

Who are the Billionaires?

Team Members

Name	Email	Phone Number
Bochao Gia Liu	gialiubartsch@gmail.com	(408) 421 1886
Navjeet Ghuman	navjeet.singh.ng@gmail.com	(510) 598 5672
Kevin Tian	kevinxtian@gmail.com	(650) 430 1524
Ezra Salomon	ezsalo@gmail.com	(510) 926 0089

Background

Elon Musk has become the wealthiest billionaire according to Forbes Real Time Billionaires List. Except for one person, the top ten billionaires are all from the United States and most of them are from the technology industry. People talk about billionaires all the time, but not many people have statistical insights on billionaires around the world.

Motivation

When we think about billionaires, we can always name a dozen of names, but how many billionaires are there in this world? Do we have any statistical knowledge of billionaires? With this question in mind, we are interested in digging deeper into the distribution features of billionaires, and we want to know if we were billionaires, how much we would be worth!

Questions to answer

General Information

- What is the billionaires distribution of the world?
- What is the billionaires distribution of the United States?
- What industries have more billionaires?
- How about the age group within each industry?
- What is the gender proportion of billionaires?

Correlation to billionaires

- Correlation of country population vs. number of billionaires
- Correlation of country GDP vs. number of billionaires
- Correlation of age vs. number of billionaires
- Correlation of gender vs. number of billionaires

Prediction

- Create a multi linear regression model that predicts the worth of billionaire

Tools/Modules to use

- Python
- Pandas
- Matplotlib
- NumPy
- SciPy
- etc.

Data sets to use

List all possible datasets you'll use

- Billionaires Statistics Dataset 2023:
<https://www.kaggle.com/datasets/nelgiriyeewithana/billionaires-statistics-dataset>

Tasks Breakdown

- Collect/Clean the data.
- Github and Analysis.
- Visualization and Summary.
- Presentation and Q&A.

Presentation

Divide your presentation steps to tasks and assign it to members.

- Gia : Introduction + Prediction Model (Beginning + Ending)
- Navjeet : Distribution in the US (First Part)
- Kevin : Distribution in the US (Second Part) + Correlation Analysis

Tasks and timeline

	Date	Task	Notes
	10/12	Finalize an idea, submit a proposal, find a dataset, etc.	check
	10/16	Start cleaning data, etc.	check
	10/17	Cleaning data	check
	10/19	Analyzing & Visualizing	check
	10/23	Practice presenting	check
	10/24	Presentation Day	check

Suggested by:
Khaled Karman