

Portraits of Nobel Laureates

Shulei Zang

Information Visualization (INFSCI 2415)

shz131@pitt.edu



Legend

This project is to look at the characteristics of groups of Nobel laureates. We will analyze the winners' group from various perspectives, such as quantities, countries, genders, ages, cities, etc., to have some available portraits of the winners.

Figure 1. Nobel Prizes by Year

The first figure shows the quantity of the Nobel Prize in terms of time. In the figure above, the x-axis is time series ranging from 1901 to 2016, and the year is chosen as the granularity. And the y-axis is how many prizes were awarded in that year. This picture can reflect the relationship between the year and the number of winners.

Figure 2. Distribution of year and age with Joint Plot

This figure shows the distribution of laureates' age in terms of time. In the figure in the middle, the x-axis is time series ranging from 1901 to 2016, and the year is chosen as the granularity. And the y-axis is age ranging from 10 to over 92. The green dots represent the winner's age. The green line stands for the logistic regression line. The blue bar graph on the top stands for the number of winners in each year range. Each bar stands for the number of winners during

that year range. The bar graph on the right side stands for the age distribution of all winners from 1901-2016. The length of the bar stands for the number of laureates in each age range. The blue line is the logistic regression line.

Figure 3. Male and Female

The figure above shows male and female laureate numbers. The blue column represents the total number of male awards, and the orange column represents the total number of female awards. There are about 50 female laureates and 893 male laureates. We can use this picture to see the ratio of male to female winners.

Figure 4&5. Birth Country and Organization University

The figure on the left represents the laureate's birth country. The x-axis represents the number of winners, and the y-axis represents different countries. The figure on the right represents the winners' organization university. The x-axis represents the number of winners, and the y-axis represents the various organizations. Colored columns indicate the number of recipients owned by different countries and organizations. The number of winners is arranged from big to small, from top to bottom. The top bar represents the most significant number of laureates. These two pictures can help us understand the comparison of the number of winners of different institutions in different countries.

Findings:

1.

From 1902 to 2016, the number of Nobel Prize winners showed an upward trend, but in some years, the number decreased significantly. If we look at the data in combination with history, we can find that during the war years, the number of awards dropped significantly.

2.

From the figure in the upper middle, we can find that the purple dots generally become denser and denser with time, which means the number of winners increases yearly. From the purple regression line, we can conclude that the winners' ages are mainly scattered between 50 to 60 years old, and this tendency increases with time. More details are in the bar graph on the top and the side. We can find from the bar graph on the top that the number of winners stays the same, in some years, the amount is less than in previous, maybe because of wars. The bar graph on the right shows that the maximum number of winners is around 60 years old. The farther away from this age, the fewer people, but it is not absolute. We can find that after the age of 60, there is an age group with low numbers of winners.

3.

From the figure in the upper right, we can find that the male number is almost 18 times than female number in total. We can conclude that the male number is far greater than the number of females.

4.

By analyzing the laureate's birth country and organization, we can find that most laureates come from the United States of America. The second is the United Kingdom. And then Germany, France, and Sweden. The organization university with the most laureates is the

University of California, then Harvard University, Stanford University, and MIT. Most of these universities are in the United States.

Data and Methods

The primary data source for this project is an archive.csv dataset document. I mainly analyze the dataset in this document. This file contains data on all winners from 1901 to 2016. They are the category, the prize, the motivation, the laureate type, the gender of the winner, the date of birth, the name, the place of birth, the institution's name, etc.

Many steps were followed to generate the data for these figures. First, installation of the Jupyter notebook, and then an account on the Jupyter notebook was created to store the CSV data. After this was accessed, I imported the following packages.

```
import pandas as pd
pd.plotting.register_matplotlib_converters()
import seaborn as sns
import numpy as np
import matplotlib.pyplot as plt
sns.set()
```

And then, I wrote down python code to filter data and get the figure I needed. Meanwhile, I changed the default color to make the picture more diverse. I mainly analyzed winners' quantities, countries, organizations, genders, ages, etc., to have a general portrait of the laureate.

Significance

The Nobel Prize is a group of yearly international honors presented in many categories by institutions in Sweden and Norway to appreciate achievements in academia, culture, and science. The five Nobel prizes were founded in 1895 by the Swedish physicist Alfred Nobel's testament. In 1901, the first prizes in physiology or medicine, chemistry, literature, peace, and physics were given out. The Nobel Prize is recognized as the most distinguished honor in literature, medicine, physics, chemistry, economics, and campaigning for peace. This project is to look at the characteristics of groups of Nobel laureates. We will analyze the winners' group from various perspectives, such as quantities, countries, genders, ages, cities, etc., to have some available portraits of the winners.

Github link:

<https://github.com/Shirlycherie/2415-Project>