

The shared files include the Forbes Billionaire List dataset from 2001 to 2023, as well as Python code for processing the data.

The data source of Dataset is:

<https://www.forbes.com/billionaires/>

[https://stats.areppim.com/stats/links\\_billionairexlists.htm](https://stats.areppim.com/stats/links_billionairexlists.htm)

[https://www.kaggle.com/datasets/guillemservera/forbes-billionaires-1997-2023/?select=billionaires\\_2019.csv](https://www.kaggle.com/datasets/guillemservera/forbes-billionaires-1997-2023/?select=billionaires_2019.csv)

<https://www.piie.com/publications/working-papers/origins-superrich-billionaire-characteristics-database?ResearchID=2917>

and data collected by the authors.

The code processes the data as follows:

1. Fit Pareto and calculate the alpha value, and calculate the corresponding p and q values.
2. Classify and count billionaires according to continent, select the data of Asia, Europe and North America to fit Pareto and calculate their respective alpha values.
3. Calculate the median age of billionaires in each year,
4. Classify billionaire by gender, calculate the proportion of women, and the alpha values after fitting Pareto for men and women respectively.
5. Classify billionaire by industry and calculate the proportion of each industry in each year.