

Compass

My Queries

Data Modeling

CONNECTIONS [2]

Search connections

cluster0.rqommr.mongodb.net

localhost:27017

admin

config

employees

local

schooldb

socialmedia

reddit

x

test

localhost:27017 > socialmedia > reddit

Documents 7.2K

Aggregations

Schema

Indexes 1

Validation

>_ Open MongoDB shell

Type a query: { field: 'value' } or [Generate query](#)

Explain

Reset

Find

Options

ADD DATA

EXPORT DATA

UPDATE

DELETE

25

1 - 25 of 7194

↺

↻

↷

⋮

{ }

⌘

```

_id: ObjectId('691029a9e6013e37e8fac9dc')
author: "CuriousWanderer567"
likes: 21170
post_score: 21170
down_votes: 0
title: "This tortoise's reaction to the darker colored material of this person..."
subreddit: "interesting"
url: "https://v.redd.it/1v12ugaws40g1"
location: "MISC."

```

```

_id: ObjectId('691029a9e6013e37e8fac9dd')
author: "Flat-Limit5595"
likes: 14176
post_score: 14176
down_votes: 0
title: "Mom with the cat she didn't want"
subreddit: "cats"
url: "https://www.reddit.com/gallery/1os6mpn"
location: "Humor"

```

```

_id: ObjectId('691029a9e6013e37e8fac9de')
author: "montageidiots"
likes: 8620
post_score: 8620
down_votes: 0
title: "[Highlight] De'Andre Hunter breaks Giddey's ankles and then put Patric..."
subreddit: "nba"
url: "https://streamable.com/dy092f"
location: "Highlight"

```

```

_id: ObjectId('691029a9e6013e37e8fac9df')
author: "TrustMeIaLawyer"
likes: 16419

```

SEARCH

Search

Replace

SOCIAL MEDIA ANALYTICS

backendsrc

__pycache__

database_adddata.py

database.py

logs.txt

test.py

backend.py

logs.txt

OUTLINE

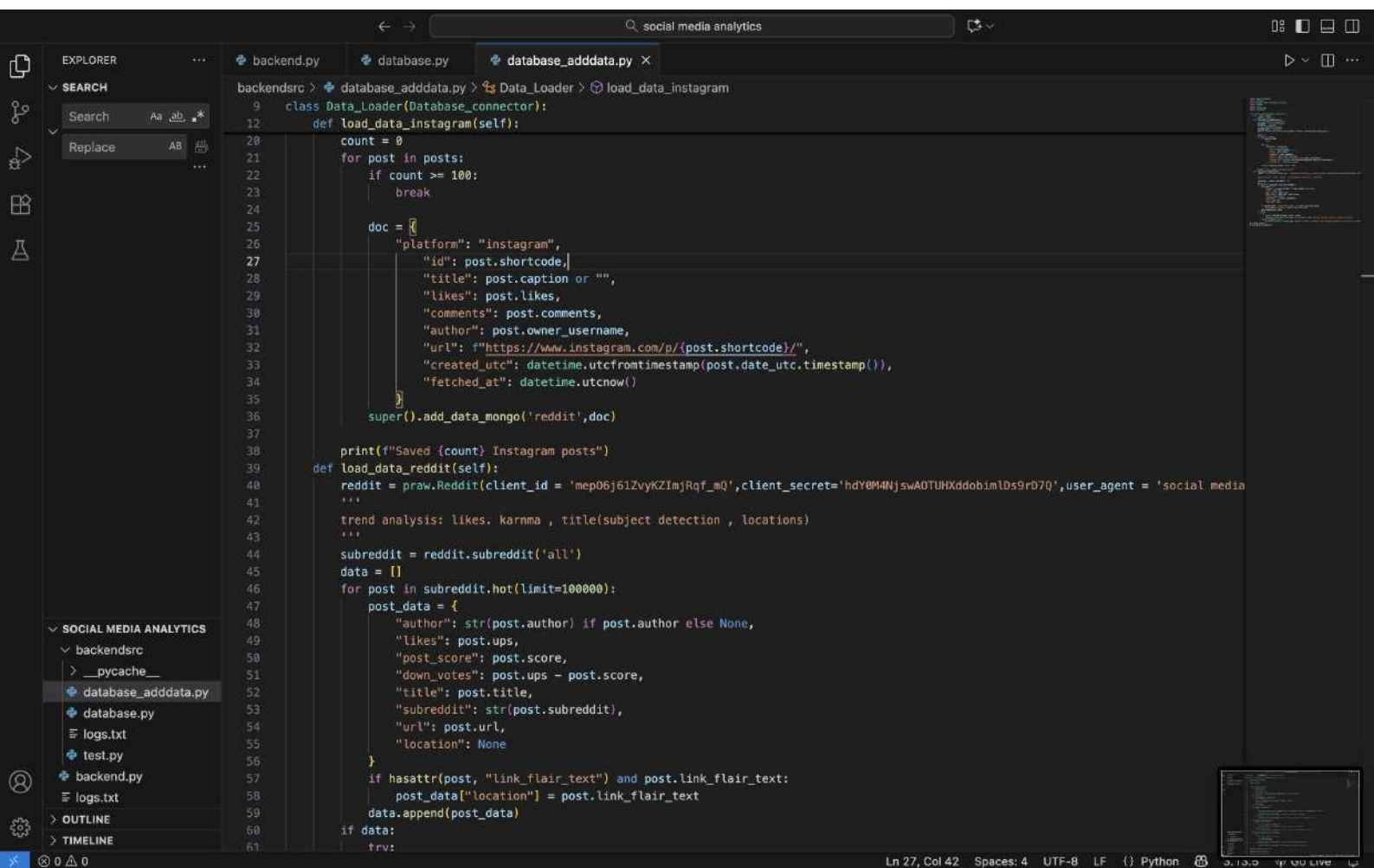
TIMELINE

backend.py database.py database_adddata.py

backendsrc > database.py > Database_connector > __init__

```
1 import mysql.connector
2 import pymongo
3 import datetime
4
5 class Database_connector:
6     def __init__(self):
7         self.conn = None
8         self.cursor = None
9         self.client = pymongo.MongoClient("mongodb://localhost:27017/")
10
11     @staticmethod
12     def write_logs(text , error=None):
13         file = open('logs.txt','a')
14         write = f'({datetime.datetime.now()}) {text} {error} \n'
15         file.write(write)
16         file.close()
17
18     #connecting to mysql database
19     def connect_sql_db(self):
20         try:
21             self.conn = mysql.connector.connect(user='root',password='harshiljain',database='socialmedia')
22             self.cursor = self.conn.cursor()
23             Database_connector.write_logs(' successfully connected to database')
24         except Exception as e:
25             Database_connector.write_logs(' error occured in connecting to the database: ',e)
26
27     #Connecting to mongodb database
28     def connect_mongo_db(self):
29         try:
30             db = self.client['socialmedia']
31             Database_connector.write_logs('Successfully connected to mongodb')
32         except Exception as e:
33             Database_connector.write_logs('Not able to connect to the mongodb database ',e)
34
35     #function to add documents in mongodb
36     def add_data_mongo(self,collection,data):
37         try:
38             db = self.client['socialmedia']
39             col = db[collection]
40             col.insert_many(data)
41             Database_connector.write_logs('successfully added to mongodb database')
42         except Exception as e:
43             Database_connector.write_logs('cannot add data to mongodb collection ',e)
44
45 database = Database_connector()
46 database.connect_sql_db()
47 database.connect_mongo_db()
48 database.add_data_mongo('x',f
```

Ln 9, Col 72 Spaces: 4 UTF-8 LF {} Python 3.13.5 Go Live



SEARCH

Search Aa

Replace AB

SOCIAL MEDIA ANALYTICS

backendsrc

> __pycache__

database_adddata.py

database.py

logs.txt

test.py

backend.py

logs.txt

> OUTLINE

> TIMELINE

backend.py > ...

1 import pymongo

2 import pandas as pd

3 import re

4 import nltk

5 from nltk.corpus import stopwords

6 from nltk.stem import WordNetLemmatizer

7 from sklearn.feature_extraction.text import TfidfVectorizer

8 from sklearn.decomposition import LatentDirichletAllocation

9 import matplotlib.pyplot as plt

10 import datetime

11

12 def download_nltk_assets():

13 try:

14 nltk.data.find('corpora/stopwords')

15 except LookupError:

16 print("Downloading NLTK 'stopwords'...")

17 nltk.download('stopwords')

18 try:

19 nltk.data.find('corpora/wordnet')

20 except LookupError:

21 print("Downloading NLTK 'wordnet'...")

22 nltk.download('wordnet')

23 try:

24 nltk.data.find('tokenizers/punkt')

25 except LookupError:

26 print("Downloading NLTK 'punkt'...")

27 nltk.download('punkt')

28

29 try:

30 nltk.data.find('tokenizers/punkt_tab')

31 except LookupError:

32 print("Downloading NLTK 'punkt_tab'...")

33 nltk.download('punkt_tab')

34

35 def fetch_reddit_data(mongo_uri="mongodb://localhost:27017/", db_name="socialmedia", collection_name="reddit"):

36 print(f"Connecting to MongoDB at {mongo_uri}...")

37 try:

38 client = pymongo.MongoClient(mongo_uri)

39 db = client[db_name]

40 collection = db[collection_name]

41

42 cursor = collection.find({}, {"title": 1})

43

44 df = pd.DataFrame(list(cursor))

Ln 159, Col 55 Spaces: 4 UTF-8 LF {} Python 3.13.5 Go Live