Project - Final Report

Top 2500 Billinories Dataset

Team Members (Name with Roll.No.):

1. M.Navadeep(22bds040)
2. Kamal Das(22bds035)
3. A.Tulasi(22bds004)

Github Repo Link: https://github.com/NavadeepGITHUB/Top-2500-Billionaires.git

Objective:

- **Analyze Wealth Patterns:** By examining wealth distribution, this project aims to reveal trends in how wealth is accumulated and concentrated among a small, influential population. This can inform discussions on economic inequality and wealth distribution on a global scale.
- **Identify Key Demographic Insights:** Understanding the demographic characteristics (age, gender, origin of wealth) of billionaires can shed light on socio-economic mobility, generational wealth transfer, and emerging global wealth centers.

- These insights are intended to contribute to academic and economic research on wealth dynamics and the socio-economic impacts of high net worth individuals globally.

Dataset Description:

- **Dataset Source:**

Billionaires dataset obtained from reliable sources covering the global billionaire population.

- **Dataset Timeline:**

Covers data from recent years (exact dates to be checked based on the dataset details).

- **Feature Description:**

This dataset includes features such as:

- Name
- Net worth
- Industry
- Country of origin
- Age
- Source of wealth (self-made, inherited, etc.)
- Gender

- **Dataset Format:**

CSV (Comma-Separated Values)

- **Data Usage Rights and Licensing:**

Confirm the license type associated with the dataset to ensure compliance for data usage.

Tools / Technologies used:

- **Data Collection:**

Python for automated data ingestion and processing.

- **Data Processing:**

Pandas to clean, organize, and analyze the data, particularly for handling missing values and performing grouping and summarization tasks.

- **Data Visualization:**

Matplotlib and **Seaborn** to create insightful visualizations.
Example visualizations include:
Histogram of billionaires' net worth to examine wealth distribution.
Heatmaps to show regional density of billionaires.
Pie charts to depict gender and wealth source distribution.

- **Jupyter Notebook:**

Used for exploratory data analysis and documenting the analysis process.

Results and Discussion:

Key Milestones and Tasks:

- **Data preprocessing:**

Cleaned the dataset by handling missing values and standardizing formats.

- **Data analysis:**

Performed descriptive analysis to identify trends in wealth, age, and industry distribution among billionaires.

- **Visualizations:**

Created charts to depict wealth distribution, regional differences, and industry dominance.

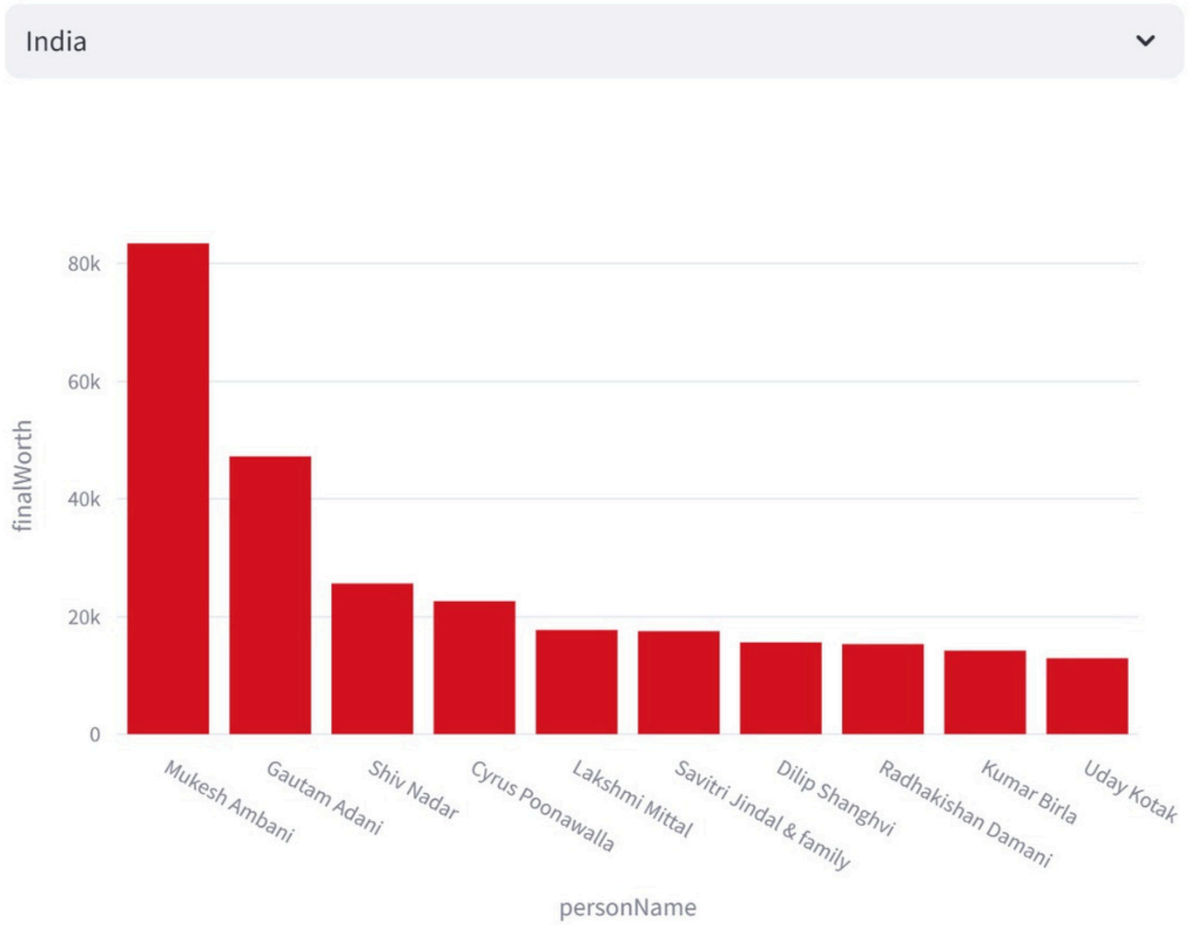
- **Challenges Faced and Solutions:**

Handling missing data in key columns (such as age or net worth) was a challenge. Applied imputation techniques and removed rows with excessive missing values.

Outliers in net worth skewed the analysis. Implemented log transformations for certain visualizations to improve interpretability.

screenshots of results:

1.Top 10 billionaires country wise:



2.Countries on map

select primary parameter

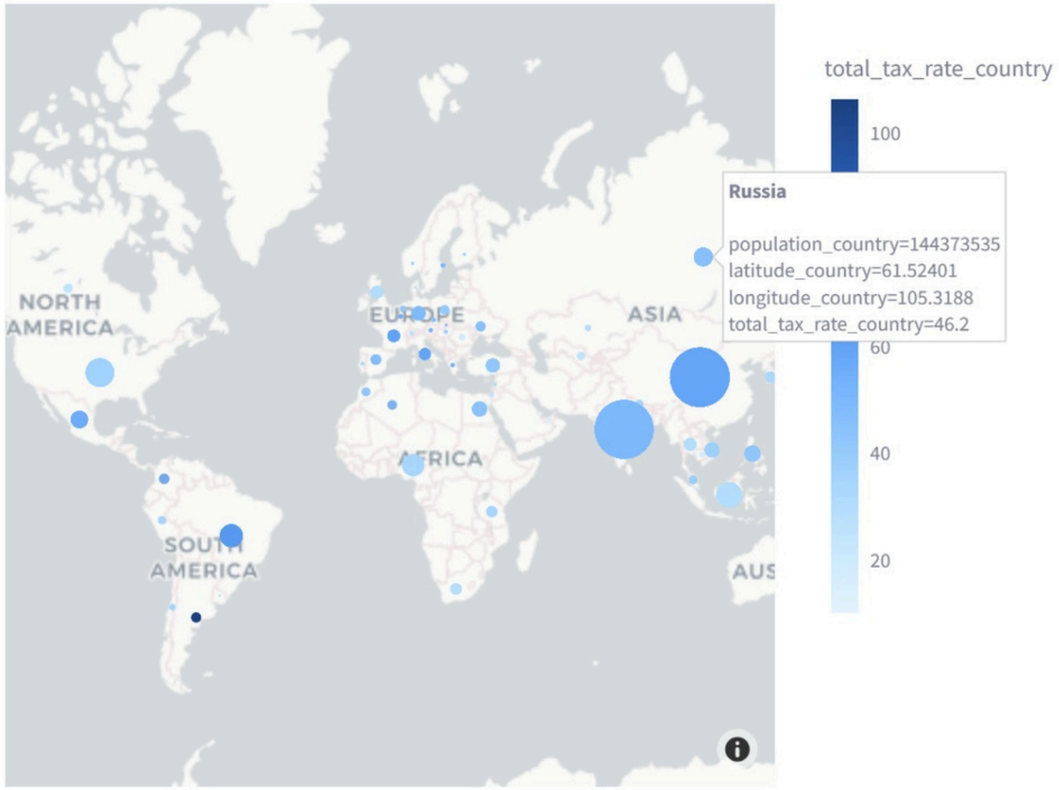
population_country

select secondary parameter

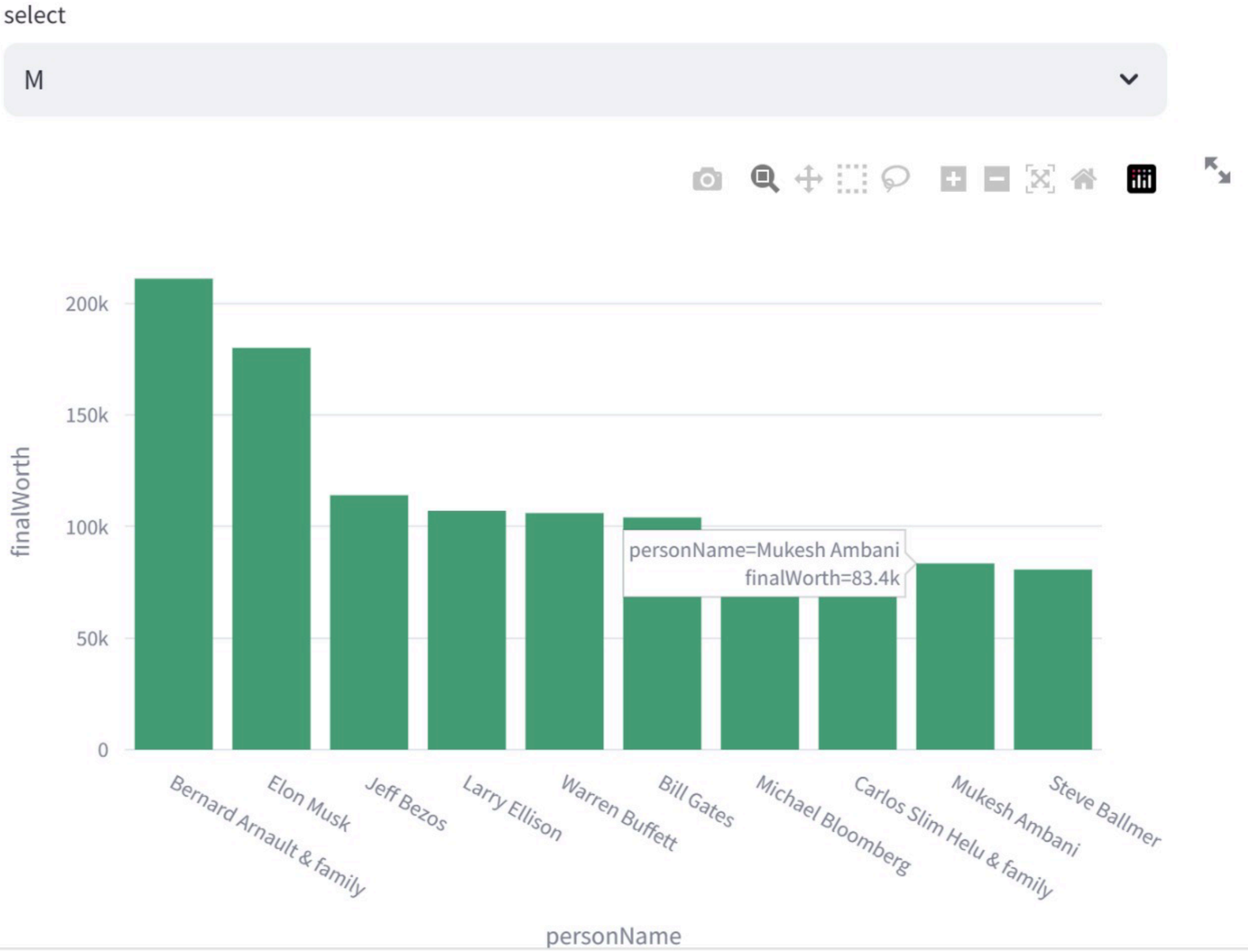
total_tax_rate_country

size represents primary parameter

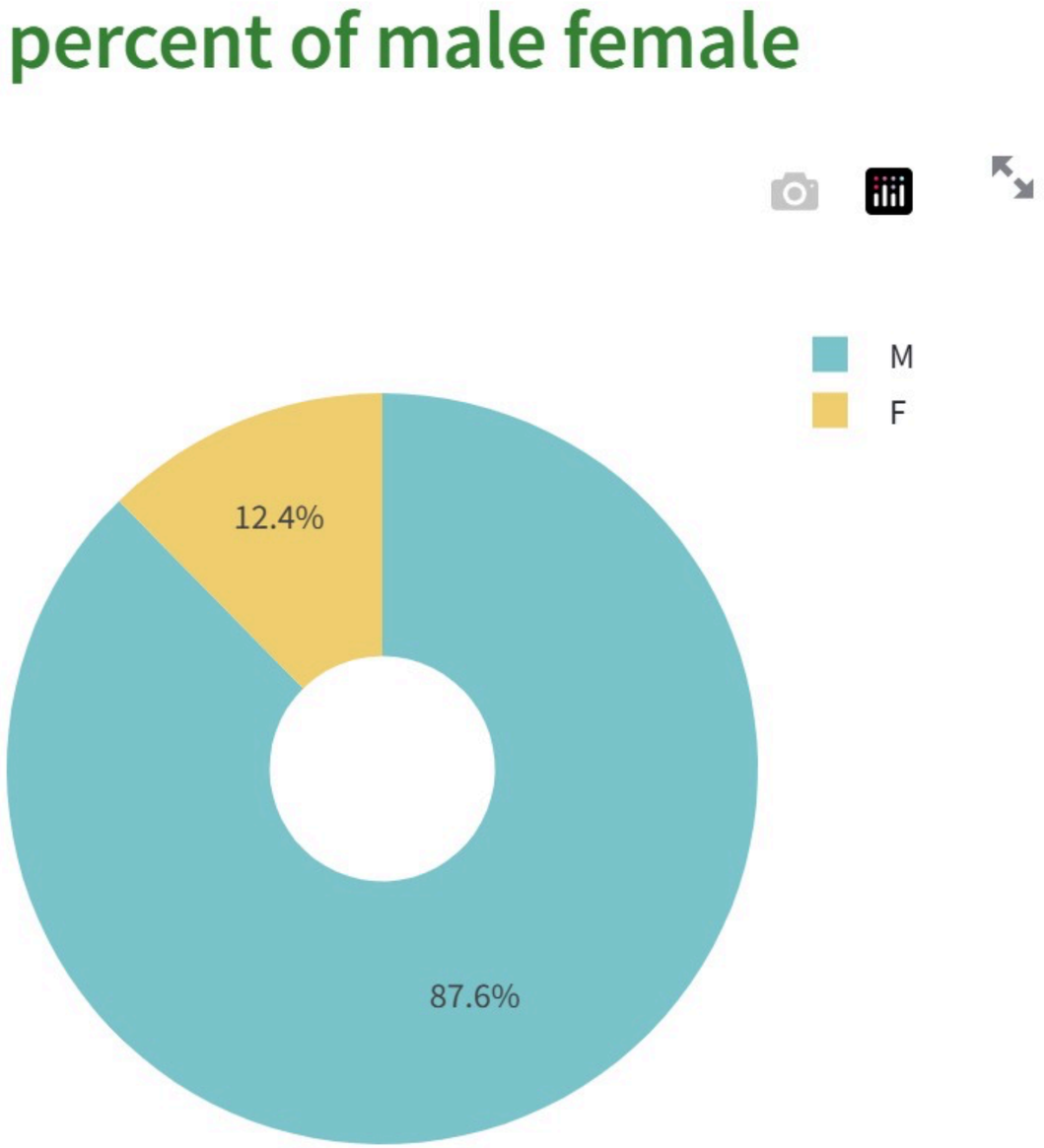
color represents secondary parameter



3.Top 10 male/female billionaires



4.Percentage of male female



Scope for future:

- **Integrate Economic and Social Indicators:**

Link the billionaire dataset with economic indicators such as GDP per capita, unemployment rates, and education levels to analyze how macroeconomic conditions correlate with billionaire wealth.

- **Time Series Analysis:**

If more data becomes available over time, conduct a time series analysis to study trends in wealth accumulation and industry dominance, examining factors like the rise of tech billionaires.

- **Expand Demographic Insights:**

Deepen the demographic analysis by examining the family background of billionaires, educational attainment, and other factors to understand broader socio-economic mobility.

- **Automated Dashboard Creation:**

Build an interactive dashboard to allow users to explore the dataset dynamically, enabling filtering by region, industry, age, and wealth.

Conclusion / References:

- This project offers a comprehensive analysis of the billionaire population, revealing key insights into the demographic and economic characteristics of the wealthiest individuals globally. The findings underscore the dominance of certain industries, such as technology and finance, in wealth accumulation, and highlight geographic trends that reflect regional economic power centers. This analysis not only informs discussions around wealth inequality but also provides a foundation for further research into socio-economic mobility and the implications of concentrated wealth.

References:

Include specific references such as:

- Forbes Billionaire List (if applicable)
- Any publications or datasets that informed your analysis

Contribution:

Member 1(Navadeep) - Lead on data preprocessing and exploratory data analysis.

Member 2(Kamal) - Focused on data visualization and reporting.

Member 3(Tulasi) - Handled the development of the interactive dashboard.

Thank You !!!