horizontal line

KANINI

Internship Documentation

**Intern Name :** Karthikeyan S

**Internship Duration :** May 13, 2024, to July 12, 2024

**Mentor :** Mr Jamal Asif

**Department :** Data Analysis

**Company :** Kanini Software Solutions

#### 

#### 

#### **Objectives**

**Primary Objective:**The primary objective of this internship was to gain practical experience in data analysis and contribute to real-world projects within the company.

**Secondary Objectives:**The secondary objectives included:

* Enhancing my data analysis skills and learning new tools and techniques.
* Understanding the data analysis process from start to finish.
* Developing teamwork and communication skills.
* Utilizing my college leave effectively to become more proficient in my field.

**Projects and Tasks**

##### **Project 1: Data Visualization and Reporting**

**Description:**This project involved the creation of interactive dashboards and reports to visualize data insights for the marketing team.

**Objectives:**The main objective was to provide clear and concise visualizations of key performance indicators and trends.

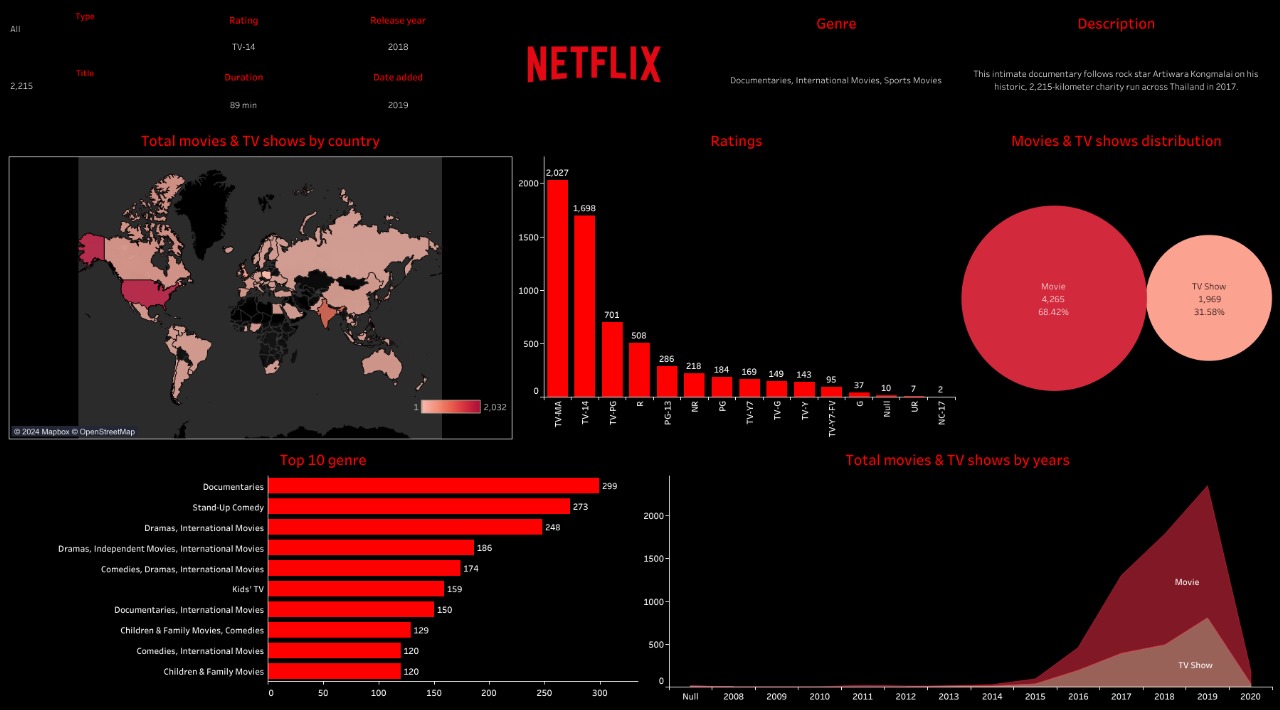
**Technologies Used:**

* Tableau
* Power BI

**Role and Contributions:**My role included designing and developing dashboards to meet stakeholder requirements and provide actionable insights.

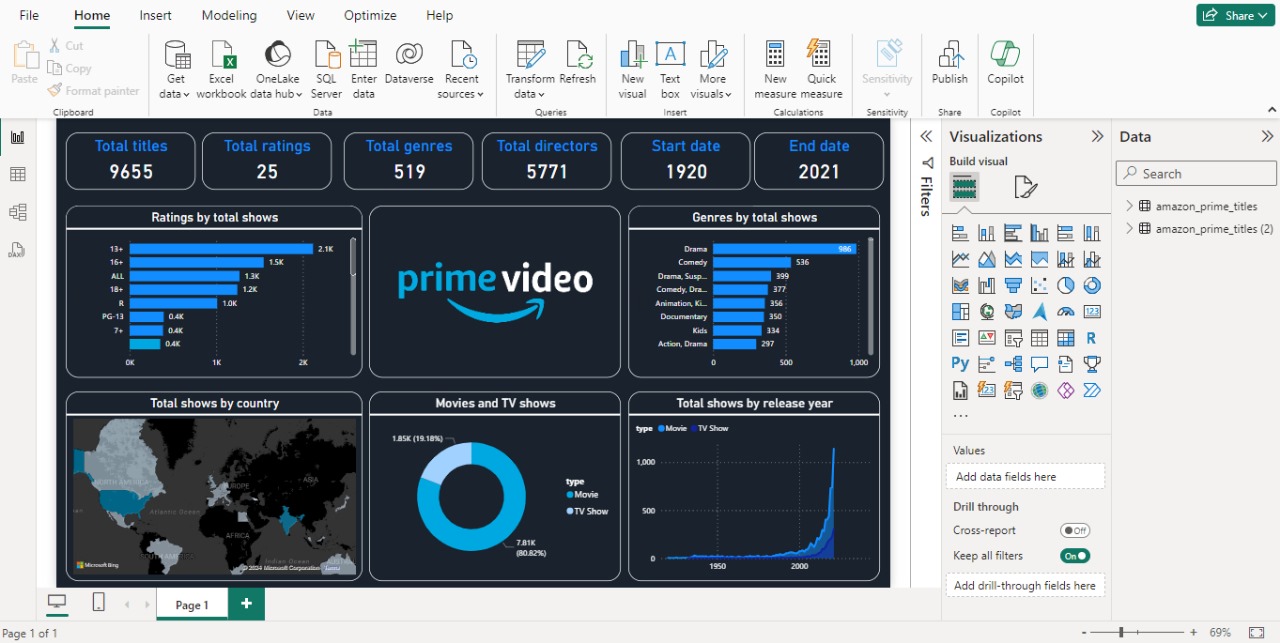
**Outcomes:**The project resulted in several dashboards that enhanced the marketing team's ability to monitor and analyze performance. These dashboards were praised for their clarity and usability.

**Images :**

****

**Project Link :** [**https://public.tableau.com/app/profile/karthikeyan.s4921/viz/karthikeyannetflix/Netflix**](https://public.tableau.com/app/profile/karthikeyan.s4921/viz/karthikeyannetflix/Netflix)

Created interactive Tableau dashboards with Netflix demo data, offering insights into viewer preferences and popular genres, praised for clarity and impact



Developed interactive Power BI dashboards with Amazon Prime demo data, providing insights into viewer preferences and popular genres, praised for clarity and impact.

##### **Project 2: Data Analysis and Machine Learning**

**Description:**This project focused on analyzing large datasets to identify patterns and build predictive models.

**Objectives:**The goals were to improve data processing efficiency and develop predictive models to support decision-making.

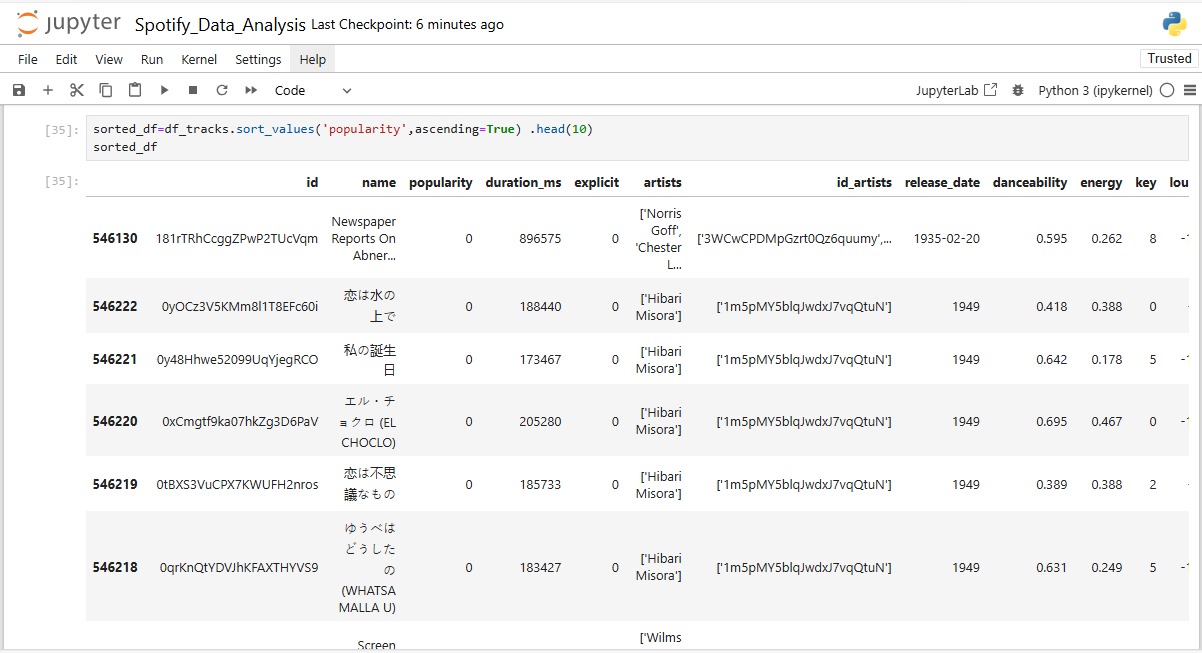
**Technologies Used:**

* Pandas
* Scikit-learn
* Matplotlib

**Role and Contributions:**I was responsible for data cleaning, preprocessing, exploratory data analysis, and building machine learning models.

**Outcomes:**The project led to the development of several predictive models that accurately forecasted key metrics, aiding in more informed business decisions.

**Images :**

****

Conducted Spotify sample data analysis using Pandas, NumPy, Matplotlib, Seaborn, and Scikit-learn. Created insightful visualizations to identify trends and patterns in the data. Leveraged Seaborn for advanced statistical plots and Scikit-learn for predictive modeling. This comprehensive analysis provided valuable insights into user behavior and music preferences.

##### **Project 3: Mixed Media Modeling**

**Description:**This theoretical project involved understanding and implementing mixed media modeling techniques using R Robyn.

**Objectives:**The objective was to learn the theory behind mixed media modeling and apply it to optimize marketing spend across various channels.

**Technologies Used:**

* R Robyn

**Role and Contributions:**I studied the theory behind mixed media modeling and explored how R Robyn could be used to apply these concepts practically.

**Outcomes:**This project enhanced my understanding of mixed media modeling and prepared me to apply these techniques in future practical scenarios.

##### **Project 4: Resume Analysis and SQL Query Generation**

# Description: In this project, I utilized the Gemini API for resume analysis and SQL query generation.

# Objectives: The objective was to automate the resume screening process and streamline data retrieval through SQL queries.

# Technologies Used:

# Gemini API (for resume analysis and SQL query generation)

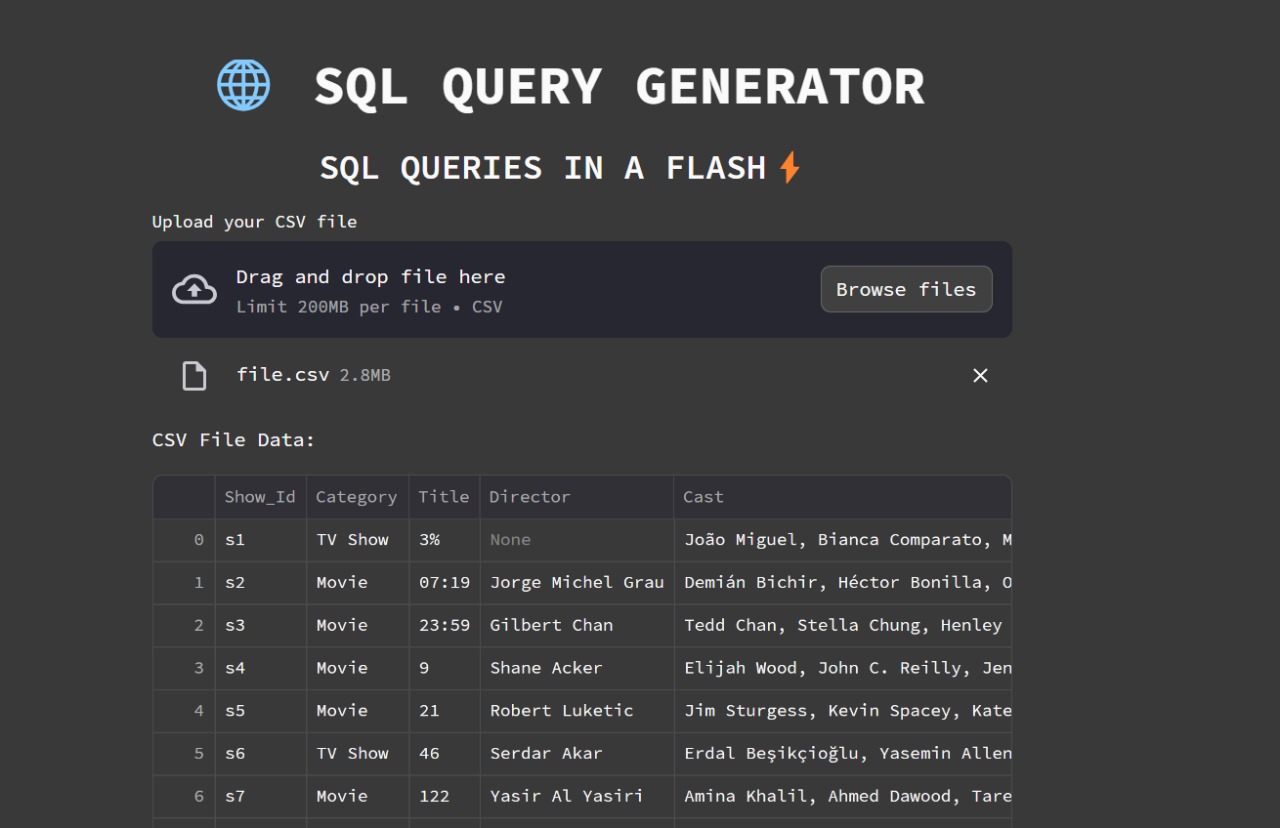
# Django

# React

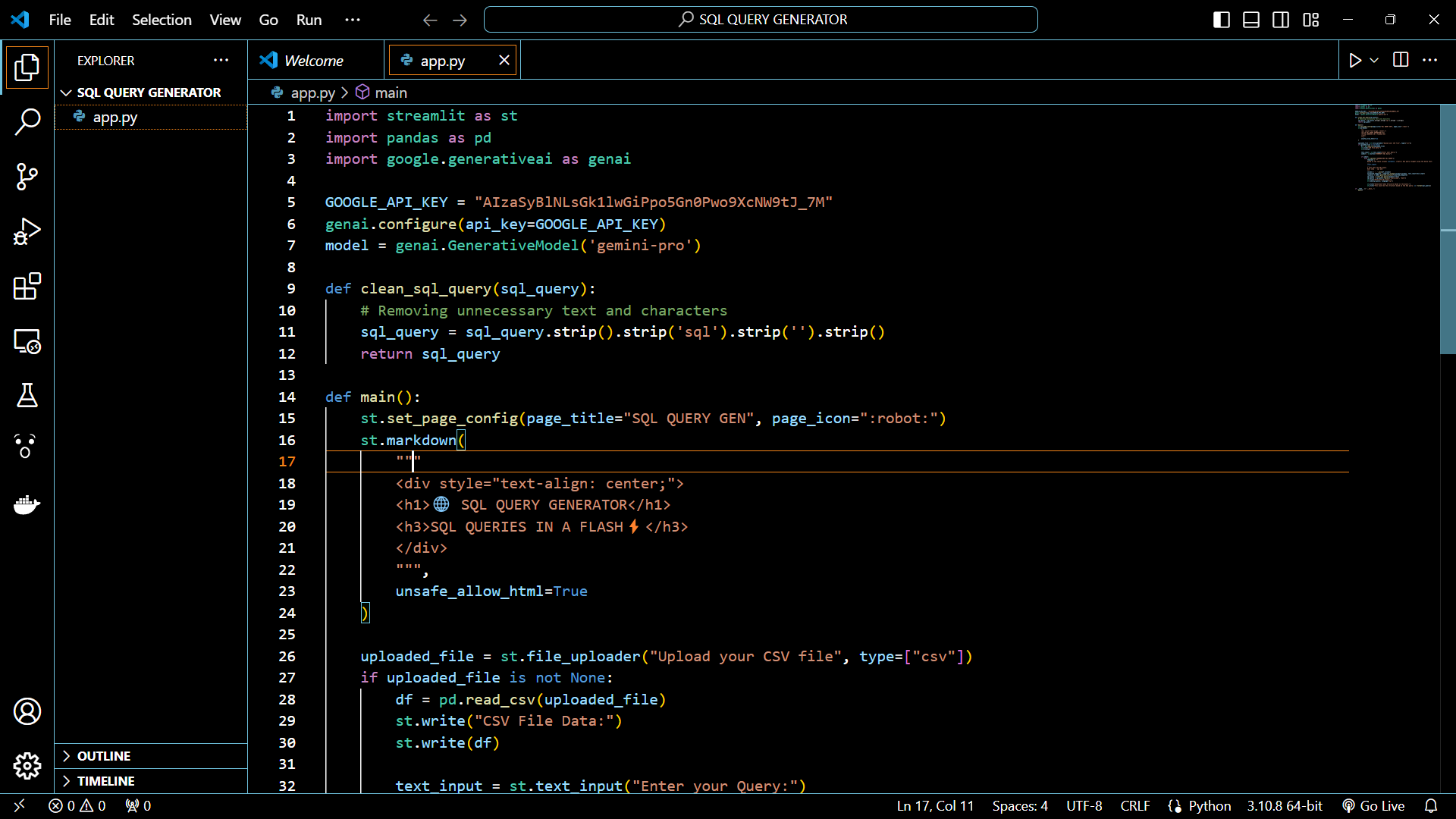
# Role and Contributions: I integrated the Gemini API into our workflow to automate resume analysis tasks and facilitate SQL query generation for data retrieval and analysis purposes.

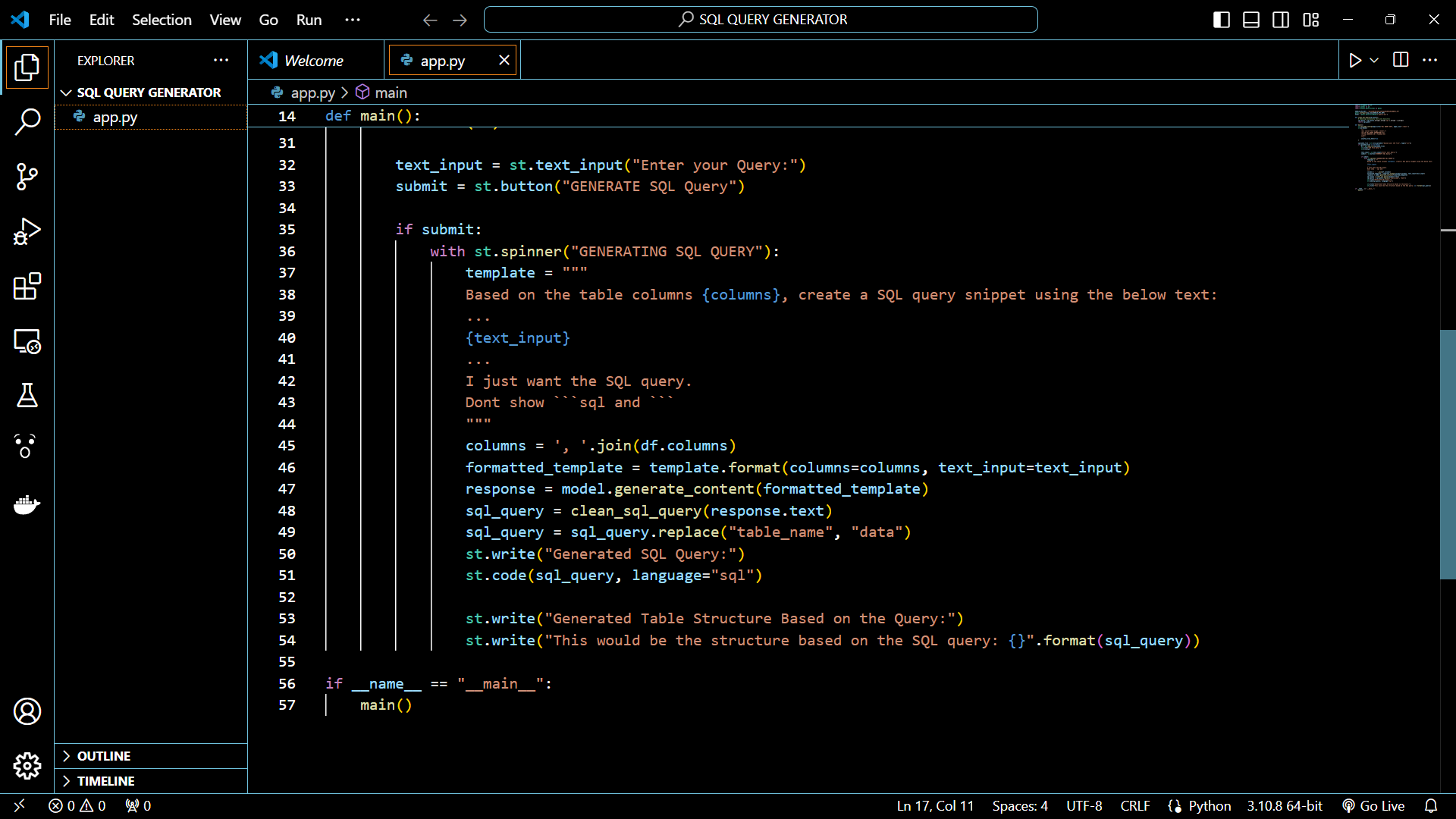
# Outcomes: This project significantly improved the efficiency of resume screening and data extraction processes, demonstrating the practical application of API integration in data-centric projects.

# Images :

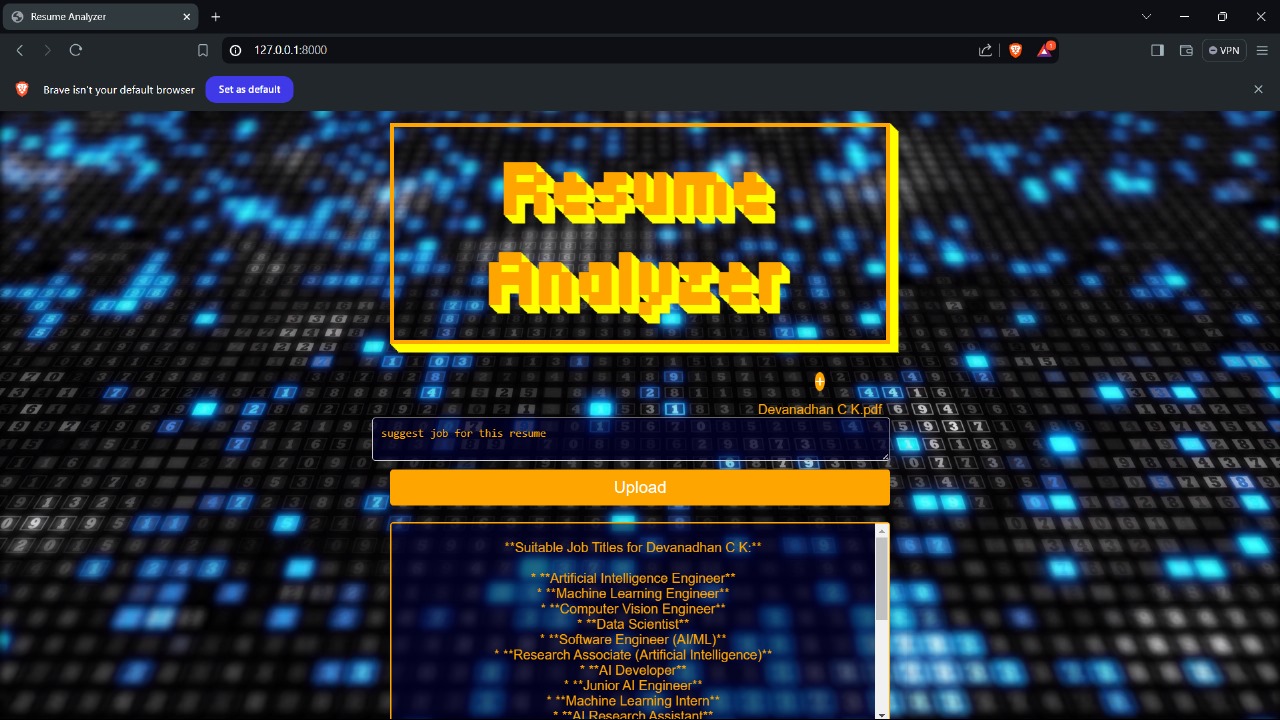


# 

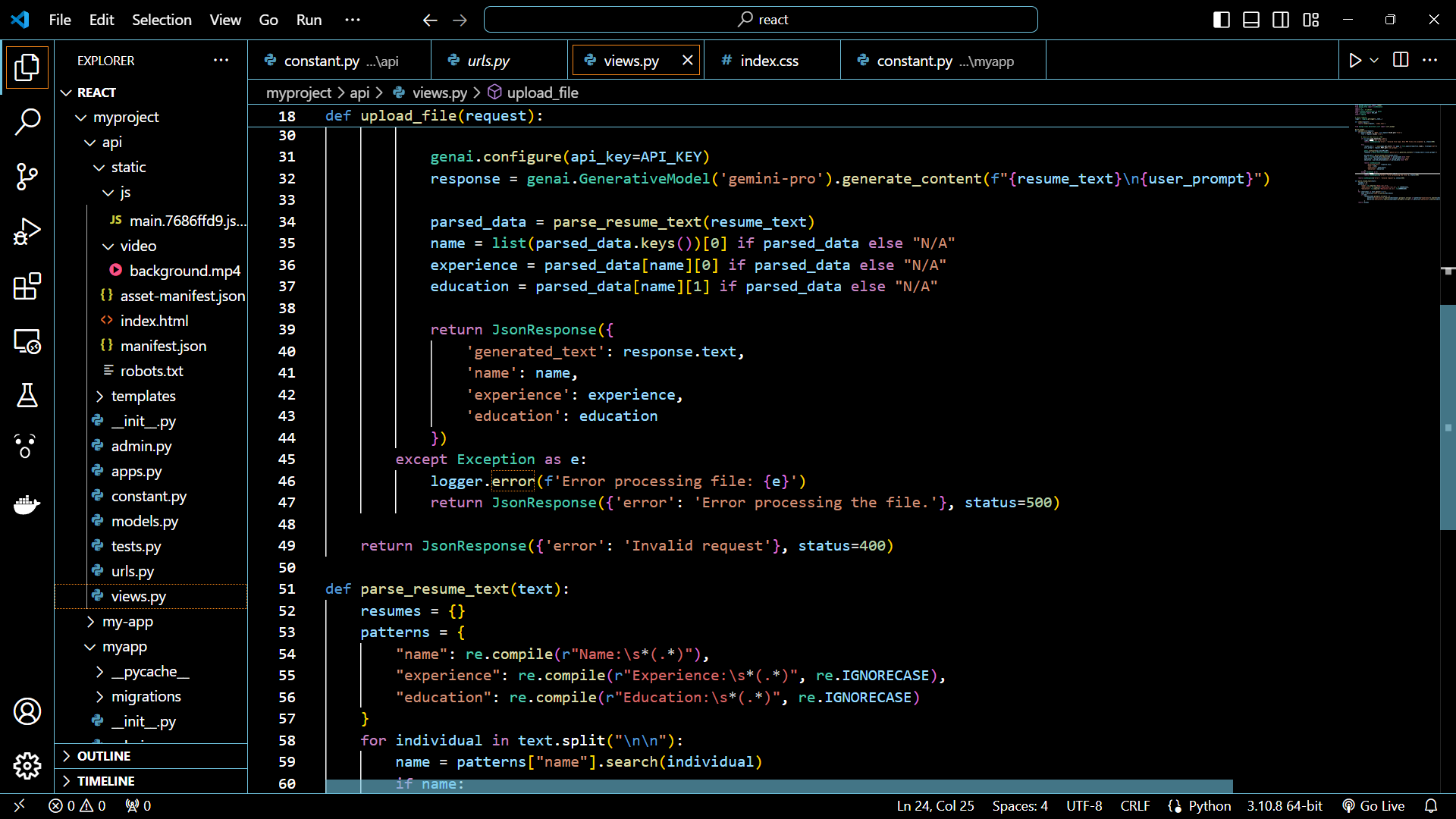




Developed an SQL query generator using the GenAI API with Streamlit, enhancing automated data retrieval capabilities. This tool streamlined the data analysis process, providing efficient and accurate query generation.









Developed a resume analyzer website using Django and React, enabling efficient and automated resume screening.

**Project 5: Speech Recognition and Speaker Diarization**

**Description:**

This project focused on implementing speech recognition and speaker diarization techniques using deep learning models to transcribe and segment audio recordings.

**Objectives:**

The goals were to develop efficient speech recognition models and accurately segment speakers in audio recordings to enhance transcription accuracy.

**Technologies Used:**

• Whisper

• Pyannote Audio

• Transformers

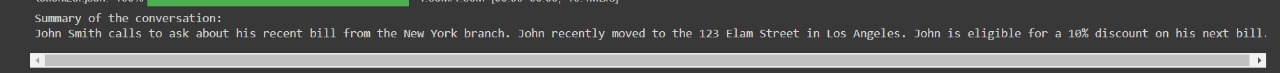
**Role and Contributions:**

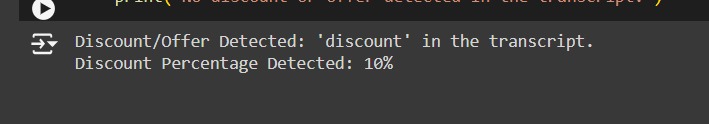
I was responsible for implementing the speech recognition pipeline, optimizing model parameters, and integrating speaker diarization techniques for accurate segmentation.

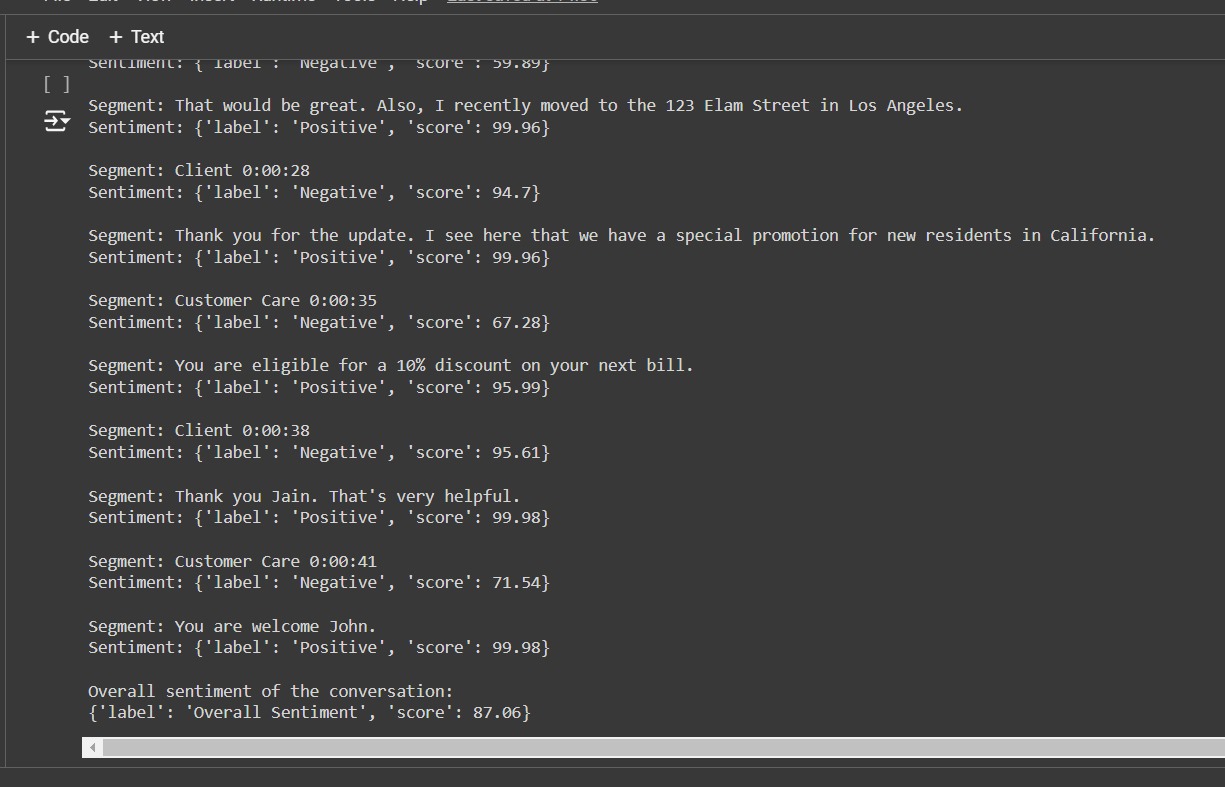
**Outcomes:**

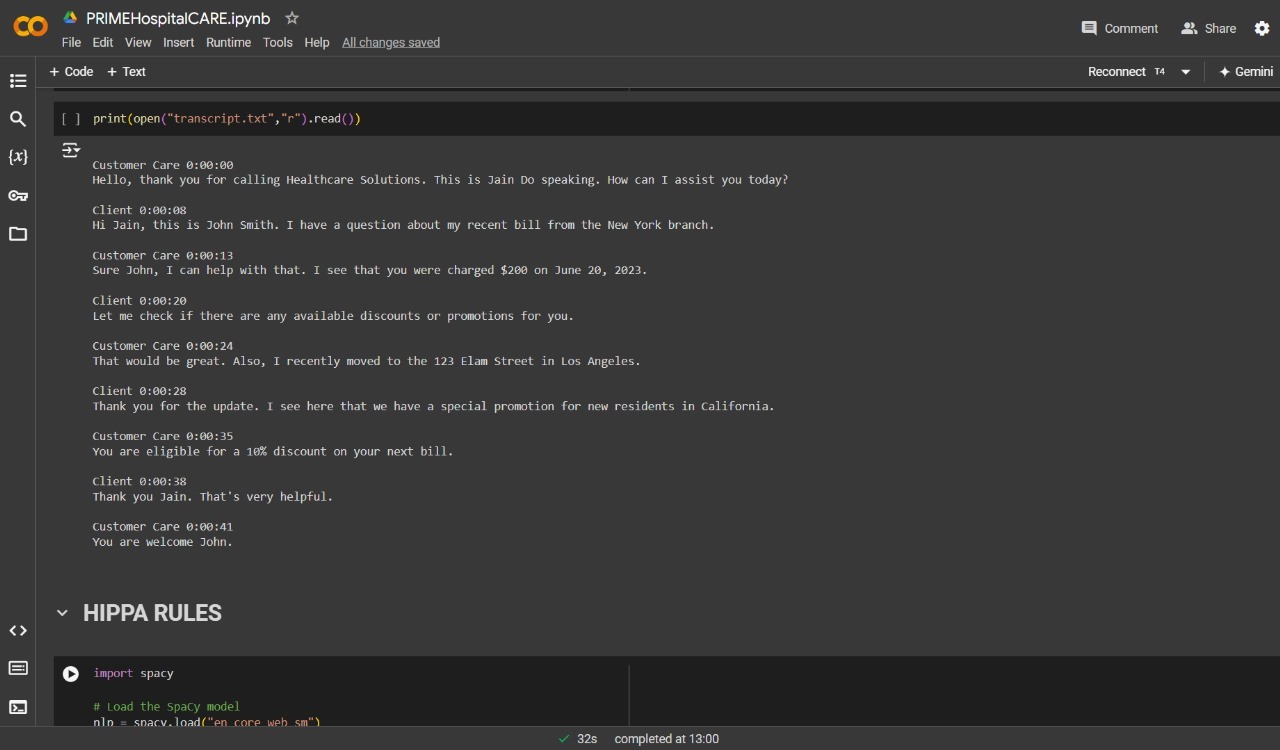
The project successfully transcribed audio files, segmented speakers, and generated formatted transcripts, enabling further analysis and application of natural language processing techniques.

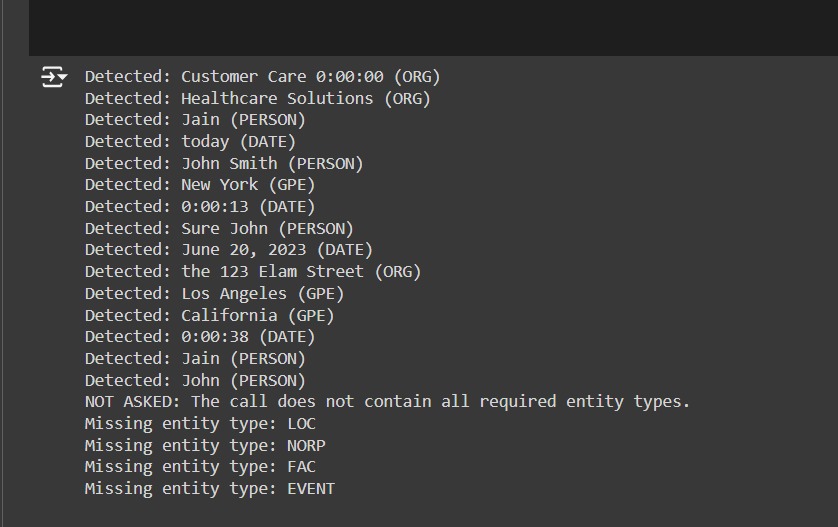
**Images:**

****

****

****

****

****

**Project Link :** [**https://colab.research.google.com/drive/13Wj\_wPqIQQvfde6A8AAhzPBvdMCZjjhG?usp=sharing**](https://colab.research.google.com/drive/13Wj_wPqIQQvfde6A8AAhzPBvdMCZjjhG?usp=sharing)

#### **Key Learnings**

**Technical Skills:**

* Proficiency in Tableau and Power BI for data visualization.
* Advanced data analysis using Pandas and Scikit-learn.
* Visualization techniques using Matplotlib.
* Theoretical understanding of mixed media modeling using R Robyn.
* Practical application of API integration using Gemini API, Django, and React.
* Understanding of Pipeline and whisper module

**Soft Skills:**

* Improved teamwork and collaboration.
* Enhanced problem-solving skills.
* Strengthened communication skills for presenting data insights.

**Domain Knowledge:**

* Best practices in data visualization.
* Machine learning algorithms and their applications.
* Concepts of mixed media modeling.
* Automated resume screening and SQL query generation.

#### **Conclusion**

**Summary of Experience:**My internship at Kanini Software Solutions was a highly enriching experience. I gained hands-on experience with industry-standard tools and techniques, worked on meaningful projects, and developed both technical and interpersonal skills.

**Internship references in LinkedIn:**

**Linkedin Profile :** [Karthikeyan S](http://www.linkedin.com/in/karthikeyan-s-657338291)