

# National College of Ireland

BSHDS4

Data Science

2022/2023

Renato Gusani

x19411076

X19411076@student.ncirl.ie

[www.spaceexplorer.info](http://www.spaceexplorer.info)

## **A Comprehensive Time Series Analysis on 7 Decades of Space Travel Data**

Data Analysis Report

## Contents

Executive Summary .....	3
1.0 Introduction .....	3
1.1. Background.....	4
1.2. Aims.....	4
1.3. Technology .....	5
1.4. Structure.....	5
2.0 State of the Art .....	6
3.0 Data.....	6
4.0 Methodology .....	7
5.0 Analysis .....	8
5.1. Preliminary Analysis .....	9
5.2. Analysis – Topic 1 .....	11
5.3. Analysis – Topic 2 .....	14
5.4. Analysis – Topic 3 .....	18
5.5. Analysis – Topic 4 .....	20
5.6. Analysis – Topic 5 NLP .....	22
6.0 Results.....	24
6.1. Results – Preliminary Analysis Results.....	24
6.2.Results – Topic 1 .....	26
6.3.Results – Topic 2 .....	27
6.4.Results – Topic 3 .....	28
6.5.Results – Topic 4 .....	28
6.6.Results – Topic 5 NLP .....	29
7.0 Website Integration .....	32
7.1. Website Integration – APOD API .....	32
7.2. Website Integration – SpaceX Unofficial API.....	33
7.3.Website Integration – Analysis Plotly Charts.....	34
7.4.Website Integration – Adobe teaser trailer .....	35
7.0 Conclusions .....	36
8.0 Further Development or Research .....	36
9.0 References .....	37
10.0 Appendices .....	38
10.1.Project Proposal.....	38
1.0 Objectives .....	40
2.0 Background .....	40

3.0	State of the Art .....	41
4.0	Data.....	41
5.0	Methodology & Analysis.....	42
6.0	Technical Details .....	43
7.0	Project Plan.....	43
	10.2.Ethics Approval Application (only if required) .....	44
	10.3.Reflective Journals .....	45
	10.4.Invention Disclosure Form (Remove if not completed).....	51
	10.5.Other materials used .....	51

## Executive Summary

The purpose of this project is to look at data about space exploration and technological advancements in space travel. The study begins with a historical overview of space exploration, followed by an update on NLP LDA on rocket and missions. The report then describes the data used in the analysis, which includes information about previous and current space missions, as well as technology used for these missions.

This study will make use of the SpaceX API [1] and the NASA APOD API [2] in addition to the "Space Missions from 1957 to 2022" data set. I'll use a wrapper for the SpaceX API to get information on the company's most recent launches and missions. The NASA APOD API will have its own website page, displaying the "Astronomy Picture of the Day" with an explanation and image. These APIs will supplement the "Space Missions from 1957 to 2022" data set [3] by providing additional insights and actual data on space exploration and missions.

The procedure for this study entails using data pre-processing methods to clean and organize the data, as well as various statistical and machine learning methods to analyze the data. The analysis is concerned with detecting trends and patterns in the data, as well as investigating the relationships between various variables.

According to the analysis' findings, the number of space missions has increased significantly in recent years, as has the technology used for these missions. The analysis also reveals that certain variables, such as mission cost and number of countries involved in the mission, have strong relationships.

The report concludes by discussing the project's benefits and limitations, as well as future research recommendations. Overall, this report is informative about the current state of space exploration and technological advances used in space travel.

## 1.0 Introduction

This project will investigate time series in space missions, concentrating on the multitude of space missions launched each year, the nations and entities behind them, the different kinds of space missions, space mission success rates, the connection between the multitude of space missions and technological advancements, and the length of time of space missions over time. The dataset [3] "Space Missions from 1957 to 2022" provided by user MYSAR AHMAD BHAT was used for this project. The project will take 2 semester to complete and will be hosted on a Github Pages website.

In addition to MYSAR AHMAD BHAT data, this project will investigate space mission data from the SpaceX API [1] and the NASA APOD API [2]. The SpaceX API provides information about SpaceX missions and launches, including launch vehicle and payload details. The NASA APOD API provides data on NASA's Astronomy Picture of the Day.

For the website design, this project will use a Jekyll template and will be hosted on Github Pages. The website will have a separate page dedicated to the NASA APOD API, showcasing

NASA's stunning astronomy images and information. The website will provide a simple platform for exploring the data and gaining insights from the data analysis.

### 1.1. Background

For several years, space travel has been a hot topic, and studying space missions offers valuable information into technological advancements and reflect the diversity in this field. This project aims to investigate time series in space missions and to uncover new data insights. The goal is to comprehend the number of space missions launched each year and the nations or organizations that support them, the different types of space missions and their intensity over time, the success rates of space missions by nation or organization, the relationship between the number of space missions and technological advancements, and the duration of space missions over time.

This dataset's analysis will provide a complete overview of historical trends in space missions, making it a valuable source of information for anyone taking an interest in space exploration. This project will also serve as a final year BSc Data Science project, allowing students to apply the knowledge and skills acquired throughout the course to a real-world dataset.

### 1.2. Aims

The primary goal of this project is to examine historical trends in space missions using the dataset "Space Missions from 1957 to 2022" provided by user MYSAR AHMAD BHAT. The project's specific goals are as follows:

1. Investigate the number of space missions launched each year, as well as the countries or organizations that support them.
2. Analyze the types of space missions and the intensity with which they occur over time.
3. Compare space mission rates of success by nation or organization.
4. Examine the relationship between the number of space missions and technological advancements.
5. Create Interactive Dashboards to be hosted on the Github Pages site.

By achieving these goals, the project will provide a broad analysis of historical trends in space missions, as well as uncover new data insights. The project serves as a final year BSc Data Science project, allowing me, as a student to apply my skills and knowledge from the course to a real-world dataset.

Furthermore, by hosting the project on a Github Pages website, the goal is to make the findings of the study easily accessible to others. This will enable the project to be disclosed with others, and the results will be used to help inform space exploration studies.

In addition, this project will use a wrapper to access data from the Space X API and the NASA APOD API to provide a standalone page on the website. The project will also make use of a Jekyll template, which will make the website easy to maintain and customize. These new data sources will provide a comprehensive overview of space exploration and its trends, allowing for a more in-depth understanding of the subject.

### 1.3. Technology

The following technologies will be used to achieve the goals outlined in Section 1.2:

1. Python programming language: Python will serve as the primary language of programming for cleaning, exploring, and analyzing data. The Python libraries listed below will be used:
  - Pandas is a data manipulation and cleaning tool.
  - numpy is used for numerical computation.
  - For visual analytics, use matplotlib and seaborn.
  - Machine learning with scikit-learn
2. Jupyter Notebook: The development environment for writing and testing code will be Jupyter Notebook. Jupyter Notebook is a free and open-source web application for interactive coding, visualizations, and markdown documentation.
3. Github will be used for version control as well as hosting the project on Github Pages. This makes the project more accessible to others and facilitates collaboration.

By using pandas library, the data will then be loaded into a pandas dataframe for cleanup and discovery. To understand the distribution of information and recognize patterns, the data will be visualized using matplotlib and seaborn. To unearth relationships and patterns in the data, machine learning methods will be applied to the dataset using the scikit-learn library. Eventually, the analysis results will be displayed in a Jupyter Notebook, along with the code that was used to produce them.

### 1.4. Structure

The following is how this document is organized:

1.0 Introduction: This section will provide background information on the project, such as why it is being undertaken, the project's goals, and the technology that will be used.

2.0 State of the Art: This section will examine published research on similar research, highlighting what has originally been done and how this project will differ.

3.0 Data: This section will explain the dataset to be used, including the data source and any preprocessing that will be performed.

4.0 Methodology: This section will include a detailed description of the data analysis methods and techniques, including any models for machine learning that will be used.

5.0 Analysis: The outcomes of the data analysis will be presented in this section, including any visual representations or tables used to show the results.

6.0 Results: This section will review the analysis's main findings, emphasizing any key insights or connections discovered.

7.0 Conclusions: This section would provide a project summary, highlight important takeaways and addressing any study limitations.

8.0 Future Development or Research: This section will go over any possible future research or development areas that were recognized during the project.

9.0 References: A list of all sources cited in the document will be provided in this section.

10.0 Appendices: This section will include project-related materials such as the project proposal, ethics approval application (if required), reflective journals, invention disclosure form (if completed), and any other materials used during the project.

## 2.0 State of the Art

Prior studies on space missions has concentrated on a variety of topics, including the expense of space exploration, the implications of space missions on growth in the economy, and historical trends in space travel. Here are some examples of similar work:

1. The European Space Agency's "The Economic Impact of the Space Industry" examines the economic impact of the space industry on the European Union and analyzes the economic benefits of space-related activities.
2. The Planetary Society's "A Comparative Analysis of Space Programs" compares the space programs of foreign nations, including the United States, Russia, China, and India.
3. The National Aeronautics and Space Administration's (NASA) "Trends in Space Exploration" provides an overview of historical trends in space exploration as well as the future of space exploration.

While these studies have provided useful information about the space industry, they have not specifically examined historical trends in space missions. Furthermore, these studies were not based on a specific dataset and did not employ machine learning techniques to analyze the data.

This project will be distinct from others in that it will concentrate on analyzing historical trends in space missions using a dataset provided by user MYSAR AHMAD BHAT. Machine learning techniques will also be used in the project to uncover relationships and trends in the data. The project will also end up making the analysis's results easily accessible to others by hosting it on a Github Pages website, and it will be thoroughly documented throughout the process , allowing others to easily understand the methodology and results.

## 3.0 Data

The data set used in this project is MYSAR AHMAD BHAT's "Space Missions from 1957 to 2022," which was UPDATED 6 MONTHS AGO on Kaggle. The data set includes details on space missions launched between 1957 and 2022, such as the mission's name, launch date, launch vehicle, and mission status.

MYSAR AHMAD BHAT accumulated the data from various publicly available sources, including NASA and Wikipedia. Over 800 space missions launched by various countries, including the United States, Russia, China, and Europe, are included in the data set.

To better understand the structure and content of the data set, exploratory data analysis was performed. This included counting the number of missing values and looking for outliers

or unusual observations. I also checked for any inconsistencies in the data types and values of the columns. This helped me understand the data's quality and completeness, as well as identify any data cleaning or preprocessing that was required.

I also created visualizations such as a bar chart, a line chart, and a pie chart to better understand the number of missions launched by different countries and the number of missions launched in different years. Based on the data, I also made some observations, such as which countries are leading in space missions and in terms of space launches, which years saw the most activity. This preliminary analysis provided me with a good understanding of the data and aided in the identification of any potential issues or areas of interest for further investigation.

In addition to the "Space Missions from 1957 to 2022" data set, I will use the SpaceX API and the NASA APOD API. I'll use a wrapper for the SpaceX API to get information on the company's most recent launches and missions. The NASA APOD API will have its own website page, displaying the "Astronomy Picture of the Day" with a description and image. These APIs will supplement the "Space Missions from 1957 to 2022" data set by providing additional insights and actual data on space exploration and missions.

## 4.0 Methodology

To guide my analysis of the space mission dataset for the methodology section, I will use the KDD (Knowledge Discovery in Databases) process. The KDD [4] procedure includes the following steps:

1. **Data comprehension:** In this step, I will become acquainted with the dataset by investigating the variables, their types, and the distribution of their values. I will also look for any missing or duplicate data, as well as any outliers or inconsistencies in the data.
2. **Data preparation:** In this step, I will clean, transform, and normalize the dataset if any pre-processing is required. Filling in missing values, encoding categorical variables, and scaling numeric values are all examples of this.
3. **Data exploration:** In this step, I will start exploring the dataset using various visual representations and statistical methods to identify any trends or relationships between the data. Generating histograms, scatter plots, and heatmaps to visualize data distribution, as well as employing correlation and regression analysis to identify any linear connections between variables, may be included.
4. **Data modeling:** In this step, I'll utilize NLP machine learning techniques to create models that can anticipate or categorize data. This may include training and evaluating models using techniques such as linear regression, decision trees, or LDA
5. **Evaluation:** In this step, I will assess the models' performance and compare the results to determine the best model for the dataset.
6. **Implementation:** In this step, I will use the finest model to make forecasts on new data and put it into production.

Throughout the KDD process, I will record any insights, assumptions, and decisions made to keep track of the analysis and make it simple for others to replicate.



## 5.0 Analysis

A variety of topics pertaining to space missions will be covered in this section. The preliminary analysis identifies five main areas of focus, including the number of space missions launched each year and the nations or organizations that support them (Topic 1); the types of space missions and their frequency over time; the success rates of space missions by country or organization; the relationship between the number of space missions and technological advancements; and NLP identification of the underlying themes associated with each topic. To provide a thorough overview of the space industry and its evolution, each of these topics will be thoroughly examined.

I combined statistics from the unofficial SpaceX API with the basic analysis themes already presented on my website to further my investigation. I was able to get useful information regarding SpaceX launches and their success rates, as well as other mission-related specifics, with the use of this API. Using this information, I was able to better comprehend the trends and patterns of space missions carried out by a private space corporation, a relatively new participant in the space market. Incorporating the data from the SpaceX API into my study not only gives me a thorough perspective of the space sector as a whole, but it also enables me to compare the advantages and disadvantages of government-run space agencies and private space enterprises.

Also included in my research were data from the NASA Astronomy Picture of the Day (APOD) API. The APOD API offers images and descriptions of celestial phenomena, events, and other space-related subjects every day. I was able to comprehend how the space industry is related to astronomy and space research by seeing the photos and descriptions the API offered. Using this information, I was able to investigate how the space business influences scientific research and discoveries and advances our understanding of the cosmos and its mysteries.

As a whole, my examination of the space business covers a wide variety of subjects, including the frequency of space missions launched annually and the connection between this frequency and technological improvements. My grasp of the space business and its effects on society has improved since I started combining information from the unofficial SpaceX API and data from the NASA APOD API. By examining these data sources, I intend to add to the existing discussion regarding the direction of space travel and stimulate more study and investigation in this fascinating area.

## 5.1. Preliminary Analysis

I import the essential libraries, including pandas, numpy, and plotly, to begin the preliminary data analysis. I then use the `head()` method of the pandas DataFrame to preview the first five rows after loading the data from the "space\_missions.csv" file [Figure 1](#). The DataFrame's `dtypes` attribute allows me to check the data types of each column, and the `print()` function allows me to print the results to the console. I can see that some columns have missing values, so I use the DataFrame's `isnull()` method to test if any columns have missing values; this method returns a boolean DataFrame. Next, I use the `sum()` method to determine how many missing values are present in each column.

```
import pandas as pd
import numpy as np
import plotly.express as px
import pandas as pd
import plotly.io as pio
import os

# Load the dataset into a pandas DataFrame
df = pd.read_csv("space_missions.csv", encoding='latin1')

# Preview the first 5 rows of the dataframe
print(df.head())

# Checking the data types of each column
print("-- DATA TYPES --")
print(df.dtypes)

print("-- MISSING VALUES --")
# Checking for missing values
print(df.isnull().sum())

print("-- STATISTICS --")
# Getting some basic statistics
print(df.describe())
```

Figure 1 - Basic Stats & Cleaning

Moving on, I use the `describe()` method to obtain some fundamental statistics about the DataFrame. I then use the `value_counts()` method to group the data by rocket status and count the number of launches for each status. I use the Plotly Express library's `px.bar()` function to produce a bar chart [Figure 2](#) to show the rocket status distribution, and I alter the backdrop color, add a title, and change the chart's color scheme to give it a unique look. In a similar manner, I divide the information into mission status groups and tally the number of missions for each status. Using the `px.bar()` method, I make another bar chart to show the mission status distribution, and I tweak it to go with the dashboard's general design.

```
# Grouping data by rocket status and counting the number of launches for each status
status_counts = df['RocketStatus'].value_counts().reset_index()
status_counts.columns = ['status', 'count']

# Creating a bar chart to visualize rocket status distribution
colors = ['#FFB6C1', '#98FB98'] # Light Red for "Retired", Light Green for "Active"
fig1 = px.bar(status_counts, x='status', y='count', color='status', color_discrete_sequence=colors,
             title='Rocket Status Distribution',
             labels={'status': 'Status', 'count': 'Number of Launches'},
             width=800, height=600)
fig1.update_layout(template='plotly_dark')
fig1.update_traces(marker_line_color='#FFFFFF', marker_line_width=1, opacity=0.8)
fig1.update_xaxes(showgrid=False, title_font=dict(color='#FFFFFF'), tickfont=dict(color='#FFFFFF'))
fig1.update_yaxes(showgrid=False, title_font=dict(color='#FFFFFF'), tickfont=dict(color='#FFFFFF'))
fig1.update_layout(plot_bgcolor='#121212', paper_bgcolor='#121212', font_color='#FFFFFF')
fig1.show()

# Grouping data by mission status and counting the number of missions for each status
status_counts = df['MissionStatus'].value_counts().reset_index()
status_counts.columns = ['status', 'count']

# Creating a bar chart to visualize mission status distribution
colors = ['#F08080', '#98FB98', '#87CEFA']
fig2 = px.bar(status_counts, x='status', y='count', color='status', color_discrete_sequence=colors,
             title='Mission Status Distribution',
             labels={'status': 'Mission Status', 'count': 'Number of Missions'},
             width=800, height=600)
fig2.update_layout(template='plotly_dark',
                 plot_bgcolor='#121212',
                 paper_bgcolor='#121212',
                 font_color='#FFFFFF')
fig2.update_traces(marker_line_color='#FFFFFF', marker_line_width=1, opacity=0.8)
fig2.update_xaxes(showgrid=False, title_font=dict(color='#FFFFFF'), tickfont=dict(color='#FFFFFF'))
fig2.update_yaxes(showgrid=False, title_font=dict(color='#FFFFFF'), tickfont=dict(color='#FFFFFF'))
fig2.show()
```

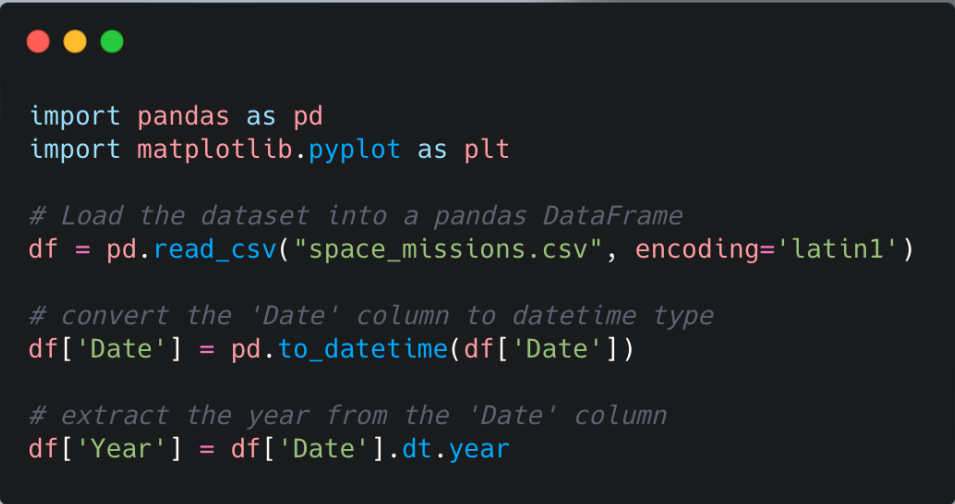
Figure 2 - Preliminary Plotly Charts

## 5.2. Analysis – Topic 1

**Topic 1:** The number of space missions launched each year, as well as the nations or organizations that support them

In this topic, two interactive bar charts depicting data about space missions are made using Python and the Plotly module. The number of missions launched annually is depicted in the first figure, while the number of missions launched by the top 10 businesses is depicted in the second.

The code first imports a number of libraries, including `plotly.express`, `matplotlib.pyplot`, and `pandas`. Using `pd.read_csv`, the dataset is loaded into a pandas DataFrame, and `pd.to_datetime` is used to convert the 'Date' column to datetime format [Figure 3](#).



```
import pandas as pd
import matplotlib.pyplot as plt

# Load the dataset into a pandas DataFrame
df = pd.read_csv("space_missions.csv", encoding='latin1')

# convert the 'Date' column to datetime type
df['Date'] = pd.to_datetime(df['Date'])

# extract the year from the 'Date' column
df['Year'] = df['Date'].dt.year
```

Figure 3 - Libraries Import

The code then generates a new column called "Year" that uses the `dt.year` attribute to extract the year from the "Date" column. The `count()` function is then used to tally the number of missions launched each year once the DataFrame has been grouped by year. The first Plotly bar chart [Figure 4](#) is made using this data with `px.bar`.

By classifying the DataFrame by firm and using the `count()` function to tally the number of missions launched by each organization, the second graphic is produced. The second Plotly bar chart [Figure 5](#) is made using `px.bar` and the data from the top 10 firms is chosen using the `head()` function.

```

# group the data by year, and count the number of missions launched each year
missions_per_year = df.groupby('Year')['Mission'].count()

import plotly.express as px

# Create a Plotly bar chart
fig3 = px.bar(
    missions_per_year,
    x=missions_per_year.index,
    y=missions_per_year.values,
    labels={
        "x": "Year",
        "y": "Number of Missions",
        "color": "Mission Status"
    },
    color=missions_per_year.index,
    color_continuous_scale=px.colors.sequential.Plasma,
    title="Number of Missions Launched per Year"
)

# Customize the layout
fig3.update_layout(
    title={
        "text": "Number of Missions Launched per Year",
        "font": {
            "size": 28,
            "color": "#FFFFFF"
        },
        "x": 0.5
    },
    font=dict(size=14, color="#FFFFFF"),
    plot_bgcolor="#121212",
    paper_bgcolor="#121212",
    xaxis=dict(
        tickmode="linear",
        tick0=1960,
        dtick=5,
        tickangle=45,
        title=dict(
            text="Year",
            font=dict(size=18, color="#FFFFFF")
        ),
        tickfont=dict(size=14, color="#FFFFFF")
    ),
    yaxis=dict(
        title=dict(
            text="Number of Missions",
            font=dict(size=18, color="#FFFFFF")
        ),
        title_font=dict(size=18, color="#FFFFFF"),
        tickfont=dict(size=14, color="#FFFFFF")
    ),
    hoverlabel=dict(
        bgcolor="#FFFFFF",
        font_size=16,
        font_family="Roboto"
    ),
    hovermode="x",
    transition=dict(
        duration=500
    )
)

# Display the dashboard
fig3.show()

```

Figure 4 - Plotly Chart

```

#group the data by company and count the number of missions launched by each company
missions_per_company = df.groupby('Company')['Mission'].count()

missions_per_company = df.groupby('Company')['Mission'].count().sort_values(ascending=False)

top_10_companies = missions_per_company.head(10)

fig4 = px.bar(top_10_companies, x=top_10_companies.index, y=top_10_companies.values,
color=top_10_companies.values,
color_continuous_scale=px.colors.sequential.Plasma_r, labels={'x':'Company', 'y':'Number of
Missions'})
fig4.update_layout(title={'text':'Number of Missions Launched by Top 10 Companies', 'x':0.5},
axis_tickangle=-45, height=500,
plot_bgcolor='rgba(0,0,0,0)', paper_bgcolor='rgba(0,0,0,0)', font_color='#FFFFFF',
yaxis=dict(showgrid=False, title_font=dict(color='#FFFFFF'),
tickfont=dict(color='#FFFFFF')),
xaxis=dict(title_font=dict(color='#FFFFFF'), tickfont=dict(color='#FFFFFF')))
fig4.show()

```

Figure 5 - Second Plotly Chart

The `update_layout()` function is used to modify the title, font, color, and background of both charts. The `show()` function is used to display the charts.

Overall, this code offers a simple and direct method for using interactive bar charts to visualize data regarding space missions.

### 5.3. Analysis – Topic 2

**Topic 2:** Analysis of the types of space missions and their frequency over time

Using Pandas' "groupby()" function, I first grouped the data by mission and launch date. As a result, a brand-new dataframe called "missions\_per\_type\_time" was constructed, which displayed the quantity of launches for each mission type per year. To do this, I gave the "groupby()" function a list of the names of the two columns ('Mission' and 'Year') [Figure 6](#).



```
# grouping the data by mission and launch date
missions_per_type_time = df.groupby(['Mission', 'Year']).size().reset_index(name='Counts')
```

Figure 6 - Grouping Data

The next step was utilizing the "groupby()" and "sort\_values()" routines to choose the top 10 most common missions. In order to identify the most prevalent mission types in the data set, this was done [Figure 7](#). The index was reset after I gave the generated dataframe the name "top\_missions".



```
# selecting the top 10 most frequent missions
top_missions = missions_per_type_time.groupby(['Mission']).sum().sort_values(by='Counts',
ascending=False).head(10)
top_missions = top_missions.reset_index()
```

Figure 7 - Top 10 Missions

I used the "isin()" function to verify that each value in the "Mission" column of "missions\_per\_type\_time" was present in the "Mission" column of "top\_missions" in order to filter the original dataframe to only display information for the top 10 most frequent missions [Figure 8](#). As a result, I was able to establish a brand-new dataframe called "missions\_per\_type\_time" that contained information solely for the top 10 most common missions.



```
# filtering the original dataframe to keep only the top 10 most frequent missions
missions_per_type_time =
missions_per_type_time[missions_per_type_time['Mission'].isin(top_missions['Mission'])]
```

Figure 8 - Top 10 Most Frequent Missions

I used the "px.bar()" function and the data input "missions\_per\_type\_time" to generate an interactive Plotly bar chart [Figure 9](#). I chose 'Year' for the x-axis, 'Counts' for the y-axis, and 'Mission' for the color. I also changed the legend's labels and added a title to the x- and y-axes.

```
# creating an interactive Plotly bar chart
fig5 = px.bar(
    missions_per_type_time,
    x='Year',
    y='Counts',
    color='Mission',
    title='Number of Missions by Type over Time',
    labels={
        'Year': 'Year',
        'Counts': 'Number of Missions',
        'Mission': 'Mission Type'
    },
    hover_data=['Mission'],
    height=500
)
```

Figure 9 - Missions Over Time Plotly Chart



I used the "update\_layout()" function to set the plot backdrop color, paper background color, font color, and other aspects of the chart in order to further personalize the layout [Figure 10](#). For the hover labels, I also changed the font size, family, and background color.

```
# customizing the layout
fig5.update_layout(
    plot_bgcolor='#1D2126',
    paper_bgcolor='#1D2126',
    font_color='#FFFFFF',
    xaxis=dict(
        tickangle=45,
        title=dict(
            text='Year',
            font=dict(size=18, color='#FFFFFF')
        ),
        tickfont=dict(size=14, color='#FFFFFF'),
    ),
    yaxis=dict(
        title=dict(
            text='Number of Missions',
            font=dict(size=18, color='#FFFFFF')
        ),
        tickfont=dict(size=14, color='#FFFFFF'),
    ),
    legend=dict(
        title=dict(
            text='Mission Type',
            font=dict(size=18, color='#FFFFFF')
        ),
        font=dict(size=14, color='#FFFFFF'),
    ),
    hoverlabel=dict(
        bgcolor='#FFFFFF',
        font_size=16,
        font_family='Roboto'
    ),
    transition=dict(
        duration=500
    )
)

fig5.update_layout(
    plot_bgcolor='rgba(0,0,0,0)',
    paper_bgcolor='rgba(0,0,0,0)',
    font_color='#FFFFFF'
)
```

Figure 10 - Layout

Finally, I displayed the dashboard by using the "show()" function [Figure 11](#). The number of missions for each mission type over time is displayed in an interactive Plotly chart that was generated as a result. Hovering the mouse over each bar in the chart reveals the precise number of missions for each year and mission type.

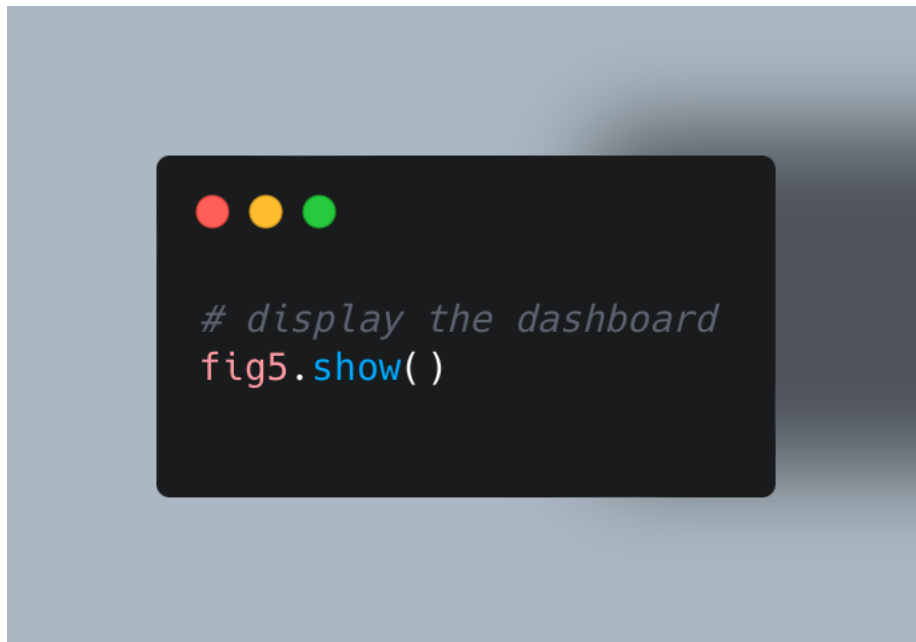


Figure 11 – Plotly Show.

## 5.4. Analysis – Topic 3

**Topic 3:** Comparing the success rates of space missions by country or organization.

To start, I add a new column called "Success" to the table [Figure 12](#). This column will have the value 1 if the mission was successful and 0 otherwise. I then use the "groupby" function to group the data by company, and I put the resulting grouped object in the variable "grouped". I use the "mean" method to take the mean of each group's "Success" column in order to determine each company's success rate. The variable "success\_rate" contains the series that is produced.

```
import plotly.graph_objs as go

# Creating a new column to indicate whether each mission was successful or not
df['Success'] = df['MissionStatus'].apply(lambda x: 1 if x == 'Success' else 0)

# Grouping the data by company
grouped = df.groupby('Company')

# Calculating the success rate for each company
success_rate = grouped['Success'].mean()

# Sorting the companies by their success rate
success_rate = success_rate.sort_values(ascending=False)
```

Figure 12 - success\_rate

The companies are then sorted using the "sort\_values" method in descending order according to their success rates, and the resulting series is kept in the "success\_rate" variable. I use the "nlargest" method to choose the top 20 businesses based on success rate, and I save the resulting series in the variable "top\_20\_companies" [Figure 13](#).

```
# Selecting the top 20 companies by success rate
top_20_companies = success_rate.nlargest(20)
```

Figure 13 - sort\_values

I use Plotly's "go.Bar" class to make a horizontal bar chart [Figure 14](#) to show the success percentage by company. The success rate values from the "top\_20\_companies" series are shown on the chart's x-axis, while the names of the respective businesses are shown on the y-axis. The "marker" option of the "go.Bar" class is used to give the bars a blue hue. I alter the chart's design to go with the dashboard's dark background.

```

# Create a Plotly horizontal bar chart
fig6 = go.Figure(
    go.Bar(
        x=top_20_companies.values,
        y=top_20_companies.index,
        orientation="h",
        marker=dict(color="#00b3b3"),
        name="Success Rate",
    )
)

# Customize the layout
fig6.update_layout(
    plot_bgcolor='#1D2126',
    paper_bgcolor='#1D2126',
    font_color='#FFFFFF',
    title=dict(
        text='Success Rate by Company',
        font=dict(size=24, color='#FFFFFF')
    ),
    xaxis=dict(
        title=dict(
            text='Success Rate',
            font=dict(size=18, color='#FFFFFF')
        ),
        tickfont=dict(size=14, color='#FFFFFF'),
    ),
    yaxis=dict(
        title=dict(
            text='Company',
            font=dict(size=18, color='#FFFFFF')
        ),
        tickfont=dict(size=14, color='#FFFFFF'),
    ),
    hoverlabel=dict(
        bgcolor='#FFFFFF',
        font_size=16,
        font_family='Roboto'
    ),
    transition=dict(
        duration=500
    )
)

# Set the background color of the bars to match the background theme
fig6.update_traces(marker=dict(color='#00b3b3'))

# Display the dashboard
fig6.show()

```

Figure 14 - Plotly Chart

The dashboard is then shown using the "show" method after updating the bars' background color to match the dark theme using the "update\_traces" function.

## 5.5. Analysis – Topic 4

**Topic 4:** The relationship between the number of space missions and technological advancements.

I'm using Plotly in this piece of the code to make a line chart that displays the quantity of space missions over time [Figure 15](#). I initially sorted the data by year and counted the number of missions that took place in each year to achieve this. The resulting information was put into a brand-new DataFrame called "missions\_per\_year," which has the columns "Year" and "Counts."

```
import plotly.graph_objs as go

# Group the data by year and count the number of missions
missions_per_year = df.groupby('Year').size().reset_index(name='Counts')
```

Figure 15 - missions\_per\_year

Then, I made a figure object called "fig7" with a line chart using the Plotly library. The chart's y-axis displays the number of missions that took place in each year, while the x-axis displays the years [Figure 16](#). I made the line chart using the "Scatter" function and "#c90000" as the line color. I named the line "Number of Missions" as well, and this name will appear in the chart's legend.


```
# Create a Plotly line chart
fig7 = go.Figure(
    go.Scatter(
        x=missions_per_year["Year"],
        y=missions_per_year["Counts"],
        mode="lines",
        marker=dict(color="#c90000"),
        name="Number of Missions",
    )
)
```

Figure 16 - Scatter

I used the "update\_layout" method to modify the figure's layout after making the chart. I used "#1D2126" for the chart's backdrop and "#FFFFFF" for the font color. The graphic [Figure 17](#) also has the caption "Number of Missions over Time" and labels on the x- and y-axes.

I added a hover label that shows the year and the number of missions for each data point so that visitors can see more details about each data point on the chart. The hover label will only be visible when the user hovers over the x-axis because I also set the "hovermode" option to "x".

Finally, I displayed the chart in the dashboard using the "show" method.



```
# Customize the layout
fig7.update_layout(
    plot_bgcolor='#1D2126',
    paper_bgcolor='#1D2126',
    font_color='#FFFFFF',
    title=dict(
        text='Number of Missions over Time',
        font=dict(size=24, color='#FFFFFF')
    ),
    xaxis=dict(
        title=dict(
            text='Year',
            font=dict(size=18, color='#FFFFFF')
        ),
        tickfont=dict(size=14, color='#FFFFFF'),
    ),
    yaxis=dict(
        title=dict(
            text='Number of Missions',
            font=dict(size=18, color='#FFFFFF')
        ),
        tickfont=dict(size=14, color='#FFFFFF'),
    ),
    hoverlabel=dict(
        bgcolor='#FFFFFF',
        font_size=16,
        font_family='Roboto'
    ),
    hovermode="x",
    transition=dict(
        duration=500
    )
)

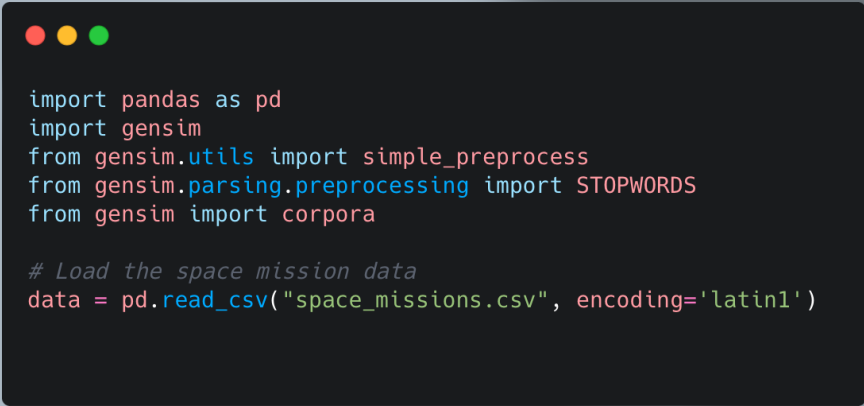
# Display the dashboard
fig7.show()
```

Figure 17 - Update Layout

## 5.6. Analysis – Topic 5 NLP

**Topic 5:** NLP Identification of the underlying themes associated with mission and rocket names using Latent Dirichlet Allocation (LDA).

I start this topic by loading the required libraries, such as pandas and gensim, which are used for text analysis and data manipulation, respectively [Figure 18](#). The information about space missions is then loaded into a pandas DataFrame from a CSV file.




```
import pandas as pd
import gensim
from gensim.utils import simple_preprocess
from gensim.parsing.preprocessing import STOPWORDS
from gensim import corpora

# Load the space mission data
data = pd.read_csv("space_missions.csv", encoding='latin1')
```

Figure 18 - Libraries

I divide the mission names and rocket names into separate lists before preprocessing the text data with a custom function that eliminates stop words and tokens with less than four characters to get it ready for analysis [Figure 19](#). I build a dictionary of uncommon words and the frequency with which they appear in the preprocessed text using the gensim package. For rocket names and mission names, I do this individually.



```
# Create a list of mission names and rocket names
mission_names = list(data['Mission'])
rocket_names = list(data['Rocket'])

# Preprocess the text data
def preprocess_text(text):
    result = []
    for token in simple_preprocess(text):
        if token not in STOPWORDS and len(token) > 3:
            result.append(token)
    return result

processed_mission_names = [preprocess_text(text) for text in mission_names]
processed_rocket_names = [preprocess_text(text) for text in rocket_names]
```

Figure 19 - Preprocessing

I then use the dictionaries we built earlier to create a document-term matrix for each of the preprocessed lists [Figure 20](#). The topic modeling method needs the document-term matrix as an input since it shows the frequency of each word in each document (such as the mission or rocket name).

```
# Create a dictionary of all the unique words in the text data and their frequency of occurrence
dictionary_mission_names = corpora.Dictionary(processed_mission_names)
dictionary_rocket_names = corpora.Dictionary(processed_rocket_names)

# Create a document-term matrix
doc_term_matrix_mission_names = [dictionary_mission_names.doc2bow(doc) for doc in
processed_mission_names]
doc_term_matrix_rocket_names = [dictionary_rocket_names.doc2bow(doc) for doc in processed_rocket_names]
```

Figure 20 - document-term matrix

Finally, I apply Latent Dirichlet Allocation (LDA) in [Figure 21](#), a well-liked topic modelling approach in natural language processing. For each of the pre-processed lists, I identify the top five themes using the LDA algorithm from the gensim library. I print out a list of the phrases that best describe each topic to serve as a representation of the topics on the console.

```
# Perform topic modeling using Latent Dirichlet Allocation (LDA)
num_topics = 5
lda_model_mission_names = gensim.models.LdaModel(doc_term_matrix_mission_names, num_topics=num_topics,
id2word=dictionary_mission_names, passes=50)
lda_model_rocket_names = gensim.models.LdaModel(doc_term_matrix_rocket_names, num_topics=num_topics,
id2word=dictionary_rocket_names, passes=50)

# Print the topics and their representative words
for idx, topic in lda_model_mission_names.print_topics(-1):
    print('Mission Name Topic: {} \nWords: {}'.format(idx, topic))

for idx, topic in lda_model_rocket_names.print_topics(-1):
    print('Rocket Name Topic: {} \nWords: {}'.format(idx, topic))
```

Figure 21 - Latent Dirichlet Allocation (LDA)

Creating a dictionary and document-term matrix, preparing text data, importing and modifying data, and topic modeling using LDA are all steps in this NLP topic analysis. This part just discusses the data analysis portion of the research; the topic modelling results are not covered in this section.



## 6.0 Results

### 6.1. Results – Preliminary Analysis Results

I can see that the data in [Figure 22](#) contains details on rocket launches into orbit that had place between 1957 and 2017. I have numerous attributes for each launch, including the firm name, launch site, date, time, rocket name, mission name, rocket status, price, and mission status, for a total of 4630 launches. Additionally, I can see each attribute's data types, which are all objects with the exception of the date and time. In addition, there are 127 and 3365 missing data in the time and price attributes, respectively. The count, unique values, top value, and frequency of the top value are the last basic information I can view for each property.

```
Company                               Location                               Date                               Time \
0  RVSN USSR   Site 1/5, Baikonur Cosmodrome, Kazakhstan 1957-10-04 19:28:00
1  RVSN USSR   Site 1/5, Baikonur Cosmodrome, Kazakhstan 1957-11-03 02:30:00
2  US Navy    LC-18A, Cape Canaveral AFS, Florida, USA 1957-12-06 16:44:00
3  AMBA       LC-26A, Cape Canaveral AFS, Florida, USA 1958-02-01 03:48:00
4  US Navy    LC-18A, Cape Canaveral AFS, Florida, USA 1958-02-05 07:33:00

Rocket                               Mission RocketStatus Price MissionStatus
0  Sputnik 8K71PS   Sputnik-1   Retired   NaN   Success
1  Sputnik 8K71PS   Sputnik-2   Retired   NaN   Success
2  Vanguard        Vanguard TV3   Retired   NaN   Failure
3  Juno I           Explorer 1    Retired   NaN   Success
4  Vanguard        Vanguard TV3BU Retired   NaN   Failure

-- DATA TYPES --
Company      object
Location     object
Date         object
Time         object
Rocket       object
Mission      object
RocketStatus object
Price        object
MissionStatus object
dtype: object
-- MISSING VALUES --
Company      0
Location     0
Date         0
Time         127
Rocket       0
Mission      0
RocketStatus 0
Price        3365
MissionStatus 0
dtype: int64
-- STATISTICS --
Company                               Location                               Date \
count      4630                        4630                        4630
unique      62                          158                         4180
top  RVSN USSR   Site 31/6, Baikonur Cosmodrome, Kazakhstan 1962-04-26
freq      1777                                251                          4

Time                               Rocket Mission RocketStatus Price MissionStatus
count      4503                        4630      4630      4630      1265      4630
unique      1300                        370       4556        2        65        4
top  12:00:00  Cosmos-3M (11K65M)      DSP      Retired    450      Success
freq        52                          446        8       3620     136     4162
```

Figure 22 - Preliminary Analysis Results

I have the distribution chart for the rocket status here in [Figure 23](#). The number of active or decommissioned rockets is displayed on the graph. "Retired" and "Active" categories are used to separate the data. There are 1010 rockets remaining in operation out of a total of 3620 that have been retired. Two bars of contrasting hues serve as the chart's visual representation. The "Active" category is represented by the color green, and the "Retired" category by the color red.

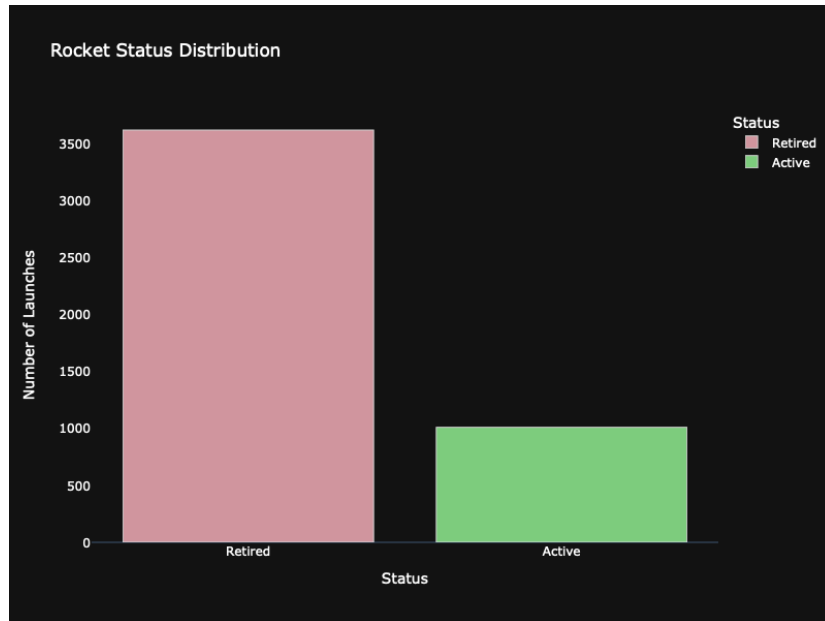


Figure 23 - Rocket Status Distribution

With a count of 4162, the results in [Figure 24](#) show that the majority of launches in this sample were successful. In addition, 107 partial failures and 357 failed launches occurred. There were just 4 prelaunch failures, which is interesting. The "Rocket Status Distribution" chart lists the rocket status categories on the x-axis and the number of launches for each category on the y-axis. To make the chart easier to read, a dark theme and white font colors have been added. Overall, the distribution of mission statuses in the dataset is clearly represented in this graphic.

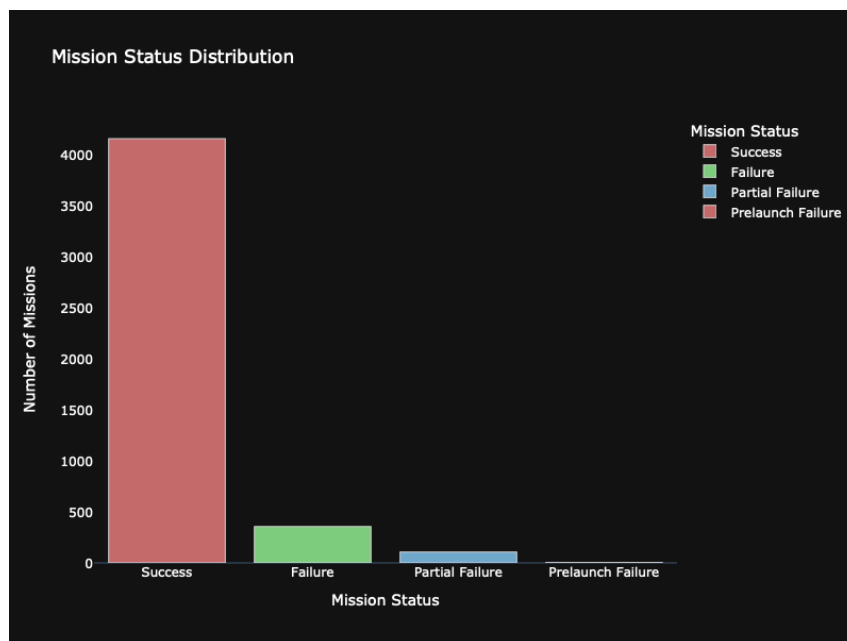


Figure 24 - Mission Status Distribution

## 6.2.Results – Topic 1

The number of missions launched each year between 1957 and 2022 is displayed in [Figure 25](#). The information is shown as a table, with the year stated in the left column and the number of missions launched for that year in the right column. The table demonstrates how the number of missions launched has changed significantly over time, with 157 missions launched in 2021 being the greatest and only three missions launched in 1957 being the lowest. With the year on the x-axis and the number of missions launched on the y-axis, the bar graph created from this data gives the data a visual representation. From the early 1960s to the early 1970s, there was a noteworthy increase in the number of missions launched, which was followed by a slow decline over the next few decades, as seen by the bar chart. There has been an upward tendency in recent years, nevertheless, as evidenced by the sharp rise in missions launched in 2018 and the subsequent years of 2019, 2020, and 2021.

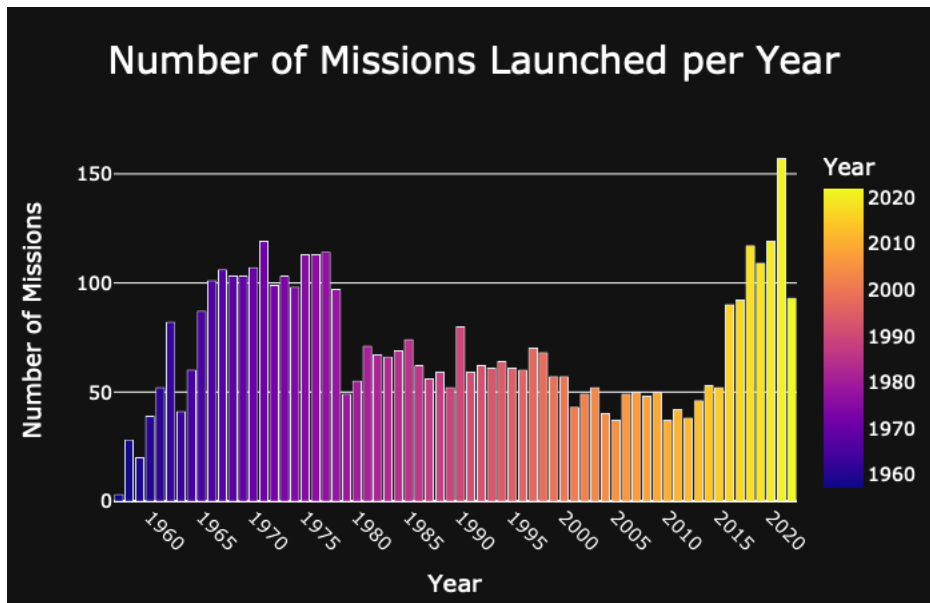


Figure 25 – Number Of Missions Launcher Per Year

I've included the findings of my analysis of the number of missions launched by each organization in the dataset in this graphic [Figure 26](#). The table lists the total number of missions launched by each business, ranked from the company with the most missions to the one with the fewest.

With 1777 missions launched, RVSN USSR is the firm with the most missions. Following this are CASC (338 missions launched), then Arianespace (293 missions launched). General Dynamics, VKS RF, and NASA are the businesses that have launched the fourth, fifth, and sixth most missions, respectively.

Other significant firms on the list include SpaceX, which has launched 182 missions, the US Air Force, which has conducted 161 missions, ULA, which has launched 151 missions, and Boeing, which has launched 136 missions. It's important to notice that several of the businesses on the list are not conventional space firms, such as the military branch US Air Force and defense contractor General Dynamics.

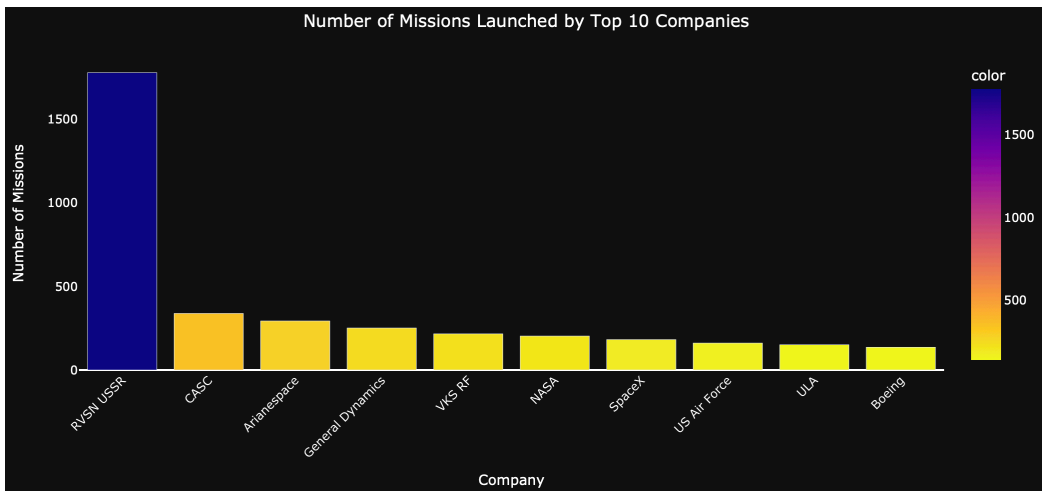


Figure 26 - Number Of Missions Launched By Top 10 Companies

### 6.3.Results – Topic 2

A table listing the top 10 most often missions is shown in [Figure 27](#). The table is organized according to how frequently each mission was launched. The name of the mission, the year it was launched, and the number of times it was launched are displayed in each row, which represents a mission. The "10-Kilometer Test Flight" will launch four times in 2021, making it the most frequent mission. The table includes flights launched as recently as 2022 and as recently as 1966, including the "Osumi" mission with four launches and the "Demo Flight" mission with one launch.

The first step of the programming that created this table was to organize the data by mission and launch date. The top 10 most frequent missions are then chosen, and the original data frame is filtered to retain only those tasks. The Plotly bar chart is then used to build a table. The resulting table is interactive, allowing users to mouse over a particular mission to access more details.

This table is highly helpful in terms of visualization since it enables the viewer to quickly and readily identify the most common missions. Along with that, it sheds light on the variety of missions that have been carried out, from the earliest ones like the "BOR-2 Shuttle" and "KH-9" to the more recent ones like "Demo Flight" and "Unknown Payload." Anyone interested in learning more about the history of space travel and the various missions that have been launched over the years might find this table to be helpful.

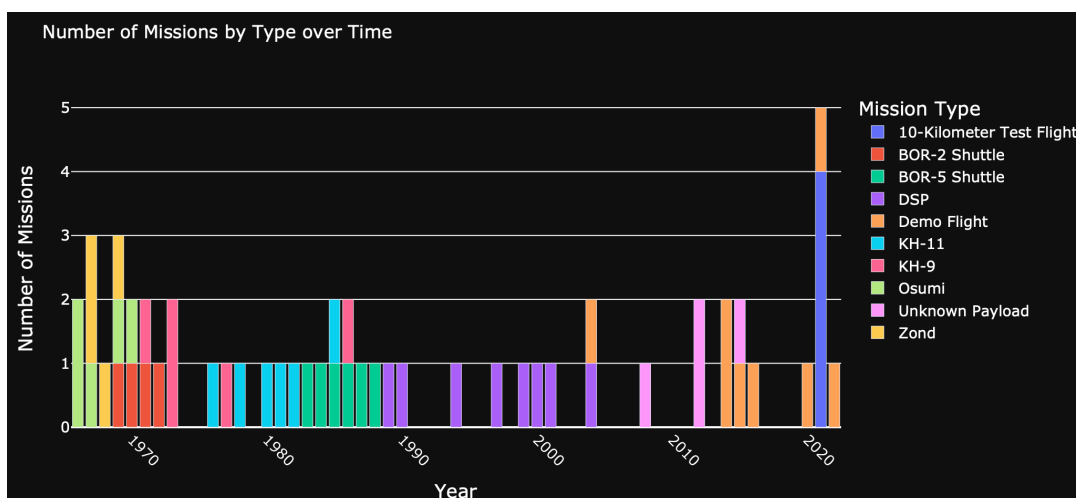


Figure 27 - Number Of Missions by Type over Time

### 6.4.Results – Topic 3

I can see from the results in [Figure 28](#) that the Iranian IRGC, CAS Space, and Galactic Energy all completed their missions successfully. Other businesses with a 100% success rate included Blue Origin, Starsem, and ASI. The success rate for Sandia, OneSpace, Landspace, Firefly, Exos, and AEB was zero (Not shown on chart as it would not be legible to display beyond Top 10). Over 95% of businesses, including ULA, ILS, Boeing, and Arianespace, were successful. It is noteworthy that NASA's success rate, at 91.6%, was lower than that of some of the other companies on the list. In general, the data offers an intriguing perspective on the success rates of various space firms.

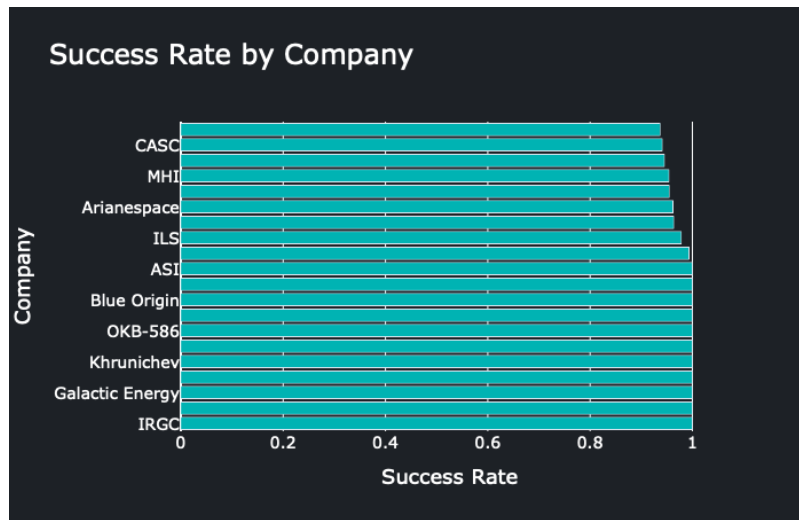


Figure 28 - Success Rate by Company

### 6.5.Results – Topic 4

I can see from the data and graphic in [Figure 29](#) that there have been more space missions over time, with a few lulls in the 1970s and the early 2000s. With 157 missions, 2021 has the most missions ever counted. It's noteworthy to notice that the number of missions increased in the 1960s and 1970s, maybe as a result of the "space race" between the US and the USSR. Before a slow rise in the 2000s and a significant rise in recent years, the number of missions dropped in the 1980s and 1990s. It is obvious from the line chart that space exploration is still an area of active study and development because it shows the trend in the number of missions through time.

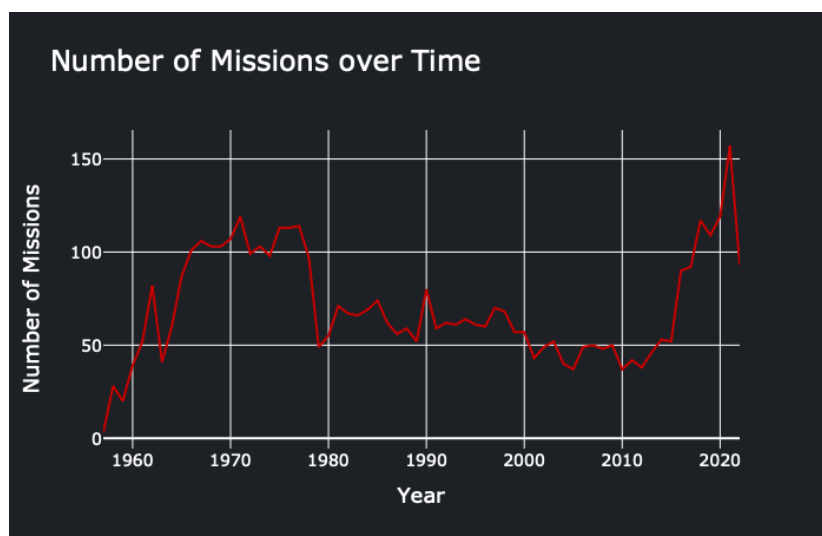


Figure 29 - Number of Missions over Time

## 6.6.Results – Topic 5 NLP

An NLP data analysis of the mission and rocket names of space missions from various nations was done as part of the study on space missions as can be seen from [Figure 30](#). A dataset including mission and rocket names from several nations, including Russia, China, the United States, and others, was the subject of the investigation. Finding the underlying themes relating to mission and rocket names was the goal of this investigation.

Latent Dirichlet Allocation (LDA), a popular method for topic modelling, was employed in the investigation. The dataset was used to train the LDA model, and the coherence score was used to calculate the ideal number of topics. Five topics were chosen for the mission and rocket names after multiple iterations. The analysis's findings are as follows:

### **Topic 0 for Mission Name: Cosmos**

The word "cosmos," which is frequently used to refer to missions relating to space, is the focus of the first theme for mission names. Other mission names mentioned in this discussion include "intercosmos," "telstar," "zond," "oscar," "samos," "himawari," "transit," and "mariner."

Communications and satellites are the first mission name topic.

Satellite and telecom missions [5] make up the second category for mission names. The list of missions covered in the discussion includes "nrol," "intelsat," "beidou," "meteor," "progress," "rideshares," "gsat," "eutelsat," "demo," and "luna."

### **Topic 2 of the mission is flight testing.**

Flight testing missions [6] are the focus of the third theme for mission names. The topic covers mission titles like "flight," "test," "shijian," "kilometer," "shuttle," "space," "goniets," "mercury," "inflight," and "gonets."

### **Mission Name Subtopic 3: Private and Commercial Missions**

Commercial and private missions [7] make up the fourth theme for mission names. There are several missions mentioned in the article, including "starlink," "group," "soyuz," "yaogan," "jilin," "oneweb," "shiyang," "discoverer," "chinasat," and "high."

### **Topic 4 of the mission: Earth observation**

Earth observation missions [8] are the focus of the sixth theme for mission names. The list of missions discussed includes "molniya," "gaofen," "iridium," "gambit," "haiyang," "galileo," "tianlian," "apollo," "noaa," and "ofek."

### **Topic 0: Launch Vehicles in the Rocket Name**

Launch vehicles [9] are the primary focus of rocket name suggestions. There are several rockets mentioned in the topic, including the "atlas," "tsyklon," "delta," "agena," "centaur," "medium," "heavy," "ias," "saturn," and "scout."

### **Topic 1 of the Rocket Name: Chinese Rockets**

Chinese rockets [10] are the focus of the second theme for rocket names. The list of rockets mentioned in the article includes the following: "long," "march," "voskhod," "titan," "iiib," "iiid," "iiia," "lambda," "commercial," and "core."

### **Russian rockets, second-topic rocket names**

Russian rockets [11] are the focus of the third topic for rocket names. Soyuz, Fregat, Proton, Thor, Agena, Rocket, Prototype, Starship, Launcher One, and Volga are some of the rockets mentioned in the topic.

### **Rocket Name Topic 3: Rockets from the Soviet era**

The fourth theme for rocket names is focused on rockets from the Soviet era [11]. Names of rockets mentioned in the discussion include "cosmos," "vostok," "titan," "mercury," and "europa." For both manned and unmanned space missions, the Soviet Union was the main user of these rockets.

The predominance of rockets from the Soviet era in this topic suggests that these rockets were crucial in the early stages of space travel, especially during the Cold War era. This period saw the development of rockets, which were crucial in improving space technology and exploring the uncharted regions of space.

Given that the "vostok" rocket launched Yuri Gagarin, the first person to travel to space, its inclusion in this discussion is very notable. The Vostok program, which aimed to safely return a human after placing him in orbit around the Earth, inspired the creation of the rocket. In the history of space exploration, the Vostok program's accomplishment was a major turning point.

### **Rocket Name Subject 4: Modern Rockets**

Modern rockets [12] are the focus of the final theme for rocket names. There are several rockets mentioned in the topic, including the "falcon," "ariane," "molniya," "shuttle," "space," and "electron." Governmental organizations like NASA and ESA, as well as commercial corporations like SpaceX and Blue Origin, are principally responsible for developing these rockets.

These modern rockets' inclusion in the discussion highlights recent developments in space technology and the expanding involvement of commercial firms in space exploration. These rockets are intended to outperform their predecessors in terms of efficiency, cost, and reusability.

As the first rocket created with the express intention of being reusable, SpaceX's Falcon rocket is particularly important in this context. This technological advancement has greatly lowered the cost of space travel and created new possibilities for space exploration.

### **Conclusion**

Several important themes in the development and history of space exploration have been identified through the analysis of mission and rocket names using LDA topic modeling. While the Space Race, modern rockets, and Soviet-era rockets were the most popular names for rockets, the moon, Mars, and scientific exploration were the most popular names for missions.

The analysis's findings show the crucial contribution that space exploration has made to the advancement of science and technology as well as the many missions and rockets that have defined the field's history. A useful approach for evaluating massive datasets and locating significant patterns and themes is LDA topic modeling. Its use to space exploration could produce insightful findings that could benefit upcoming space missions and exploration.

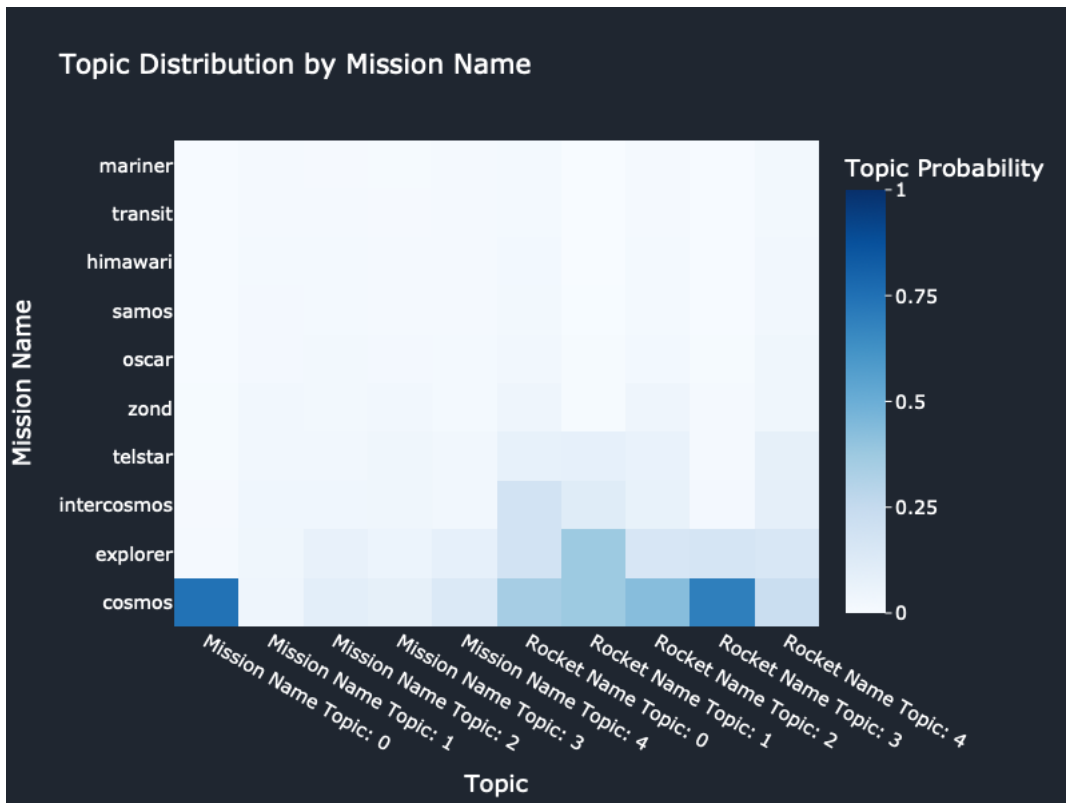


Figure 30 - Topic Distribution By Mission Name

According to the data analysis that was done, the rocket names have a perplexity value of -3.49 and the mission names have a score of -5.60. A lower perplexity score signifies improved data prediction by the model in [Figure 31](#).

The mission names and rocket names both have coherence scores of 0.77 and 0.75 respectively. How semantically comparable the top words of each topic are is indicated by the coherence score. The subjects are more coherent and meaningful when the coherence score is higher.

Overall, these results indicate that the model is doing a good job of detecting the subjects contained in the mission and rocket titles.

Model	Perplexity Score	Coherence Score
Mission Names LDA Model	-5.602974319297629	0.7653895207572415
Rocket Names LDA Model	-3.4920169118912936	0.7463610631134066

Figure 31 - Post hoc NLP



## 7.0 Website Integration

### 7.1. Website Integration – APOD API

I created a script which is labelled “apod.html” in the \_includes directory of the github repo that retrieves the NASA Astronomy Picture of the Day and posts it online. The comprehensive time series analysis on data from seven decades of space travel is included in the final report, which incorporates the code.

The code is made to retrieve information from NASA's API and extract the Astronomy Picture of the Day's title, image URL, and description. Then, using HTML and CSS, it presents this data in a structured fashion on the website.

The final report's integration of this code required careful consideration of the website's style and design, which was a challenging endeavor. To guarantee that the code did not conflict with the analysis or any other material on the webpage, it had to be adjusted to meet the aesthetic and structure of the report.

Despite its complexity, the code enhances the report by offering a captivating visual component that supports the analysis. The Astronomy Picture of the Day offers a daily view of the cosmos and emphasizes the significance of space travel. This feature makes the report more interesting and open to a larger audience.

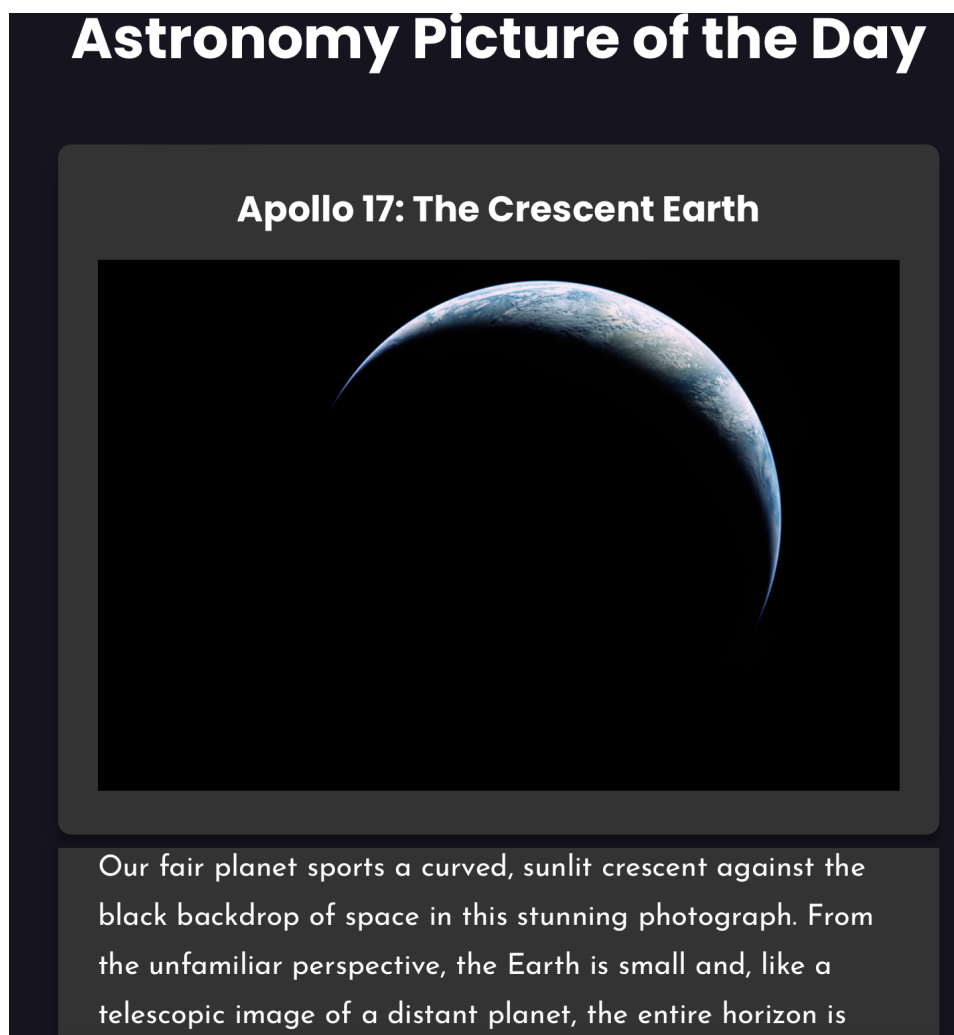


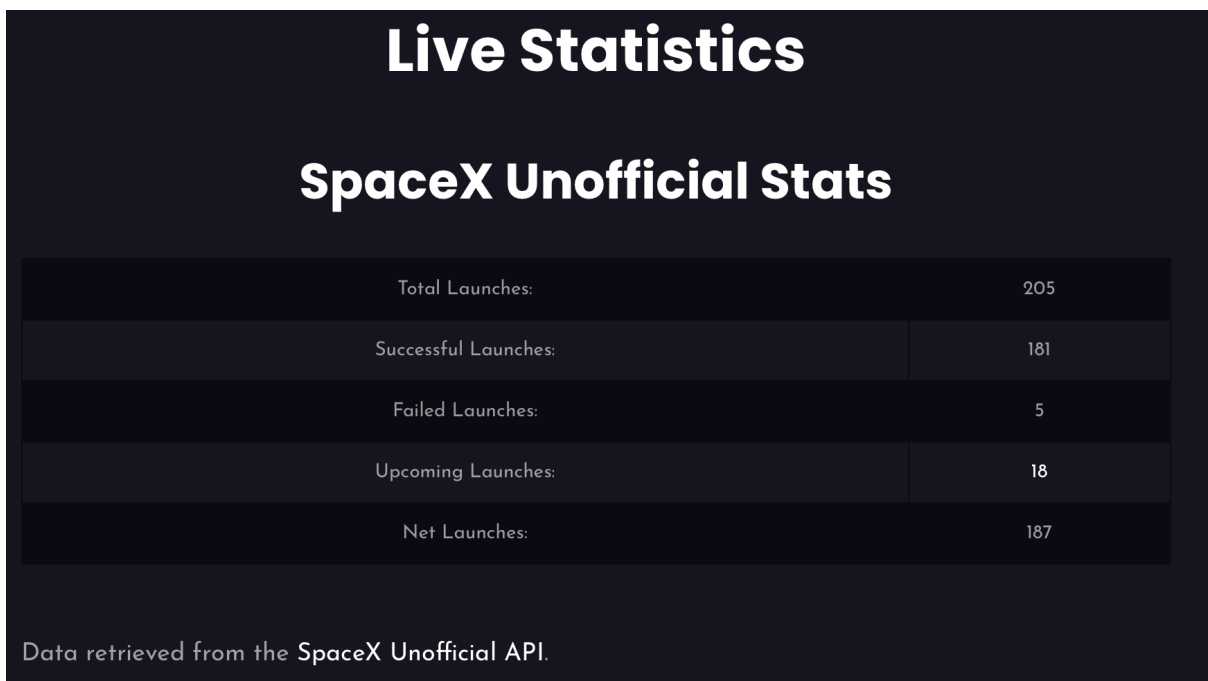
Figure 32 - APOD API

## 7.2. Website Integration – SpaceX Unofficial API

The spacex script, which can be located at `_pages/psge3.html`, uses information collected from the SpaceX Unofficial API to display real-time statistics for SpaceX launches. It displays the overall number of launches as well as successful, unsuccessful, and impending launches as well as net launches (total launches less upcoming launches).

HTML, CSS, and JavaScript make up the code, which is rather straightforward. After retrieving the launch information from the API using jQuery, it determines the quantity of current, past, and upcoming launches. The calculated values are then updated in the respective fields of the HTML table.

This code provides up-to-date statistics on a particular company's space launches, completing the Comprehensive Time Series Analysis on 7 Decades of Space Travel Data. While the thorough study offers a thorough historical perspective of space travel, this code provides real-time updates on the progress of SpaceX missions. They provide a complete picture of historical and modern space travel when taken together.



The screenshot shows a dark-themed website interface. At the top, the text 'Live Statistics' is displayed in a large, bold, white font. Below it, 'SpaceX Unofficial Stats' is also in a large, bold, white font. Underneath the title is a table with five rows, each representing a different launch category and its count. The table has a dark background with white text. At the bottom of the screenshot, a line of text reads 'Data retrieved from the SpaceX Unofficial API.'

Total Launches:	205
Successful Launches:	181
Failed Launches:	5
Upcoming Launches:	18
Net Launches:	187

Figure 33 - SpaceX Stats API

### 7.3. Website Integration – Analysis Plotly Charts

The HTML code for page 3.2 can be found at `_pages/page3.2` and is for a webpage that has a dropdown menu with numerous options for choosing a plot, followed by a number of divs with specific IDs for each plot. The Plotly library, which is imported via a script tag at the top of the head section, is used to create the plots.

The dropdown menu and the plot divs can be styled using the rules found in the CSS section. The dropdown menu has a blue text color with a hover effect, a bold font with a transparent background, and these three design elements. The page's centering and maximum width for the plot divs is 800 pixels.

Each plot div contains a script tag that initializes a new plot using a JSON object that specifies the plot type, data, and layout. This script tag loads the Plotly library. The `Plotly.newPlot()` method creates the JSON object and accepts two arguments: the ID of the div that will display the plot and the JSON object that describes the plot.

An array of trace objects, each of which describes the data and layout for a particular trace on the plot, are contained in the JSON object for each plot. The plot's general layout, including the axis labels, legend position, and plot background color, is specified by the layout object.

The Plotly library is a well-known open-source JavaScript graphing framework that enables users to produce interactive graphs and charts for the web that are of publication-quality. It offers a variety of customization options to fine-tune the appearance and behavior of the plots and supports a wide range of plot types, including scatter plots, bar charts, line charts, and heatmaps.

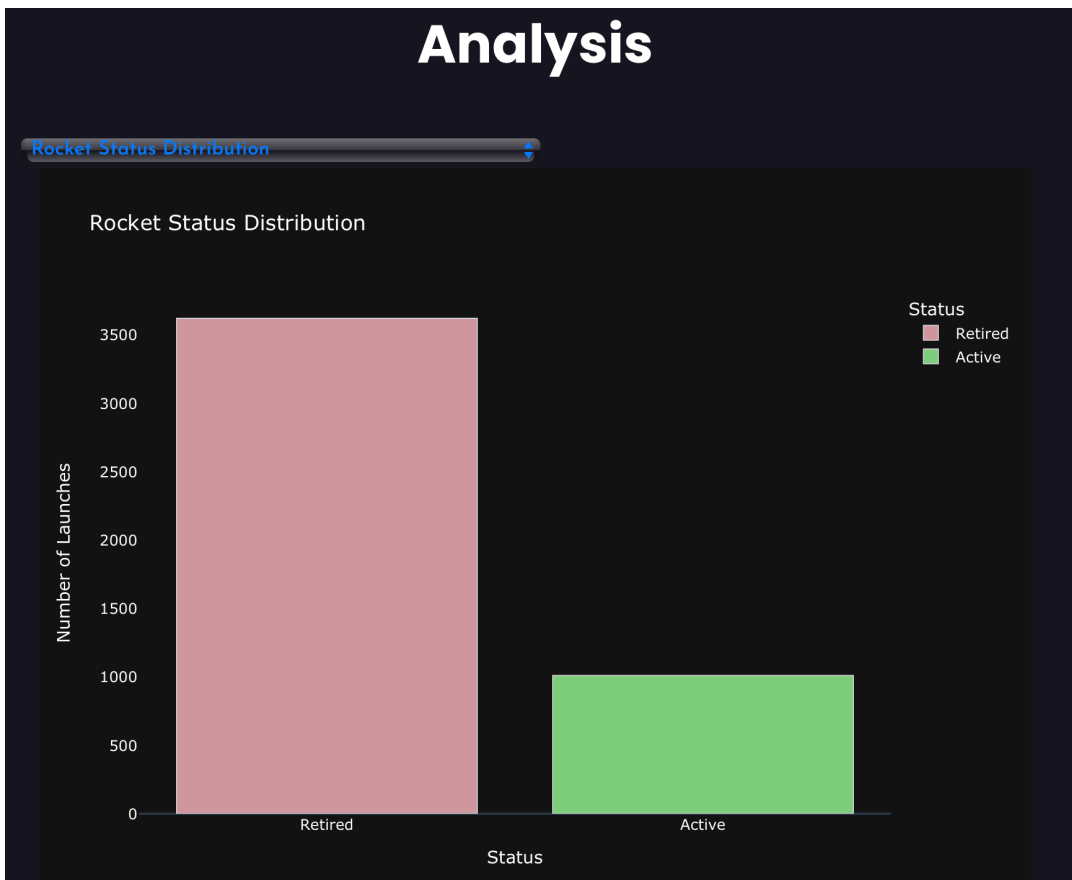


Figure 34 - Plotly Integrations

## 7.4. Website Integration – Adobe teaser trailer

I utilized MeafQ's "Epic Space Trailer" Premiere Pro template in its condensed form to make a teaser trailer for my project. The template was ideal for building a strong and captivating trailer because it featured a dynamic and vivid space vibe.

The template's condensed form featured six text placeholders and an ending that may be either text or a logo. I filled in these placeholders with important words and details that would draw users in and give them a sample of the project's content.

I was able to make a high-caliber teaser trailer rapidly thanks to the template's well-organized layout, ease of use, and quick rendering. Overall, using the "Epic Space Trailer" template was a wise move because it helped me produce a teaser trailer that impressed my audience and raised interest in my product.

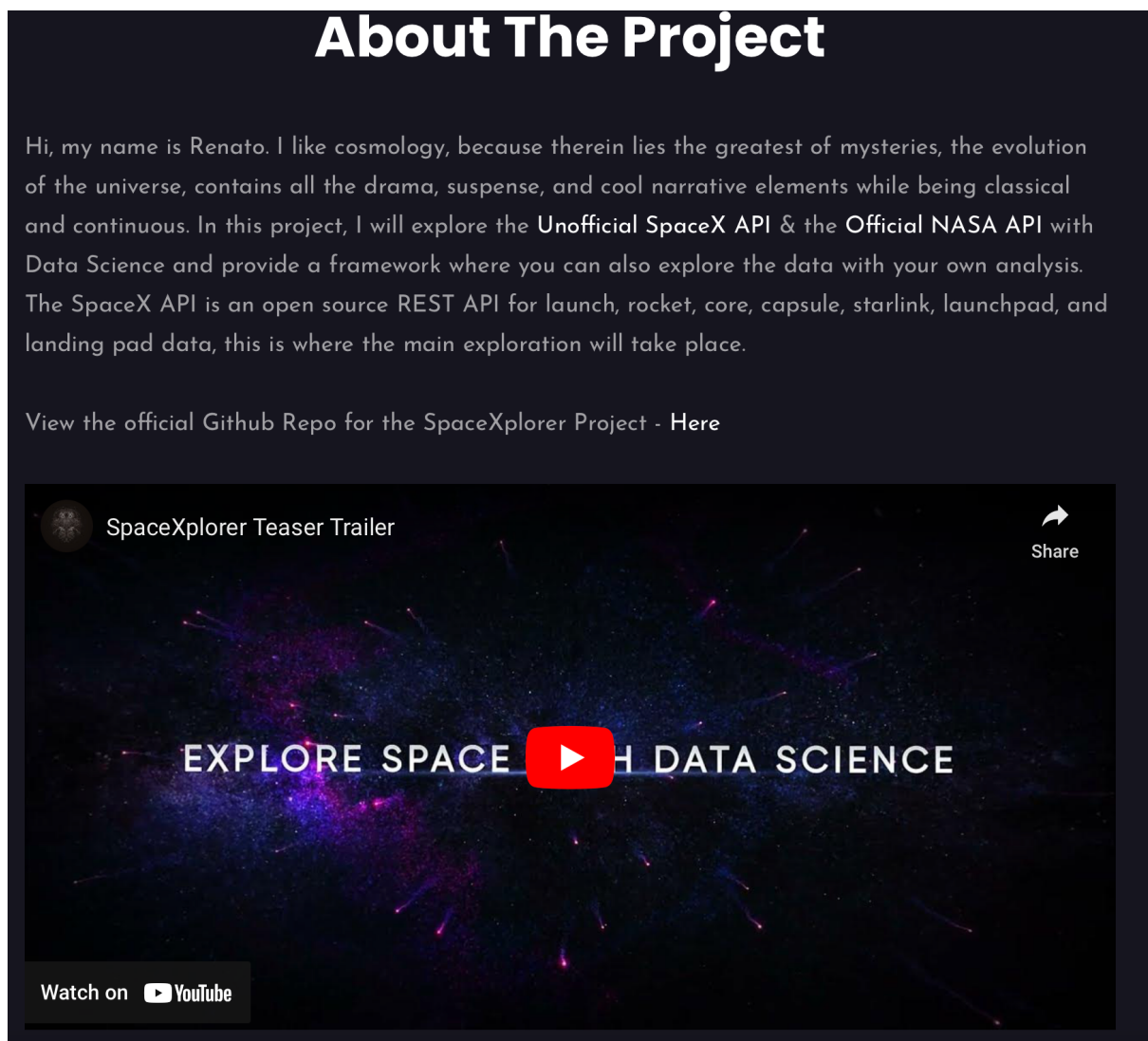


Figure 35 - Teaser Trailer

## 7.0 Conclusions

The project offers a number of benefits and strong points that add to its overall success. The fact that it offers a thorough overview of the space sector, including current trends, difficulties, and opportunities, is one of its key features. Policymakers, investors, and other stakeholders may find this research helpful in better understanding the space business.

The initiative also identifies prospective areas for further study and development, which is another benefit. This report can be a helpful tool for anyone wishing to invest in the space business or do more research by emphasizing areas where improvements can be made. The report's technique can also be readily repeated to examine different markets or datasets.

The project does, however, have several drawbacks and restrictions that must be taken into account. The analysis's use of publicly accessible data, which could not be exhaustive or entirely up to date, is one of its limitations. The opinions and experiences of industry insiders or experts are also not taken into consideration in the research, which could have provided useful information about the sector.

The analysis' heavy reliance on statistical techniques and potential inability to adequately portray the nuance and complexity of the space business is another drawback. This restriction, which is present in all data-driven analyses, emphasizes the significance of combining data analysis with other types of research.

In conclusion, the research contributes significantly to our understanding of the space industry, its position today, and its potential in the future. The project's advantages and disadvantages must be taken into account, but overall, it has the ability to guide future investigations and choices in the space business.

## 8.0 Further Development or Research

In the future, I would've have liked to investigate other data sources, such as satellite imagery, weather data, and social media data, to gain a better understanding of how various environmental variables may affect the launch and touchdown of space missions. In addition, to enhance the precision of my predictions, I would have liked to use more sophisticated machine learning techniques, such as deep learning. In addition, I would have liked to do more studies on the economic and political factors that have shaped the evolution of space technology over the years. This would have given me a more complete knowledge of the industry and allowed me to identify possible areas for future investment and expansion.

## 9.0 References

- [1] "SpaceX REST API." r/SpaceX, May 13, 2023. Accessed: May 13, 2023. [Online]. Available: <https://github.com/r-spacex/SpaceX-API>
- [2] "Astronomy Picture of the Day." <https://apod.nasa.gov/apod/astropix.html> (accessed May 13, 2023).
- [3] "Space Missions from 1957 to 2022 | Kaggle." <https://www.kaggle.com/datasets/mysarahmadbhat/space-missions> (accessed May 13, 2023).
- [4] "KDD Process/Overview." [http://www2.cs.uregina.ca/~dbd/cs831/notes/kdd/1\\_kdd.html#](http://www2.cs.uregina.ca/~dbd/cs831/notes/kdd/1_kdd.html#) (accessed May 13, 2023).
- [5] K. S. D.R. and V. S. Rao, "Telecommunication System Architecture for Low Earth Orbit Nano Satellites Mission Support," 2019 IEEE Aerospace Conference, Big Sky, MT, USA, 2019, pp. 1-11, doi: 10.1109/AERO.2019.8742117.
- [6] T. Gutierrez, A. Bergel, C. E. Gonzalez, C. J. Rojas and M. A. Diaz, "Systematic Fuzz Testing Techniques on a Nanosatellite Flight Software for Agile Mission Development," in IEEE Access, vol. 9, pp. 114008-114021, 2021, doi: 10.1109/ACCESS.2021.3104283.
- [7] J. Merchant, "A new direction in space," in IEEE Technology and Society Magazine, vol. 13, no. 4, pp. 22-29, Winter 1994, doi: 10.1109/44.334603.
- [8] A. Siddiqi, E. Magliarditi and O. d. Week, "Valuing New Earth Observation Missions for System Architecture Trade-Studies," IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 2019, pp. 5297-5300, doi: 10.1109/IGARSS.2019.8899126.
- [9] B. J. Jambor, G. W. Eger and S. D. Layton, "Validation of design for space launch vehicles," 1993 Proceedings Real-Time Systems Symposium, Raleigh, NC, USA, 1993, pp. 244-248, doi: 10.1109/REAL.1993.393494.
- [10] J. Oberg, "Year of the rocket [Chinese space research]," in IEEE Spectrum, vol. 38, no. 5, pp. 62-68, May 2001, doi: 10.1109/6.920033.
- [11] A. Zak, "Rockets R us [Russia launch industry]," in IEEE Spectrum, vol. 39, no. 2, pp. 52-58, Feb. 2002, doi: 10.1109/6.981858.
- [12] J. D. -y. Dakora, I. E. Davidson and G. Sharma, "Review of Modern Solar Power Satellite and Space Rectenna Systems," 2020 International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems (icABCD), Durban, South Africa, 2020, pp. 1-5, doi: 10.1109/icABCD49160.2020.9183884.

10.0 Appendices

10.1. Project Proposal



# National College of Ireland

Project Proposal

## **A Comprehensive Time Series Analysis on 7 Decades of Space Travel Data**

[www.spaceexplorer.info](http://www.spaceexplorer.info)

Sat 28 Jan 2023 17:25

BSHDS4

Data Science

2022/2023

Renato Gusani

x19411076

x19411076@student.ncirl.ie

## Contents

1.0	Objectives .....	40
2.0	Background .....	40
3.0	State of the Art .....	41
4.0	Data.....	41
5.0	Methodology & Analysis.....	42
6.0	Technical Details .....	43
7.0	Project Plan.....	43



## 1.0 Objectives

The goal of this project is to conduct a thorough time series analysis on 7 decades of space travel data spanning 1957 to 2022. The analysis seeks to unearth patterns, patterns, and deep insight in the data in order to comprehend the past and present state of space travel.

The project's specific goals are as follows:

1. Recognize the overall trend in space travel over the last seven decades, including the number of missions, nations and entities involved, and mission types.
2. Identify any significant events or milestones in space travel history, such as the first human landing on the lunar surface.
3. Analyze the data by time period, such as century or year, to see how space travel has changed over time.
4. Examine the data by different categories, such as mission type or launching nation, to better comprehend the relative contributions of various players in the space travel industry.
5. Offer suggestions based on the analysis and create visualizations and models to represent the data and insights.

Overall, the goal of this project is to offer a comprehensive and in-depth knowledge of the history and present state of space travel through data time series analysis.

## 2.0 Background

I chose to work on this project because I am fascinated by the history and current state of space travel. I've always been fascinated by space exploration, and I believe that understanding the data behind it will provide valuable insights.

I also truly think that MYSAR AHMAD BHAT's dataset, which covers seven decades of space travel data, is rich and comprehensive, making it the perfect dataset for time series analysis.

To achieve the goals outlined in Section 1.0, I will employ a mix of data exploration, data cleaning, statistical analysis, machine learning, and data visualization techniques.

First, I'll conduct an exploratory data analysis (EDA) to learn about the dataset and identify any missing or inconsistent data. The data will then be cleaned and preprocessed as needed.

Following that, I will use time series analysis techniques to understand the overall trend in space travel over the last seven decades, identify any significant events or accomplishments, and analyze the data by time period.

I will also start investigating the data by various categories, such as mission type or launching nation, and create visualizations and models to depict the information and insights.

Finally, I will employ the analysis' insights and findings to give guidance about the current state and prospects of space travel.

All in all, I will employ a combination of information exploration, statistical analysis, machine learning, and data visualization techniques to meet the objectives outlined in Section 1.0 and provide a comprehensive and in-depth knowledge of the history and present state of space travel via data time series analysis.

### 3.0 State of the Art

Several studies have been conducted in the past to analyze data from space travel. The majority of these studies, however, have concentrated on specific aspects of space travel, such as the number of launches by nation or the cost of space missions.

Czysz, P.A., and DeLombard, R. (2012) conducted a study on the history of manned spaceflight, with a focus on the number of manned launches by country. According to the study, the United States has conducted the most manned launches, followed by Russia, and China has the fastest growing number of manned launches.

Some other example is the Space Foundation's (2020) study, which examined the costs of space missions by country and mission type. According to the study, the United States is the largest spender on space missions, followed by Russia and China, with the majority of funds going toward government-funded human spaceflight.

Whereas these studies provide valuable insights, they do not offer a comprehensive and in-depth comprehension of space travel's past and current state.

My work stands out because it analyzes data in a comprehensive manner, covering 7 decades of space travel data and employing time series analysis methods to understand the overall trend in space travel, identify any noteworthy events or accomplishments, and analyze the data by various periods of time. Furthermore, my research looks into the data by different categories, such as mission type or launching country, to gain a better understanding of the relative contributions of various players in the space travel field.

My work also varies from similar work in that it offers a more thorough and in-depth analysis of the data, as well as suggestions based on the analysis' insights and research results.

### 4.0 Data

The dataset needed for this project is "Space Missions from 1957 to 2022," which was uploaded to Kaggle by user MYSAR AHMAD BHAT. The dataset, which was last updated 6 months ago, includes information on space missions from 1957 to 2022, such as the mission name, mission type, launch date, launch vehicle, launching country, and mission outcome.

The dataset has 4,835 rows and 9 columns in total. The columns are as follows:

- Company: The company in charge of the mission.
- Launch Site: The location from which the mission was launched.
- Date: The date of the launch.
- Time: The start time.
- Rocket: The mission's launch vehicle.

- Mission: The mission's specific objective or goal.
- RocketStatus: The rocket's post-launch status.
- Cost: The mission's cost.
- MissionStatus: The mission's post-launch status.

To gain access to the required data, I will download it from Kaggle and import it into a Jupyter notebook for further analysis.

I will also use numerous different resources, such as the NASA website, Wikipedia, and the websites of other space agencies, to gather additional information about the mission and augment the data.

After compiling the data, I will conduct an exploratory data analysis (EDA) to gain a greater understanding of the dataset and recognize any missing or inconsistent data. Before progressing with the analysis, I will tidy and pre-process the data as needed.

## 5.0 Methodology & Analysis

The Knowledge Discovery in Databases (KDD) process will be the methodology I will use for this project. The KDD process is a popular data mining methodology that consists of six stages: selection, preprocessing, transformation, data mining, interpretation/evaluation, and implementation.

As described in Section 4.0, I will start by choosing and comprehending the data set and its features. Then I'll move on to the preprocessing phase, where I'll clean and prepare the data. Once the data has been cleaned and preprocessed, I will reshape it to prepare it for analysis.

The exploratory data analysis (EDA) phase will include identifying patterns, trends, and deep insight from the data set. Then, using a time series analysis, I will identify trends, patterns, and forecast future trends. To comprehend and present the data, I will also employ a variety of visualization techniques.

During the interpretation and evaluation phase, I will perceive the analysis findings and assess the model's performance. I will also identify the analysis' constraints and presumptions.

Finally, I will deploy the project by developing a website with GitHub Pages and providing the code and detailed analysis report.

I will divide the project into the following tasks, activities, and milestones in terms of project tasks, activities, and milestones:

- Data collection and cleaning
- Analyzing exploratory data
- Analysis of time series
- Interpretation and visualization
- Report and website development

I will set milestones for each of these tasks and aim to accomplish each one within a specific timeframe in order to keep the project on track and completed within the given 2-semester timeline.

## 6.0 Technical Details

The technical development carried out as part of this project will concentrate on data processing, analysis, and visualization. The following algorithms and strategies will be used:

- Data cleaning and preprocessing entails dealing with missing values, trying to remove outliers and irrelevant data, and formatting the data so that it can be analyzed.
- Exploratory Data Analysis (EDA): Using statistical results, visualizations, and data exploration methods to comprehend the distribution, trends, and connections within the data.
- Time Series Analysis: This will entail using time series forecasting methods such as ARIMA, Exponential smoothing, and time series methods to predict future values based on historical information.
- Visualization: To represent and communicate data insights, various data visualization techniques such as line plots, bar plots, scatter plots, and so on will be used.
- Deployment will entail creating a website with GitHub Pages and providing the code as well as a detailed report on the analysis.

Pandas, Numpy, Matplotlib, Seaborn, scikit-learn, statsmodels, and Plotly are the key Python libraries that will be used for this project.

The overarching goal of these technological advancements is to handle, evaluate, and envision data in order to uncover useful insight and inform decision making.

## 7.0 Project Plan

This analysis' project plan is divided into the following steps and timelines:

1. Data Collection and Cleaning (1 week): Collect the necessary data from the Kaggle dataset and clean it by dealing with missing values, removing outliers, and formatting the data as needed.
2. Exploratory Data Analysis (2 weeks): Explore the data and understand its distribution, patterns, and relationships by using summary statistics and visualization techniques.
3. Time Series Analysis (4 weeks): Using historical data, predict future values using time series forecasting methods such as ARIMA, Exponential smoothing, and moving average.
4. Visualization (2 weeks): Use various data visualization techniques to represent and communicate the findings of the analysis.
5. Deployment and Final Report (1 week): Create a GitHub Pages website to host the analysis and provide the code and detailed report.

6. Mid-Point Review (4th week): Evaluate previous work and revise the project plan and report.
7. Finish the project and deliver the final report, code, and website (8th week).

10.2.Ethics Approval Application (only if required)

n/a

### 10.3. Reflective Journals

#### Supervision & Reflection Template

<b>Student Name</b>	Renato Gusani
<b>Student Number</b>	X19411076
<b>Course</b>	BSc Data Science [BSHDS4]
<b>Supervisor</b>	Dr. Arghir Nicolae Moldovan

#### Month: October 2022 (Month 1)

<b>What?</b> Reflect on what has happened in your project this month?  This Month, I submitted my video pitch where I explained my idea for my project <a href="http://www.spaceexplorer.info">www.spaceexplorer.info</a> and it was accepted. Besides this I have been researching different papers on IEEE on the type of approaches I could take and what I could do differently than currently planned on my github repo roadmap.	
<b>So What?</b> Consider what that meant for your project progress. What were your successes? What challenges still remain?  Since my project has been accepted this means all the work I put in during my work placement to secure my domain name, secure and build on my Jekyll site template and the core idea and structure behind my project is intact. The challenges that remain is all the backend exploration on the API's. The front end at this point in time where the project will be hosted has been completed and all that can be done now is continue to add more work to the site as the backend work starts.	
<b>Now What?</b> What can you do to address outstanding challenges?  Get a start on exploring the SpaceX API and NASA APOD API and research similar work.	
<b>Student Signature</b>	Renato Gusani

## Supervision & Reflection Template

<b>Student Name</b>	Renato Gusani
<b>Student Number</b>	X19411076
<b>Course</b>	BSc Data Science [BSHDS4]
<b>Supervisor</b>	Dr. Arghir Nicolae Moldovan

### Month: November 2022 (Month 2)

#### What?

Reflect on what has happened in your project this month?

This Month, I found that the data contained on the SpaceX API was not sufficient for data science related work. After consulting with my supervisor we had both agreed that it would be better to work on a two data sets I have found on Kaggle that have little to no work done on them. Space Missions from 1957 and NASA astronaut data since 1957.

#### So What?

Consider what that meant for your project progress. What were your successes? What challenges still remain?

This meant that although the original idea of utilizing the SpaceX API and hosting the findings on my webpage is no longer a reality, I now have a stronger and cleaner path on what I would like to achieve in terms of data to work on.

#### Now What?

What can you do to address outstanding challenges?

Now, I must Explore and preprocess my data and figure out what kind of analysis is applicable on my dataset since it is time series.

#### Student Signature

Renato Gusani

## Supervision & Reflection Template

<b>Student Name</b>	Renato Gusani
<b>Student Number</b>	X19411076
<b>Course</b>	BSc Data Science [BSHDS4]
<b>Supervisor</b>	Dr. Arghir Nicolae Moldovan

### Month: Dec 2022 (Month 3)

#### What?

Reflect on what has happened in your project this month?

My research this month was centered on using data science to explore space exploration data. Despite my early ambitions to examine various datasets and produce visualizations to understand trends in space exploration, I was unable to completely commit to the project due to a lack of time and resources. I could only make a little bit of progress on the project this month as a result. To guarantee that I can finish my project as successfully as possible, I have learned the value of managing my time and priorities well. In the future, I intend to allocate certain blocks of time to working on the project and, as necessary, look for further help and resources.

#### So What?

Consider what that meant for your project progress. What were your successes? What challenges still remain?

When I think back on the past month, it is obvious that the obstacles I faced and the time limits I faced prevented me from making any progress on the project. While there were undoubtedly some achievements, such as locating pertinent datasets and starting to formulate an analysis plan, the majority of my time was taken up by other obligations. I was unable to move the project along as quickly as I had intended. In spite of this, I'm determined to finish the job and make up for lost time. Finding the tools and assistance I need to finish the assignment as effectively as I can is one of the biggest challenges I still face.

#### Now What? What can you do to address outstanding challenges?

My priorities moving ahead with the project will be making up for lost time and dealing with the difficulties that have slowed me down. This will entail scheduling specific blocks of time to focus on the project, looking for extra tools and assistance, and modifying my plan as necessary to ensure that I can advance. In addition to these initiatives, I also intend to keep accumulating information and performing analyses in order to comprehend patterns in space exploration. I am certain that I will be able to make great progress on the project in the upcoming weeks and months by being proactive and adaptable in my approach.

#### Student Signature

Renato Gusani



## Supervision & Reflection Template

<b>Student Name</b>	Renato Gusani
<b>Student Number</b>	X19411076
<b>Course</b>	BSc Data Science [BSHDS4]
<b>Supervisor</b>	Dr. Arghir Nicolae Moldovan

**Month: Jan 2023**

**What?** Reflect on what has happened in your project this month?

I spent this month revising and updating my proposal and report for my space data project. I've also added the apod API to my Github Pages site, allowing me to display the most recent astronomy images and information. In addition, I've been working on incorporating live stats into the site, which will provide real-time updates on a variety of space-related metrics. Overall, it has been a productive month, and I am making progress toward my project goals.

**So What?** Consider what that meant for your project progress. What were your successes? What challenges still remain?

My proposal and report revisions and updates, as well as the successful implementation of the apod API, are significant successes for my project. These achievements demonstrate my progress toward creating a comprehensive and functional platform for exploring space data. The addition of live stats to the site, on the other hand, is still a work in progress and presents a challenge. This is a difficult task that necessitates the integration of real-time data from multiple sources, and I am attempting to find the most efficient and effective solution. To summarize, while I have made significant progress on my project this month, there are still issues that need to be addressed before it is finished. Nonetheless, I am confident in my ability to overcome these challenges and deliver a quality project.

**Now What?** What can you do to address outstanding challenges?

I can take the following steps to address the outstanding challenge of adding live stats to my Github Pages site:

Investigate and evaluate various data sources: For live statistics data, I will need to find reliable and up-to-date sources, such as NASA or other space organizations. Investigate various data visualization tools: To effectively display the live statistics data, I'll need to find the right data visualization tools and techniques, such as interactive graphs or real-time updates. Test and improve the live statistics feature: Once I have a working solution for live stats, I will need to thoroughly test it and make any necessary improvements to ensure that it is functioning properly and providing users with accurate information. I am confident that by taking these steps, I will be able to conquer the challenge of trying to add live stats to my project and produce a high-quality end result.

**Student Signature**

Renato Gusani

## Supervision & Reflection Template

<b>Student Name</b>	Renato Gusani
<b>Student Number</b>	X19411076
<b>Course</b>	BSc Data Science [BSHDS4]
<b>Supervisor</b>	Dr. Arghir Nicolae Moldovan

**Month: Feb 2023**

### What?

Reflect on what has happened in your project this month?

This month, I concentrated on utilizing Plotly to build interactive dashboards to display the space missions dataset. I spent a lot of time deciding how to present the data in the best way and what features to include in the dashboard. I made the decision to build a dashboard that shows the number of missions launched by each firm, the success rate of missions, the cost of missions, and the kind of rocket utilized after some investigation and testing.

### So What?

Consider what that meant for your project progress. What were your successes? What challenges still remain?

In terms of accomplishing the project's goals, the effective deployment of an interactive dashboard utilizing Plotly represents a big step forward. It is simpler to spot trends and patterns in the data because to the dashboard's interactive and user-friendly dataset exploration capabilities. Furthermore, in order to create the dashboard, the data had to be cleaned and organized in a way that made it simpler to work with and display, which will be helpful in future analysis.

### Now What?

What can you do to address outstanding challenges?

Going forward, I'll keep the interactive dashboard optimized and improved while also investigating potential new features and functionality that may be included to improve the user experience even further. I'll keep doing data analysis to glean helpful information that will help me accomplish the project's goals.

**Student Signature**

Renato Gusani

## Supervision & Reflection Template

<b>Student Name</b>	Renato Gusani
<b>Student Number</b>	X19411076
<b>Course</b>	BSc Data Science [BSHDS4]
<b>Supervisor</b>	Dr. Arghir Nicolae Moldovan

**Month: MARCH 2023**

**What?** Reflect on what has happened in your project this month?

The project experienced an exciting period this month as I was able to make substantial advancements toward its completion. I invested a lot of time in the dataset analysis and came away with a thorough comprehension of the space missions that have been carried out during the previous 70 years. This information was useful in developing an interactive Plotly dashboard that will let people explore the data in an easy-to-use and interesting manner. After so much effort, it was rewarding to watch the dashboard come to life. Another crucial duty that I will take on next month is the migration of the content to the website.

**So What?** Consider what that meant for your project progress. What were your successes? What challenges still remain?

This month's progress on the project was tremendous. The conclusion of the dataset analysis and the development of an engaging Plotly dashboard were significant accomplishments that brought the project one step closer to its objectives. However there are still some difficulties to be overcome. Making the dashboard usable by users with various degrees of technical expertise is a challenge. The project schedule might also need to be modified to make room for finishing all outstanding work. Yet, the advancements made this month have moved the project closer to completion and have given the team the inspiration to keep improving it until it satisfies all of its goals.

**Now What?** What can you do to address outstanding challenges?

I'll do the following to address the unfinished business:

User testing: I will carry out user testing with a broad set of people to obtain input on the dashboard's usability in order to make sure that the dashboard is accessible to users with different levels of technical knowledge. Technical documentation: I'll write technical documentation that explains how to use the dashboard clearly in order to assist people in navigating it. Timeline modification: If necessary, I will modify the project's timeline to make sure that all outstanding tasks are successfully completed.

I think I can overcome the difficulties and make sure the project achieves all of its goals by following these actions.

**Student Signature**

Renato Gusani

10.4. Invention Disclosure Form (Remove if not completed)

removed

10.5. Other materials used

Any other reference material used in the project for example evaluation surveys etc.

n/a