Table of contents

Chatbot Documentation

GitHub Repository

Overview

This simple chatbot was developed by **Elton S** as par of the BCG GenAI (https://www.theforage.com/simulations/bcg/gen-ai-anlo) job simulation at Forage (https://www.theforage.com/). It aims at assisting with f nancial inquiries at BCG (Boston Consulting Group) by providing responses to user queries related to f nancial data.

Getting Star ed

. Introduction (Introduction)

. Environment Setup (Setup)

. Data Extraction & Analysis with Jupyter (Data Extraction & Analysis with Jupyter)

. Chatbot (Chatbot)

About Me

Elton S

I'm a passionate data professional with a strong interest in programming, machine learning, and ar if cial intelligence.

LinkedIn GitHub

Introduction

The chatbot serves as a user-friendly inter ace for individuals seeking information about various f nancial metrics, such as revenue, income, assets, liabilities, and cash flow.

Feature improvements Over Original Project Requirements

. Enhanced User Inter ace

- Dynamic Greeting: Personalized greetings based on the time of day improve user engagement.
- Interactive Input: Users can input queries directly, streamlining the interaction.
- Enhanced Formatting and Visual Appeal: The chatbot is formatted in a way that enhances visual appeal and readability

. Natural Language Processing

- Flexible Query Handling: The chatbot understands a wider range of natural language queries.
- Keyword Recognition: Maps user inquiries to specif c f nancial metrics for intuitive interactions.

. Enhanced Formatting and Visual Appeal The chatbot's responses are formatted in a way that enhances visual appeal and readability

. Case-Insensitive Comparisons The chatbot processes user queries in a case-insensitive manner making it flexible and user-friendly.

Limitations

. The chatbot is limited to the data available in the CSV f le.

. The chatbot does not understand complex or ambiguous queries.

Setup

This project can be executed using any integrated development environment (IDE) that suppor s Python and Jupyter Notebook. Visual Studio Code (https://code.visualstudio.com/)

Basic Requirements:

Python (https://www.python.org/downloads/) Jupyter Notebook (https://jupyter.org/) pandas (https://pandas.pydata.org/)

VS Code Setup

. Install Visual Studio Code (https://code.visualstudio.com/Download)

. Install Python (https://apps.microsoft.com/search/publisher?name=Python+Software+Foundation)

. Install the necessary extensions

- Python (https://marketplace.visualstudio.com/items?itemName=ms-python.python)
- Jupyter (https://marketplace.visualstudio.com/items?itemName=ms-toolsai.jupyter)

. Install Required Libraries using pip

- Press `Ctrl+`` (backtick) or `Cmd+`` (backtick) to open the integrated terminal.
- Install pandas using the following code:

```
python pip install pandas
```

- To check if the package was installed successfully, you can run the following command:

```
python pip show pandas
```

This will display information about the installed pandas package, including its version

and location.

. Add the chatbot folder into the *VS Code Explorer* (Ctrl+E).

Data Extraction & Analysis with Jupyter

This Python code analyzes the f nancial per ormance of **Microsoft, Tesla, and Apple** using data manually extracted from the **SEC's EDGAR database** (https://www.sec.gov/edgar/search/#)). The data is entered into **Microsoft Excel** and expor ed as a CSV f le The analysis covers a period of **three years** (2021 2023, subject to data availability).

Executing the Code

. Ensure that the **Jupyter Notebook** Extension and the **pandas library** are installed.

. Open and execute the " **Task 1 - Data Extraction and Analysis.ipynb** " f le.

. Click on the `Run All` button.

What the Code Does

. Extracts data from the `extracted_data.csv` f le.

. Year-over-Year Growth Rates:

- Calculates the percentage change in key f nancial metrics (revenue, income, assets, liabilities, cash flow).
- Rounds growth rates to two decimal places.

. Groups data by company and calculates average growth rates for each metric over the period.

. Expor s processed data and summary statistics to CSV f les.

- The **Year-by-Year Growth Rate** data is extracted to `f nal_data.csv`
- The **Average Growth Rate** data is extracted to `f nal_summary.csv`

Chatbot

This chatbot is designed to assist you with f nancial inquiries based on the data provided in the f nal_data.csv and f nal_summary.csv.

Executing the Code

. Ensure you have Jupyter Python libraries like Pandas installed.

. Open and execute the " **Task 1 - Data Extraction and Analysis.ipynb** " f le.

. Press `F5` and click on " ***Python File*** *Debug the currently active Python f le* " in the dropdown menu to star the chatbot

Chatbot Syntax

The chatbot can understand queries in a flexible format, such as " `[Metric]` `[Company]` `[Year]` " or any variation of this order.

Mandatory Information:

Metric: The f nancial metric you're interested in (revenue, income, assets, liabilities, cash flow). Company or Year: At least one of these must be specif ed.

Example queries:

Display the total income for Tesla across all available years:

```
income tesla
```

Display the total assets for all companies in the year 2022:

```
asset 2022
```

Display the total revenue for Microsoft in the year 2023:

```
revenue microsoft 2023
```