QueryEase: Simplified Database Querying with LlamaIndex 🦙

(Query Engine + Retriver)

Submitted By : Swetha Srinivasan (ss9577@rit.edu) , Sagnna Doddaguni Siddaramaiah (sd7168@rit.edu)

1. We will perform text-to-SQL over a toy dataset: this will do "retrieval" (sql query over db) and "synthesis".
2. We will show you how to define a text-to-SQL retriever.

Imports and Defining the Database

```
In [102…   # Install the llama-index-llms-openai package using pip
           # This package presumably contains Llama, an indexing library for LLMS models
           # Use the %pip magic command in IPython to install the package

           %pip install llama-index-llms-openai
```

```
Requirement already satisfied: llama-index-llms-openai in /usr/local/lib/python3.10/dist-packages (0.1.15)
Requirement already satisfied: llama-index-core<0.11.0,>=0.10.24 in /usr/local/lib/python3.10/dist-packages (from
llama-index-llms-openai) (0.10.29)
Requirement already satisfied: PyYAML>=6.0.1 in /usr/local/lib/python3.10/dist-packages (from llama-index-core<0.1
1.0,>=0.10.24->llama-index-llms-openai) (6.0.1)
Requirement already satisfied: SQLAlchemy[asyncio]>=1.4.49 in /usr/local/lib/python3.10/dist-packages (from llama-
index-core<0.11.0,>=0.10.24->llama-index-llms-openai) (2.0.29)
Requirement already satisfied: aiohttp<4.0.0,>=3.8.6 in /usr/local/lib/python3.10/dist-packages (from llama-index-
core<0.11.0,>=0.10.24->llama-index-llms-openai) (3.9.3)
Requirement already satisfied: dataclasses-json in /usr/local/lib/python3.10/dist-packages (from llama-index-core<
0.11.0,>=0.10.24->llama-index-llms-openai) (0.6.4)
Requirement already satisfied: deprecated>=1.2.9.3 in /usr/local/lib/python3.10/dist-packages (from llama-index-co
re<0.11.0,>=0.10.24->llama-index-llms-openai) (1.2.14)
Requirement already satisfied: dirtyjson<2.0.0,>=1.0.8 in /usr/local/lib/python3.10/dist-packages (from llama-inde
x-core<0.11.0,>=0.10.24->llama-index-llms-openai) (1.0.8)
```

```
Requirement already satisfied: fsspec>=2023.5.0 in /usr/local/lib/python3.10/dist-packages (from llama-index-core<
0.11.0,>=0.10.24->llama-index-llms-openai) (2023.6.0)
Requirement already satisfied: httpx in /usr/local/lib/python3.10/dist-packages (from llama-index-core<0.11.0,>=0.
10.24->llama-index-llms-openai) (0.27.0)
Requirement already satisfied: llamaindex-py-client<0.2.0,>=0.1.18 in /usr/local/lib/python3.10/dist-packages (fro
m llama-index-core<0.11.0,>=0.10.24->llama-index-llms-openai) (0.1.18)
Requirement already satisfied: nest-asyncio<2.0.0,>=1.5.8 in /usr/local/lib/python3.10/dist-packages (from llama-i
ndex-core<0.11.0,>=0.10.24->llama-index-llms-openai) (1.6.0)
Requirement already satisfied: networkx>=3.0 in /usr/local/lib/python3.10/dist-packages (from llama-index-core<0.1
1.0,>=0.10.24->llama-index-llms-openai) (3.3)
Requirement already satisfied: nltk<4.0.0,>=3.8.1 in /usr/local/lib/python3.10/dist-packages (from llama-index-cor
e<0.11.0,>=0.10.24->llama-index-llms-openai) (3.8.1)
Requirement already satisfied: numpy in /usr/local/lib/python3.10/dist-packages (from llama-index-core<0.11.0,>=0.
10.24->llama-index-llms-openai) (1.25.2)
Requirement already satisfied: openai>=1.1.0 in /usr/local/lib/python3.10/dist-packages (from llama-index-core<0.1
1.0,>=0.10.24->llama-index-llms-openai) (1.20.0)
Requirement already satisfied: pandas in /usr/local/lib/python3.10/dist-packages (from llama-index-core<0.11.0,>=
0.10.24->llama-index-llms-openai) (2.0.3)
Requirement already satisfied: pillow>=9.0.0 in /usr/local/lib/python3.10/dist-packages (from llama-index-core<0.1
1.0,>=0.10.24->llama-index-llms-openai) (9.4.0)
Requirement already satisfied: requests>=2.31.0 in /usr/local/lib/python3.10/dist-packages (from llama-index-core<
0.11.0,>=0.10.24->llama-index-llms-openai) (2.31.0)
Requirement already satisfied: tenacity<9.0.0,>=8.2.0 in /usr/local/lib/python3.10/dist-packages (from llama-index
-core<0.11.0,>=0.10.24->llama-index-llms-openai) (8.2.3)
Requirement already satisfied: tiktoken>=0.3.3 in /usr/local/lib/python3.10/dist-packages (from llama-index-core<
0.11.0,>=0.10.24->llama-index-llms-openai) (0.6.0)
Requirement already satisfied: tqdm<5.0.0,>=4.66.1 in /usr/local/lib/python3.10/dist-packages (from llama-index-co
re<0.11.0,>=0.10.24->llama-index-llms-openai) (4.66.2)
Requirement already satisfied: typing-extensions>=4.5.0 in /usr/local/lib/python3.10/dist-packages (from llama-ind
ex-core<0.11.0,>=0.10.24->llama-index-llms-openai) (4.11.0)
Requirement already satisfied: typing-inspect>=0.8.0 in /usr/local/lib/python3.10/dist-packages (from llama-index-
core<0.11.0,>=0.10.24->llama-index-llms-openai) (0.9.0)
Requirement already satisfied: wrapt in /usr/local/lib/python3.10/dist-packages (from llama-index-core<0.11.0,>=0.
10.24->llama-index-llms-openai) (1.14.1)
Requirement already satisfied: aiosignal>=1.1.2 in /usr/local/lib/python3.10/dist-packages (from aiohttp<4.0.0,>=
3.8.6->llama-index-core<0.11.0,>=0.10.24->llama-index-llms-openai) (1.3.1)
Requirement already satisfied: attrs>=17.3.0 in /usr/local/lib/python3.10/dist-packages (from aiohttp<4.0.0,>=3.8.
6->llama-index-core<0.11.0,>=0.10.24->llama-index-llms-openai) (23.2.0)
```

```
Requirement already satisfied: frozenlist>=1.1.1 in /usr/local/lib/python3.10/dist-packages (from aiohttp<4.0.0,>=
3.8.6->llama-index-core<0.11.0,>=0.10.24->llama-index-llms-openai) (1.4.1)
Requirement already satisfied: multidict<7.0,>=4.5 in /usr/local/lib/python3.10/dist-packages (from aiohttp<4.0.0,
>=3.8.6->llama-index-core<0.11.0,>=0.10.24->llama-index-llms-openai) (6.0.5)
Requirement already satisfied: yarl<2.0,>=1.0 in /usr/local/lib/python3.10/dist-packages (from aiohttp<4.0.0,>=3.
8.6->llama-index-core<0.11.0,>=0.10.24->llama-index-llms-openai) (1.9.4)
Requirement already satisfied: async-timeout<5.0,>=4.0 in /usr/local/lib/python3.10/dist-packages (from aiohttp<4.
0.0,>=3.8.6->llama-index-core<0.11.0,>=0.10.24->llama-index-llms-openai) (4.0.3)
Requirement already satisfied: pydantic>=1.10 in /usr/local/lib/python3.10/dist-packages (from llamaindex-py-clien
t<0.2.0,>=0.1.18->llama-index-core<0.11.0,>=0.10.24->llama-index-llms-openai) (2.6.4)
Requirement already satisfied: anyio in /usr/local/lib/python3.10/dist-packages (from httpx->llama-index-core<0.1
1.0,>=0.10.24->llama-index-llms-openai) (3.7.1)
Requirement already satisfied: certifi in /usr/local/lib/python3.10/dist-packages (from httpx->llama-index-core<0.
11.0,>=0.10.24->llama-index-llms-openai) (2024.2.2)
Requirement already satisfied: httpcore==1.* in /usr/local/lib/python3.10/dist-packages (from httpx->llama-index-c
ore<0.11.0,>=0.10.24->llama-index-llms-openai) (1.0.5)
Requirement already satisfied: idna in /usr/local/lib/python3.10/dist-packages (from httpx->llama-index-core<0.11.
0,>=0.10.24->llama-index-llms-openai) (3.6)
Requirement already satisfied: sniffio in /usr/local/lib/python3.10/dist-packages (from httpx->llama-index-core<0.
11.0,>=0.10.24->llama-index-llms-openai) (1.3.1)
Requirement already satisfied: h11<0.15,>=0.13 in /usr/local/lib/python3.10/dist-packages (from httpcore==1.*->htt
px->llama-index-core<0.11.0,>=0.10.24->llama-index-llms-openai) (0.14.0)
Requirement already satisfied: click in /usr/local/lib/python3.10/dist-packages (from nltk<4.0.0,>=3.8.1->llama-in
dex-core<0.11.0,>=0.10.24->llama-index-llms-openai) (8.1.7)
Requirement already satisfied: joblib in /usr/local/lib/python3.10/dist-packages (from nltk<4.0.0,>=3.8.1->llama-i
ndex-core<0.11.0,>=0.10.24->llama-index-llms-openai) (1.4.0)
Requirement already satisfied: regex>=2021.8.3 in /usr/local/lib/python3.10/dist-packages (from nltk<4.0.0,>=3.8.1
->llama-index-core<0.11.0,>=0.10.24->llama-index-llms-openai) (2023.12.25)
Requirement already satisfied: distro<2,>=1.7.0 in /usr/lib/python3/dist-packages (from openai>=1.1.0->llama-index
-core<0.11.0,>=0.10.24->llama-index-llms-openai) (1.7.0)
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.10/dist-packages (from requests>
=2.31.0->llama-index-core<0.11.0,>=0.10.24->llama-index-llms-openai) (3.3.2)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.10/dist-packages (from requests>=2.31.
0->llama-index-core<0.11.0,>=0.10.24->llama-index-llms-openai) (2.0.7)
Requirement already satisfied: greenlet!=0.4.17 in /usr/local/lib/python3.10/dist-packages (from SQLAlchemy[asynci
o]>=1.4.49->llama-index-core<0.11.0,>=0.10.24->llama-index-llms-openai) (3.0.3)
Requirement already satisfied: mypy-extensions>=0.3.0 in /usr/local/lib/python3.10/dist-packages (from typing-insp
ect>=0.8.0->llama-index-core<0.11.0,>=0.10.24->llama-index-llms-openai) (1.0.0)
```

```
Requirement already satisfied: marshmallow<4.0.0,>=3.18.0 in /usr/local/lib/python3.10/dist-packages (from datacla
sses-json->llama-index-core<0.11.0,>=0.10.24->llama-index-llms-openai) (3.21.1)
Requirement already satisfied: python-dateutil>=2.8.2 in /usr/local/lib/python3.10/dist-packages (from pandas->lla
ma-index-core<0.11.0,>=0.10.24->llama-index-llms-openai) (2.8.2)
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.10/dist-packages (from pandas->llama-index-c
ore<0.11.0,>=0.10.24->llama-index-llms-openai) (2023.4)
Requirement already satisfied: tzdata>=2022.1 in /usr/local/lib/python3.10/dist-packages (from pandas->llama-index
-core<0.11.0,>=0.10.24->llama-index-llms-openai) (2024.1)
Requirement already satisfied: exceptiongroup in /usr/local/lib/python3.10/dist-packages (from anyio->httpx->llama
-index-core<0.11.0,>=0.10.24->llama-index-llms-openai) (1.2.0)
Requirement already satisfied: packaging>=17.0 in /usr/local/lib/python3.10/dist-packages (from marshmallow<4.0.0,
>=3.18.0->dataclasses-json->llama-index-core<0.11.0,>=0.10.24->llama-index-llms-openai) (24.0)
Requirement already satisfied: annotated-types>=0.4.0 in /usr/local/lib/python3.10/dist-packages (from pydantic>=
1.10->llamaindex-py-client<0.2.0,>=0.1.18->llama-index-core<0.11.0,>=0.10.24->llama-index-llms-openai) (0.6.0)
Requirement already satisfied: pydantic-core==2.16.3 in /usr/local/lib/python3.10/dist-packages (from pydantic>=1.
10->llamaindex-py-client<0.2.0,>=0.1.18->llama-index-core<0.11.0,>=0.10.24->llama-index-llms-openai) (2.16.3)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.10/dist-packages (from python-dateutil>=2.8.2->p
andas->llama-index-core<0.11.0,>=0.10.24->llama-index-llms-openai) (1.16.0)
```

In [103…
```python
# Install the llama-index package using pip
# This package contains the Llama indexing functionality

!pip install llama-index
```

```
Requirement already satisfied: llama-index in /usr/local/lib/python3.10/dist-packages (0.10.29)
Requirement already satisfied: llama-index-agent-openai<0.3.0,>=0.1.4 in /usr/local/lib/python3.10/dist-packages (
from llama-index) (0.2.2)
Requirement already satisfied: llama-index-cli<0.2.0,>=0.1.2 in /usr/local/lib/python3.10/dist-packages (from llam
a-index) (0.1.11)
Requirement already satisfied: llama-index-core<0.11.0,>=0.10.29 in /usr/local/lib/python3.10/dist-packages (from
llama-index) (0.10.29)
Requirement already satisfied: llama-index-embeddings-openai<0.2.0,>=0.1.5 in /usr/local/lib/python3.10/dist-packa
ges (from llama-index) (0.1.7)
Requirement already satisfied: llama-index-indices-managed-llama-cloud<0.2.0,>=0.1.2 in /usr/local/lib/python3.10/
dist-packages (from llama-index) (0.1.5)
Requirement already satisfied: llama-index-legacy<0.10.0,>=0.9.48 in /usr/local/lib/python3.10/dist-packages (from
llama-index) (0.9.48)
```

```
Requirement already satisfied: llama-index-llms-openai<0.2.0,>=0.1.13 in /usr/local/lib/python3.10/dist-packages (
from llama-index) (0.1.15)
Requirement already satisfied: llama-index-multi-modal-llms-openai<0.2.0,>=0.1.3 in /usr/local/lib/python3.10/dist
-packages (from llama-index) (0.1.5)
Requirement already satisfied: llama-index-program-openai<0.2.0,>=0.1.3 in /usr/local/lib/python3.10/dist-packages
(from llama-index) (0.1.5)
Requirement already satisfied: llama-index-question-gen-openai<0.2.0,>=0.1.2 in /usr/local/lib/python3.10/dist-pac
kages (from llama-index) (0.1.3)
Requirement already satisfied: llama-index-readers-file<0.2.0,>=0.1.4 in /usr/local/lib/python3.10/dist-packages (
from llama-index) (0.1.19)
Requirement already satisfied: llama-index-readers-llama-parse<0.2.0,>=0.1.2 in /usr/local/lib/python3.10/dist-pac
kages (from llama-index) (0.1.4)
Requirement already satisfied: openai>=1.14.0 in /usr/local/lib/python3.10/dist-packages (from llama-index-agent-o
penai<0.3.0,>=0.1.4->llama-index) (1.20.0)
Requirement already satisfied: PyYAML>=6.0.1 in /usr/local/lib/python3.10/dist-packages (from llama-index-core<0.1
1.0,>=0.10.29->llama-index) (6.0.1)
Requirement already satisfied: SQLAlchemy[asyncio]>=1.4.49 in /usr/local/lib/python3.10/dist-packages (from llama-
index-core<0.11.0,>=0.10.29->llama-index) (2.0.29)
Requirement already satisfied: aiohttp<4.0.0,>=3.8.6 in /usr/local/lib/python3.10/dist-packages (from llama-index-
core<0.11.0,>=0.10.29->llama-index) (3.9.3)
Requirement already satisfied: dataclasses-json in /usr/local/lib/python3.10/dist-packages (from llama-index-core<
0.11.0,>=0.10.29->llama-index) (0.6.4)
Requirement already satisfied: deprecated>=1.2.9.3 in /usr/local/lib/python3.10/dist-packages (from llama-index-co
re<0.11.0,>=0.10.29->llama-index) (1.2.14)
Requirement already satisfied: dirtyjson<2.0.0,>=1.0.8 in /usr/local/lib/python3.10/dist-packages (from llama-inde
x-core<0.11.0,>=0.10.29->llama-index) (1.0.8)
Requirement already satisfied: fsspec>=2023.5.0 in /usr/local/lib/python3.10/dist-packages (from llama-index-core<
0.11.0,>=0.10.29->llama-index) (2023.6.0)
Requirement already satisfied: httpx in /usr/local/lib/python3.10/dist-packages (from llama-index-core<0.11.0,>=0.
10.29->llama-index) (0.27.0)
Requirement already satisfied: llamaindex-py-client<0.2.0,>=0.1.18 in /usr/local/lib/python3.10/dist-packages (fro
m llama-index-core<0.11.0,>=0.10.29->llama-index) (0.1.18)
Requirement already satisfied: nest-asyncio<2.0.0,>=1.5.8 in /usr/local/lib/python3.10/dist-packages (from llama-i
ndex-core<0.11.0,>=0.10.29->llama-index) (1.6.0)
Requirement already satisfied: networkx>=3.0 in /usr/local/lib/python3.10/dist-packages (from llama-index-core<0.1
1.0,>=0.10.29->llama-index) (3.3)
Requirement already satisfied: nltk<4.0.0,>=3.8.1 in /usr/local/lib/python3.10/dist-packages (from llama-index-cor
e<0.11.0,>=0.10.29->llama-index) (3.8.1)
```

```
Requirement already satisfied: numpy in /usr/local/lib/python3.10/dist-packages (from llama-index-core<0.11.0,>=0.
10.29->llama-index) (1.25.2)
Requirement already satisfied: pandas in /usr/local/lib/python3.10/dist-packages (from llama-index-core<0.11.0,>=
0.10.29->llama-index) (2.0.3)
Requirement already satisfied: pillow>=9.0.0 in /usr/local/lib/python3.10/dist-packages (from llama-index-core<0.1
1.0,>=0.10.29->llama-index) (9.4.0)
Requirement already satisfied: requests>=2.31.0 in /usr/local/lib/python3.10/dist-packages (from llama-index-core<
0.11.0,>=0.10.29->llama-index) (2.31.0)
Requirement already satisfied: tenacity<9.0.0,>=8.2.0 in /usr/local/lib/python3.10/dist-packages (from llama-index
-core<0.11.0,>=0.10.29->llama-index) (8.2.3)
Requirement already satisfied: tiktoken>=0.3.3 in /usr/local/lib/python3.10/dist-packages (from llama-index-core<
0.11.0,>=0.10.29->llama-index) (0.6.0)
Requirement already satisfied: tqdm<5.0.0,>=4.66.1 in /usr/local/lib/python3.10/dist-packages (from llama-index-co
re<0.11.0,>=0.10.29->llama-index) (4.66.2)
Requirement already satisfied: typing-extensions>=4.5.0 in /usr/local/lib/python3.10/dist-packages (from llama-ind
ex-core<0.11.0,>=0.10.29->llama-index) (4.11.0)
Requirement already satisfied: typing-inspect>=0.8.0 in /usr/local/lib/python3.10/dist-packages (from llama-index-
core<0.11.0,>=0.10.29->llama-index) (0.9.0)
Requirement already satisfied: wrapt in /usr/local/lib/python3.10/dist-packages (from llama-index-core<0.11.0,>=0.
10.29->llama-index) (1.14.1)
Requirement already satisfied: beautifulsoup4<5.0.0,>=4.12.3 in /usr/local/lib/python3.10/dist-packages (from llam
a-index-readers-file<0.2.0,>=0.1.4->llama-index) (4.12.3)
Requirement already satisfied: pypdf<5.0.0,>=4.0.1 in /usr/local/lib/python3.10/dist-packages (from llama-index-re
aders-file<0.2.0,>=0.1.4->llama-index) (4.2.0)
Requirement already satisfied: striprtf<0.0.27,>=0.0.26 in /usr/local/lib/python3.10/dist-packages (from llama-ind
ex-readers-file<0.2.0,>=0.1.4->llama-index) (0.0.26)
Requirement already satisfied: llama-parse<0.5.0,>=0.4.0 in /usr/local/lib/python3.10/dist-packages (from llama-in
dex-readers-llama-parse<0.2.0,>=0.1.2->llama-index) (0.4.1)
Requirement already satisfied: aiosignal>=1.1.2 in /usr/local/lib/python3.10/dist-packages (from aiohttp<4.0.0,>=
3.8.6->llama-index-core<0.11.0,>=0.10.29->llama-index) (1.3.1)
Requirement already satisfied: attrs>=17.3.0 in /usr/local/lib/python3.10/dist-packages (from aiohttp<4.0.0,>=3.8.
6->llama-index-core<0.11.0,>=0.10.29->llama-index) (23.2.0)
Requirement already satisfied: frozenlist>=1.1.1 in /usr/local/lib/python3.10/dist-packages (from aiohttp<4.0.0,>=
3.8.6->llama-index-core<0.11.0,>=0.10.29->llama-index) (1.4.1)
Requirement already satisfied: multidict<7.0,>=4.5 in /usr/local/lib/python3.10/dist-packages (from aiohttp<4.0.0,
>=3.8.6->llama-index-core<0.11.0,>=0.10.29->llama-index) (6.0.5)
Requirement already satisfied: yarl<2.0,>=1.0 in /usr/local/lib/python3.10/dist-packages (from aiohttp<4.0.0,>=3.
8.6->llama-index-core<0.11.0,>=0.10.29->llama-index) (1.9.4)
```

```
Requirement already satisfied: async-timeout<5.0,>=4.0 in /usr/local/lib/python3.10/dist-packages (from aiohttp<4.
0.0,>=3.8.6->llama-index-core<0.11.0,>=0.10.29->llama-index) (4.0.3)
Requirement already satisfied: soupsieve>1.2 in /usr/local/lib/python3.10/dist-packages (from beautifulsoup4<5.0.
0,>=4.12.3->llama-index-readers-file<0.2.0,>=0.1.4->llama-index) (2.5)
Requirement already satisfied: pydantic>=1.10 in /usr/local/lib/python3.10/dist-packages (from llamaindex-py-clien
t<0.2.0,>=0.1.18->llama-index-core<0.11.0,>=0.10.29->llama-index) (2.6.4)
Requirement already satisfied: anyio in /usr/local/lib/python3.10/dist-packages (from httpx->llama-index-core<0.1
1.0,>=0.10.29->llama-index) (3.7.1)
Requirement already satisfied: certifi in /usr/local/lib/python3.10/dist-packages (from httpx->llama-index-core<0.
11.0,>=0.10.29->llama-index) (2024.2.2)
Requirement already satisfied: httpcore==1.* in /usr/local/lib/python3.10/dist-packages (from httpx->llama-index-c
ore<0.11.0,>=0.10.29->llama-index) (1.0.5)
Requirement already satisfied: idna in /usr/local/lib/python3.10/dist-packages (from httpx->llama-index-core<0.11.
0,>=0.10.29->llama-index) (3.6)
Requirement already satisfied: sniffio in /usr/local/lib/python3.10/dist-packages (from httpx->llama-index-core<0.
11.0,>=0.10.29->llama-index) (1.3.1)
Requirement already satisfied: h11<0.15,>=0.13 in /usr/local/lib/python3.10/dist-packages (from httpcore==1.*->htt
px->llama-index-core<0.11.0,>=0.10.29->llama-index) (0.14.0)
Requirement already satisfied: click in /usr/local/lib/python3.10/dist-packages (from nltk<4.0.0,>=3.8.1->llama-in
dex-core<0.11.0,>=0.10.29->llama-index) (8.1.7)
Requirement already satisfied: joblib in /usr/local/lib/python3.10/dist-packages (from nltk<4.0.0,>=3.8.1->llama-i
ndex-core<0.11.0,>=0.10.29->llama-index) (1.4.0)
Requirement already satisfied: regex>=2021.8.3 in /usr/local/lib/python3.10/dist-packages (from nltk<4.0.0,>=3.8.1
->llama-index-core<0.11.0,>=0.10.29->llama-index) (2023.12.25)
Requirement already satisfied: distro<2,>=1.7.0 in /usr/lib/python3/dist-packages (from openai>=1.14.0->llama-inde
x-agent-openai<0.3.0,>=0.1.4->llama-index) (1.7.0)
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.10/dist-packages (from requests>
=2.31.0->llama-index-core<0.11.0,>=0.10.29->llama-index) (3.3.2)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.10/dist-packages (from requests>=2.31.
0->llama-index-core<0.11.0,>=0.10.29->llama-index) (2.0.7)
Requirement already satisfied: greenlet!=0.4.17 in /usr/local/lib/python3.10/dist-packages (from SQLAlchemy[asynci
o]>=1.4.49->llama-index-core<0.11.0,>=0.10.29->llama-index) (3.0.3)
Requirement already satisfied: mypy-extensions>=0.3.0 in /usr/local/lib/python3.10/dist-packages (from typing-insp
ect>=0.8.0->llama-index-core<0.11.0,>=0.10.29->llama-index) (1.0.0)
Requirement already satisfied: marshmallow<4.0.0,>=3.18.0 in /usr/local/lib/python3.10/dist-packages (from datacla
sses-json->llama-index-core<0.11.0,>=0.10.29->llama-index) (3.21.1)
Requirement already satisfied: python-dateutil>=2.8.2 in /usr/local/lib/python3.10/dist-packages (from pandas->lla
ma-index-core<0.11.0,>=0.10.29->llama-index) (2.8.2)
```

```
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.10/dist-packages (from pandas->llama-index-c
ore<0.11.0,>=0.10.29->llama-index) (2023.4)
Requirement already satisfied: tzdata>=2022.1 in /usr/local/lib/python3.10/dist-packages (from pandas->llama-index
-core<0.11.0,>=0.10.29->llama-index) (2024.1)
Requirement already satisfied: exceptiongroup in /usr/local/lib/python3.10/dist-packages (from anyio->httpx->llama
-index-core<0.11.0,>=0.10.29->llama-index) (1.2.0)
Requirement already satisfied: packaging>=17.0 in /usr/local/lib/python3.10/dist-packages (from marshmallow<4.0.0,
>=3.18.0->dataclasses-json->llama-index-core<0.11.0,>=0.10.29->llama-index) (24.0)
Requirement already satisfied: annotated-types>=0.4.0 in /usr/local/lib/python3.10/dist-packages (from pydantic>=
1.10->llamaindex-py-client<0.2.0,>=0.1.18->llama-index-core<0.11.0,>=0.10.29->llama-index) (0.6.0)
Requirement already satisfied: pydantic-core==2.16.3 in /usr/local/lib/python3.10/dist-packages (from pydantic>=1.
10->llamaindex-py-client<0.2.0,>=0.1.18->llama-index-core<0.11.0,>=0.10.29->llama-index) (2.16.3)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.10/dist-packages (from python-dateutil>=2.8.2->p
andas->llama-index-core<0.11.0,>=0.10.29->llama-index) (1.16.0)
```

In [104…
```python
# Import necessary libraries
import os  # For interacting with the operating system
import openai  # For accessing OpenAI's API
```

In [105…
```python
# Set the OpenAI API key in the environment variable OPENAI_API_KEY
os.environ["OPENAI_API_KEY"] = "sk-proj-XWgXa7kxYx5vwwcP69Y5T3BlbkFJ8Xrc6e3jrYunzuOIdQOd"

# Access the OpenAI API key from the environment variable and assign it to openai.api_key
openai.api_key = os.environ["OPENAI_API_KEY"]
```

In [106…
```python
# import logging
# import sys

# logging.basicConfig(stream=sys.stdout, level=logging.INFO)
# logging.getLogger().addHandler(logging.StreamHandler(stream=sys.stdout))
```

In [107…
```python
# Import the necessary module from IPython to display Markdown content
from IPython.display import Markdown, display
```

Create Database Schema

We use `sqlalchemy`, a popular SQL database toolkit, to create an empty `city_stats` Table

In [108…
```python
from sqlalchemy import (
    create_engine,
    MetaData,
    Table,
    Column,
    String,
    Integer,
    select,
)
```

In [109…
```python
# Create a SQLite in-memory database engine
engine = create_engine("sqlite:///:memory:")

# Initialize a metadata object to hold information about database tables
metadata_obj = MetaData()
```

```python
In [110…    # create city SQL table
            table_name = "city_stats"
            city_stats_table = Table(
                table_name,
                metadata_obj,
                Column("city_name", String(16), primary_key=True),
                Column("population", Integer),
                Column("country", String(16), nullable=False),
            )
            metadata_obj.create_all(engine)
```

Define SQL Database

We first define our `SQLDatabase` abstraction (a light wrapper around SQLAlchemy). This abstraction aims to simplify common database operations and provide a more user-friendly way to interact with the database. This abstraction can improve code readability, maintainability, and ease of use when working with databases in Python applications.

```python
In [111…    from llama_index.core import SQLDatabase
            from llama_index.llms.openai import OpenAI
```

```python
In [112…    llm = OpenAI(temperature=0.1, model="gpt-3.5-turbo")
```

```python
In [113…    sql_database = SQLDatabase(engine, include_tables=["city_stats"])
```

We add some testing data to our SQL database.

In [114…
```python
sql_database = SQLDatabase(engine, include_tables=["city_stats"])
from sqlalchemy import insert

rows = [
    {"city_name": "Toronto", "population": 2930000, "country": "Canada"},
    {"city_name": "Tokyo", "population": 13960000, "country": "Japan"},
    {
        "city_name": "Chicago",
        "population": 2679000,
        "country": "United States",
    },
    {"city_name": "Seoul", "population": 9776000, "country": "South Korea"},
]
for row in rows:
    stmt = insert(city_stats_table).values(**row)
    with engine.begin() as connection:
        cursor = connection.execute(stmt)
```

In [115…
```python
# view current table
stmt = select(
    city_stats_table.c.city_name,
    city_stats_table.c.population,
    city_stats_table.c.country,
).select_from(city_stats_table)

with engine.connect() as connection:
    results = connection.execute(stmt).fetchall()
    print(results)
```

```
[('Toronto', 2930000, 'Canada'), ('Tokyo', 13960000, 'Japan'), ('Chicago', 2679000, 'United States'), ('Seoul', 97
76000, 'South Korea')]
```

Query Index

We first show how we can execute a raw SQL query, which directly executes over the table.

In [116…
```python
from sqlalchemy import text

with engine.connect() as con:
    rows = con.execute(text("SELECT city_name from city_stats"))
    for row in rows:
        print(row)
```

```
('Chicago',)
('Seoul',)
('Tokyo',)
('Toronto',)
```

Text-to-SQL Query Engine

We have constructed our SQL database, we can use the NLSQLTableQueryEngine to construct natural language queries that are synthesized into SQL queries.

We need to specify the tables we want to use with this query engine. If we don't the query engine will pull all the schema context, which could overflow the context window of the LLM.

In [117…
```python
from llama_index.core.query_engine import NLSQLTableQueryEngine

query_engine = NLSQLTableQueryEngine(
    sql_database=sql_database, tables=["city_stats"], llm=llm
)
```

In [118…
```python
query_str = "Which city has the highest population?"
response = query_engine.query(query_str)
```

In [119…
```python
display(Markdown(f"<b>{response}</b>"))
```

Tokyo has the highest population among all cities, with a population of 13,960,000.

Text-to-SQL Retriever

So far our text-to-SQL capability is packaged in a query engine and consists of both retrieval and synthesis.

We can use SQL retriever on its own. We show some different parameters and also show how to plug it into our `RetrieverQueryEngine` to get roughly the same results.

In [120…
```python
from llama_index.core.retrievers import NLSQLRetriever

# default retrieval (return_raw=True)
nl_sql_retriever = NLSQLRetriever(
    sql_database, tables=["city_stats"], return_raw=True
)
```

In [121…
```python
results = nl_sql_retriever.retrieve(
    "Return the top 5 cities (along with their populations) with the highest population."
)
```

In [122…
```python
from llama_index.core.response.notebook_utils import display_source_node

for n in results:
    display_source_node(n)
```

Node ID: 09b14ab8-1423-420c-bea9-8909ffccbb32

Similarity: None

Text: [('Tokyo', 13960000), ('Seoul', 9776000), ('Toronto', 2930000), ('Chicago', 2679000)]

In [123…
```python
# default retrieval (return_raw=False)
nl_sql_retriever = NLSQLRetriever(
    sql_database, tables=["city_stats"], return_raw=False
)
```

In [124…
```python
results = nl_sql_retriever.retrieve(
    "Return the top 5 cities (along with their populations) with the highest population."
)
```

In [125…
```python
# NOTE: all the content is in the metadata
for n in results:
    display_source_node(n, show_source_metadata=True)
```

Node ID: de3ac13b-6b25-4664-a74e-257e0bbc90d9

Similarity: None

Text:

Metadata: {'city_name': 'Tokyo', 'population': 13960000}

Node ID: edaffa47-a0e4-4b59-9922-c964bc8a6fbb

Similarity: None

Text:

Metadata: {'city_name': 'Seoul', 'population': 9776000}

Node ID: 3d946cbd-f3da-461f-b557-fe948c1d0607

Similarity: None

Text:

Metadata: {'city_name': 'Toronto', 'population': 2930000}

Node ID: 920a7170-c09d-443c-8013-a4fd2c5053f3

Similarity: None

Text:

Metadata: {'city_name': 'Chicago', 'population': 2679000}

Plug into `RetrieverQueryEngine`

We compose our SQL Retriever with our standard `RetrieverQueryEngine` to synthesize a response.

In [126…
```python
from llama_index.core.query_engine import RetrieverQueryEngine

query_engine = RetrieverQueryEngine.from_args(nl_sql_retriever)
```

In [127…

```python
response = query_engine.query(
    "Return the top 5 cities (along with their populations) with the highest population."
)
```

In [128…

```python
print(str(response))
```

```
Tokyo: 13,960,000
Seoul: 9,776,000
Toronto: 2,930,000
Chicago: 2,679,000
```