Project Journal

Data Collection and Preparation

I used a dataset of tobacco taxes for this task. So, first I loaded this dataset and then inserted it into the MongoDB.

Next, I conducted exploratory data analysis (EDA) to give insights on patterns, missing values, and outliers in the data..

Time Spent: 2 Hours

Time Series Forecasting and Inserting structured data to DB

I created a time forecasting model, Prophet, to forecast the tax. And inserted the structured data into database.

Time Spent :1 Hour

Integrating my work to code artefact:

After completing my individual work and ensuring its accuracy. I integrated my code in a structured manner into the code artefact along with teammates, ensuring seamless functionality and alignment with the overall.

Time Spent: 1 Hour

Report:

For the project report, I wrote the Methodology, Related Work, and Results sections for tobacco data and studies I referred to built the prophet model, then I ensured these sections were detailed and accurately reflected the analysis, models, and findings. Once I made my initial drafts, I confirmed content and accuracy with my teammates to ensure that the document provided a perspective, and I compiled the sections in the final report while taking care of formatting and alignment with the overall structure of the paper.

Time Spent: 6 Hours

Challenges Faced:

In integrating my code into the code artifact, I faced small issues, such as variable names not matching and the dataset structure not being consistent.

Choosing the appropriate ML algorithm to predict tax trends was challenging due to the variety of available models and the unique nature of the data. I had to explore multiple

options and evaluate their performance before finalizing the Prophet model for time forecasting.

Solution:

With the help of teammates, I debugged the errors, ensured consistent variable usage, and verified data compatibility to resolve the integration issues.

I researched time forecasting techniques and evaluated models like ARIMA and Prophet. After testing these models with our dataset, I selected Prophet because it provided the best results and offered an intuitive way to handle time-series data.

Learing Outcomes:

I learned how to manage large data sets, integrate with Python, and work with MongoDB. I learned how to use and predictusing prophet time models for accurate forecasts. I also developed skills in analytical data analysis, especially in pattern identification Deal with missing values and effective visualization . In terms of working together I learned how to integrate my work with teammates to incorporate feedback to improve both technical usability and content reporting.