

Exploratory Data Analysis on the Billionaire Dataset

IBU 512 Advanced Topics in Database Systems

Ajla Korman



Introduction

- Billionaire wealth is not distributed equally across the world
- This research looks at:
 - a. who billionaires are
 - b. where they come from
 - c. which industry they operate in
 - d. age and gender demographics
- Aim of the research was:
 - a. Find out what factors influence billionaire growth
 - b. Helps understand why some places have more billionaires than others

Methods & Material

- [Billionaire Statistics Dataset](#)
- Contains 35 attributes on billionaires worldwide
- Includes information on net worth, industries, age and gender, country of citizenship, and economic indicators
- Used for EDA (Exploratory Data Analysis)
- Python libraries used: Pandas, NumPy, Matplotlib, and Seaborn for data analysis and visualization

Data Preprocessing & Cleaning

- Removed unnecessary columns (e.g., birth details, specific locations, first and last name)
- Handled missing values: Used OpenAI API & World Bank API for estimation
- Dropped null & duplicate records to improve dataset quality
- Final dataset contained 15 key columns and 2457 records

rank	personName	finalWorth	age	gender
industries	countryOfCitizenship	selfMade	gdp_country	tax_revenue_country
total_tax_rate_country	population_country	source	organization	city

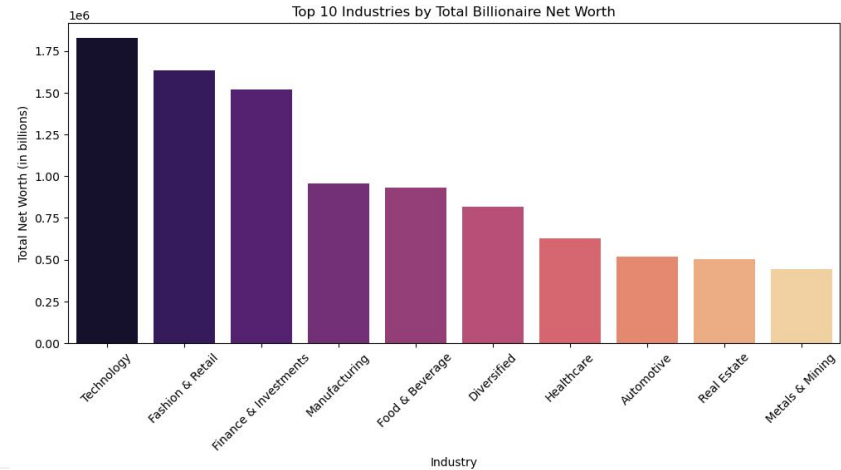
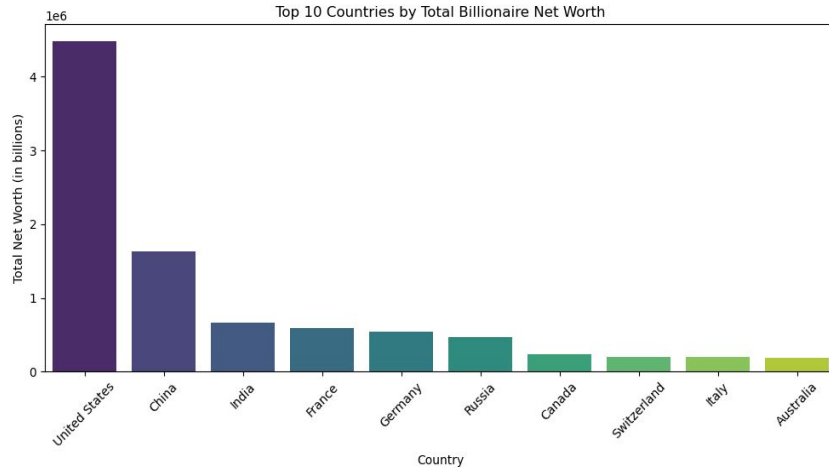
Net Worth Analysis

- Calculated key statistics: Mean, median, and outliers using IQR method.
- Plotted histogram: Showed billionaire net worth distribution.
- Identified industries with highest net worth: Tech, Finance, Manufacturing.

	personName	industries	countryOfCitizenship
0	Bernard Arnault & family	Fashion & Retail	France
1	Elon Musk	Automotive	United States
2	Jeff Bezos	Technology	United States
3	Larry Ellison	Technology	United States
4	Warren Buffett	Finance & Investments	United States
5	Bill Gates	Technology	United States
6	Michael Bloomberg	Media & Entertainment	United States
7	Carlos Slim Helu & family	Telecom	Mexico
8	Mukesh Ambani	Diversified	India
9	Steve Ballmer	Technology	United States

Net Worth Analysis

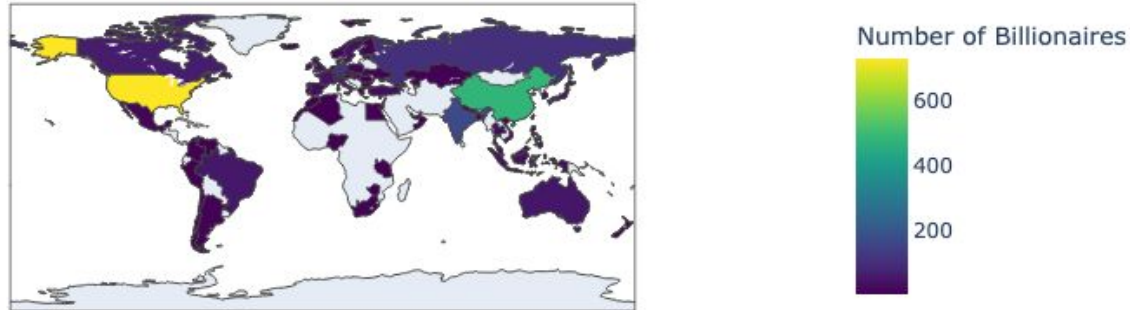
- The United States has the highest total billionaire net worth
- Technology, Finance & Retail are the most profitable industries for billionaires
- Wealth concentration is highly uneven, with a few countries and industries dominating



Geographical Distribution

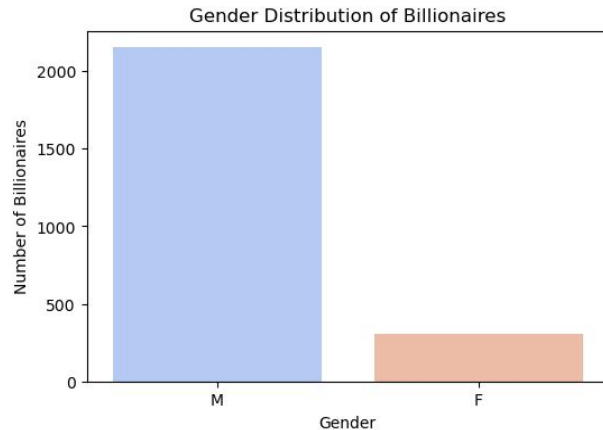
- Mapped billionaires per country using a choropleth map
- Shows unequal distribution in the whole world

Global Distribution of Billionaires



Gender Analysis

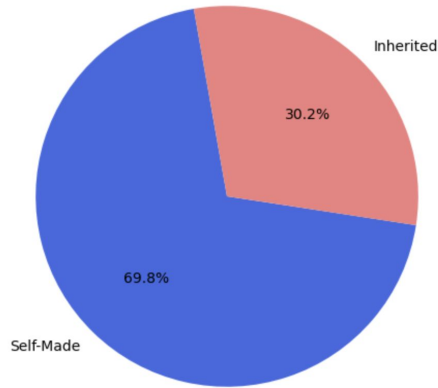
- Significant gender gap: Male billionaires greatly outnumber female billionaires
- Industries with higher female representation: Fashion, Retail, Food & Beverage
- Socioeconomic & historical barriers limited female wealth accumulation



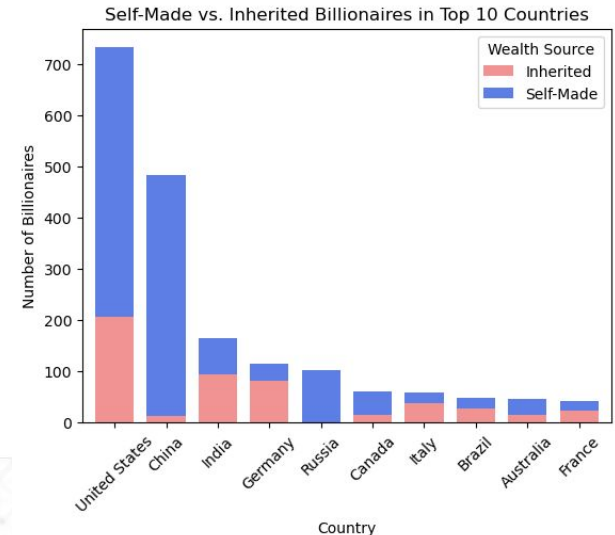
	Industry	Female Billionaires	Male Billionaires
0	Automotive	6.0	64.0
1	Construction & Engineering	6.0	36.0
2	Diversified	29.0	150.0
3	Energy	9.0	86.0
4	Fashion & Retail	36.0	214.0
5	Finance & Investments	25.0	318.0
6	Food & Beverage	39.0	160.0
7	Gambling & Casinos	3.0	19.0
8	Healthcare	28.0	167.0
9	Logistics	6.0	27.0

Wealth Source Analysis

- Most billionaires are self-made (69.8%)
- Industries with highest self-made billionaires: Tech, Manufacturing, Finance
- Regional differences: U.S. has highest share of self-made billionaires

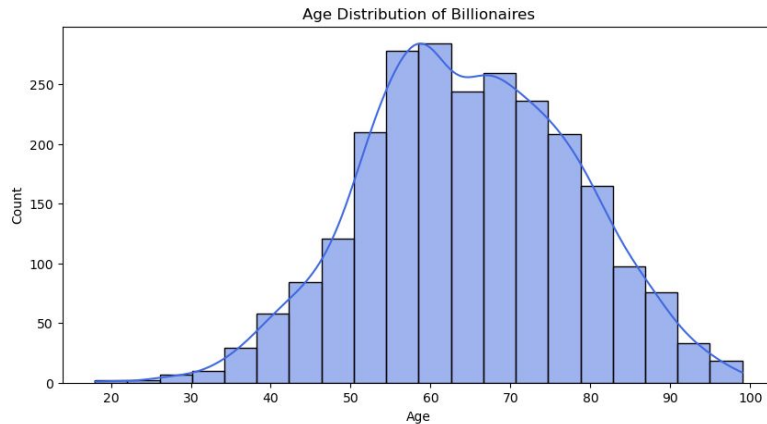


	industries	Inherited	Self-Made
0	Automotive	33	37
1	Construction & Engineering	19	23
2	Diversified	101	78
3	Energy	25	71
4	Fashion & Retail	101	149
5	Finance & Investments	78	265
6	Food & Beverage	95	105
7	Gambling & Casinos	4	18
8	Healthcare	52	143
9	Logistics	6	27
10	Manufacturing	80	216
11	Media & Entertainment	24	62
12	Metals & Mining	19	53
13	Real Estate	46	114
14	Service	19	29
15	Sports	14	24
16	Technology	22	277
17	Telecom	4	24



Age Distribution

- Histogram of billionaire ages: Most are between 50-70 years old
- Younger billionaires dominate Tech industry, while older ones are in Manufacturing & Energy



	Industry	Average Net Worth	Average Age
0	Automotive	7425.71	65.77
1	Telecom	6789.29	67.21
2	Logistics	6730.30	66.15
3	Fashion & Retail	6534.40	66.67
4	Metals & Mining	6130.56	61.51
5	Technology	6107.69	56.61
6	Media & Entertainment	4815.12	65.74
7	Food & Beverage	4681.41	66.48
8	Energy	4628.42	67.62
9	Gambling & Casinos	4627.27	68.77

Key Findings & Discussion

1. **Wealth is Highly Concentrated in a Few Countries:**

- The United States dominates in total billionaire net worth, followed by China and India. European and emerging economies contribute at lower levels.

2. **Technology, Finance, and Retail Drive Billionaire Wealth:**

- The technology sector, in particular, has created more self-made billionaires, unlike inherited wealth in industries such as real estate or manufacturing.

Key Findings & Discussion

3. **Persistent Gender Disparity in Billionaire Distribution:**

- The number of male billionaires vastly outnumbers female billionaires.
- Female billionaires are concentrated in industries like Fashion & Retail and Food & Beverage, often linked to inherited wealth.

4. **Self-Made Billionaires are the Majority:**

- Most billionaires accumulate wealth through entrepreneurship, rather than inheritance.
- This is especially evident in sectors like Technology and Finance, where innovation plays a crucial role.

Conclusion & Future Work

- Research confirmed global wealth concentration patterns.
- Findings align with past studies on billionaire demographics & economics.
- Future work:
 - a. Time-series analysis of billionaire growth over decades.
 - b. Machine learning models for billionaire wealth prediction.

Thank you for your attention