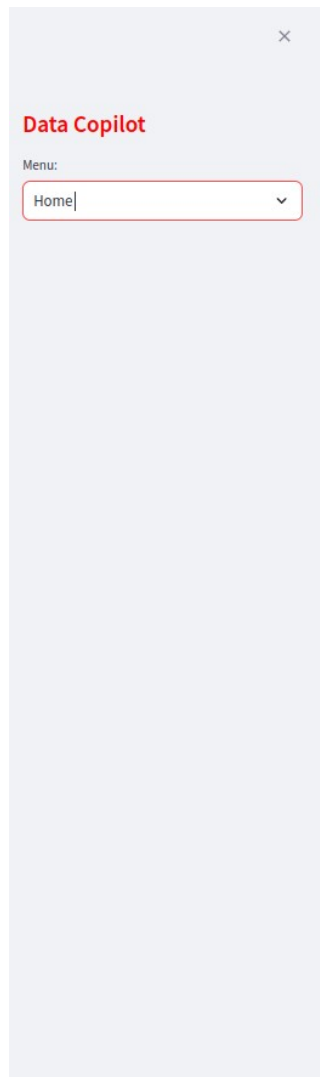


Self-Service Analytics



Deploy

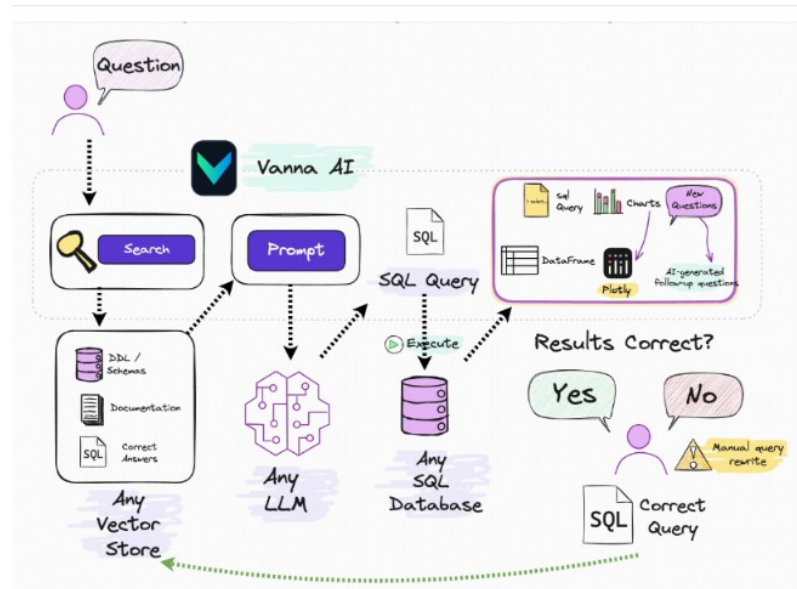
What is Data Copilot?

Data Copilot is a game-changer that streamlines the data-to-insight life-cycle. It is an AI-powered assistant, built on cutting-edge LLM models, empowers data scientists, engineers, and analysts to unlock insights from data faster than ever, allows them to focus on deeper analysis and strategic decision-making. Imagine asking questions in plain English, having them translated into SQL query and python script on the fly, then receiving results in informative texts and visual plots.

Tech-stack:

- [RAG](#)
- [Vanna.AI](#)
- [Ollama](#)
- [Streamlit](#)

How Vanna works



Data Copilot

Menu:

Train

Model 'starcoder2' selected

Deplo

Manage Training Data

Show Training data

Show

	id	question	content	training_data_type	
0	044ba63a-a15d-5339-b4f1-950e93c541be-ddl	None	CREATE TABLE [Invoice] ([InvoiceId] INTEGER NOT NULL, [CustomerId] INTEGER	ddl	
1	291a67cf-a386-5e8a-b858-dee5a9060a31-ddl	None	CREATE TABLE [Artist] ([ArtistId] INTEGER NOT NULL, [Name] NVARCHAR(120),	ddl	
2	3337d9c1-6447-541d-b8d3-9a90f3f85fc8-ddl	None	CREATE INDEX [IFK_AlbumArtistId] ON [Album] ([ArtistId])	ddl	
3	344578ab-d80f-52d4-976e-78a3e54d3c6d-ddl	None	CREATE TABLE [Employee] ([EmployeeId] INTEGER NOT NULL, [LastName] NVAR	ddl	
4	3858dfec-459a-53b5-afc9-8d854748161f-ddl	None	CREATE TABLE [MediaType] ([MediaTypeId] INTEGER NOT NULL, [Name] NVARCH	ddl	

Add Training data

DDL script

```
CREATE TABLE IF NOT EXISTS t_person (  
  id INT PRIMARY KEY,  
  name text,  
  email text
```

Add DDL script

SQL query

```
select * from t_book;
```

Add SQL query

Documentation

```
table "t_book" stores information on book title and author
```

Data Copilot

Menu:

Ask AI

Model 'starcoder2' selected

Output Settings

- ☒ Show SQL Query
- ☒ Show Dataframe
- ☒ Show Python Code
- ☒ Show Plotly Chart
- ☒ Show Summary

Reset

Ask AI



What are the top 5 customers by sales



```
1 SELECT c.FirstName, c.LastName, SUM(i.Total) AS TotalSales
2 FROM Invoice i
3 JOIN Customer c ON i.CustomerId = c.CustomerId
4 GROUP BY c.FirstName, c.LastName
5 ORDER BY TotalSales DESC
6 LIMIT 5
```



	FirstName	LastName	TotalSales
0	Helena	Holý	49.62
1	Richard	Cunningham	47.62
2	Luis	Rojas	46.62
3	Hugh	O'Reilly	45.62
4	Ladislav	Kovács	45.62



```
1 import plotly.express as px
2 import plotly.graph_objects as go
3
4 fig = px.bar(df, x='FirstName', y='TotalSales')
5 fig.update_layout(title='Top 5 Customers by Sales')
6
7 # Check if there is only one value in the dataframe
8 if df.shape[0] == 1:
9     fig = go.Figure(data=[go.Indicator(
10         mode="number+delta",
11         value=df['TotalSales'].values[0],
12         title="Top Customer"
```

Ask me a question about your data



Data Copilot

Menu:

Ask AI

Model 'starcoder2' selected

Output Settings

- ☒ Show SQL Query
- ☒ Show Dataframe
- ☒ Show Python Code
- ☒ Show Plotly Chart
- ☒ Show Summary

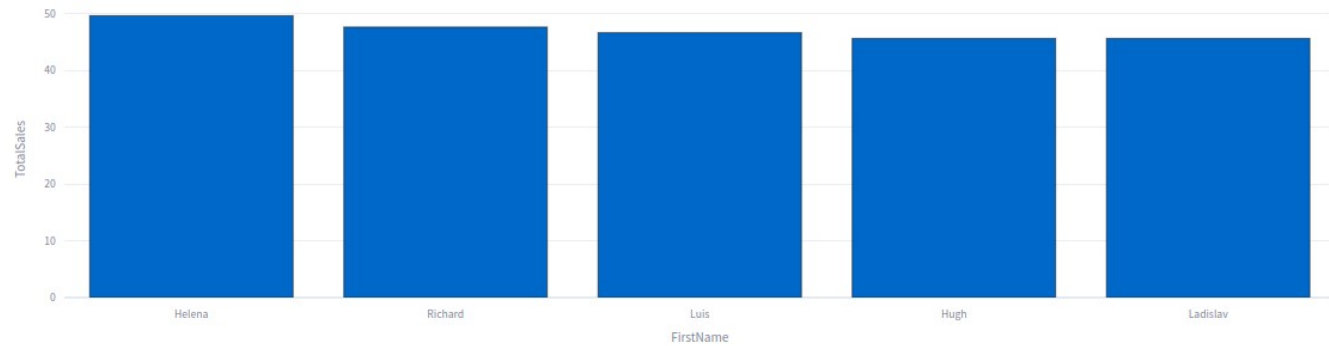
Reset



```
1 import plotly.express as px
2 import plotly.graph_objects as go
3
4 fig = px.bar(df, x='FirstName', y='TotalSales')
5 fig.update_layout(title='Top 5 Customers by Sales')
6
7 # Check if there is only one value in the dataframe
8 if df.shape[0] == 1:
9     fig = go.Figure(data=[go.Indicator(
10         mode="number+delta",
11         value=df['TotalSales'].values[0],
12         title="Top Customer"
13     )])
```



Top 5 Customers by Sales



The top 5 customers by sales are Helena Holý, Richard Cunningham, Luis Rojas, Hugh O'Reilly, and Ladislav Kovács.

Ask me a question about your data



Data Copilot

Menu:

Configure



Model 'starcoder2' selected



Deploy

Configure

Select an LLM Model:

starcoder2



starcoder2

sqlcoder

duckdb-nsql

wizardcoder

llama3

gemma:7b

gemma:2b

Resources

- Vanna.AI (<https://github.com/vanna-ai/vanna>) : Chat with Data
- Ollama (<https://ollama.com/>) : private open-source LLM
- Streamlit (<https://streamlit.io/>) : build data apps faster
- Number Stations (<https://numbersstation.ai/>) : AI-native Data Analytics
- DuckDB (<https://duckdb.org/>) : Fast in-memory OLAP