Cuiyan Hua, Yanfei Li, William Li, Zhehong Ren
November 14th, 2023
QST BA 222
Professor Leder-Luis

Analyzing Trends Among Billionaires Around the World

The annual release of the billionaire rankings has consistently captured the attention of the public. Billionaire statistics provide real-world examples of entrepreneurial strategies, industry dynamics, and various elements contributing to wealth creation. We aim to answer several key questions that can help us better understand recent trends in the market. For example, which industries have the most billionaires, and what factors are associated with the final worth of billionaires? We can analyze if “age”, “gender”, “education level”, or “gdp_country” correlate with “finalWorth”. These questions can help guide our career choices towards industries that may offer a greater likelihood of achieving financial success. The following report will provide more details on analyzing those factors using various visualization plots and regression models.

We chose the most recent statistic called Billionaires Statistic Dataset in 2023 from Kaggle: https://www.kaggle.com/datasets/nelgiriyewithana/billionaires-statistics-dataset. There are a total of 35 variables and 2640 observations in this dataset. After loading the dataset, we recognized the need to refine the data to facilitate smoother and more insightful analysis. We began by selecting relevant variables and dropping unused ones from the dataset. To simplify the analysis, we created a new column called “continent” using dictionary mapping to group countries into their continents. Then, we checked for missing values for all the variables. For age and continent, with approximately 2% of missing values, we think dropping them would be more appropriate than inputting mean values. We also excluded GDP and education enrollment from our analysis due to their substantial 7% missing values that may negatively impact our analysis quality and results.

After cleaning the dataset, we first looked at the gender distribution of billionaires. We were able to see the significant disparity between the number of male and female billionaires. Our cleaned

dataset comprised 2,549 billionaires: 80.2% of the billionaires are male, and 19.8% are female. Furthermore, across all industries, males outnumber females, with the “Finance & Investments” and “Energy” sectors showcasing the lowest female-to-male ratio, 7.1% and 8.9%. The top-ranked billionaire is Bernard Arnault (CEO of LVMH), counted as a male in the “Fashion & Retail” industry, with a final worth of $211 billion U.S. Dollars. Meanwhile, the richest female in 2023 is Francoise Bettencourt Meyers (Vice Chairman of L'Oréal), who is also in the “Fashion & Retail” sector and sits at rank 11.

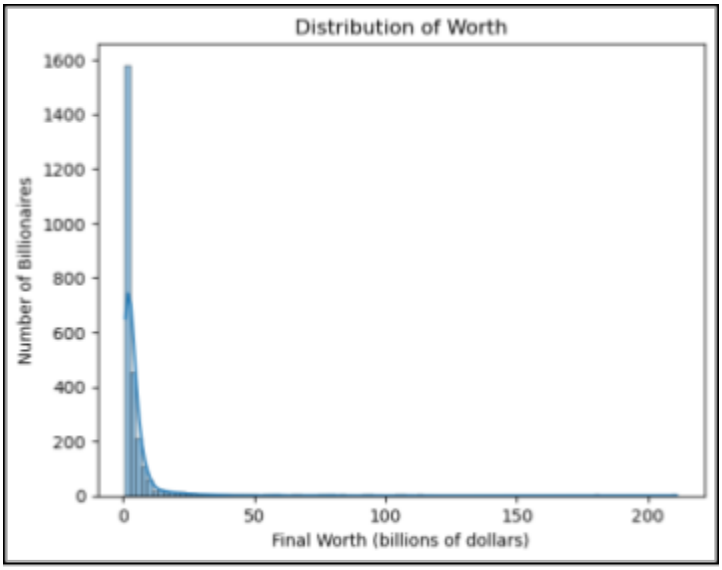

Figure 1: Distribution of Billionaires by Final Worth (billions of dollars)

This histogram illustrates a left-skewed distribution in the worth of billionaires. With a mean worth of 4.7 billion and a median of 2.7 billion, the data suggests a skewness toward the lower end of the wealth. The histogram reflects the concentration of billionaires within the first bin is from 1 billion to 3.1 billion with a count of 1579 individuals. This concentration highlights the considerable worth disparity even among the billionaire class. The standard deviation of 9.99 billion also emphasizes the variability of “finalorth” in this dataset, further illustrating the diverse range of worth among the billionaires.

In order to further examine “finalWorth”, we created a violin plot comparing inherited and self-made billionaires, exploring the relationship between wealth origin and final worth. The

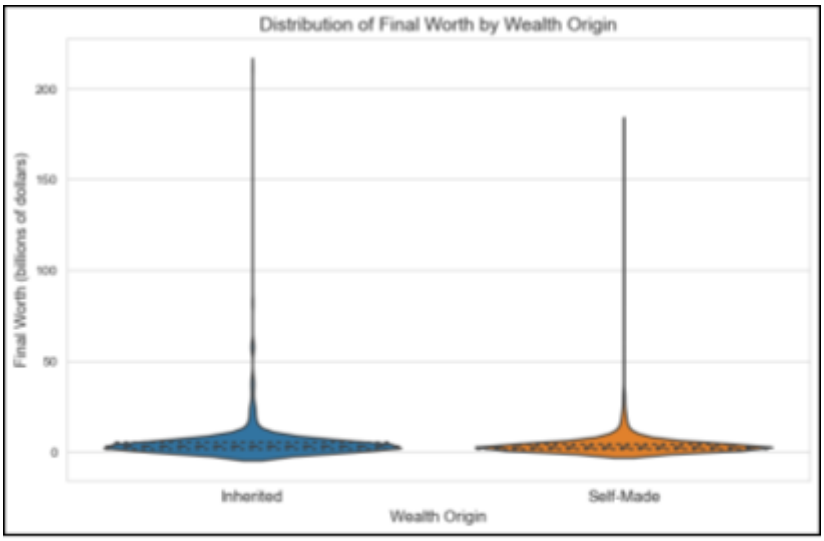

Figure 2: Violin Plot Comparing the Wealth Distribution of Inherited vs. Self-Made Billionaires

broader base for the self-made group, which comprises 1,780 individuals, suggests that their wealth is more uniformly distributed across the lower and middle ranges, up to under $200 billion. This contrasts with the inherited group's plot, which shows a pronounced upper bulge despite only having 769 individuals. This reflects their higher average net worth of $5.15 billion compared to the $4.51 billion of self-made billionaires. The extreme values on the higher end notably stretch the inherited wealth plot, skewing the average. The median net worth for inherited billionaires, at $2.7 billion, only slightly exceeds the $2.2 billion median of the self-made cohort. This suggests that the central tendency of wealth among the inherited is higher. However, the significant two-thirds collective wealth share of self-made billionaires points to a more expansive spread of wealth within this group, as depicted by the width of the plot at different wealth levels.

Analyzing the age distribution of billionaires is important as it can provide insights into the relationship between age and wealth accumulation among the world's wealthiest individuals. We observed an approximately normal distribution with a slightly right skew, suggesting a concentration of billionaires in the higher age spectrum. The skewness from the distribution might indicate the

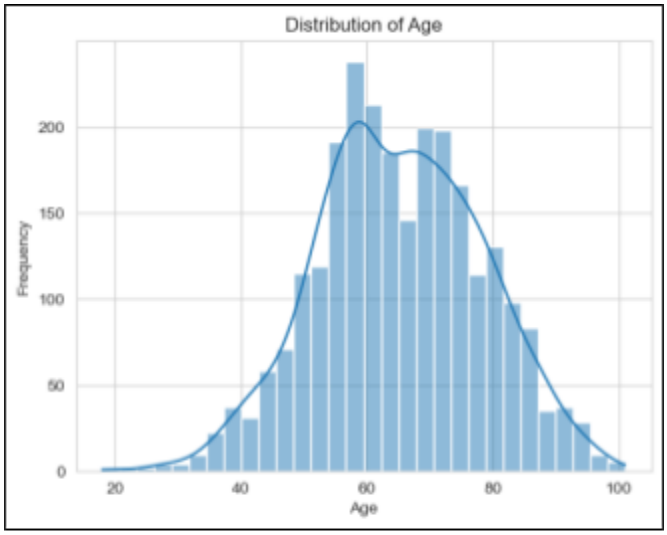

Figure 3: Frequency of Billionaires per Age Group

effect of wealth accumulation, which often increases with age. From the histogram, we can observe two peaks. The first and more pronounced peak is around the late 50s to early 60s, implying the most common age for billionaires in this dataset. The second peak around the 70s indicates another common age of billionaires. The summary statistic also shows that the average age of billionaires is approximately 65 years old, ranging from age 18 to 101, which aligns with the histogram.

Delving into the industries for younger billionaires aged 18 to 45, the “Technology” sector leads with 60 individuals, followed by “Finance & Investments” (22), “Fashion & Retail” (16), “Media & Entertainment” (12), and “Food & Beverage” (10). This trend likely reflects the recent surge in wealth generation through tech innovation and entrepreneurial ventures in younger demographics. In contrast, among the older age group of 45 and above, “Finance & Investments” (339), “Manufacturing” (305), and “Technology” (247) dominate, followed by “Fashion & Retail” (239) and “Food & Beverage” (190). This suggests a prevalence of more traditional wealth accumulation pathways in the older age group.

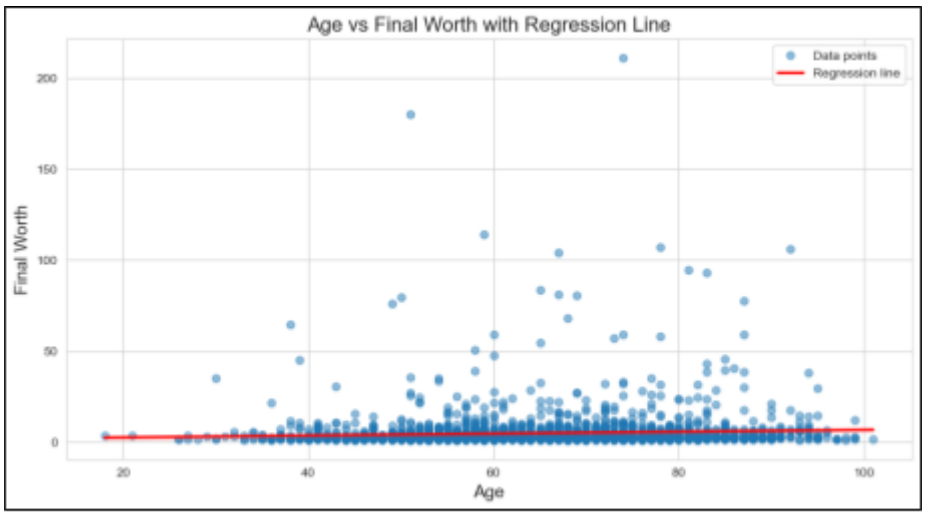

Figure 4: Regression Analysis on Age and Final Worth

To examine our dataset of billionaires more comprehensively, we conducted a regression analysis to understand the influence of age on an individual’s final worth. The coefficient for age was found to be 0.0506, suggesting a statistically significant but relatively small influence on “finalWorth", with a p-value of 0.001. While this positive coefficient predicts an average increase

in “finalWorth” of approximately $50.6 million for each additional year of age, the relatively flat regression line implies that age does not translate into a strong factor to determine “finalWorth”. Overall, the regression analysis infers a predictive relationship, where an increase in age is associated with an increase in an individual’s “finalWorth” from the dataset.

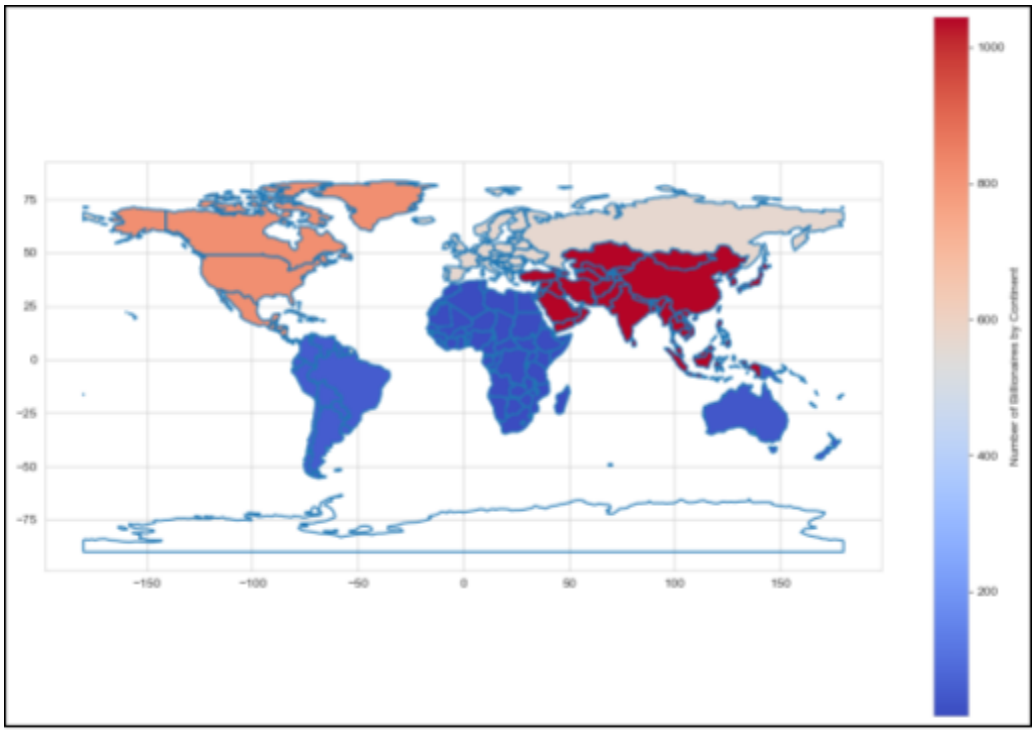

Figure 5: Heat Map of Billionaires by Continent

Lastly, using geopandas to create a geospatial heatmap to depict the global spread of billionaires by continent, we can observe the intense red hue over Asia signifies its majority share, accommodating 1,045 billionaires, which constitutes 41% of the dataset's totality. These billionaires are involved, notably in “Manufacturing” (211 billionaires), “Technology” (130), and “Healthcare” (110), signifying a robust economic diversity. North America, second in density, stands out by its “Finance & Investments” sector, home to 204 billionaires, and a substantial “Technology” sphere, with 146 billionaires. Europe, painted with a milder shade, represents 23% of the dataset with 570 billionaires, primarily from the “Fashion & Retail” and “Finance & Investments” sectors, suggesting they lean towards more traditional enterprises. Contrastingly, Africa's more modest presence is highlighted in blue, indicating fewer billionaires, and has a notable emphasis on “Metals & Mining”. The visual representation underscores the stark

economic contrasts between continents, with Asia and North America at the forefront in both numbers and wealth, while other continents display distinctive sectoral wealth concentrations.

Our study of the 2023 Billionaires Statistic Dataset revealed interesting trends about billionaires worldwide. We noticed significant differences in gender distribution and in wealth accumulation between self-made or inherited billionaires. Although our regression analysis indicated a statistically significant correlation between age and wealth, the influence of age on final worth was found to be relatively minor, as indicated by the nearly flat regression line. Geographically, Asia has the highest number of billionaires, followed by North America, revealing variation in economic strengths and sector dominants between continents. These insights not only provide us with an exciting understanding of the billionaire landscape but also encourage further exploration of the journey of a billionaire.

To put together the project, each team member attended meetings to find the dataset and discuss our findings. William Li composed the majority of the coding and created sophisticated graphs for our report. Cuiyan Hua, Zhehong Ren, and Yanfei Li interpreted the code and analyzed data to verbalize the final report.