Visualization of spacecraft launches from 1951-2022

University of Southern Denmark

Faculty of Engineering

Maersk Mc-Kinney Moller Institute

Data Visualization

Group 16

Course code - T520040102

Anders Lindhardt Madsen - andem20@student.sdu.dk
Ásbjörn Skarphéðinsson - usska18@student.sdu.dk
Victor Johannes Larsen Røer - viroe20@student.sdu.dk
Marc Lindegård Weller Bertelsen - berte20@student.sdu.dk

Abstract

Background. Spaceflight has become an integral part of the modern world and it's relevance is only increasing, given the multitude of theoretical and practical applications. A wealth of available data on this topic exists, and it is therefore of great interest to aid the historical understanding of the field by presenting the data in a visual and intuitive way.

Goal. The goal of this project was to formulate a set of questions pertaining to space launches. To answer these questions, a dataset was to be obtained, inspected and processed. The data was then to be used to create an interactive data visualization dashboard to be published online.

Method. An open source space launch dataset was obtained and used to create visualizations. The data was processed using JavaScript and published as a dashboard using the Apache ECharts framework.

Results. The work resulted in visualizations that provides an insight into the history of spaceflight as they relate to historical events, shows the biggest national actors in space launches and the most common mission types as the primary motivating factor throughout history. It also provides information on the geographical location of these launches. The dashboard can be found at https://datavis.space, with source code and more at https://github.com/Victor4X/datavis-project

Conclusion. The resulting dashboard presented the data in a visual way that could answer the questions presented in the project objectives. Improvements to the current visualizations were suggested, as well as non-essential features.

Editorial

The table shows the responsibilities and contributions of the writing process.

Section	**Responsible**	**Contributes**	**Controlled by**
Frontpage	All	All	All members
Title	All	All	All members
Abstract	All	All	All members
Preface	All	All	All members
Readers guide	All	All	All members
Editorial	All	All	All members
Introduction	All	All	All members
Analysis	All	All	All members
Design	All	All	All members
Implementation	All	All	All members
Discussion	All	All	All members
Conclusion	All	All	All members
Bibliography	All	All	All members
Appendices	All	All	All members

Table 1 Responsibilities

Table of contents

1 Introduction

1.1 Background and motivation

On October 4, 1957, the Sputnik artificial satellite was successfully launched into orbit, marking the first time that mankind overcame the gravitational pull of earth and entered the space age. Ever since, countless objects, both manned and unmanned have been launched into space, the rate of which has been rapidly growing in recent years due to the rise of commercial spaceflight industry. Spaceflight has since become an integral part of the modern world as we rely on positioning systems, imaging, communications and a slew of other functions provided by satellites in orbit around the earth. Important scientific research is conducted at the International Space Station and incredibly intricate space telescopes further our understanding of the universe and the possibility of extraterrestrial life.
This topic is especially interesting and due to wealth of data available online, there was an opportunity to explore it in a visual and compelling way. Specifically, the evolution of spaceflight as it relates to historical and geopolitical events is explored, as these can only be understood in context. We also want to understand who the main actors behind these incredibly difficult and costly operations are, their motivations and the distribution of different launch or mission types. By this effort we hope to provide others with insights into this recent – and ongoing – era in history.

1.2 Project objectives

Building on the reasons and motivations outlined in the previous section, the goal of this project is to create a data visualization dashboard that utilizes space launch data. The dashboard will serve to answer the following questions:

1. What is the history of spaceflight?
2. Who are the biggest actors in the space race?
3. What are the primary motivating factors behind space travel?
4. What are the most preferred locations for space launches?
5. What is the correlation between launch location and mission type?

2 Data

The data set contains data of all governmental and commercial spacecraft launches since the beginning of human spaceflight until present day, including future planned flights. It has been compiled by a non-profit group of developers called The Space Devs. According to their website, *"[Space Devs] is a group of space enthusiast developers working on a range of services, united in a common goal to improve public knowledge and accessibility of spaceflight information.*

The dataset and accompanying API documentation can be found at the Space Devs' https://ll.thespacedevs.com/2.2.0/launch/.

2.1 Description

The data set is in JSON format and is composed of multiple nested objects which makes it easy to select specific parts to work with due everything being a key/value pair. The data set takes up 78 MB of storage containing 6721 entries with each entry containing in-depth details about the specific launch including launch service provider, rocket, mission and launch pad.

2.1.1 Relevant variables

`id:` `string` | Launch ID
`status:` `object` | Contains ID, Success status and description of launch
`net:` `string` | Datetime of launch
`launch_service_provider:` `object` | Contains ID, name and type of launch service provider.
`rocket:` `object` | Contains rocket ID and configuration information.
`mission:` `object` | Contains mission ID, name, description and orbit information.
`pad:` `object` | Contains launch pad ID, latitude and longitude coordinates and country code.

2.2 Data Processing

The data in the data set is mostly consistent and has the same types and shape. Therefore no substantial cleaning was needed, though the data set also contains data concerning future flights which in case of some of the visualizations is not desired. Therefore the future planned flights has been removed from the data set for given visualizations.

Furthermore due to many of the visualisations being related to the total time period, a modified data set has been made where flights are indexed based on the year. Because of the relative big size and depth of nested objects in the data set, smaller specified data sets has been extracted to make it faster to work with.

Some topics in the data set includes up to 19 different countries where some countries might be grouped. Additionally the countries can have a big difference in contributions making some of the countries vanish in some visualizations. To counter that the countries has been grouped into four major regions - Europe, USA, Russia and China.

The launch pads are represented individually with their dedicated coordinates. Often launch pads are clustered at specific global locations which can make some vanish when doing visualizations representing the world. Therefore the launch pads are grouped based on a radius of 500km which concatenate the clusters of launch pads in to single points of data, as in most cases launch pads within close proximity of each other are associated with the same organization or group of organizations.

3 Visualization/Dashboard

3.1 Design

3.1.1 Visualization library

We chose to host the visualizations of our data on a website for anyone to access, and decided that the easiest way to accomplish this, was by using a client-side JavaScript library. We ended up using Apache ECharts, as it looked very approachable and powerful. Before ending with this solution, we tried out Dash (part of Plotly), with Python as the programming language. We originally went with this setup because we all are comfortable working with data in Python, but ultimately found the resulting visualizations and user experience lacking.

3.1.2 Choosing the colors and gradients

Our dashboard has a dark background, as we found that the most fitting with the space theme. This choice does however mean that we have to carefully consider how we use color properly in our visualizations. We did not see this as an obstacle in choosing a good color scheme, but rather as an exercise in figuring out what is and isn't important when it comes to clarity. One of these exercises was figuring out how to best highlight the weight of each data point - in other words, large vs. small numbers. A good example of this is in our correlation matrix visualization of the different mission types vs. their orbits. In order to convey which overlaps were most common, we needed to have the higher numbers stand out more to the reader. We accomplished this by utilizing contrasting colors and lightness - when the background is darker, the values we want to jump out at the reader should be lighter. We later found many examples of this exact technique being used across the web (e.g. GitHub's activity tracker in Appendix A), and we feel that the resulting visualizations correctly highlight the important messages in the data.

3.2 Features

The most important features of the dashboard are the visualizations that help answer the questions outlined in the project objectives. In the following section the research questions are enumerated and the accompanying visualizations are demonstrated and explained, after which possible improvements and non-essential features are discussed.

3.2.1 What is the history of spaceflight?

For this question an animated line graph, (see fig 3.1), was used to visualize the amount of successful launches of each actor in the space race from the earliest records to the present day. In an attempt to add context to the graph, significant events in spaceflight history were added as markers to highlight the year of the event. Please note that these events are not represented in the dataset and are added at the groups discretion.

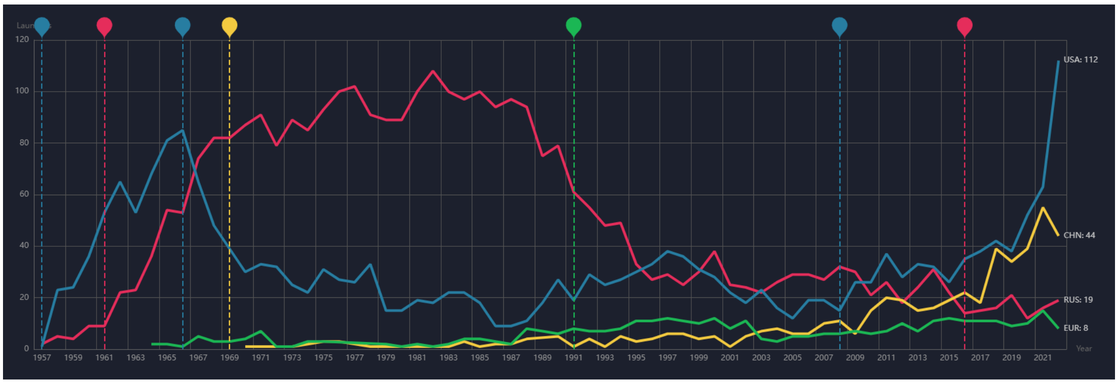

Figure 3.1 Line chart showing successful launches by nation from 1957-2021. European actors are grouped under 'EUR'

3.2.2 Who are the biggest actors in the space race?

According to the data, each actor can be labelled by their nationality or organization. Grouping actors by nationality is the most accurate choice for the Cold War era and onward, whereas grouping by organization only becomes more relevant in recent history. Therefore, an actor in this case is the nation state in which the launch was executed, regardless of state affiliation of the launch organization.
To determine the biggest actor the amount of launches, both in total and each year, was calculated for each actor. Two synchronized graphs are used to visualize the findings. The first is an animated bar chart, fig 3.2, that shows all successful launches of the actors each year. The other is an animated world map, fig 3.3, where color intensity indicates the cumulative launches within that country. With these graphs, both the yearly and total actor size can be visualized.

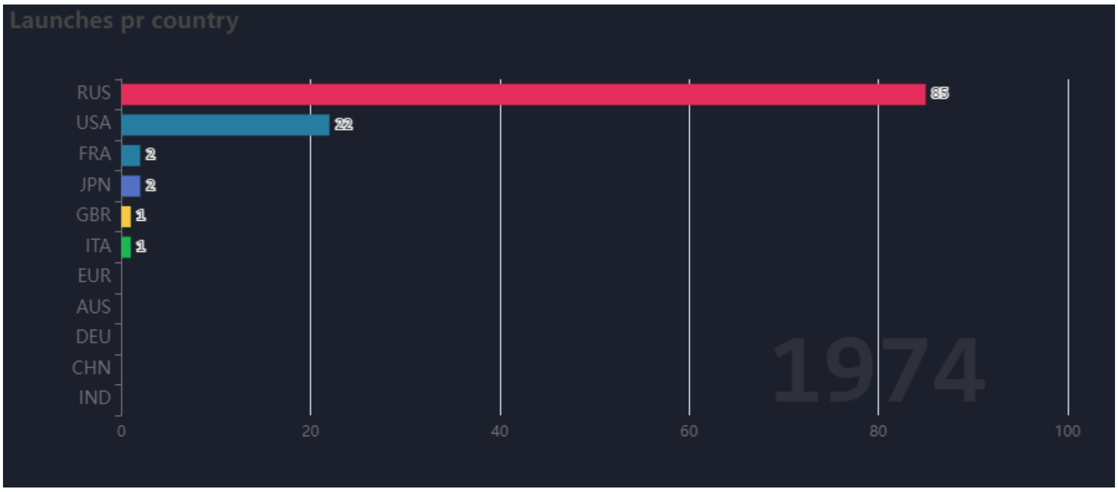

Figure 3.2 A still frame from the animated bar chart showing the distribution of launches by country

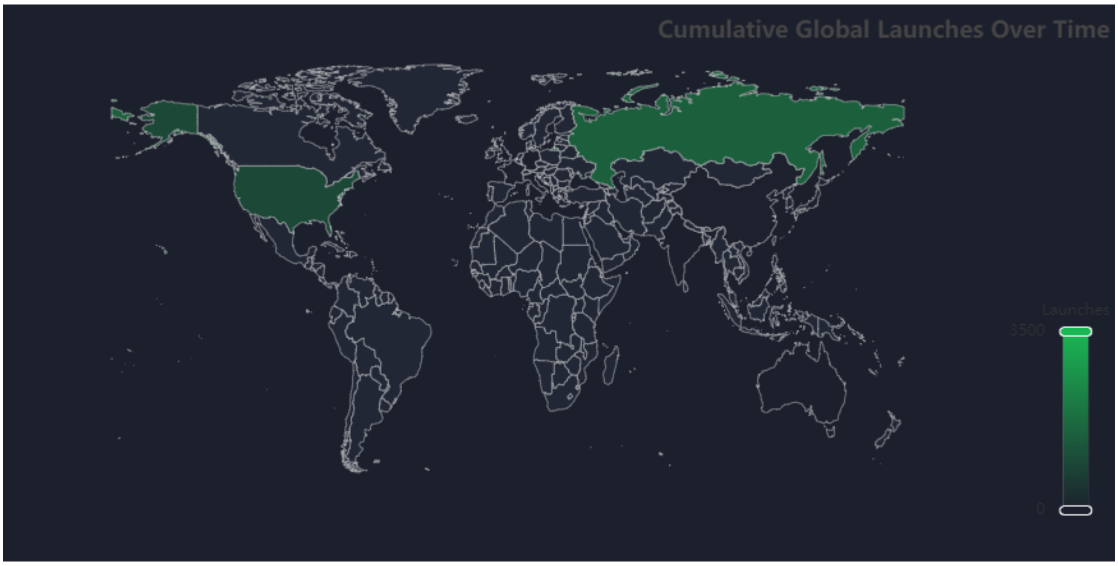

Figure 3.3 A still frame from the animated world map showing frequency of launches as a heatmap

3.2.3 What are the primary motivating factors behind space travel?

A way to explore the motivating factors behind spaceflight, the mission types are evaluated, as the most common missions may indicate the biggest motivations for space launch. To add more granularity, the mission types can be further grouped by the orbit types in which they operate. To visualize this, a correlation matrix, fig 3.4, with mission types correlated to different types of orbit is used. In addition, a bar chart of the total count of each orbit type is attached on the vertical axis and another showing the total count of each mission type is attached on the horizontal axis. The graph includes the option to normalize the distribution of launches either by mission type or orbit type. Orbit types that had < 25 launches were mostly extraterrestrial orbit missions and were grouped under 'Extraterrestrial and other' to make the visualization more concise.

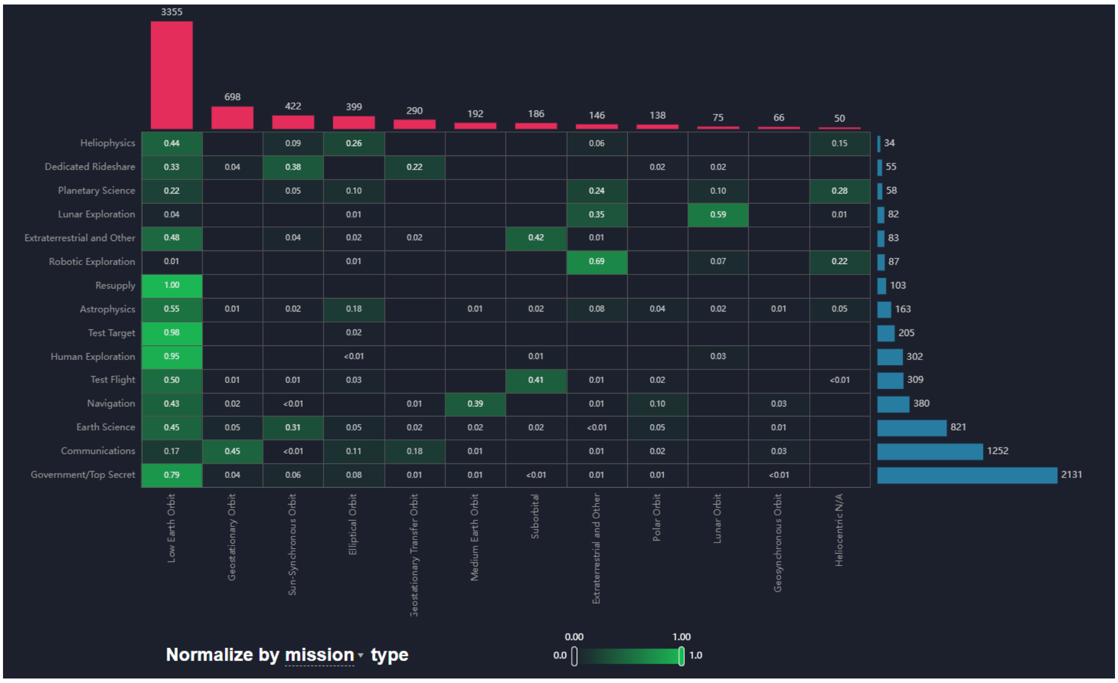

Figure 3.4 A correlation matrix of the mission types in relation to orbit types

Looking at motivating factors from a different angle, the second visualization addresses the historical market share of space travel, defining launch organization types as *Government, Commercial* or *Private*. This is visualized as a stacked area chart, fig 3.5 as this can help identifying the market share ratio over time. The chart is a time series with years on the x-axis and the total amount of launches for each year on the y-axis.

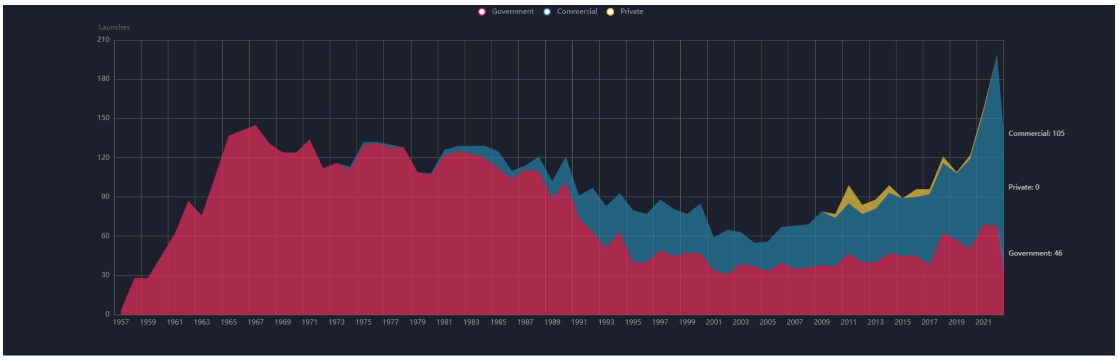

Figure 3.5 A stacked area chart showing the evolution of spaceflight market share

3.2.4 What are the most preferred locations for space launches?

We can approach this question by making a visualization of where all the different launch pads are located. This is done by making a bubble chart overlaid on a map of the world, fig 3.6. The data set contains launch pads that are located very close to each other, so to make the chart more clear and readable, we grouped pads in a 500km range. The bubbles represent the location of the different launch locations, with the size of the bubble being the total amount of launches from that location. When hovering each bubble, information about that particular location is presented including the total count of launches on that location and the most occurring mission type.

Figure 3.6 A bubble chart over world map, showing locations of launch pads

3.2.5 What is the correlation between launch location and mission type?

Trying to understand which pads were most commonly used for which mission types, we used the same bubble chart from the previous question, fig 3.6. The size of the bubble is the amount of launches from that location. Hovering on a bubble reveals is most common mission type and the numerical value of the bubble size. Attempting to gain additional insight into this question, an area chart is placed over a world map, fig 3.7. The visualization is made to investigate whether any relation between location of the launch pads and the mission type exists.

When hovering the areas the launch pad locations are revealed to substantiate the visualized areas.

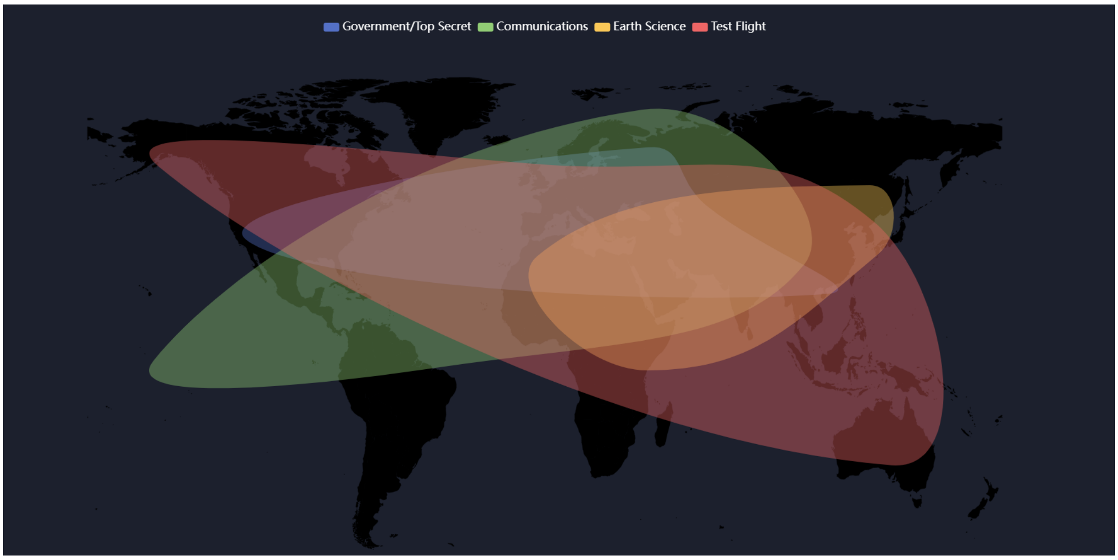

Figure 3.7 An area chart over a world map, showing the correlation between mission type and launch location

3.2.6 Rocket production

For the purpose of visualizing rocket capability evolution over time, we made a bubble chart comparing mass to LEO (low earth orbit) with the date of first flight for that rocket type, fig 3.8. Keep in mind that the y-axis is logarithmic. The chart can be set to show if the rocket type is still in active use or discontinued, in which country the rocket was produced and if the producer is government or corporation. The size of each bubble is the total amount of launches of the rocket type. Hovering over each bubble shows additional information about the vehicle.

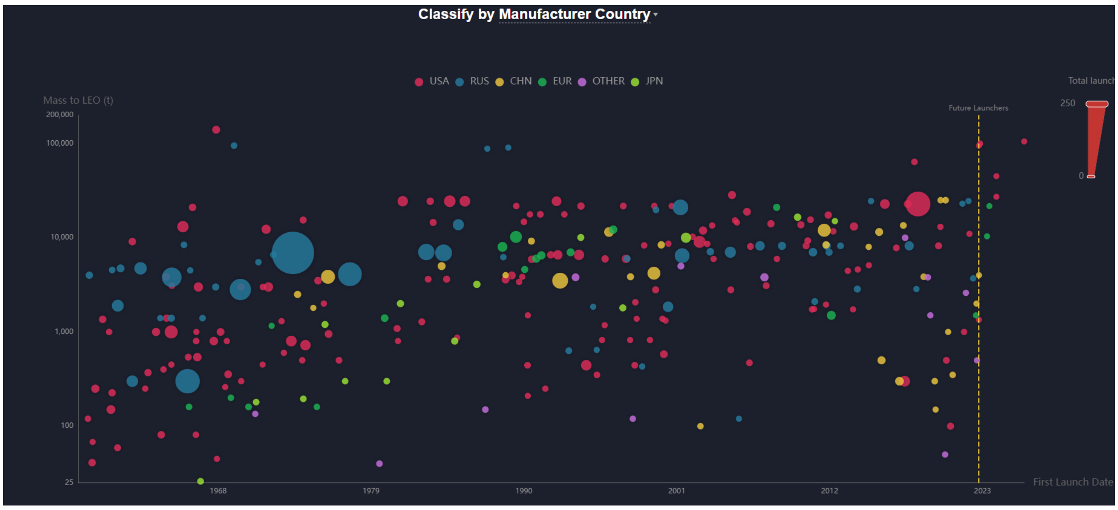

Figure 3.8 Rocket capability bubble chart, showing rocket production by country

3.2.7 Visualizing the rockets

The dashboard mainly concerns itself with visualizing the action of countries/corporations, the different missions and locations used for space flight, but every aspect visualized on the dashboard would not have been possible without the rockets to do the work. So to give viewers an idea of what arguably is the most important part of space flight history, the actual rockets. We want to giver viewers an idea of the scale and comparable size of the different rockets used for space flight. For this we uses a bar chart, fig 3.9, where each bar represents a different rocket, and the height of the bar represents the relative height of the rocket. To make the chart more interesting we used a picture of the rocket being represented as the bar. Only the 8 most commonly used rocket families were selected to be used on the chart, to keep it simple and clear. The diameter of the rocket is displayed above each bar. It is possible to hover over each rocket to get additional information such as *launch mass* and *low earth orbit capacity*.

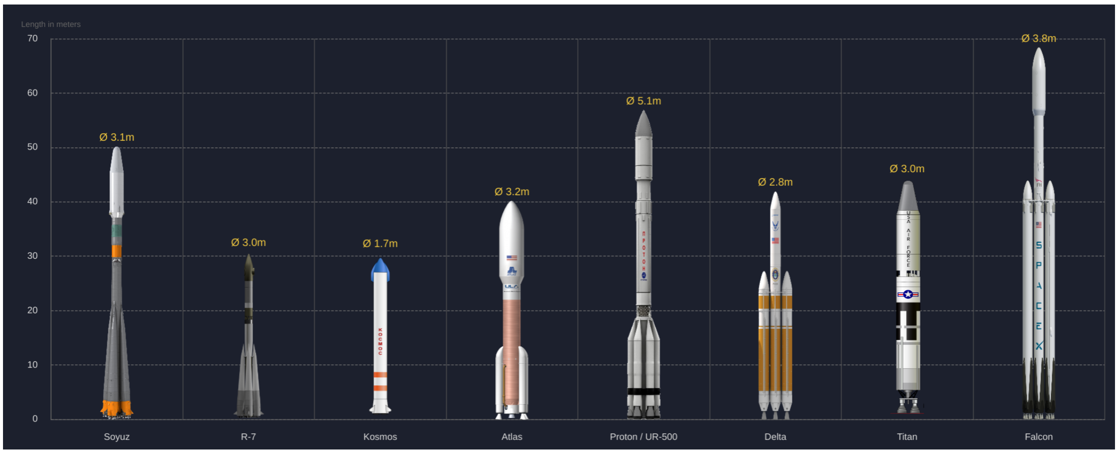

Figure 3.9 Visualization of 8 most used rocket families

All visualizations are integrated in a single page dashboard which can be found here: https://datavis.space
Demo video is available on YouTube: https://youtu.be_yBXvVh-Jrs
The dashboard has many interactive visualizations where the user can hover specific elements in the graph to see additional information or select specific parts of a graph to be shown.

3.2.8 Optional features

There are a few non-critical features that have not been implemented, but could provide additional value to the visualization dashboard. One such feature pertaining to fig. 3.1 would be the ability to toggle between actor types, where one would be the existing actors grouped by nationality and the other would group the actors by organization.

Distributing the visualizations between different pages could also help keep the messages tidy, and to group similar charts. The current dashboard works well, but this feature could improve usability if done right.
Another possible feature would be to synchronize all of the animations that visualize over years and add common playback controls, making it easier to compare different visualizations at specific points in time.
Additionally, with all the rocket information available in the data set, it would be possible to make improved hover animations that display further information on the rocket configurations, with images that would be displayed when hovering over a specific launch. Alternatively this information could be used to create a gallery view of the rockets in the data set.
All of these features would provide additional interactivity and improve the user experience.

4 Results

4.1 The story in the data

Examining the line chart, seen on fig. 3.1, the history of spaceflight can be characterized by three distinct periods of spaceflight. The first period is the space race, from the launch of Sputnik 1 in 1957, to the fall of the Soviet Union in 1991. The second period is a spaceflight "winter", where the interest in spaceflight was low, as evident by the drop in launch count. The latest period is a spaceflight renaissance, starting around 2016, where renewed interest in spaceflight can be witnessed as a result of increasing commercialization of the sector, with both the USA and China rising on the chart.

Looking into who the biggest actors were in the space race, we need to look at the bar chart, seen on fig. 3.2. Here we need to examining the the animation frames of the years, from 1957 to 1991. What we notice is the Soviet Union dominated during most of the space race, in terms of amount of successful launches, with the USA occasionally pulling ahead. Looking at the years after the space race, 1991 and onward, we notice the biggest actor changing more frequently.

At first glance one might only see the noble goal of the space race, a goal of advancing science and technology. However an underlying purpose of the space race was also one of dominance and power. While highly advanced propulsion- and guidance-systems are good at sending people to space, the very same technology can equally be used to develop Intercontinental Ballistic Missiles (ICBM), which would be useful should the cold war heat up. After the end of the space race we would then expect the goal of developing rocket technology would change. Looking at fig. 3.5, we can observe that early on it is only governments that are launching rockets, we can also notice that the amount of launches performed by governments started to decrease, as the cold war slowly came to an end. As the tensions between the 1. and 2. world slowly faded near the end of the cold war, we see the emergence of commercial spaceflight, from this we can say that the goal has shifted. The goal from the commercial aspect is to make spaceflight more cost effective and marketable.

Space flight was limited to two different countries in the beginning, but with time this gave way to a very global industry. From and fig. 3.6, we see what launches happen

all across the world, combine this with what we can observe from fig. 3.7, the different mission types can mostly be seen in every part of the world. A notable exception here is the concentration of *Earth Science* missions being launched from the eastern hemisphere. The marked areas has the most intensity around the equator which might be due to the faster rotation of the earth a that location.
By making the bubble chart, fig. 3.6, we are looking to find some pattern in the placement of the bubbles, and with that try to find out if there is a optimal location for space launches. Initial impression from looking at the chart, we might conclude that the most optimal location would be in Russia and Kazakhstan, but if we think about the whole context of the data, we realize that this is only the case because Russia, at the time Soviet Union, has a longer history of space flight than other places, like China.

Even though the space race was famously a very active period for spaceflight, the amount of launches in pre-1991 period compared to the 2000's was unexpected. Similarly, the data also revealed how it is only in very recent years that the record for most space launches a year was broken, a record originally made in the 1960's.
The correlation matrix, fig. 3.4, show that far most spacecrafts are launches to the lower earth orbit (LEO). This orbit is well suited for navigation satellites such enabling GPS due to the relative low altitude and possibility for satellites to orbit a wider range of routes. LEO is also the orbit used for the international space station (ISS), which also could explain why most of the missions are targeting the low earth orbit.
Communications satellites are launched into geostationary orbit instead of LEO, due to this orbit enabling the satellites to retain a sort of stationary location above earth.

When looking at the rocket capability bubble chart, fig. 3.8, it seems like the demand for higher capability launch vehicles plateaus in around 1980 with the introduction of the Space Shuttles, and a mass to LEO of around 24000 tonnes. There are heavier lift vehicles around, but they get used rarely. The trend in more recent years actually points to advantages in specializing the launch vehicle to lighter payloads, with an explosion of new launchers with LEO launch capabilities below 2000 tonnes.
We can also observe from fig. 3.8, that the most used rocket is by far the Soyuz U, with 714 launches dating from 1973 all the way to 2017, showing how capable the platform has been. It is notable that this rocket outlasted its original country of origin. Also worth noting, is the relative large size of the Falcon 9 Block 5, which already has been used for 129 launches, starting in 2018.

4.2 Improvements

In the initial iterations of the correlation matrix, no grouping was done on orbit types. This resulted in a slightly more granular view with many values, as seen in fig. 4.1. The increased granularity does not contribute to the insights intended with the visualization and were replaced for a more concise view with grouping.

Figure 4.1 Correlation matrix graph before grouping orbit types

Grouping the actors by organization resulted in a very cluttered graph after the end of the space race era, as many actors with few launches emerged alongside the nation-state actors. This development became very unclear on the line graph (fig. 3.1), even with highlighting and a zoomed-in view to better fit the scale, and the impression of the actual amount of launches per year became hidden within the multiple overlapping lines. This resulted in the current grouping of actors by nation.

5 Conclusion

The work done in completing the project objectives resulted in a online data visualization dashboard. This dashboard answers questions about space flight and its evolution throughout its history. Many of the visualizations are either animated or intractable, which provides more information on the given visualization.
A challenge posed from the data set was inconsistency in the data, with missing values for some fields, this would cause the visualizations to be wrong or not work. Because of this it took extra effort to make some visualizations and some visualization ideas were dropped.
Initially we started with Python for data processing and visualization, but later in the process the decision was made to switch to a JavaScript based approach for visualization. This did mean some of the already created visualizations had to be remade with the new JavaScript library.

From the data set used, a lot of interesting visualizations are possible, and the dashboard covers only a few. For some of the questions asked in for the projects objectives, it would be possible to get a more complete answers by using and comparing with additional data sets about the financial and political landscape throughout history.

To gain even more insight in the history of space flight and for more intractability with the visualization, it would be desirable to have a control panel, where a year could be selected. This would set all graphs to only visualize the data for that year or up til the year. This would make it easier to compere all the information from the visualizations in the same context.

Appendices

A - GitHub Activity Tracker

Figure A.1 Comparison between activity tracker on https://github.com/Victor4X in light mode and dark mode