

***DATA SCIENCE REPORT***

**SPACE MISSIONS EXPLORATORY**  
**DATA ANALYSIS**

**Nicolas DESBORDES**

*01/28/2021*

## I) INTRODUCTION

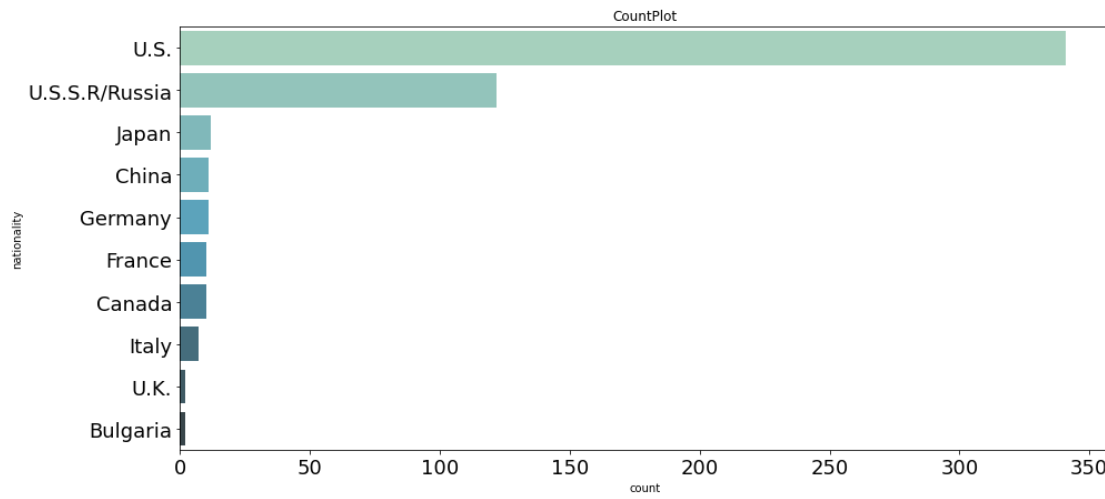
We are given the task of exploring two datasets. The first one describes the astronauts, their identity and the different missions they conducted from 1959 to 2018. The other table gives details about the space missions (company name, status mission, date...). The main objective of this analysis is to explore the data and draw some meaningful conclusions.

To solve this problem, we firstly conducted some data preparation and cleaning, in order to optimise the data quality. Secondly, we derived the key insights given at the Astronauts level, focusing for instance on the differences between the two genders as well as between both military and civilian backgrounds. Lastly, we deeply analysed the data at the Missions level, trying to find out what makes a mission succeed or fail.

## II) KEY INSIGHTS AT THE ASTRONAUTS LEVEL

Exploring the Astronauts table, we draw the key following insights:

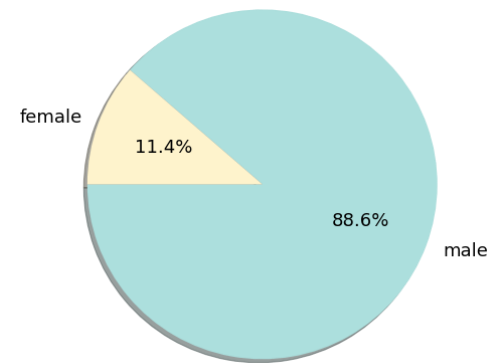
- More than 80% of astronauts are from the USA or from Russia, with a high dominance of American astronauts (61%).
- Only 34 nationalities are represented in the space field, which is surprising.



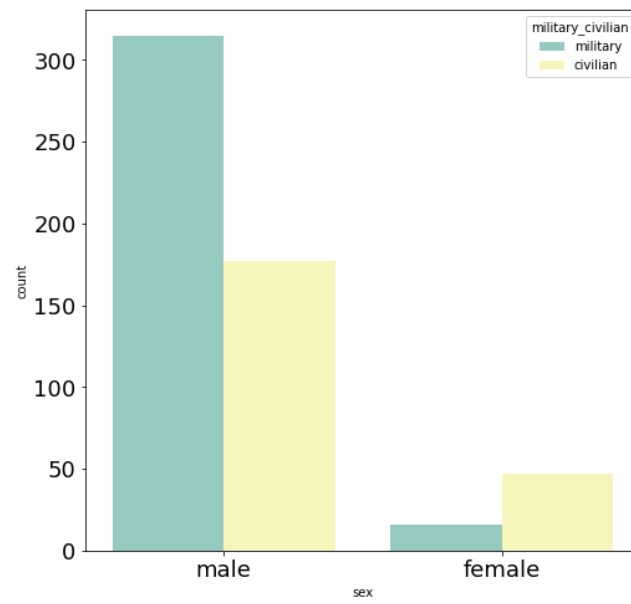
**Figure 1: Number of astronauts by nationality**

- The average age of astronauts is around 45 years old, and in average, an astronaut has done 2 missions with around 10 years of experience. This can be explained either by the fact that the missions are very spaced or by the fact that they are very long. Usually, astronauts have done 0 extra vehicular activities.

- 60% of the astronauts have a military background.
- With a ratio of 7 to 1, female are very under represented in the space field. They have been under represented through all of space history. There was no female selected in the first programs before 1963, first female in space was a Russian cosmonaut named Valentina Tereshkova in 1963. And we had to wait until 1978 to have another woman selected with Sally K. Ride, first female American Astronaut in space in 1978.
- Another thing on gender distribution is when you look at the distribution of Civilian/Military. There are more civilian female astronauts than military ones, and this is the opposite for male astronauts.



**Figure 2: Repartition of the male and female across the astronauts**

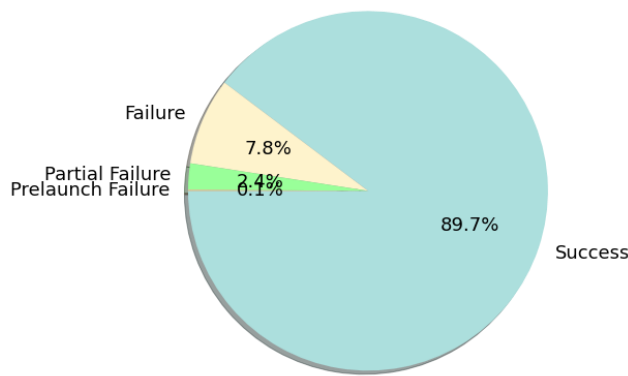


**Figure 3: Background (civilian or military) of the astronauts for each gender type**

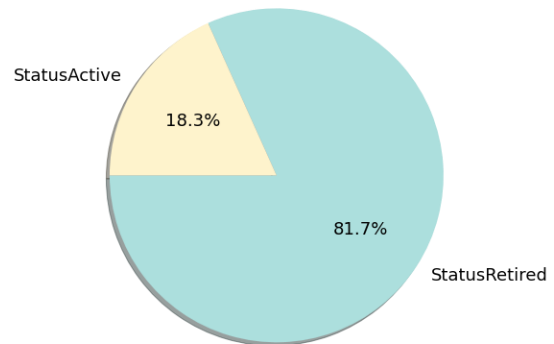
### **III) KEY INSIGHTS AT THE MISSIONS LEVEL**

Exploring both Missions and Astronauts tables, we draw the key following insights:

- In average, there are 5 passengers per mission.
- A vast majority of missions succeed (almost 90%), and given all past missions, 18% are active missions.

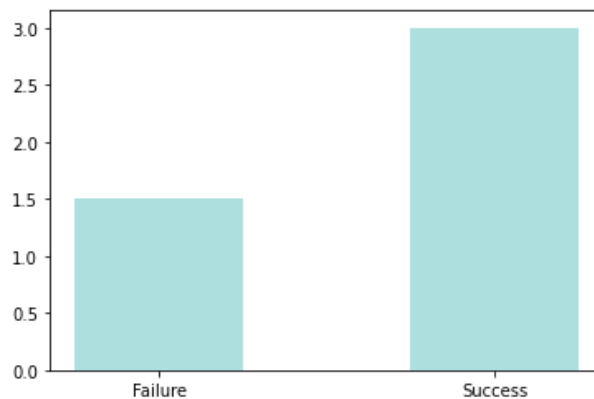


**Figure 4: Proportion of each status mission**



**Figure 5: Proportion of each status rocket**

- There seems to have no relation with EVA hours and hours of flight.
- Failed missions tend to have less EVA and duration than the successful one. Which I find curious. Indeed, I would imagine that longer missions and missions with the more EVA would be more complicated, and more difficult to succeed. This result could be explain by the fact that there are only a few missions that have failed.



**Figure 6: Distribution of EVA for each status mission**

#### **IV) CONCLUSIONS & NEXT STEPS**

To put it in a nutshell, the exploration of these 2 datasets helped us retrieve meaningful insights about the astronauts and the missions they conducted. Nevertheless, we could go deeper and try to:

- Determine why a mission fails using Machine learning algorithms for example
- Predict the cost of each mission according to its characteristics (given the fact that we only have 30% of information regarding the cost)