

Project Report

Exploratory Data Analysis (EDA) of Forbes 2022 Billionaires Dataset

1. Introduction

The objective of this project is to perform an exploratory data analysis (EDA) on the 2022 Forbes Billionaires dataset. This dataset provides detailed information about the world's richest individuals, including their rank, name, net worth, age, country, source of wealth, and industry. The analysis aims to uncover patterns, insights, and relationships between various factors such as age, net worth, industry, and geographical distribution of billionaires.

2. Data Overview

The dataset consists of 2,600 entries and 8 columns.

Key columns include:

Rank: Ranking of the billionaire.

Name: Name of the billionaire.

Networth: The net worth in billions of dollars.

Age: Age of the billionaire.

Country: Country of origin.

Source: Source of the billionaire's wealth.

Industry: Industry in which the billionaire operates.

3. Data Cleaning

The dataset was analyzed for missing values and duplicates:

▫No missing values were found (`rich.isnull().sum()` resulted in all zeros).

▫No duplicated rows were present (`rich.duplicated().sum()` returned 0).

The dataset is clean, with 2,600 entries and 8 columns.

4. Data Types

The dataset consists of both numerical and categorical data:

Numerical columns: rank, age, unnamed: 0

Categorical columns: name, networth, country, source, industry

5. Descriptive Statistics

Descriptive statistics for numerical columns (rank, age) reveal:

Average age of billionaires: 64.27 years.

Minimum age: 19 years, **maximum age:** 100 years.

Standard deviation in age: 13.22 years, indicating a wide spread in age distribution.

6. Exploratory Analysis

Age Distribution of Billionaires

The age distribution was segmented into age ranges (0-20, 21-30, ..., 91+), and the frequency of billionaires in each age group was visualized using a bar plot. The majority of billionaires fall between the ages of 51-70, with the 61-70 age group being the most prominent.

Net Worth vs Frequency

The net worth distribution was visualized with a bar plot showing the frequency of billionaires with different net worth levels. Due to the wide range in net worth (from \$1 billion to \$219 billion), many billionaires cluster at the lower end of the wealth spectrum.

Industry Representation

A bar plot of industry versus frequency shows which industries produce the most billionaires. The technology sector dominates, followed by finance and investments. Other significant industries include fashion & retail, automotive, and healthcare.

Net Worth Variation with Age and Industry

Using a categorical plot (catplot), we analyzed how net worth varies across age groups and industries. Technology billionaires tend to have high net worths even at younger ages compared to other industries like fashion or healthcare.

Top 10 Richest Billionaires

A horizontal bar plot highlighted the top 10 richest people in the world. Elon Musk leads the list with a net worth of \$219 billion, followed by Jeff Bezos, Bernard Arnault & family, and Bill Gates.

Richest People in India

A plot was created to visualize the top 10 richest billionaires in India, revealing names like Mukesh Ambani and Gautam Adani leading the list.

Young Billionaires (≤ 50 years)

A subset of the data was created to show billionaires who are 50 years old or younger. Elon Musk, Mark Zuckerberg, and Larry Page are among the youngest in this group, primarily from the technology sector.

Industry with Most Billionaires

A bar plot was created to show the industry distribution among billionaires. The technology industry has the largest number of billionaires, followed by finance & investments, and fashion & retail.

7. Key Insights

Age Distribution: Most billionaires are aged between 51-70, indicating that wealth accumulation takes time.

Technology Dominance: The technology sector produces the most billionaires, with the youngest average age. This indicates that the tech industry has created immense wealth, especially for younger entrepreneurs.

Geographical Focus: The United States dominates the list of billionaires, followed by countries like China, India, and Germany.

Industry Distribution: Besides technology, other key industries include finance & investments, fashion & retail, and healthcare.

Top Billionaires: Elon Musk is the richest person in the world, followed by Jeff Bezos and Bernard Arnault & family, each with massive fortunes accumulated in tech, retail, and fashion industries, respectively.

8. Conclusion

This EDA provided a comprehensive view of the wealth distribution across different factors such as age, industry, and geography. The technology sector is clearly the leader in creating billionaires, and younger individuals are making their mark in this industry. The data also suggests that while wealth is concentrated in a few industries, there is diversity across regions and sectors.