

1 ABSTRACT

Utilizing data gathered about the backgrounds of the 357 NASA Astronauts from 1959-2013, an analysis of the path to becoming an Astronaut can be performed. The most common path is to be born a male in the United States, join the military, and then to study Aerospace for both Undergraduate and Graduate Degrees. Investigating the birthplace further shows that California, New York, and Texas historically produce the most Astronauts. Calculating the number of Astronauts per million inhabitants shows a more uniform distribution of Astronauts across the United States, however, there is also some correlation between the number of Astronauts and the access to prestigious Universities in the state of birth. This correlation requires further investigation in order to evaluate if there is a link between birthplace and the chances of becoming an Astronaut. The change in the proportion of female Astronauts over time is also analyzed. After 19 years, NASA selected its first female Astronauts in 1978, however the percentage of women in a given selection year has not reached 50% as of the end of this dataset. Additionally, women are not as successful as their male counterparts when using the percentage of spacewalks as a measure. The path has changed over the years with the inclusion of women in the program, however, females remain underrepresented and underutilized in the NASA Space Program.

2 INTRODUCTION

The National Aeronautics and Space Administration (NASA) is the United States government agency responsible for U.S. space exploration, space technology, Earth and space science, and aeronautics research. NASA has an objective to expand human knowledge through new scientific discoveries and to extend human presence deeper into space. The goal of this analysis is to investigate the “path” to becoming a NASA astronaut and how that path has changed over time. First, the most common path is presented with a deeper look into some of the background information about the Astronauts. Next, the temporal changes in the make-up of the Space Program will be presented to specifically determine how the path has improved for women over time.

3 METHODS

The dataset used for this analysis was found using Google’s Datasetsearch tool and was obtained from Kaggle. This dataset was distributed using the Creative Commons CC0 license which dedicates the data to the public domain.

Five attributes about the Astronauts’ backgrounds were used to determine the path. These attributes are Gender, Birthplace, Military Service, Undergraduate Degree, and Graduate Degree. To provide an easily digestible graphic displaying the most common path, binning was performed on four of the five attributes to provide higher-level categories. The binned data were then aggregated across all Astronauts and plotted using a Sankey plot.

Further analysis of the Birthplace data was performed, and the frequency of Astronauts by US State was plotted on a map to find areas where the number of Astronauts exceeded the expected value based on the state population.

Next, the inclusion of women in the Astronaut program was further investigated by plotting the counts of males and females for each selection year provided in the dataset. The number of flight hours and space walks are then used as a metric for comparing the success of women in the program with the success of their male counterparts.

4 RESULTS

4.1 MOST COMMON PATH

The data for all Astronauts in the dataset (1959 – 2013) were aggregated and plotted in the Sankey plot below.

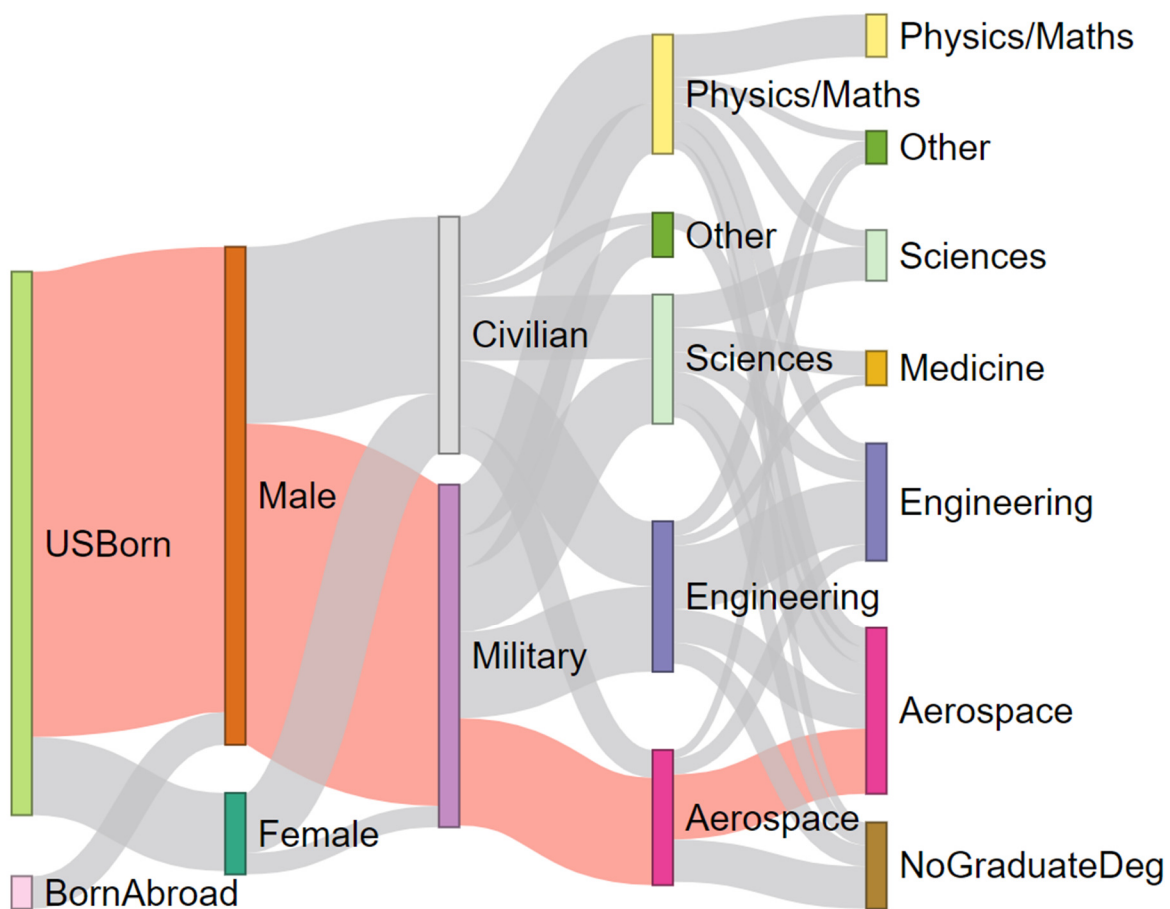


Figure 4-1 - Most Common Path

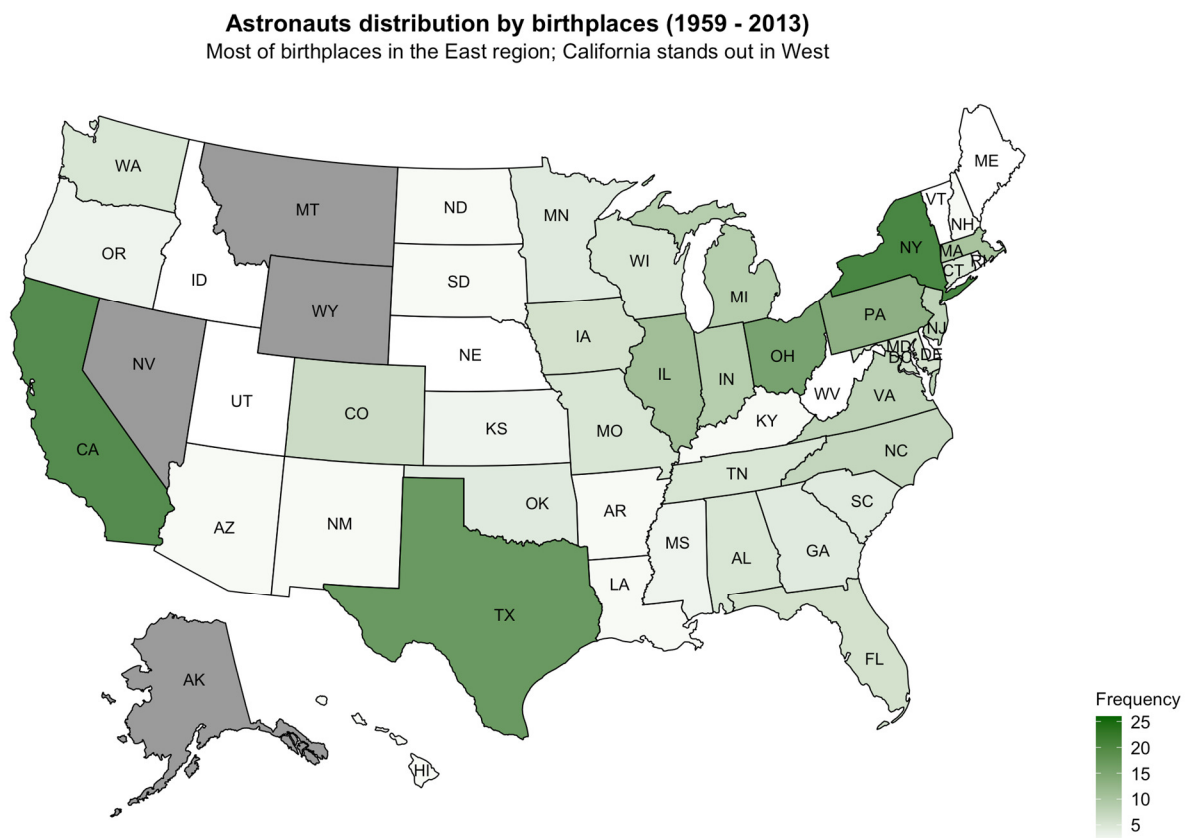
The highlighted path was chosen by determining the 'link' to the next level in the plot that had the highest value (i.e. the path that most Astronauts took). It is apparent from the plot that studying Aerospace after joining the military is the most common path to becoming an Astronaut. Studying other types of Engineering would also be a good choice if one were aspiring to become an Astronaut.

The prevalence of Medical degrees in the Graduate Majors was an interesting outcome as these degrees do not seem to be related to space flight. However, in the context of medical research on the International Space Station, it is understandable that there would be a relatively large group of Astronauts in the medical field. Also interesting is the apparent disparity in the numbers of women and men serving as NASA Astronauts. This disparity will be investigated further in this analysis.

4.2 MOST COMMON BIRTHPLACES

Drilling into the birthplace information from the dataset, it is possible to analyze the regions which have produced the most Astronauts. It is interesting because some hypotheses could be proposed trying to find the reason for some patterns. Some key findings in the following plot are:

- Concentration points in California, Texas, and the North East region
- Alaska, Nevada, Wyoming, and Montana have not produced any astronauts
- New York with 26 and California with 25 are those with most astronauts selected
- Texas, which is the home of NASA's Johnson Space Centre (est. 1961), is also very strong



Data Source: Kaggle NASA Astronauts

Figure 4-2 - Astronauts by Birthplace

One potential explanation for these hotspots could be access to more prestigious education. The plot below shows the location of some of the top-ranked Universities in the United States. This map resembles

21. Jul. 2020

the plot above. NASA's selection process is very strict, and an important characteristic is a strong education in several areas, so a probable important filter is an education from a renowned university. Important universities as University of Stanford and Berkeley in California, United States Military Academy in New York, University of Texas in Texas and several others are in those states with more astronauts selected.



Figure 4-3 - Prestigious Universities of the US

Another explanation for the higher number of Astronauts from California, Texas, and New York could be the higher population in those states. The plot below shows the number of astronauts by birthplace relative to each state's proportion of the total US population.

Astronauts distribution by birthplaces (1959 - 2013)

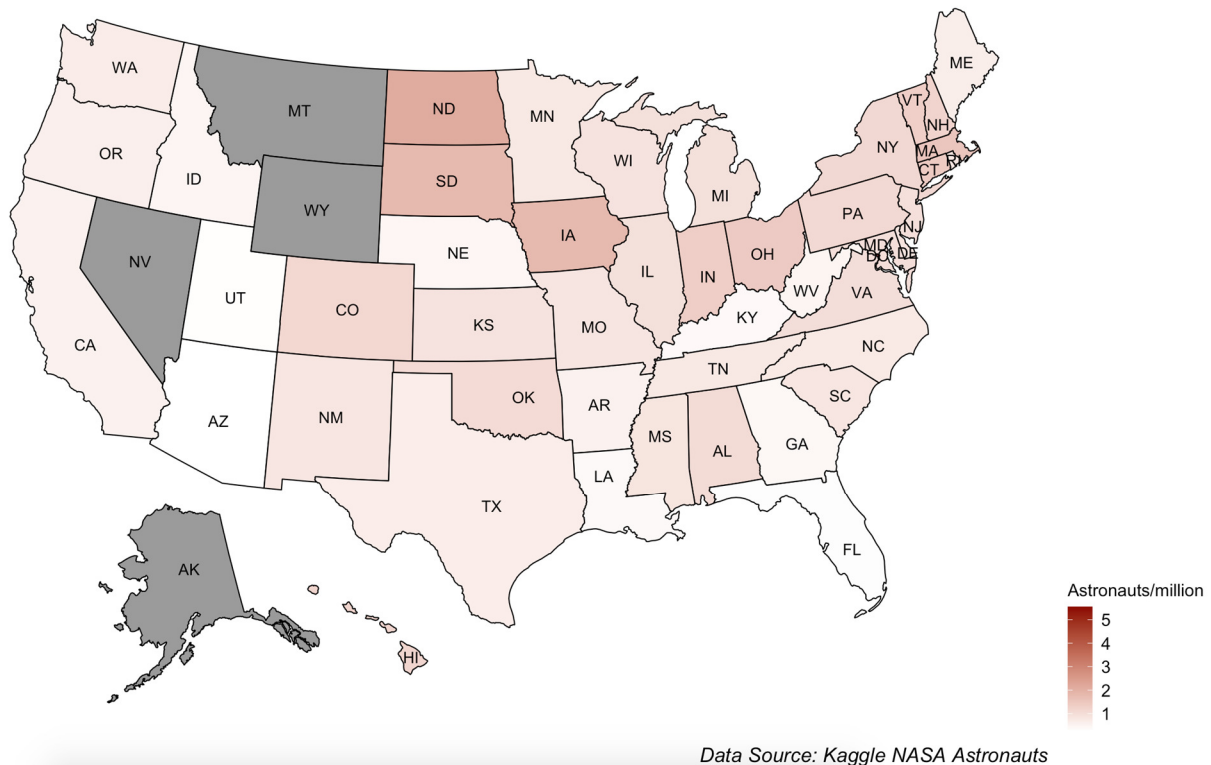


Figure 4-4 - Astronauts per Million Inhabitants

Each state shows by color its proportion of astronauts per million inhabitants, so a fairer comparison is performed. Now, a more balanced distribution by birthplaces is depicted in the plot. States in the Midwest region show a bit more astronauts per million inhabitants, but this is likely due to the lower population of these states.

4.3 INCLUSION OF WOMEN

The number of men and women chosen for each selection year is provided in the plot below. Interestingly, there were no female astronauts selected by NASA until 1978, 19 years after the start of the Space Program. As an aside, Sally Ride was selected to be an astronaut as part of 1978's astronaut group. A curious fact is that she applied after seeing an advertisement in the Stanford student newspaper and was one of only 35 people selected out of the 8000 applications. The 1978 selection also included Shannon W. Lucid; the first mother to be hired as an astronaut.

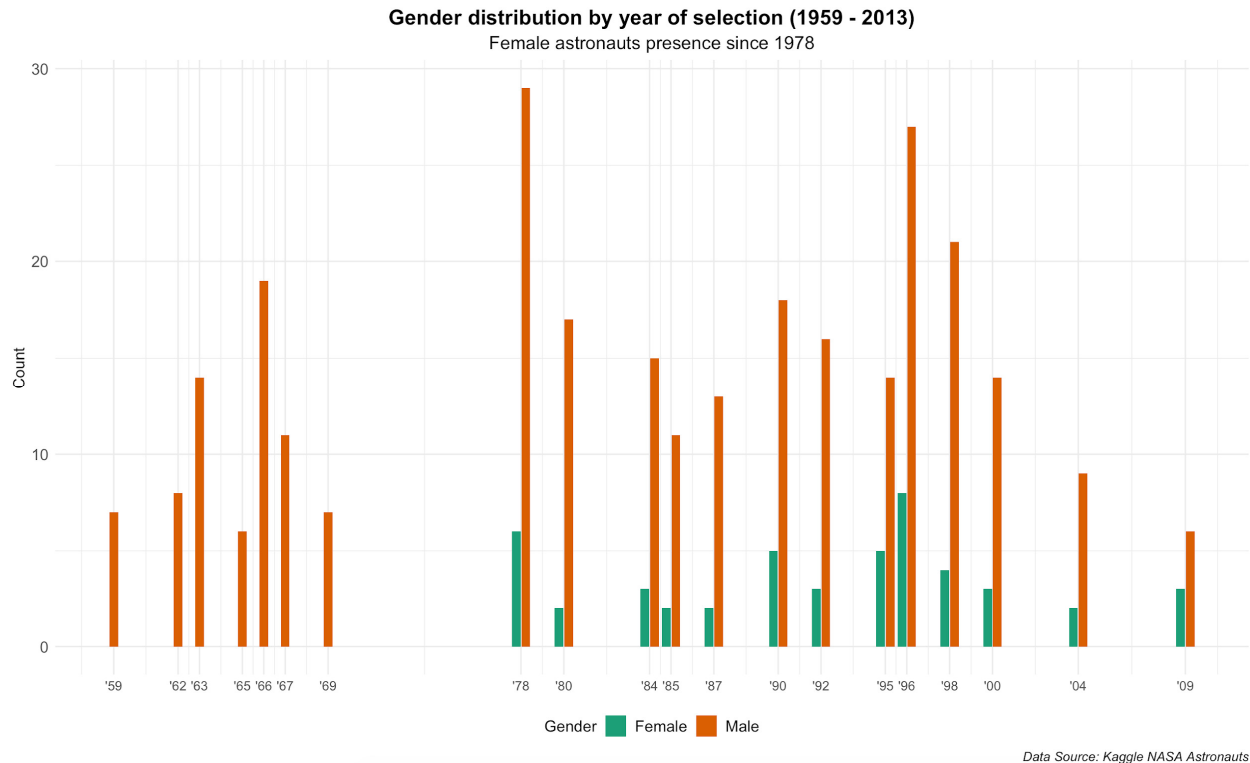


Figure 4-5 - Gender By Selection Year

It is interesting to see what happened in 1978 and 1996, there are obvious peaks in these years. An investigation into these years showed that there are important events that could have affected the astronaut's selection. The NASA Astronaut Group 8 was a group of 35 astronauts selected on January 16, 1978. The previous Astronaut selection took place in 1969, almost 10 years before the first group including females. One possible reason for this is the long delay between the last Apollo lunar mission (1972) and the first Space Shuttle mission (1981). Similarly, in 1996, construction was started on the International Space Station, which required a lot of resources and astronauts.

4.4 THE SUCCESS OF WOMEN

Since beginning to include women in the NASA Astronaut program in 1978, the percentage of women in each selection year has seen a gradual rise but has not reached 50% as of the last year presented in this data (2013). The graph in Figure 4-6 below shows the gradual increase and also shows that the share of flight hours has roughly tracked the enrollment percentage of women.

A similar plot comparing the percentage of women for each selection year versus the percentage of total spacewalks is shown in the Figure 4-7 below. This shows that even with gradually increasing selection numbers, women still do not receive as many opportunities for spacewalks. One potential reason for this is that the roles women perform do not require spacewalks. For example, Sally Ride was the first woman in space in 1983 and had more than 300 hours in space but her missions involved deployment of communication satellites which did not require as many spacewalks. In any case, there is a definite gap between males and females in this aspect of the Space Program.

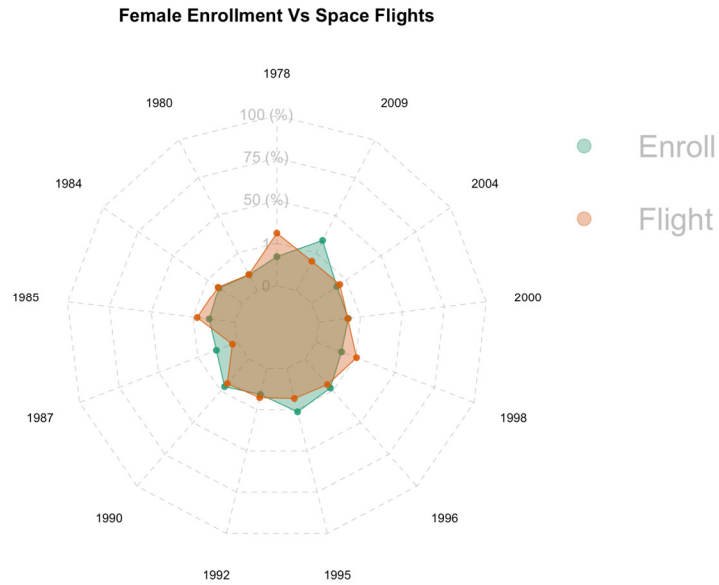


Figure 4-6 - Female Enrollment vs. Space Flights

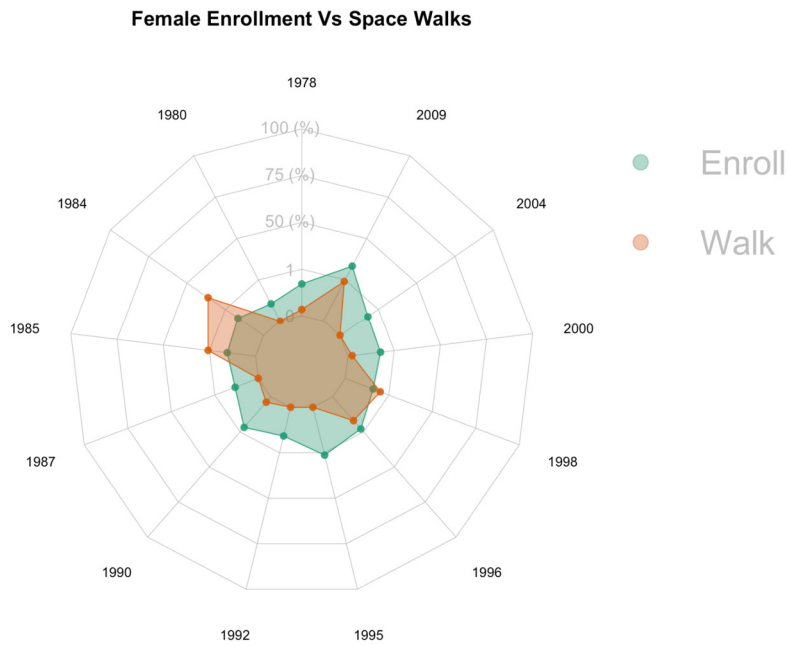


Figure 4-7 - Female Enrollment vs Percentage of Spacewalks

5 CONCLUSIONS

The traditional path to becoming an Astronaut comes as no surprise. The vast majority of Astronauts have been US-born males with a military background and education in the field of Aerospace.

Taking a look at birthplaces, most of the astronauts were born in those regions with a strong influence from excellent higher education institutions. The North-East region, California and Texas are well identified concentration points. Using a proportion of astronauts per million inhabitants makes a fairer comparison between states and yields a map that is relatively uniform. One potential future analysis in this area could be to determine if Astronauts attended the prestigious universities near their birthplaces or if they more often went to different states for their education. This would help to further determine if there is any connection between birthplace and the chances of becoming an Astronaut.

The inclusion of women was a good first step by NASA in 1978, however, as of the end of this data, women still have not reached 50% of a NASA induction class. Additionally, they have also been underrepresented in the percentage of spacewalks, showing that even those women inducted are being overshadowed by their male counterparts.

6 REFERENCES

- NASA Astronauts Dataset, 1959-Present. <https://www.kaggle.com/nasa/astronaut-yearbook>.
- Wikipedia contributors. (2020, July 17). List of astronauts by year of selection. In Wikipedia, The Free Encyclopedia. Retrieved 21:21, July 20, 2020, from https://en.wikipedia.org/w/index.php?title=List_of_astronauts_by_year_of_selection&oldid=968078165
- Wikipedia contributors. (2020, June 16). Sally Ride. In Wikipedia, The Free Encyclopedia. Retrieved 21:21, July 20, 2020, from https://en.wikipedia.org/w/index.php?title=Sally_Ride&oldid=962830910
- NASA website. <https://www.nasa.gov/>
- Top Universities in USA. <http://www.mapsofworld.com/usa/universities/>