1. **File Structure**

Folders are bolded, files are underlined

Final Project

- get_data.py : helper functions for the project
- nyt_api.ipynb: test the helper functions and loads MongoDB with books from the API
- readme.md : contains information about the project
- config.py : contains the API key
- books_historical (1).json: local data that we exported from Mongodb
- neo4j project apan5400.ipynb: loads Neo4j with the data in a particular schematic format
- Main_Page_-_Streamlit_App.py : contains the readme for the streamlit and acts as the main page for the app
- **pages**
  - pages/1 Feature_Analysis.py
  - pages/2 Publisher_Analysis.py
  - pages/3 Seasonal_Analysis.py
  - pages/4 Sentiment_Analysis.py
  - pages/5 Lifespan_Analysis.py
  - pages/6 Co-Author_Analysis.py
  - pages/7 Neo4J Visualization.py
  - pages/books_historical (1).json

2. **Setup**

a. Open docker
b. MongoDB
   i. Create a connection called nyt_bestsellers. Use the information below to create a database.

```
# connecting to MongoDB
client = MongoClient("mongodb://localhost:27017/")
db = client['nyt_bestsellers']
books_collection = db['books']
historical_collection = db['books_historical']
```

   ii. Run the notebook titled nyt_api.ipynb. This will load MongoDB with the necessary records from the API. This will take approximately 1 hour because of API limitations.

C. Neo4j

   iii. Create a local instance called **project**. Then create a database called booksnyt. Start the instance and connect to it. Use the information below for reference.

```
database_name = "booksnyt"
```

```
username = "neo4j"
password = "12345678"
uri = "neo4j://127.0.0.1:7687"
```

iv. Run the notebook titled `neo4j project apan5400.ipynb`. This will load Neo4j with the graph schema.

D. Go to the NYT Developer portal and get an API key and put it inside the `config.py` file.

E. You are ready to run the project. Ensure your file setup is the same as the one in step 1. Go to the terminal and run: **`streamlit run Main_Page_-_Streamlit_App.py`** in the command line.