Change of the Demographic Profile and the Objectives of the Space Missions 1959 – 2018

Fernando J. Aponte-Aponte[1], Víctor M. Hernández-Denis[2], Alanís Barriera-Rodríguez[3], Yaritza Colon-García[4], Mario Signoret-Rodríguez[5], John P Gallegos Matos[6]

---

[1] 3rd Year Student, Computer Science Program; Inter-American University, San Germán Campus: FAponte2354@intersg.edu

[2] 5th Year Student, Industrial Engineering Program; Inter-American University, Bayamón Campus: VHernandez1706@interbayamon.edu

[3] 2nd Year Student, Industrial Engineering Program; Inter-American University, Bayamón Campus: Abarriera2828@interbayamon.edu

[4] 3rd Year Student, Natural Sciences Program; Inter-American University, San Germán Campus: ycolon4369@intersg.edu

[5] Faculty, Business Administration Program: Inter-American University, Bayamón Campus: msignoret@bayamon.inter.edu

[6] Faculty, Department of Natural Sciences: Inter-American University, Bayamón Campus: johngallegos1958@gmail.com

Abstract

This exploratory research is the final product of the hands-on workshops offered by the Institute of Data Sciences at the Inter-American University of Puerto Rico during the 2021 Summer Session. It includes a statistical and descriptive analysis of a dataset that incorporates demographical and operational characteristics of 1,278 space flights that were carried out during the 1959 to 2018 period. The goal of this study was to discover, through the use of the R Studio software, how changes in the strategic direction of the space programs had an impact in the operational and demographic characteristics of the space flights missions. This endeavor required the application of the Dplyr and Ggplot2 packages for data exploration and visualization. The findings of this research indicates that, in the early 1980s, there were major changes in the inclusion of women and other minority groups in these spaceflights. Similarly, we found strong evidence that confirms the advancement of the astronaut selection programs in achieving a certain degree of ethnic, professional, national and gender diversification. Finally, it was possible to verify, how change in the strategic direction of the space programs (from a militaristic to a scientific scope) had an impact in the duration of extravehicular activities and of space flights.

Introduction:

On October 4, 1957 the Soviet Union launched the first artificial satellite into space. This milestone in human history propelled a race between the two most powerful post WWII nations to conquer space(Onion et al., 2020). In the midst of the cold war, this race to space had immediate military and geopolitical implications. The two nations then had to select the most capable and trained personnel available for this mission in the shortest period of time(Collins, 2015; Rodríguez, 2019). The first astronauts were selected among the military forces for their piloting skills and their ability to undergo dangerous and physically demanding activities of these first expeditions(Lathers, 2019). This first cadre of space travelers were composed exclusively of white men with military test pilot experience (NASA,2010). "'Women have fought an uphill battle to, first, participate in spaceflight and then, to do so on an equal playing field with their male counterparts'' (Healey, 2018. p1). Up until 1978, NASA selected only white men. Then a combination of shifting in the space mission objectives, cultural attitudes and new requirements to fly the space shuttle gave women and other minority groups the opportunity to join the astronaut corps (Treat et al.,2020).

From the early 1980s the interest in space in both, the United States and the Soviet Union, changed drastically. It was no longer a race to conquer it, but as a possible platform to conduct scientific experiments. As a result, the United States adopted the Spacelab and Space Shuttle Programs. Similarly, in 1986 the Soviet Union adopted an analogous approach with the space station MIR. This strategic shift in both space programs had its impacts on both the duration of space missions and the demographic profile of astronauts(Treat et al.,2020). During the past 60 years, the demographics characteristics of the astronaut selection program and the types and durations of missions have changed. This has facilitated a more diverse group of space travelers with a variety of educational and scientific backgrounds.

This study aims to explore whether or not the demographic profile of the astronauts and the characteristics of the space flight has changed systematically over time. For this task we analyzed a dataset of all 1,278 space mission flights from 1959 to 2018.The 24 variables included in this dataset describe the dates, mission, gender, education, experience (civilian or military), mission roles (pilot, payload specialist, mission specialist, etc.) date of selection and duration of all missions during the above mentioned period.

Methodology:

The dataset we used to gather the information was brought to us by instructor Rolando J. Acosta over the summer in the Data Science Summer Institute. The data set can be found under his Github profile as austronauts.csv inside a project-datasets folder. Then, when we opened the file inside R Studio and modified the data to optimize the search between the columns, we used the **as_tibble()**

function to manage the data in a more efficient way taking out null values and unwanted, incomplete and unreadable data.

After we loaded the austronauts.csv to our RStudio as a markdown file for easy manipulation we had the file called demographics.rmd inside a git repository called Astronauts-Demographics. The data was actually read in RStudio through the import method in the platform here. We added the parameters necessary to possibly read the data in the best way, we encoded the data to **UTF8** since there was some data in Russian, certain fields were difficult to interpret. As row names we used the first line since that was each column title, then we used "tab" as our separators and double quote as our quote to let the interpreter know what was a string or a single quote in a name in particular.

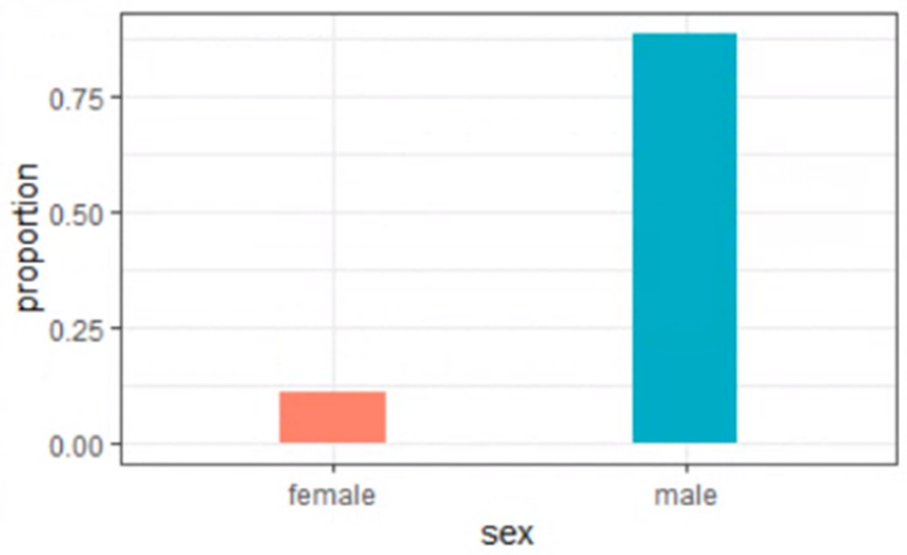
We prepared the R markdown files with the necessary libraries and other setup. We used "astronauts" as the title inside the demographics.Rmd and **pdf_document** as the output for later use. We called our libraries: **tidyverse**, **dplyr**, **ggplot2**, **dslabs**, **ggridges.** Tidyverse being the main basis of our project which is a collection of R packages for data science. Dplyr is the library for data manipulation, making it easy to manage the data. Ggplot2 for the graphic part of the run which actually works as a system to create the graphics of each of our analyses. Dslabs worked perfectly too, since it helped us out at the moment of making our data analysis at the time of practicing beforehand. Finally ggridges which provided us with more graphics to better visualize our data as ridges.

In our code lines, we transferred the data of the csv to an object called astronauts since the actual csv file was called ("*austronauts*" a particular ***typo***) as a **tbl_df** which worked as our first line of the pipelining method, this subclass of data.frame helped in the manipulation of data making it more clear for the eye. In another line of code we implemented a **summary** of the new object but even at that point the data was kind of hard to read, so instead the **glimpse** function was used; transposing our data and making it possible to better understand what we were working on. After this our dataset had 1277 observations of 23 variables (rows).

Then, we stated that we would focus on the demographics of the astronauts so we graph all the information on various sections so that we could interpret it. We used a **mutate** function to **sort** the dataset by nationality to have a better view of the graph, in this way gather the information needed to actually know who had the more missions per year. After analyzing the data we saw that the median of missions was in a gap between 1980 to the beginning of 2000. We also analized the demograpgics by *sex* and his status of *military/civilian*. To reach this pouint we need to analyze what we where looking for especifically so first we saved the data of a **ggplot** inside a variable called **bxp** (for boxplot) we **mapped** the data as **sex** for **x** and **year_of_mission** for **y**. Doing so, we divided the information gathered in a **face_wrap** for *military/civilian* adding some **jitter** so that we could visualize how many astronauts actually had each mission and not just a simple dot per mision.

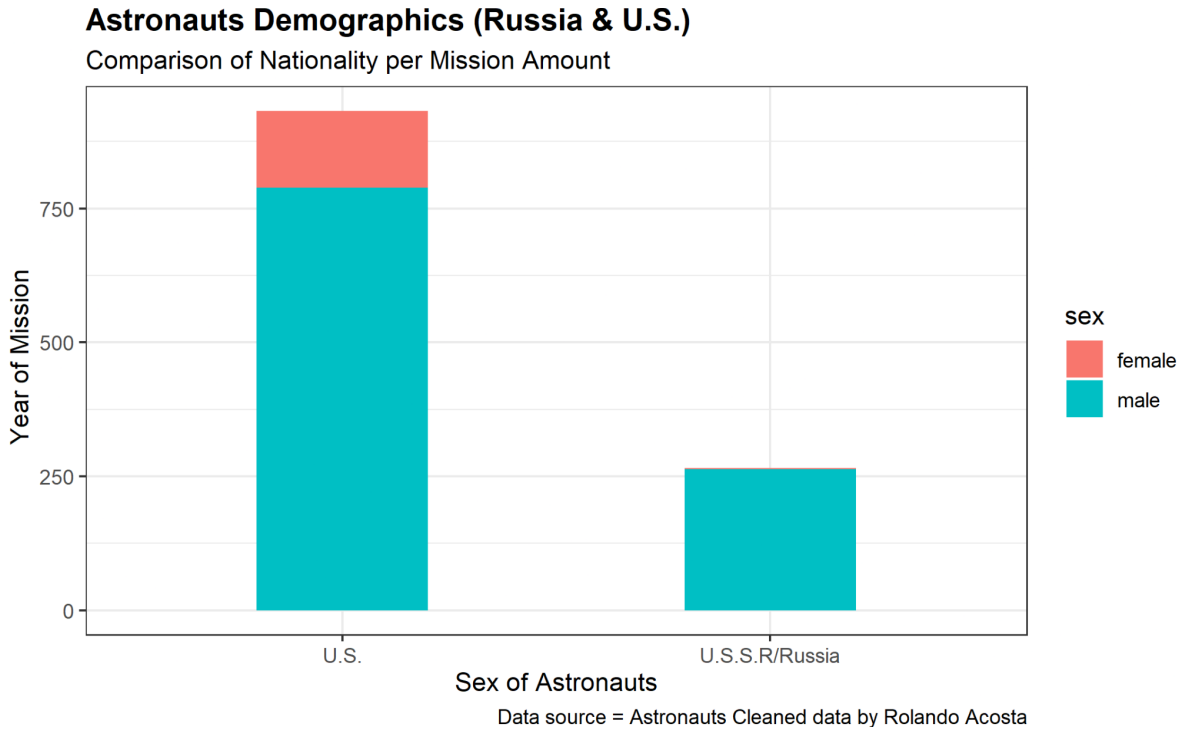
Results:

Fig. 1 - Proportion of astronauts versus there gender



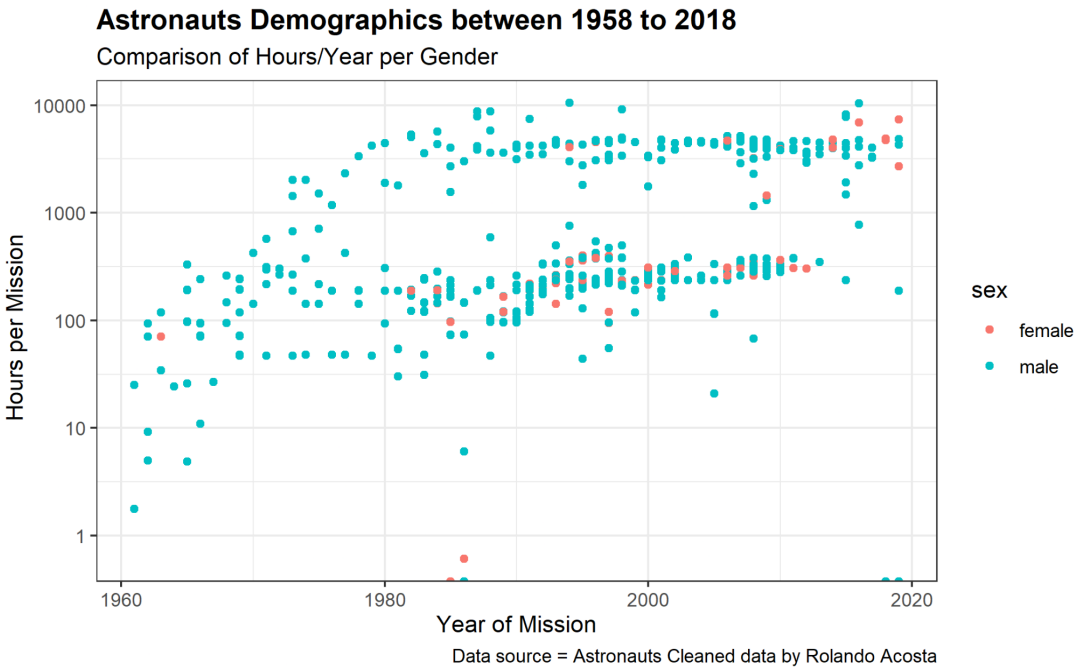
The diagram shows the difference in proportion for the astronauts of each gender for the past 60 years of space travel. We can observe how the proportion of female astronauts is located below 0.25 while the male proportion is substantially higher in comparison being its proportion over 0.75. These results are showing how the male gender has predominated the space travel mansions.

Fig. 2 - The two nations with more flights versus the number of missions by gender



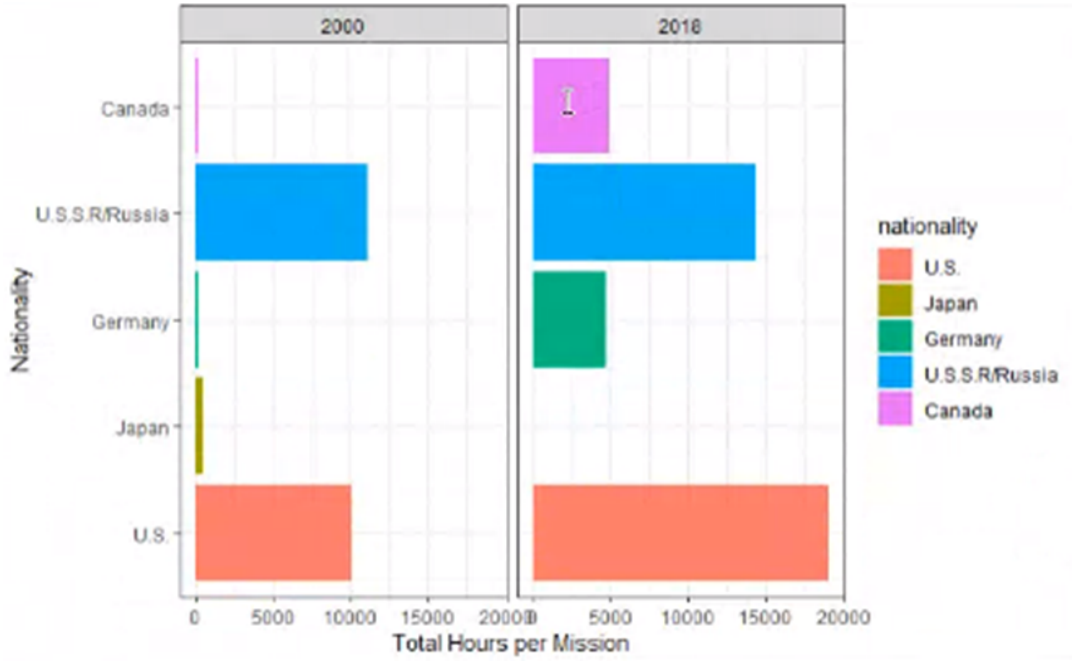
In Fig. 2 we can observe how the number of missions performed by the U.S. space program has a substantial lead on its Soviet counterpart . But it does not only have a far lead on the total number of missions but also has a far lead in the total number of missions performed by a female crew.

Fig. 3 - Year of missions versus the hours of each mission



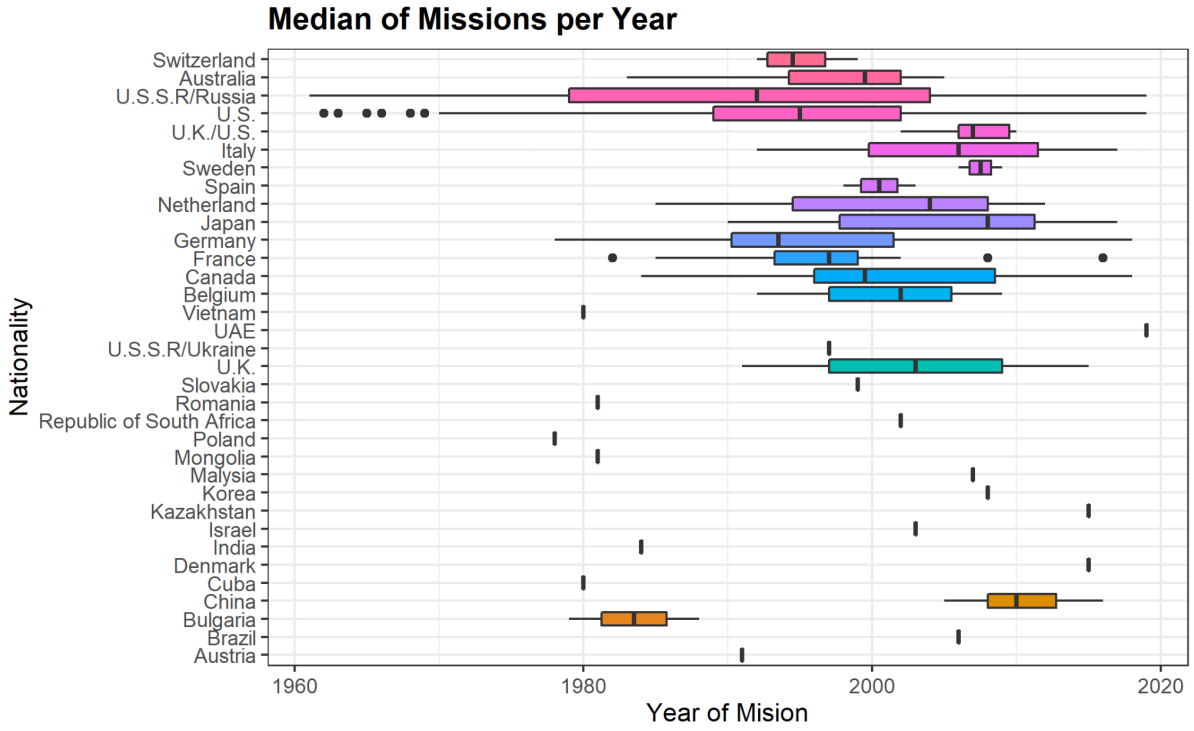
The hours of duration of each mission has extended a substantial amount during these past 60 years as you can see in Fig. 3. It can be observed the duration of a space travel mission during the 1960' s was only over an hour and how during the years these times have changed so much being over 10 thousand hours in the 2018. It also can be observed how the inclusion of females in the space travel missions have changed so much during the years.

Fig. 4 - Total of hours for each mission



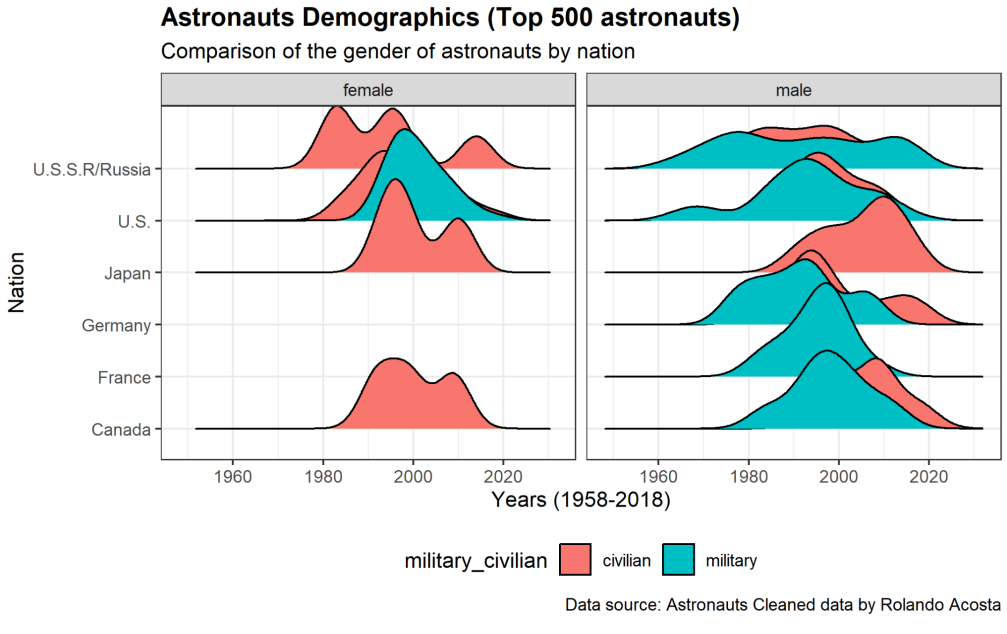
The diagram shows the total hours per mission during the 2000 and 2018 space travel missions with the nationality of the astronauts. We can see how the Soviets had the most hours during the 2000's just a little over the U.S. followed by Japan and Canada and Germany. But we can see how the total hours per mission basically doubled for almost every nationality in 2018 with the U.S. having the highest amount at almost 20 thousands hours and Japan being the lowest having almost 0 hours and being the only one to go down in total hours per mission.

Fig. 5 Year of mission versus nationality



Here you can see the median representation of each nation since the beginning of 1958 with the first flight. Some nations don't represent a big impact in the graph since they just had 1 or in some particular cases 2 astronauts. The other nations like the United States, Rusia, Japan, Germany and even Canada have a significant number of astronauts which imply their behavior for space research.

Fig. 6 Year of mission versus nationality



When segregating the data by nationality and gender whe can see in 1996 and 2003 the peak in total hours of flights performed with a crew including females in these five missions.

Fig. 7 - Table of gender of the astronaut versus hours of mission

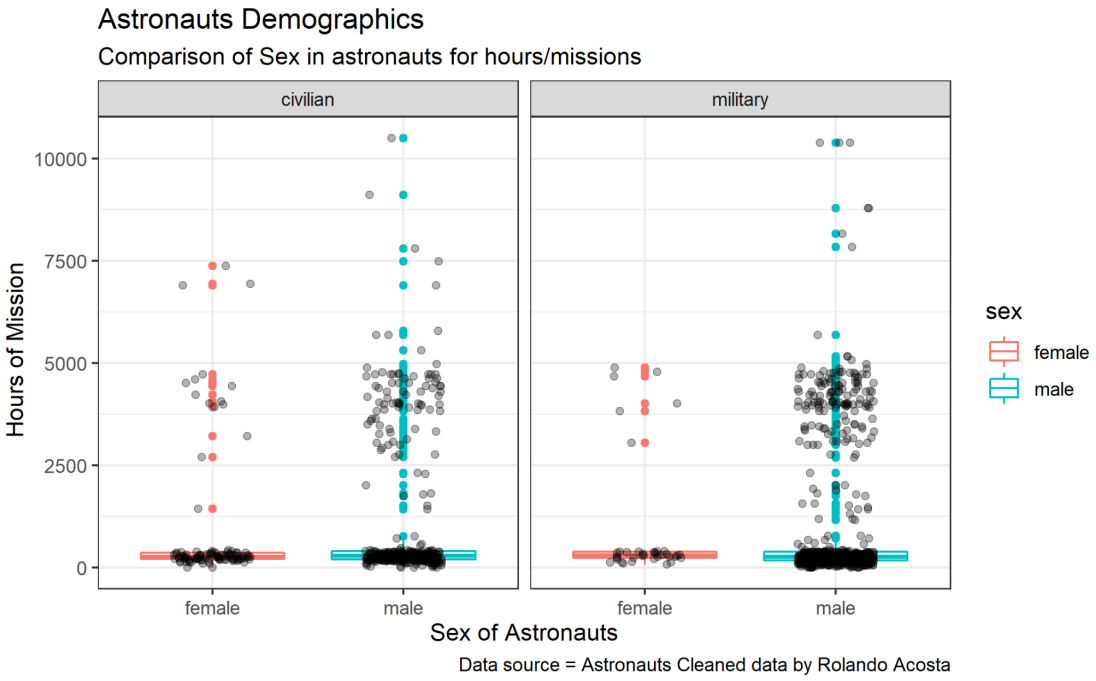
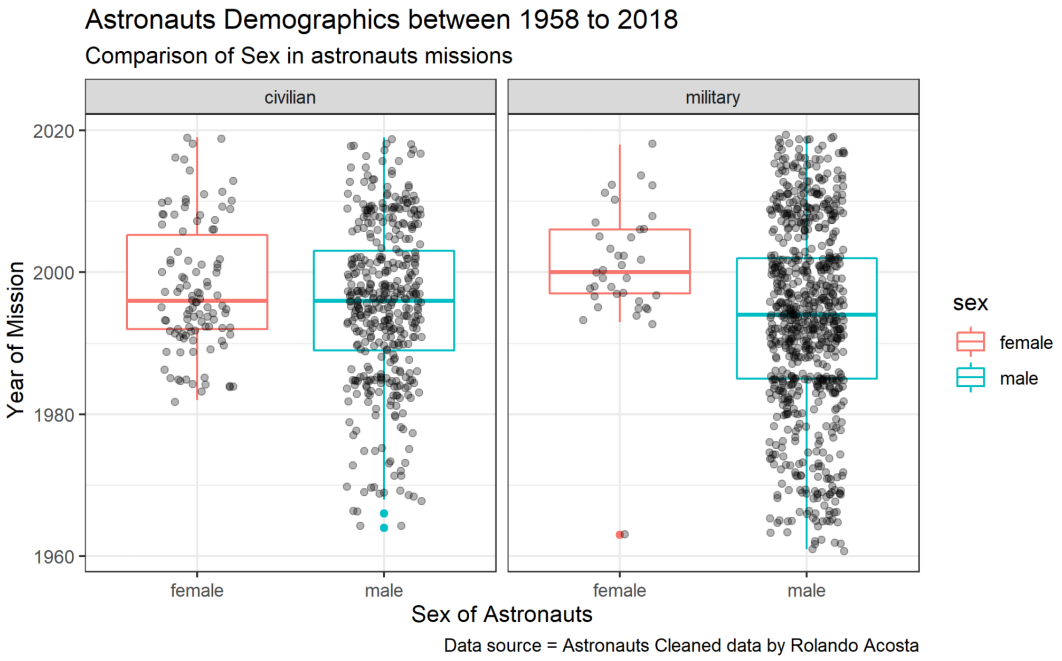


Fig. 7 shows the hours of mission performed by civilian and military astronauts depending on their gender. We can see how the data from the civilian males are very similar to the military but still are less in minority in mission hours by a really small difference. But when the female data is observed and compared we can see how the civilian female astronauts have more hours in mission then the females with a military background.

Fig. 8 - Demographic profile of the astronauts



The first astronauts to be recruited were white military males selected for their pilot skills and their ability to undergo dangerous and physically demanding tasks. The diagram shows how this recruitment method has changed drastically by showing us the amount of civilian and military astronauts performing missions during the past 60 years. It is notable to acknowledge how the astronauts are being recruited not only for their military background but also for their scientific knowledge and preparation for the mission in more recent times.

Conclusion:

The aim of this research was to explore changes in both the operational characteristics and the demographic profile of astronauts for the past 60 years of space missions (1959-2018). In this sense, the R Studio software and the Dplyr and Ggplot2 packages were crucial to this analysis, as they facilitated the manipulation and visualization of our dataset. From this analysis, the following conclusions emerge: First, even though since 1978 we begin to see the inclusion of women in space missions, there is still a long way to go in terms of inclusion and diversity. Only 10% of the total 1,278 spaceflights analyzed in this sample were performed by women. It is worth acknowledging that the U.S. space program far leads its Soviet counterpart in total flight hours performed by female crew. In this regard, from the early 1990s there has been an increase in both the number of flights and the duration of space flights that include female crew.

Also worth mentioning is the fact that when the data is examined by segregating by gender and professional experience, a significant reduction in the total hours of flights performed by crew with military experience over the years can be observed. This reduction coincides with an increase in the total number of flight hours performed by crew with a scientific background. It should be also noted that by segregating the missions according to their professional background (civil or military) you can see a greater number of flight hours of the female crew with a civilian background. In terms of astronaut nationality, the United States, Russia, Japan, Germany and Canada are the five countries that have the most flight hours. However, when comparing the data between 2000 and 2018, a significant growth can be seen in 2018 in the total number of hours flown by crew of Canadian and German origin. When segregating the data according to nationality and gender, it can be seen that between 1996 and 2003 there were peaks in the total hours of flights with female crew in these five nations.

Regarding the operational characteristics of these flights, the effect of the change of strategic direction of the Soviet and American programs can be clearly evidenced. Since the early 1980’s and coinciding with the MIR and the Space Shuttle programs, the number of flight hours per year has increased vertiginously. This increase approximates nearly 900 flight hours per year since the turn of this decade. Undoubtedly, this strategic shift towards a scientific approach has had its impacts on the operational characteristics and demographic profile of astronauts.

References:

Collins, D. (2015). Astronaut Requirements. Retrieved from https://www.nasa.gov/audience/forstudents/postsecondary/features/F_Astronaut_Requirements.html

Healey, D. (2018). There are no bras in space: How spaceflight adapted to women and how women adapt to spaceflight. *The Georgetown Journal of Gender and the Law*, *19*(3), 1–23.

Kovacs, G., & Shadden, M. (2017). Analysis of age as a factor in NASA astronaut selection and career landmarks. *PLoS One*, *12*(7). https://doi.org/DOI:10.1371/journal.pone.0181381

Lathers, M. (2009). No Official Requirement": Women, History, Time, and the U.S. Space Program. *Femenist Studies*, *14*(26).

National Aeronautics and Space Administration. (2010). Early astronaut selection and training. Retrieved from https://science.ksc.nasa.gov/history/early-astronauts.txt

Onion, A., Sullivan, M., & Mullen, M. (2020). The Space Race. Retrieved from https://www.history.com/topics/cold-war/space-race

Rodríguez, H. (2019). La Carrera Espacial paso a paso. Retrieved from https://www.nationalgeographic.com.es/llegada-del-hombre-a-la-luna/carrera-espacial-paso-a-paso_14369

Treat, J., Bennett, J., & Tuner, C. (2020). How ‘the right stuff’ has changed. *National Geographic*. Retrieved from https://www.nationalgeographic.com/science/graphics/charting-how-nasa-astronaut-demographics-have-changed-over-time