

So you want to be a billionaire?

A look at the characteristics of billionaires around the world

Data

Sources and variables

Our data primarily comes from two sources: Forbes and Business Insider. Forbes provided the main dataset for our project (data.csv), and it was accessible in both CSV and JSON format from import.io, a web data extraction platform.¹ We created the countries.csv file from a table on the Wikipedia entry for ISO 3166-1 country codes.² This file is used to map values from the Forbes dataset onto the map, since the Forbes set had billionaires listed with their country represented as a 3-letter code, while the map JSON file used a 3-digit ISO 3166-1 code for each country.

For the globe visualization, we use the following variables:

- *Country name*: the full name of a country
- *Country ID*: the 3-digit ISO 3166-1 code for each country
- *Country code*: the 3-letter abbreviation for each country that corresponds to the ISO ID
- *Number of billionaires*: the number of billionaires in each country
- *Industry*: the industry that is the source of wealth for any given billionaire

The Forbes dataset is also used in the *Net worth by industry* plot. This plot uses the following variables:

- *Net worth*: the net worth of a billionaire (in \$B)
- *Billionaire*: the name of the billionaire
- *Age*: the age of the billionaire in question
- *Industry*: the industry that made a billionaire wealthy

For the visualization about billionaires and where they went to college, we use data from Business Insider.³ This dataset uses just two variables: (1) college name and (2) the number of billionaires who went to that school.

Criteria for selection / subset of data

¹ <https://import.io/data/set/?mode=loadSet&set=f6bf4b86-ce14-4969-ba0a->

² https://en.wikipedia.org/wiki/ISO_3166-1

³ <http://www.businessinsider.com/schools-with-most-billionaires-2014-10>

All billionaires were left in the dataset, however, the industry of “Politics” was omitted from the globe visualization. This decision was made because only a single billionaire, as reported in the dataset, made their wealth from this field. Further, this individual is from a small country that is hard to spot on the map. If left in the data used for the globe, filtering by politics leaves what appears to be a blank globe, thus providing the user with very little information.

Combining data, additional data and code

To generate the map, we use the world-110m.json files from Mike Bostock’s GitHub page.⁴ Since this file represents countries by their 3-digit ISO 3166-1 code, and the Forbes dataset uses 3-letter abbreviations, we use a supplemental file (countries.csv) to map between the two. This supplemental file has 3-letter and 3-digit code variables for each country name, allowing us to easily plot values from the Forbes set onto the map.

To learn how to make an interactive globe with D3, code from Derek Watkins, Johan Sundstrom and Mike Bostock on bl.ocks.org was referenced and modified. Watkins provided good examples of globe styling (sunlight and drop-shadow appearance)⁵, while Sundstrom⁶ and Mike⁷ provided a good basis for making a rotating globe.

Several D3 extensions are also used in the creation of the globe visualization and tooltips: topojson, queue and tip. The first two come directly from d3js.org/, while tip is linked to from CloudFlare’s CDN. Skeleton.css is a CSS boilerplate used to help make the website responsive. It is from a third-party source and available at getskeleton.com.⁸

We sourced the images of billionaire faces from Bloomberg.⁹ We stored these images as individual files in the /img/ folder in our directory, and appended a column to the end of data.csv (our primary dataset) with a link to a given billionaire’s portrait, where available.

⁴ <https://github.com/mbostock/topojson/blob/master/examples/world-110m.json>

⁵ <http://bl.ocks.org/dwtkns/4686432>

⁶ <http://bl.ocks.org/johan/1392560>

⁷ <http://bl.ocks.org/mbostock/3795040>

⁸ <http://getskeleton.com/>

⁹ <http://www.bloomberg.com/billionaires/latest>

Mapping

Globe

Color and position are the primary means of mapping data on the globe. The color gradient of a country varies based on how many billionaires for a given sector are from that country. The darkest green represents the most billionaires relative to other countries with billionaires from that sector.

Billionaires are not evenly distributed around the world. A few countries tend to have a number of billionaires that is magnitudes greater than the much larger number of countries that have a few billionaires at most. As such, a log scale is used for assigning gradient. We initially used a linear scale, but saw this left a globe with a handful of noticeably colored nations, while those that had a few billionaires were nearly indistinguishable from those that had none. Since this is not conducive to communicating any kind of useful information about billionaires, we implemented a log scale.

The globe is interactive; you can drag and spin the globe, as well as click on a country to see its name and the number of billionaires from that country based on how you chose to filter the data.

Net worth by industry plot and heatmap

For the *Net worth by industry* plot, we map each individual onto either a circle, or an image of his or her face (where available – tech has a lot of them!). The location of the dot along the categorical x-axis corresponds to the industry that made the individual wealthy, and the point's location along the y-axis corresponds to that individual's net worth. Industries depicted along the x-axis can be controlled by changing the three drop-down menus above the plot. Further, data presented on this plot can be filtered by different age groups and countries. Clicking on a point brings up a tooltip with more information on the person.

The adjacent heatmap provides another view of the same data plotted on the individual net worth by industry plot. This visualization compares the total number of billionaires and total net worth of those billionaires by industry. Color gradient is used to illustrate the relative differences between industries. For example, if the cell corresponding to *Technology* and *Net worth* is darker than that of *Oils* and *Net Worth*, then we can easily tell that the overall net worth of technology billionaires is higher than that of oil billionaires. Again, clicking on a cell of the heatmap brings up a tooltip with the absolute value on which the color gradient of that cell is based.

College plot

The final visualization shows the number of billionaires that attended different universities around the world. Length of each bar corresponds to how many billionaires attended the given school.

Story

Through our visualizations, we came to see that if you want to be a billionaire, America is the place to be. At 492 total billionaires, America not only has the most, but it also has the most in technology as well as investments. Technology is a special sector because it is the industry that also has the youngest billionaires. Therefore, for anybody looking to become a billionaire (and for whom inheriting a fortune is not an option), entering the technology sector in America is the best bet. Further, Cornell has turned out the 5th highest number of billionaires out of any school around the world. We imagine that all these factors come as good news to members of this class in particular!