a valid SQL query without
any explanations for the question. \n2. If the provided context is almost su
fficient but requires knowledge of a specific string in a particular column,
please generate an intermediate SQL query to find the distinct strings in th
at column. Prepend the query with a comment saying intermediate_sql \n3. If
the provided context is insufficient, please explain why it can\'t be genera
ted. \n4. Please use the most relevant table(s). \n5. If the question has be
en asked and answered before, please repeat the answer exactly as it was giv
en before. \n'}, {'role': 'user', 'content': ' \n    There are 3 tables: art
ists, albums and tracks, where albums and artists are linked by ArtistId, al
bums and tracks are linked by AlbumId,\n    Can you find the top 10 most pop
ular artists based on the number of tracks\n'}, {'role': 'assistant', 'conte
nt': 'SELECT ar.Name AS ArtistName, COUNT(t.TrackId) as NumberOfTracks\nFROM
artists ar\nJOIN albums al ON ar.ArtistId = al.ArtistId\nJOIN tracks t ON a
l.AlbumId = t.AlbumId\nGROUP BY ar.ArtistId\nORDER BY NumberOfTracks DESC\nL
IMIT \n  10'}, {'role': 'user', 'content': '  \n    List all albums and thei
r corresponding artist names  \n'}, {'role': 'assistant', 'content': 'SELECT
a.Title, ar.Name AS ArtistName\nFROM albums a\nJOIN artists ar ON a.ArtistId
= ar.ArtistId'}, {'role': 'user', 'content': '  \n    List all genres and th
e number of tracks in each genre:\n'}, {'role': 'assistant', 'content': 'SEL
ECT g.GenreId, g.Name, COUNT(t.TrackId) AS NumberOfTracks\nFROM genres g\nLE
FT JOIN tracks t ON g.GenreId = t.GenreId\nGROUP BY g.GenreId'}, {'role': 'u
ser', 'content': '  \n    Hint: album quantity is found in invoice_items, \n
\n    Find the top 5 customers who bought the most albums in total quantity
(across all invoices):\n'}, {'role': 'assistant', 'content': 'SELECT i.Custo
merId, SUM(ii.Quantity) AS TotalAlbumsBought\nFROM invoices i\nJOIN invoice_
```

```
items ii ON i.InvoiceId = ii.InvoiceId\nJOIN albums al ON ii.TrackId IN (SEL
ECT TrackId FROM tracks WHERE AlbumId = al.AlbumId)\nGROUP BY i.CustomerId\n
ORDER BY TotalAlbumsBought DESC \nLIMIT 5'}, {'role': 'user', 'content': '
\n    Find the top 5 most expensive tracks (based on unit price):\n'}, {'rol
e': 'assistant', 'content': 'SELECT TrackId, Name, UnitPrice FROM tracks ORD
ER BY UnitPrice DESC LIMIT 5'}, {'role': 'user', 'content': '  \n    Find al
l tracks with a name containing "What" (case-insensitive)\n'}, {'role': 'ass
istant', 'content': "SELECT DISTINCT LOWER(name) AS Pattern FROM tracks WHER
E lower(Name) LIKE '%what%'"}, {'role': 'user', 'content': 'Can you list all
tables in the SQLite database catalog?'}, {'role': 'assistant', 'content':
"SELECT name FROM sqlite_master WHERE type = 'table'"}, {'role': 'user', 'co
ntent': 'How many customers are there'}, {'role': 'assistant', 'content': 'S
ELECT COUNT(*) AS NumberOfCustomers FROM customers'}, {'role': 'user', 'cont
ent': '  \n     Find the customer with the most invoices \n'}, {'role': 'ass
istant', 'content': 'SELECT c.*, COUNT(i.InvoiceId) AS InvoiceCount\nFROM cu
stomers c JOIN invoices i ON c.CustomerId = i.CustomerId\nGROUP BY c.Custome
rId\nORDER BY InvoiceCount DESC\nLIMIT 1'}, {'role': 'user', 'content': '
\n    List all invoices with a total exceeding $10:\n'}, {'role': 'assistan
t', 'content': 'SELECT * FROM invoices WHERE Total > 10.0'}, {'role': 'use
r', 'content': '  \n     Identify artists who have albums with tracks appear
ing in multiple genres:\n\n\n'}]
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "You are a SQLite expert. Please help to gene
rate a SQL query to answer the question. Your response should ONLY be based
on the given context and follow the response guidelines and format instructi
ons. \n===Tables \nCREATE TABLE \"tracks\"\r\n(\r\n    TrackId INTEGER PRIMA
RY KEY AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(200)  NOT NULL,\r\n    A
lbumId INTEGER,\r\n    MediaTypeId INTEGER  NOT NULL,\r\n    GenreId INTEGE
R,\r\n    Composer NVARCHAR(220),\r\n    Milliseconds INTEGER  NOT NULL,\r\n
Bytes INTEGER,\r\n    UnitPrice NUMERIC(10,2)  NOT NULL,\r\n    FOREIGN KEY
(AlbumId) REFERENCES \"albums\" (AlbumId) \r\n\t\tON DELETE NO ACTION ON UPD
ATE NO ACTION,\r\n    FOREIGN KEY (GenreId) REFERENCES \"genres\" (GenreId)
\r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION,\r\n    FOREIGN KEY (MediaTy
peId) REFERENCES \"media_types\" (MediaTypeId) \r\n\t\tON DELETE NO ACTION O
N UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_AlbumArtistId ON \"albums\" (Art
istId)\n\nCREATE INDEX IFK_TrackGenreId ON \"tracks\" (GenreId)\n\nCREATE IN
DEX IFK_TrackAlbumId ON \"tracks\" (AlbumId)\n\nCREATE TABLE \"albums\"\r\n
(\r\n    AlbumId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n    Title NV
ARCHAR(160)  NOT NULL,\r\n    ArtistId INTEGER  NOT NULL,\r\n    FOREIGN KEY
(ArtistId) REFERENCES \"artists\" (ArtistId) \r\n\t\tON DELETE NO ACTION ON
UPDATE NO ACTION\r\n)\n\nCREATE INDEX IFK_TrackMediaTypeId ON \"tracks\" (Me
diaTypeId)\n\nCREATE TABLE \"genres\"\r\n(\r\n    GenreId INTEGER PRIMARY KE
Y AUTOINCREMENT NOT NULL,\r\n    Name NVARCHAR(120)\r\n)\n\nCREATE INDEX IFK
_PlaylistTrackTrackId ON \"playlist_track\" (TrackId)\n\nCREATE TABLE \"arti
sts\"\r\n(\r\n    ArtistId INTEGER PRIMARY KEY AUTOINCREMENT NOT NULL,\r\n
Name NVARCHAR(120)\r\n)\n\nCREATE TABLE \"playlist_track\"\r\n(\r\n    Playl
istId INTEGER  NOT NULL,\r\n    TrackId INTEGER  NOT NULL,\r\n    CONSTRAINT
PK_PlaylistTrack PRIMARY KEY  (PlaylistId, TrackId),\r\n    FOREIGN KEY (Pla
ylistId) REFERENCES \"playlists\" (PlaylistId) \r\n\t\tON DELETE NO ACTION O
N UPDATE NO ACTION,\r\n    FOREIGN KEY (TrackId) REFERENCES \"tracks\" (Trac
kId) \r\n\t\tON DELETE NO ACTION ON UPDATE NO ACTION\r\n)\n\n\n===Additional
Context \n\nIn the chinook database invoice means order\n\n===Response Guide
```

```
lines \n1. If the provided context is sufficient, please generate a valid SQ
L query without any explanations for the question. \n2. If the provided cont
ext is almost sufficient but requires knowledge of a specific string in a pa
rticular column, please generate an intermediate SQL query to find the disti
nct strings in that column. Prepend the query with a comment saying intermed
iate_sql \n3. If the provided context is insufficient, please explain why it
can't be generated. \n4. Please use the most relevant table(s). \n5. If the
question has been asked and answered before, please repeat the answer exactl
y as it was given before. \n"}, {"role": "user", "content": " \n    There ar
e 3 tables: artists, albums and tracks, where albums and artists are linked
by ArtistId, albums and tracks are linked by AlbumId,\n    Can you find the
top 10 most popular artists based on the number of tracks\n"}, {"role": "ass
istant", "content": "SELECT ar.Name AS ArtistName, COUNT(t.TrackId) as Numbe
rOfTracks\nFROM artists ar\nJOIN albums al ON ar.ArtistId = al.ArtistId\nJOI
N tracks t ON al.AlbumId = t.AlbumId\nGROUP BY ar.ArtistId\nORDER BY NumberO
fTracks DESC\nLIMIT \n  10"}, {"role": "user", "content": "  \n    List all
albums and their corresponding artist names  \n"}, {"role": "assistant", "co
ntent": "SELECT a.Title, ar.Name AS ArtistName\nFROM albums a\nJOIN artists
ar ON a.ArtistId = ar.ArtistId"}, {"role": "user", "content": "  \n    List
all genres and the number of tracks in each genre:\n"}, {"role": "assistan
t", "content": "SELECT g.GenreId, g.Name, COUNT(t.TrackId) AS NumberOfTracks
\nFROM genres g\nLEFT JOIN tracks t ON g.GenreId = t.GenreId\nGROUP BY g.Gen
reId"}, {"role": "user", "content": "  \n    Hint: album quantity is found i
n invoice_items, \n    \n    Find the top 5 customers who bought the most al
bums in total quantity (across all invoices):\n"}, {"role": "assistant", "co
ntent": "SELECT i.CustomerId, SUM(ii.Quantity) AS TotalAlbumsBought\nFROM in
voices i\nJOIN invoice_items ii ON i.InvoiceId = ii.InvoiceId\nJOIN albums a
l ON ii.TrackId IN (SELECT TrackId FROM tracks WHERE AlbumId = al.AlbumId)\n
GROUP BY i.CustomerId\nORDER BY TotalAlbumsBought DESC \nLIMIT 5"}, {"role":
"user", "content": "  \n    Find the top 5 most expensive tracks (based on u
nit price):\n"}, {"role": "assistant", "content": "SELECT TrackId, Name, Uni
tPrice FROM tracks ORDER BY UnitPrice DESC LIMIT 5"}, {"role": "user", "cont
ent": "  \n    Find all tracks with a name containing \"What\" (case-insensi
tive)\n"}, {"role": "assistant", "content": "SELECT DISTINCT LOWER(name) AS
Pattern FROM tracks WHERE lower(Name) LIKE '%what%'"}, {"role": "user", "con
tent": "Can you list all tables in the SQLite database catalog?"}, {"role":
"assistant", "content": "SELECT name FROM sqlite_master WHERE type = 'tabl
e'"}, {"role": "user", "content": "How many customers are there"}, {"role":
"assistant", "content": "SELECT COUNT(*) AS NumberOfCustomers FROM customer
s"}, {"role": "user", "content": "  \n     Find the customer with the most i
nvoices \n"}, {"role": "assistant", "content": "SELECT c.*, COUNT(i.InvoiceI
d) AS InvoiceCount\nFROM customers c JOIN invoices i ON c.CustomerId = i.Cus
tomerId\nGROUP BY c.CustomerId\nORDER BY InvoiceCount DESC\nLIMIT 1"}, {"rol
e": "user", "content": "  \n    List all invoices with a total exceeding $1
0:\n"}, {"role": "assistant", "content": "SELECT * FROM invoices WHERE Total
> 10.0"}, {"role": "user", "content": "  \n     Identify artists who have al
bums with tracks appearing in multiple genres:\n\n\n"}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T22:30:49.840658107Z', 'messa
ge': {'role': 'assistant', 'content': 'First, we need to find the number of
distinct genres for each album and then filter those that appear in more tha
n one genre:\n\n```sql\nSELECT al.ArtistId, ar.Name AS ArtistName, COUNT(DIS
TINCT t.GenreId) as NumberOfDistinctGenres\nFROM albums al\nJOIN artists ar
ON al.ArtistId = ar.ArtistId\nJOIN tracks t ON al.AlbumId = t.AlbumId\nGROUP
BY al.ArtistId, al.AlbumId HAVING COUNT(DISTINCT t.GenreId) > 1;\n```'}, 'do
ne_reason': 'stop', 'done': True, 'total_duration': 145036046706, 'load_dura
```

~~~
tion': 3573468, 'prompt_eval_count': 1635, 'prompt_eval_duration': 102104776
000, 'eval_count': 138, 'eval_duration': 41619791000}
LLM Response: First, we need to find the number of distinct genres for each
album and then filter those that appear in more than one genre:

```sql
SELECT al.ArtistId, ar.Name AS ArtistName, COUNT(DISTINCT t.GenreId) as Numb
erOfDistinctGenres
FROM albums al
JOIN artists ar ON al.ArtistId = ar.ArtistId
JOIN tracks t ON al.AlbumId = t.AlbumId
GROUP BY al.ArtistId, al.AlbumId HAVING COUNT(DISTINCT t.GenreId) > 1;
```
Info: Output from LLM: First, we need to find the number of distinct genres
for each album and then filter those that appear in more than one genre:

```sql
SELECT al.ArtistId, ar.Name AS ArtistName, COUNT(DISTINCT t.GenreId) as Numb
erOfDistinctGenres
FROM albums al
JOIN artists ar ON al.ArtistId = ar.ArtistId
JOIN tracks t ON al.AlbumId = t.AlbumId
GROUP BY al.ArtistId, al.AlbumId HAVING COUNT(DISTINCT t.GenreId) > 1;
```
Extracted SQL: SELECT al.ArtistId, ar.Name AS ArtistName, COUNT(DISTINCT t.G
enreId) as NumberOfDistinctGenres
FROM albums al
JOIN artists ar ON al.ArtistId = ar.ArtistId
JOIN tracks t ON al.AlbumId = t.AlbumId
GROUP BY al.ArtistId, al.AlbumId HAVING COUNT(DISTINCT t.GenreId) > 1
SELECT al.ArtistId, ar.Name AS ArtistName, COUNT(DISTINCT t.GenreId) as Numb
erOfDistinctGenres
FROM albums al
JOIN artists ar ON al.ArtistId = ar.ArtistId
JOIN tracks t ON al.AlbumId = t.AlbumId
GROUP BY al.ArtistId, al.AlbumId HAVING COUNT(DISTINCT t.GenreId) > 1
    ArtistId            ArtistName  NumberOfDistinctGenres
0         81          Eric Clapton                       2
1         90           Iron Maiden                       2
2         90           Iron Maiden                       2
3         90           Iron Maiden                       2
4        100         Lenny Kravitz                       3
5        147  Battlestar Galactica                       3
6        148                Heroes                       2
7        149                  Lost                       2
8        149                  Lost                       2
9        149                  Lost                       2
10       156            The Office                       2
Info: Ollama parameters:
model=phi3:14b,
options={},
keep_alive=None
Info: Prompt Content:
[{"role": "system", "content": "The following is a pandas DataFrame that con
tains the results of the query that answers the question the user asked: '
\n     Identify artists who have albums with tracks appearing in multiple ge
~~~

```
nres:\n\n\n'\n\nThe DataFrame was produced using this query: SELECT al.Artis
tId, ar.Name AS ArtistName, COUNT(DISTINCT t.GenreId) as NumberOfDistinctGen
res\nFROM albums al\nJOIN artists ar ON al.ArtistId = ar.ArtistId\nJOIN trac
ks t ON al.AlbumId = t.AlbumId\nGROUP BY al.ArtistId, al.AlbumId HAVING COUN
T(DISTINCT t.GenreId) > 1\n\nThe following is information about the resultin
g pandas DataFrame 'df': \nRunning df.dtypes gives:\n ArtistId
int64\nArtistName                object\nNumberOfDistinctGenres     int64\nd
type: object"}, {"role": "user", "content": "Can you generate the Python plo
tly code to chart the results of the dataframe? Assume the data is in a pand
as dataframe called 'df'. If there is only one value in the dataframe, use a
n Indicator. Respond with only Python code. Do not answer with any explanati
ons -- just the code."}]
Info: Ollama Response:
{'model': 'phi3:14b', 'created_at': '2024-08-01T22:31:30.120839638Z', 'messa
ge': {'role': 'assistant', 'content': '```python\nimport plotly.express as p
x\n\n# Assuming df is your DataFrame containing the query results\nfig = px.
bar(df, x=\'ArtistName\', y=\'NumberOfDistinctGenres\')\nfig.update_layout(t
itle="Artists with Albums in Multiple Genres")\nfig.show()\n```'}, 'done_rea
son': 'stop', 'done': True, 'total_duration': 40254060743, 'load_duration':
3013169, 'prompt_eval_count': 289, 'prompt_eval_duration': 17847149000, 'eva
l_count': 75, 'eval_duration': 22269964000}
```



Out[39]:

```
('SELECT al.ArtistId, ar.Name AS ArtistName, COUNT(DISTINCT t.GenreId) as N
umberOfDistinctGenres\nFROM albums al\nJOIN artists ar ON al.ArtistId = ar.
ArtistId\nJOIN tracks t ON al.AlbumId = t.AlbumId\nGROUP BY al.ArtistId, a
l.AlbumId HAVING COUNT(DISTINCT t.GenreId) > 1',
     ArtistId            ArtistName  NumberOfDistinctGenres
 0         81          Eric Clapton                       2
 1         90           Iron Maiden                       2
 2         90           Iron Maiden                       2
 3         90           Iron Maiden                       2
 4        100         Lenny Kravitz                       3
 5        147  Battlestar Galactica                       3
 6        148                Heroes                       2
 7        149                  Lost                       2
 8        149                  Lost                       2
 9        149                  Lost                       2
 10       156            The Office                       2,
 Figure({
     'data': [{'alignmentgroup': 'True',
               'hovertemplate': 'ArtistName=%{x}<br>NumberOfDistinctGenres
=%{y}<extra></extra>',
               'legendgroup': '',
               'marker': {'color': '#636efa', 'pattern': {'shape': ''}},
               'name': '',
               'offsetgroup': '',
               'orientation': 'v',
               'showlegend': False,
               'textposition': 'auto',
               'type': 'bar',
               'x': array(['Eric Clapton', 'Iron Maiden', 'Iron Maiden', 'I
ron Maiden',
                           'Lenny Kravitz', 'Battlestar Galactica', 'Heroe
s', 'Lost', 'Lost',
                           'Lost', 'The Office'], dtype=object),
               'xaxis': 'x',
               'y': array([2, 2, 2, 2, 3, 3, 2, 2, 2, 2, 2]),
               'yaxis': 'y'}],
     'layout': {'barmode': 'relative',
                'legend': {'tracegroupgap': 0},
                'margin': {'t': 60},
                'template': '...',
                'title': {'text': 'Artists with Albums in Multiple Genre
s'},
                'xaxis': {'anchor': 'y', 'domain': [0.0, 1.0], 'title': {'t
ext': 'ArtistName'}},
                'yaxis': {'anchor': 'x', 'domain': [0.0, 1.0], 'title': {'t
ext': 'NumberOfDistinctGenres'}}}
 }))
```

Check completion time

In [ ]:

```
```

In [40]:

```
ts_stop = time()
```

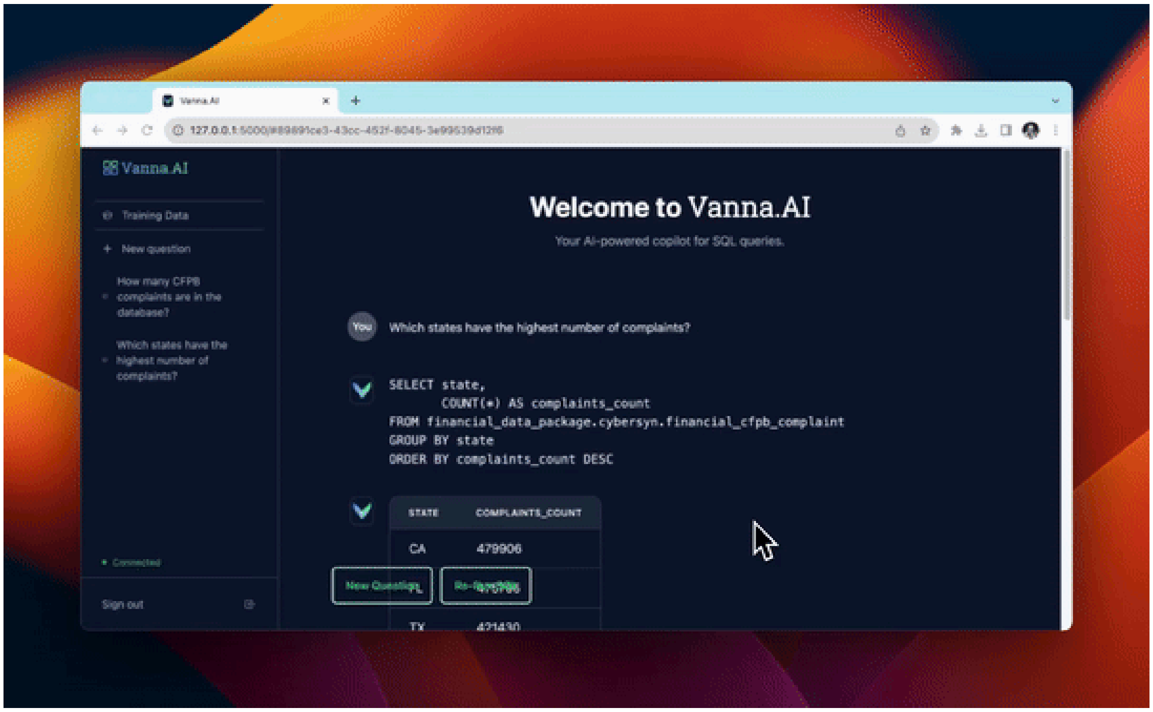
```
elapsed_time = ts_stop - ts_start
print(f"test running on '{hostname}' with '{model_name}' LLM took : {elapsed
```

```
test running on 'ducklover1' with 'phi3:14b' LLM took : 3844.80 sec
```

```
In [41]: from datetime import datetime
         print(datetime.now())
```

```
2024-08-01 18:31:30.201214
```

Launch the User Interface



Welcome to Vanna.AI

Your AI-powered copilot for SQL queries.

How many CFPB complaints are in the database?

Which states have the highest number of complaints?

Which states have the highest number of complaints?

```
SELECT state,
       COUNT(*) AS complaints_count
FROM financial_data_package.cybersyn.financial_cfpb_complaint
GROUP BY state
ORDER BY complaints_count DESC
```

STATE	COMPLAINTS_COUNT
CA	479906
FL	[illegible]
TX	421430

from vanna.flask import VannaFlaskApp app = VannaFlaskApp(vn) app.run()

Next Steps

Using Vanna via Jupyter notebooks is great for getting started but check out additional customizable interfaces like the

- Streamlit app
- Flask app
- Slackbot