

SPACE MISSION FROM 1957 DATA ANALYSIS

REPORT



Assessing Progress and Impact



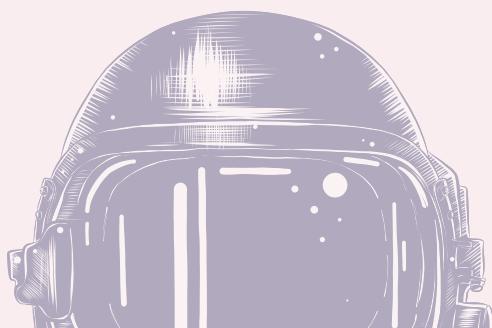
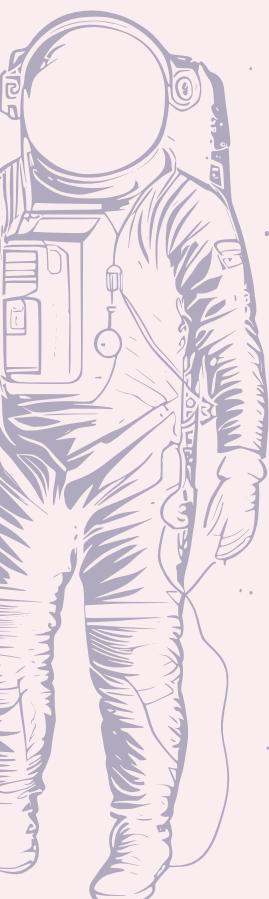
Presented By :
Team Zeta (Machine
Learning)

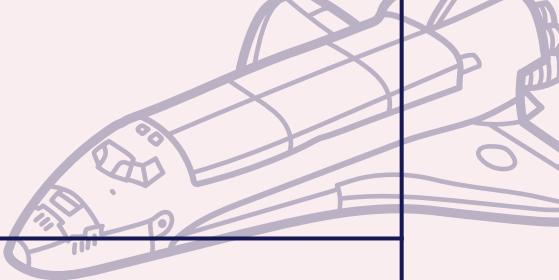


OUR TEAM



- Mahnoor Khawar (TEAM LEAD)
- Fahad Abbas
- Amna Khan
- Jawad Khan
- Tayyaba Ishfaq
- Akif Mehmood
- Muhammad Talha

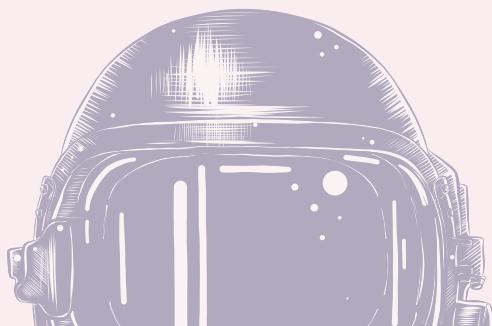




INTRODUCTION



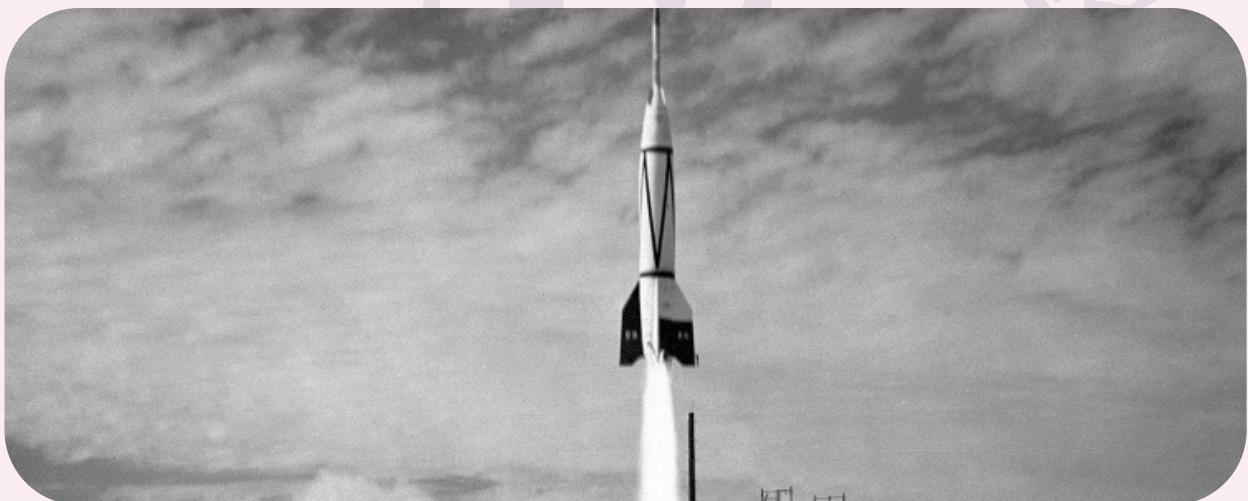
The Space Race was a pivotal 20th-century competition between the Soviet Union (USSR) and the United States (US), centered around their efforts to achieve superior spaceflight capability. Emerging from the post-World War II ballistic missile-based nuclear arms race, the Space Race became a critical element of national security, symbolism, and ideology. It spurred pioneering achievements, including the launch of artificial satellites, robotic space probes to the Moon, Venus, and Mars, and human spaceflight. On October 4, 1957, the USSR successfully launched Sputnik 1, the first Earth-orbiting satellite, followed by Sputnik 2 on November 3, 1957, carrying a dog named Laika, the first living organism to orbit the Earth.



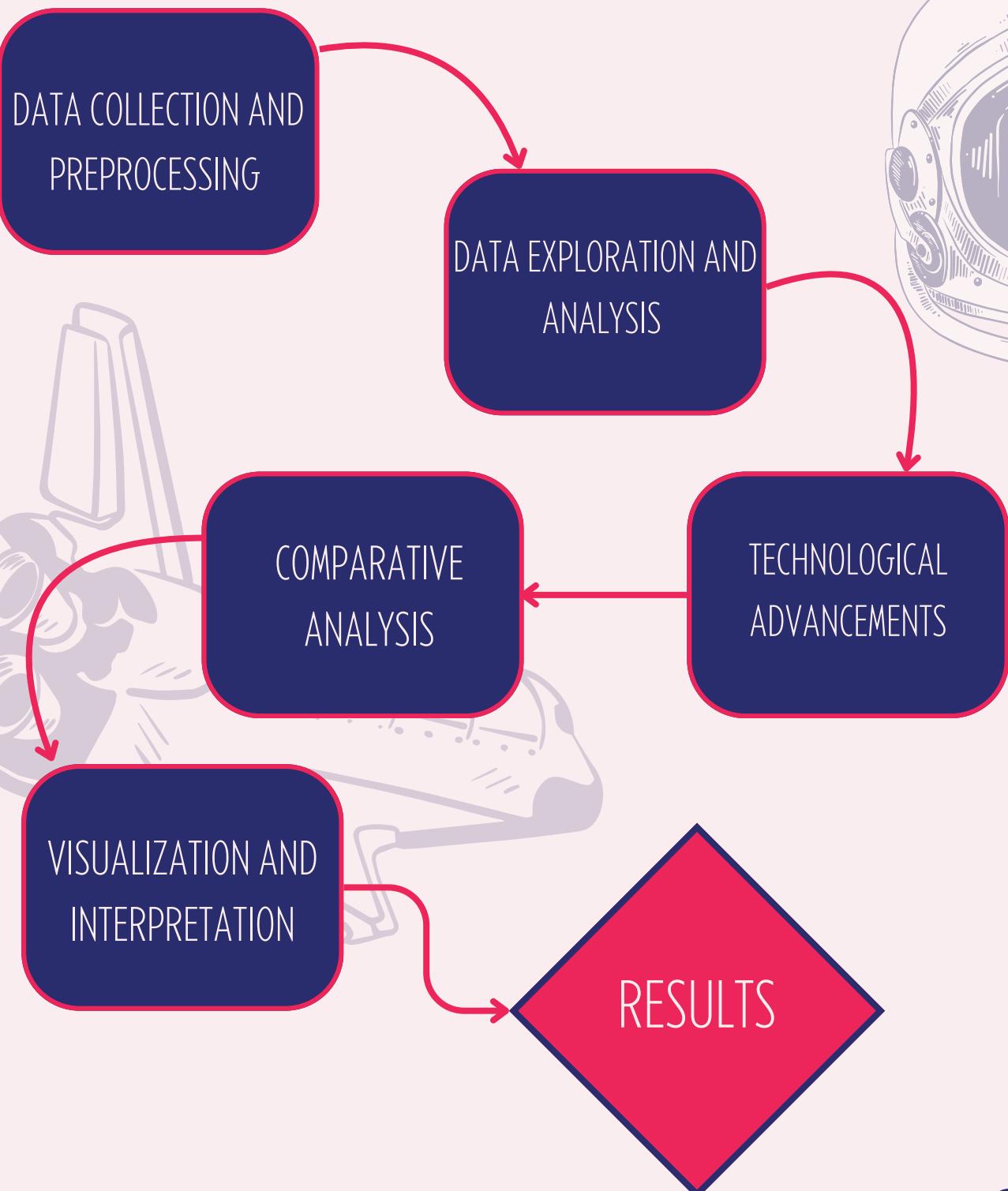
PROJECT OVERVIEW



The project titled "Space Mission from 1957 Data Analysis" by ML Team Zeta at ITSOLERA focuses on examining the pivotal Space Race between the Soviet Union (USSR) and the United States (US). This analysis aims to uncover key events and milestones of space missions from 1957 onwards, understand the technological advancements and their implications, and compare the achievements of the USSR and the US during the early years of the Space Race. The project employs a comprehensive methodology, including data collection and preprocessing, exploratory data analysis, examination of technological developments, comparative analysis, and data visualization. Tools such as Python, Pandas, NumPy, Matplotlib, Seaborn, and Jupyter Notebook are used for data manipulation, analysis, and visualization. Expected outcomes include detailed insights into space missions, technological advancements, comparative achievements, and visualizations that effectively communicate trends and milestones. The project's objective is to provide a deeper understanding of the competitive dynamics and progression of space exploration, highlighting the impact of early space missions on subsequent advancements in spaceflight technology.



METHODOLOGY



DATA COLLECTION AND PREPROCESING



The data collection phase involved sourcing historical data on space missions starting from 1957. The primary focus was on significant launches by both the USSR and the US during the early years of the Space Race.

- Mission Names: The official names of space missions.
- Launch Dates: The specific dates when the missions were launched.
- Countries Involved: The nations that conducted the space missions.
- Mission Types: The categorization of missions (e.g., satellite, probe, human spaceflight).
- Technological Milestones: Key technological advancements and achievements associated with each mission.
- Success Rates: Information on whether the missions were successful, partially successful, or failed.

After collecting the raw data, the preprocessing phase was initiated to clean and organize the data into a structured format suitable for analysis. The steps involved were:

· Data Cleaning:

- **Handling Missing Values:** Identifying and addressing missing or incomplete data entries.

· Data Formatting:

- **Standardizing Date Formats:** Ensuring all dates are in a consistent format for accurate temporal analysis.
- **Normalizing Country Names:** Standardizing country names to avoid discrepancies caused by different naming conventions.

· Data Structuring:

- **Organizing Data into Tables:** Structuring the cleaned data into well-organized tables with clearly defined columns and rows.

· Data Transformation:

- **Creating New Variables:** Generating additional variables or metrics derived from the existing data to aid in analysis (e.g., calculating the duration between launches).
- **Aggregating Data:** Summarizing data at different levels (e.g., yearly totals) to facilitate trend analysis

LINK:<https://www.kaggle.com/datasets/agirlcoding/all-space-missions-from-1957>

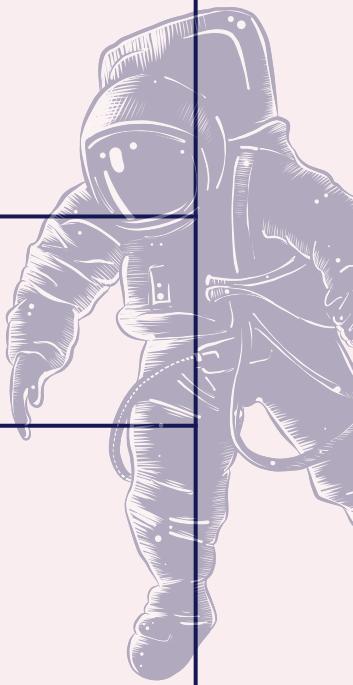


DATA EXPLORATION AND ANALYSIS:

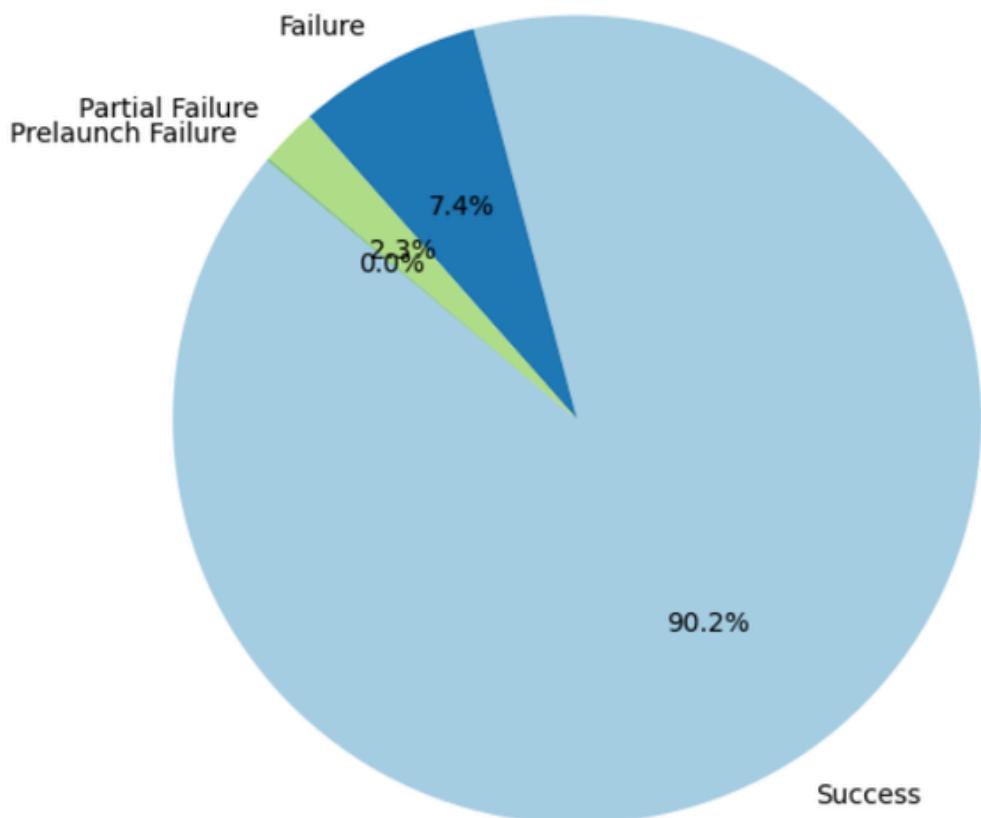


The section on Data Exploration and Analysis outlines several methods for understanding space mission data starting in 1957. First, it calculates the global mission success and failure rates, creating counts and displaying them as a pie chart. After that, the research changes to looking at the percentage of missions carried out by various nations, once more displaying the results with a pie chart. Success Rate and Weighted Performance Index (WPI), which takes into account both the success rate and the total number of missions, are two new measures that are designed in order to evaluate the performance of specific countries. A scatter plot is used to visualize the performance of each country, with the x-axis representing the total number of missions, the y-axis representing the success rate, and the size of the points indicating the Weighted Performance Index. Finally, a function plots the mission status distribution for each country, displaying individual pie charts for each country's mission outcomes, along with relevant mission statistics in the legend. This comprehensive exploration provides deep insights into the success rates and contributions of different countries in space missions.

DATA EXPLORATION AND ANALYSIS:



Frequency Distribution of Mission Status from 1957 onwards



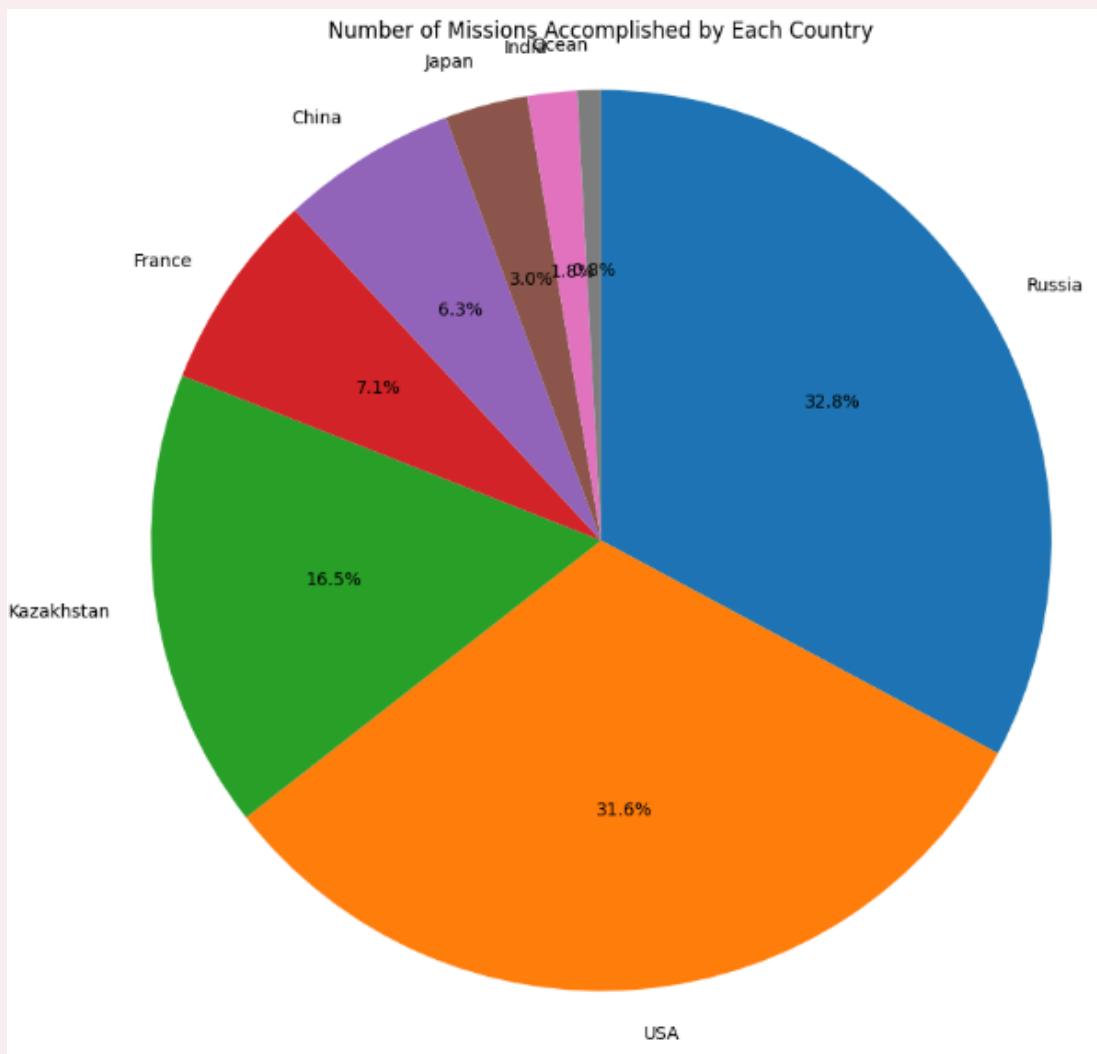
Total number of missions: 4249
Success: 3834
Failure: 314
Prelaunch Failure : 2
Partial Failure: 99

CONFIDENTIAL





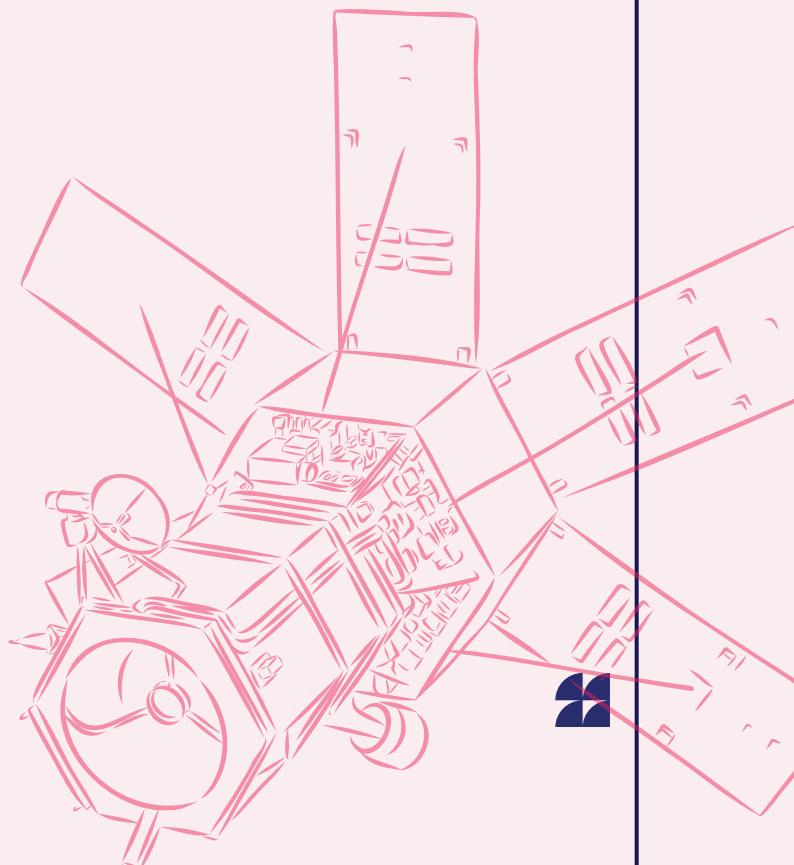
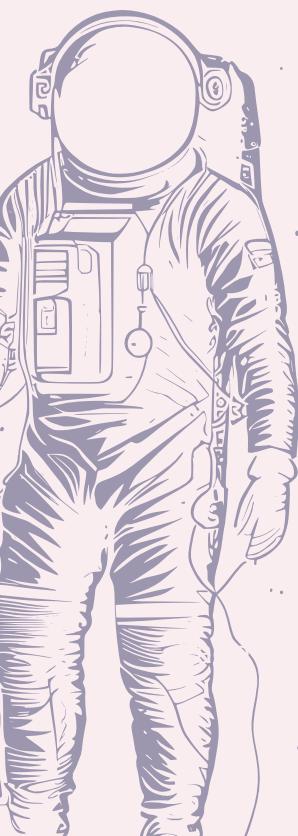
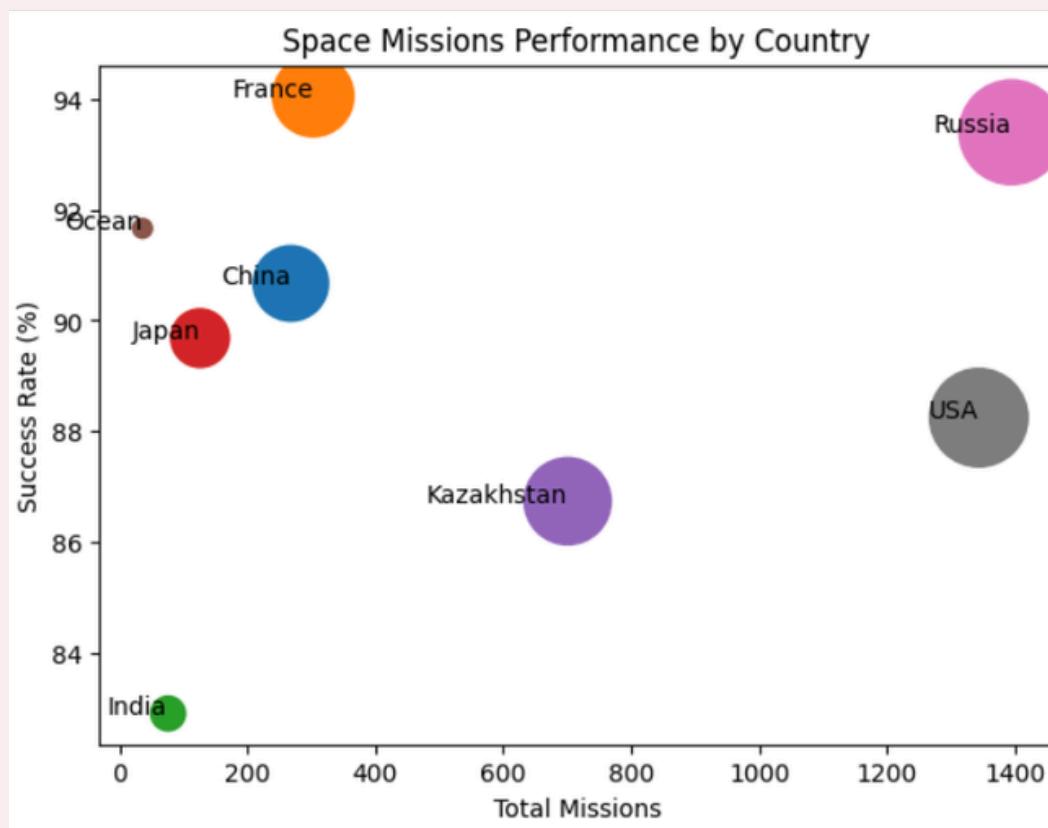
DATA EXPLORATION AND ANALYSIS:



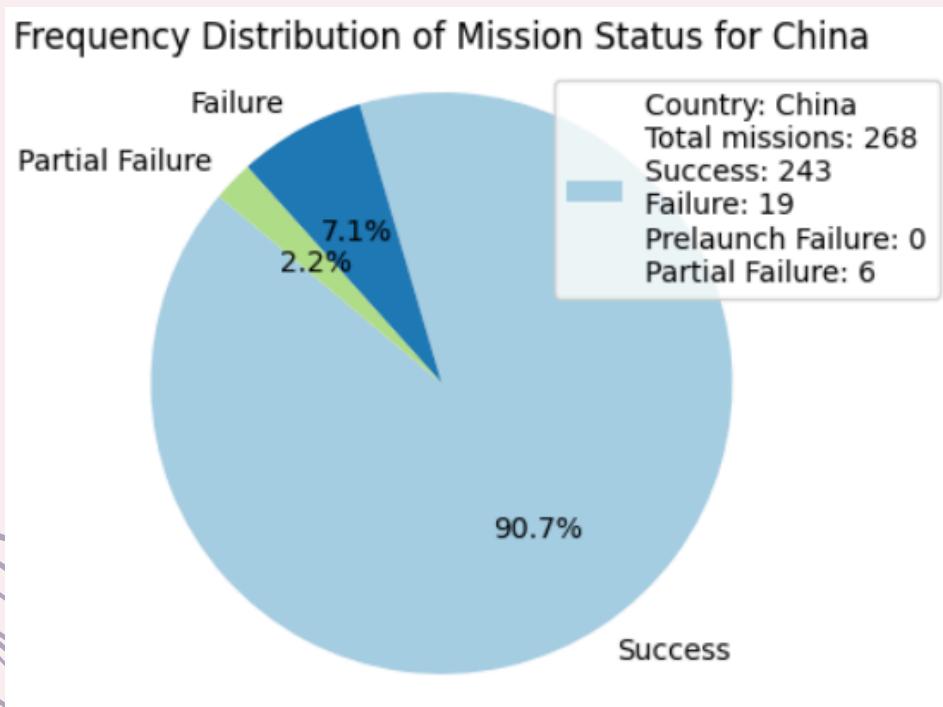
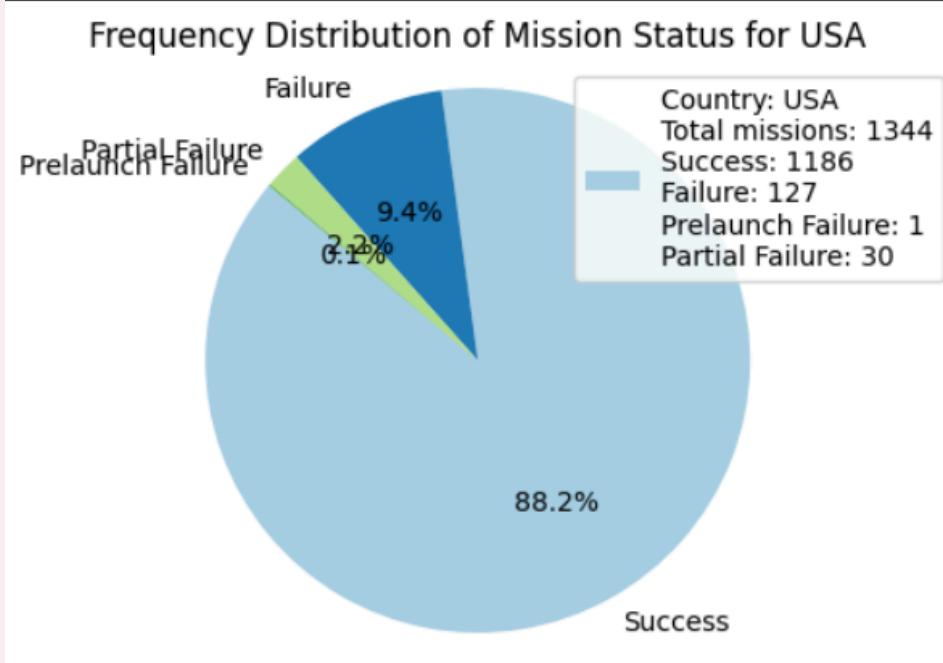
	Country Name	Success	Failure	Prelaunch Failure	Partial Failure	Total Missions	Success Rate	WPI
0	China	243	19	0	6	268	90.671642	507.281666
1	France	285	13	0	5	303	94.059406	537.740229
2	India	63	8	0	5	76	82.894737	360.078607
3	Japan	113	10	0	3	126	89.682540	434.439001
4	Kazakhstan	608	72	1	20	701	86.733238	568.443867
5	Ocean	33	3	0	0	36	91.666667	331.000809
6	Russia	1303	62	0	30	1395	93.405018	676.379947
7	USA	1186	127	1	30	1344	88.244048	635.723293



DATA EXPLORATION AND ANALYSIS:

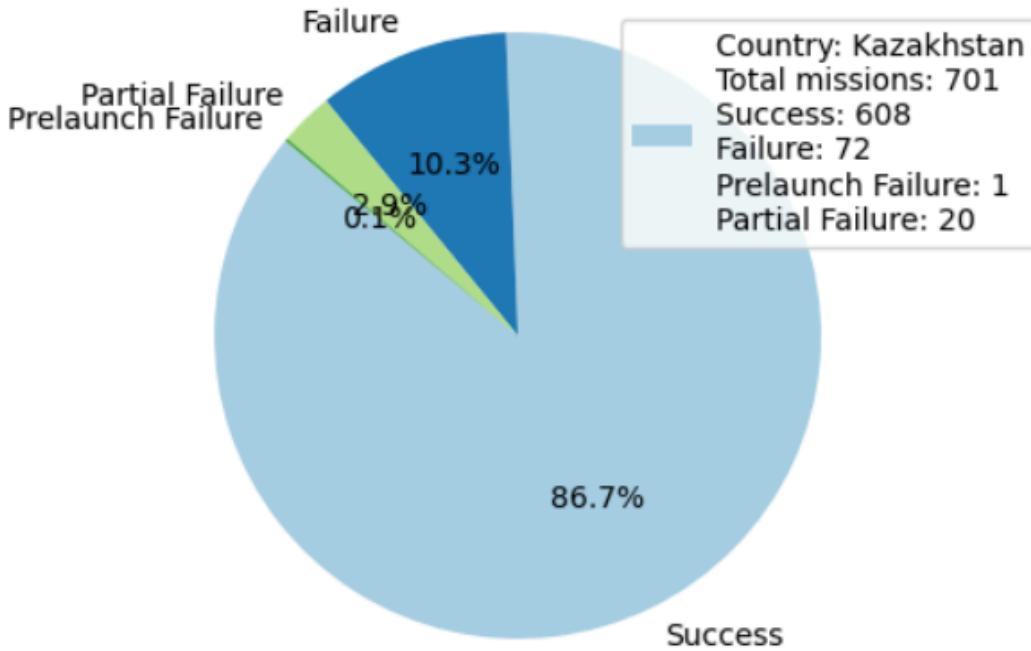


FREQUENCY DISTRIBUTION

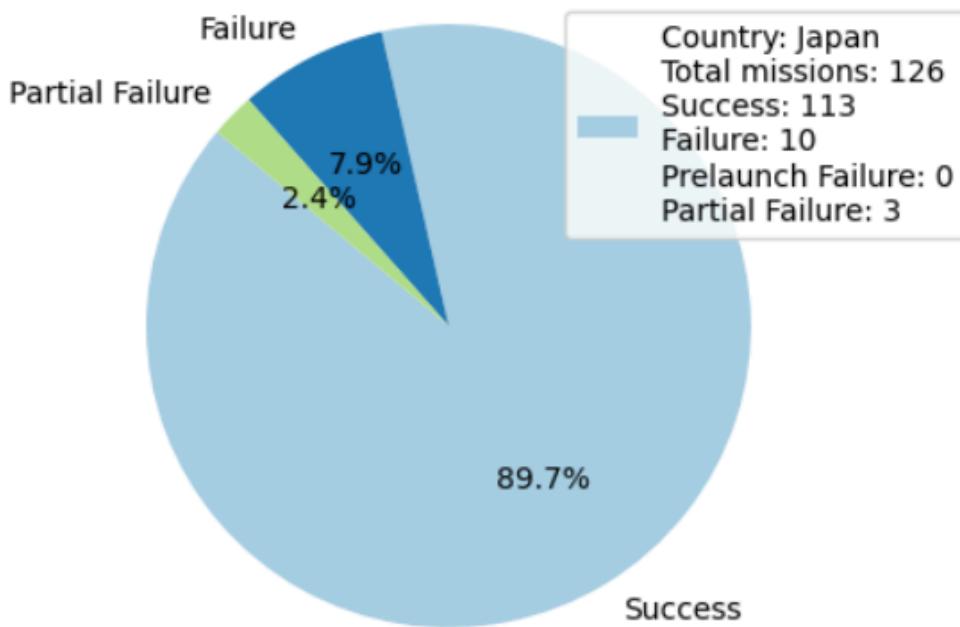


FREQUENCY DISTRIBUTION

Frequency Distribution of Mission Status for Kazakhstan

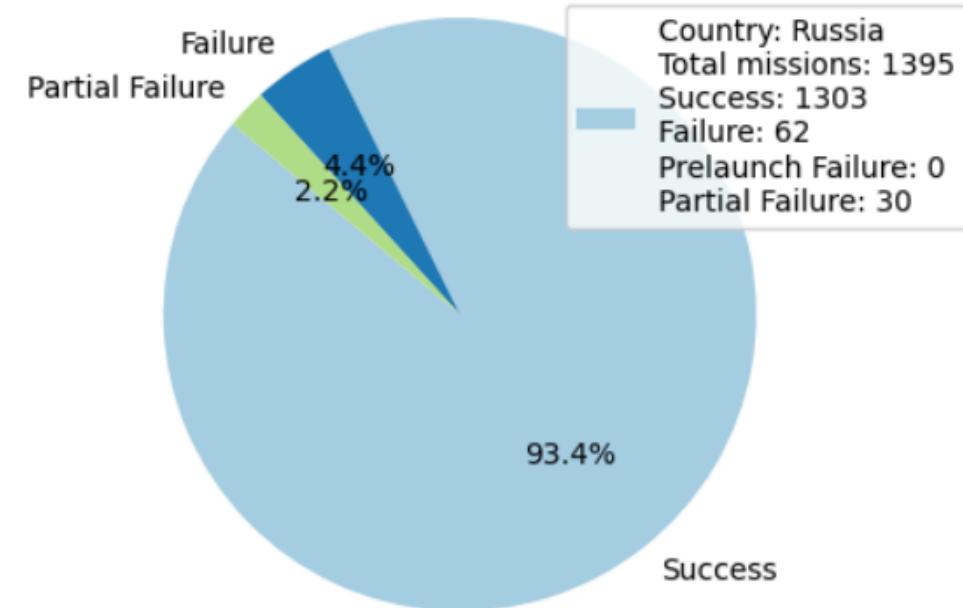


Frequency Distribution of Mission Status for Japan

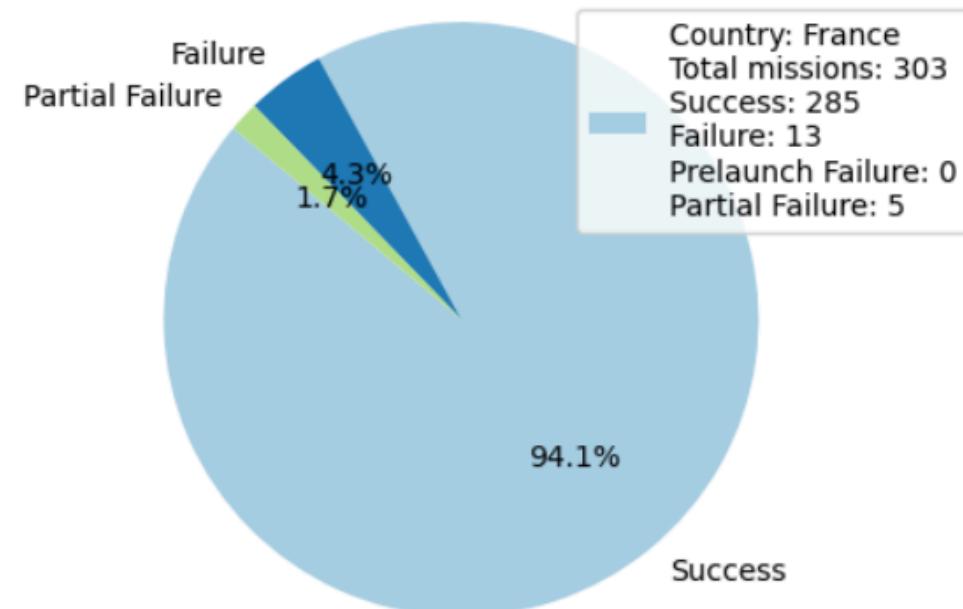


FREQUENCY DISTRIBUTION

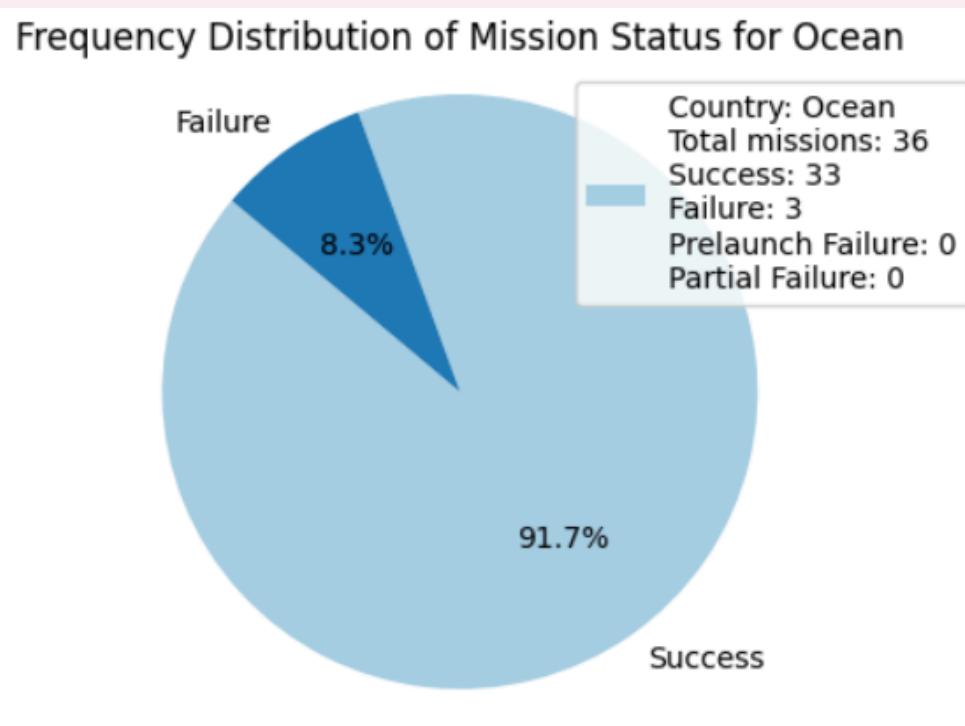
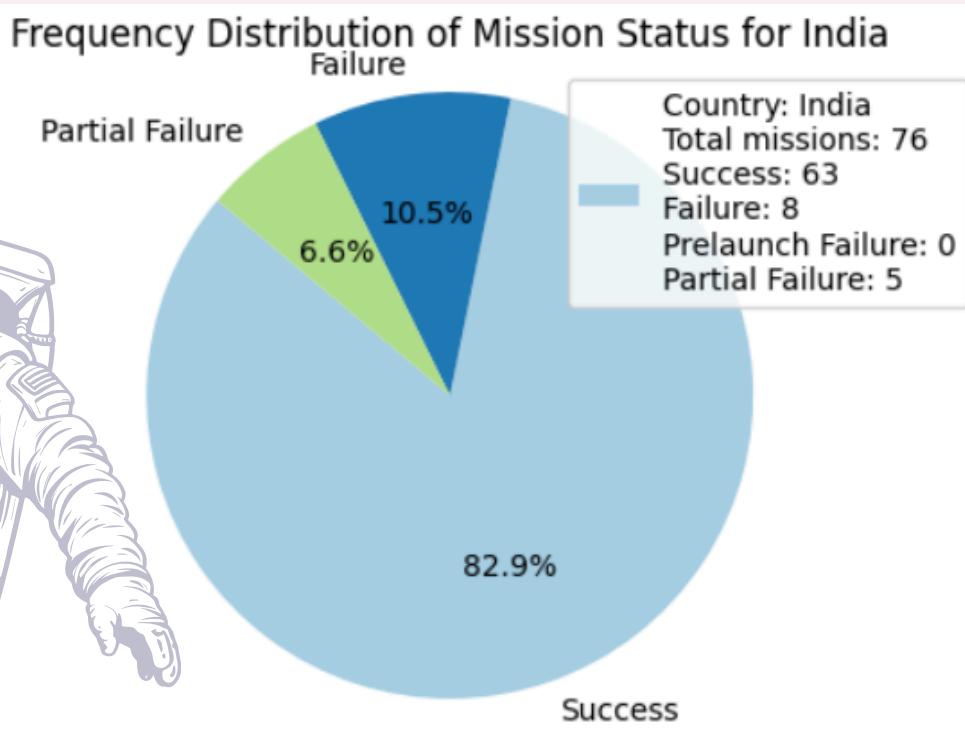
Frequency Distribution of Mission Status for Russia



Frequency Distribution of Mission Status for France



FREQUENCY DISTRIBUTION



TECHNOLOGICAL ADVANCEMENTS



The code's Technological Advancements section thoroughly examines the advancements in technology and the contributions made by different nations and businesses to space missions starting in 1957. Here's a thorough explanation:

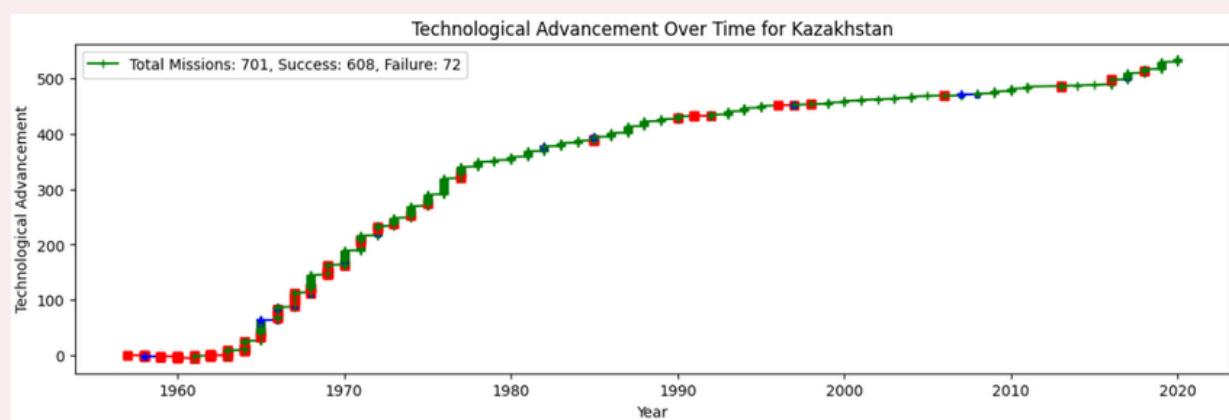
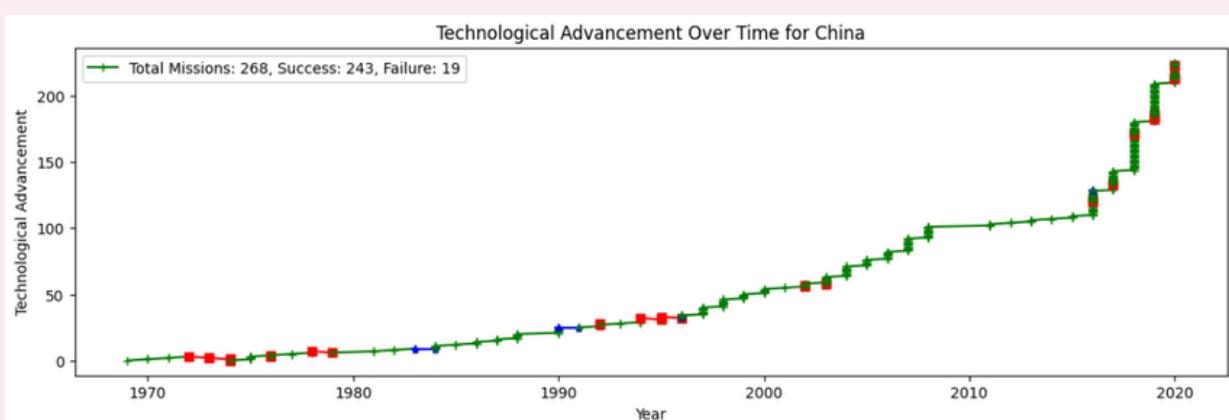
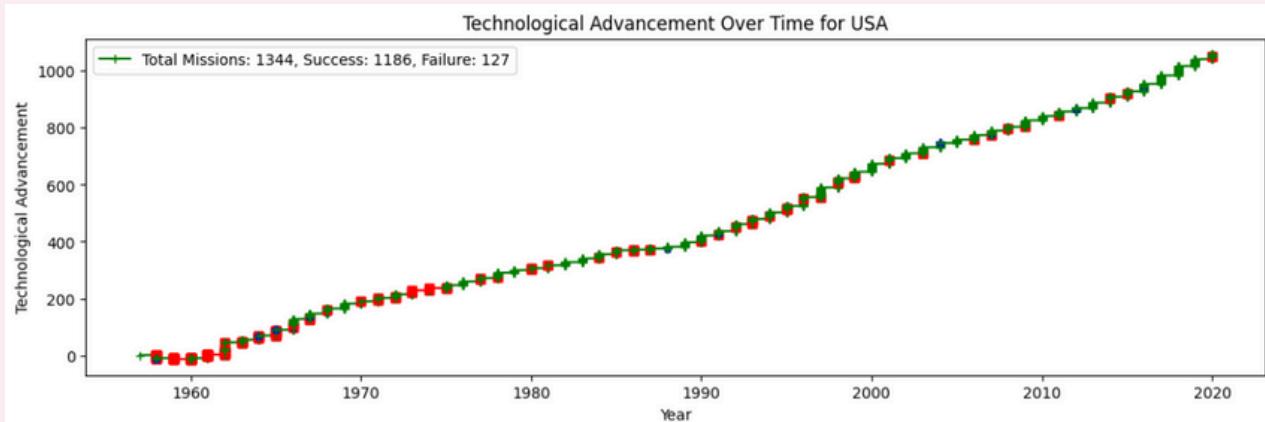
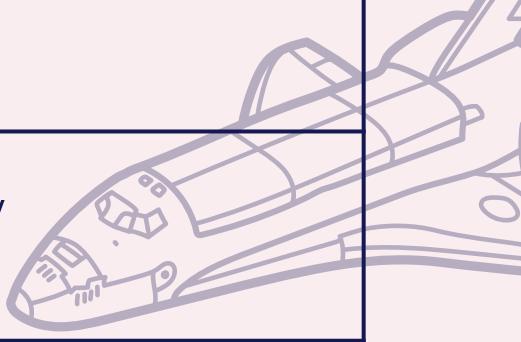
TECHNOLOGICAL DEVELOPMENT BY COUNTRY

Function: plot_technological_advancement(country_name). The goal is to illustrate how every country has progressed technologically over time.

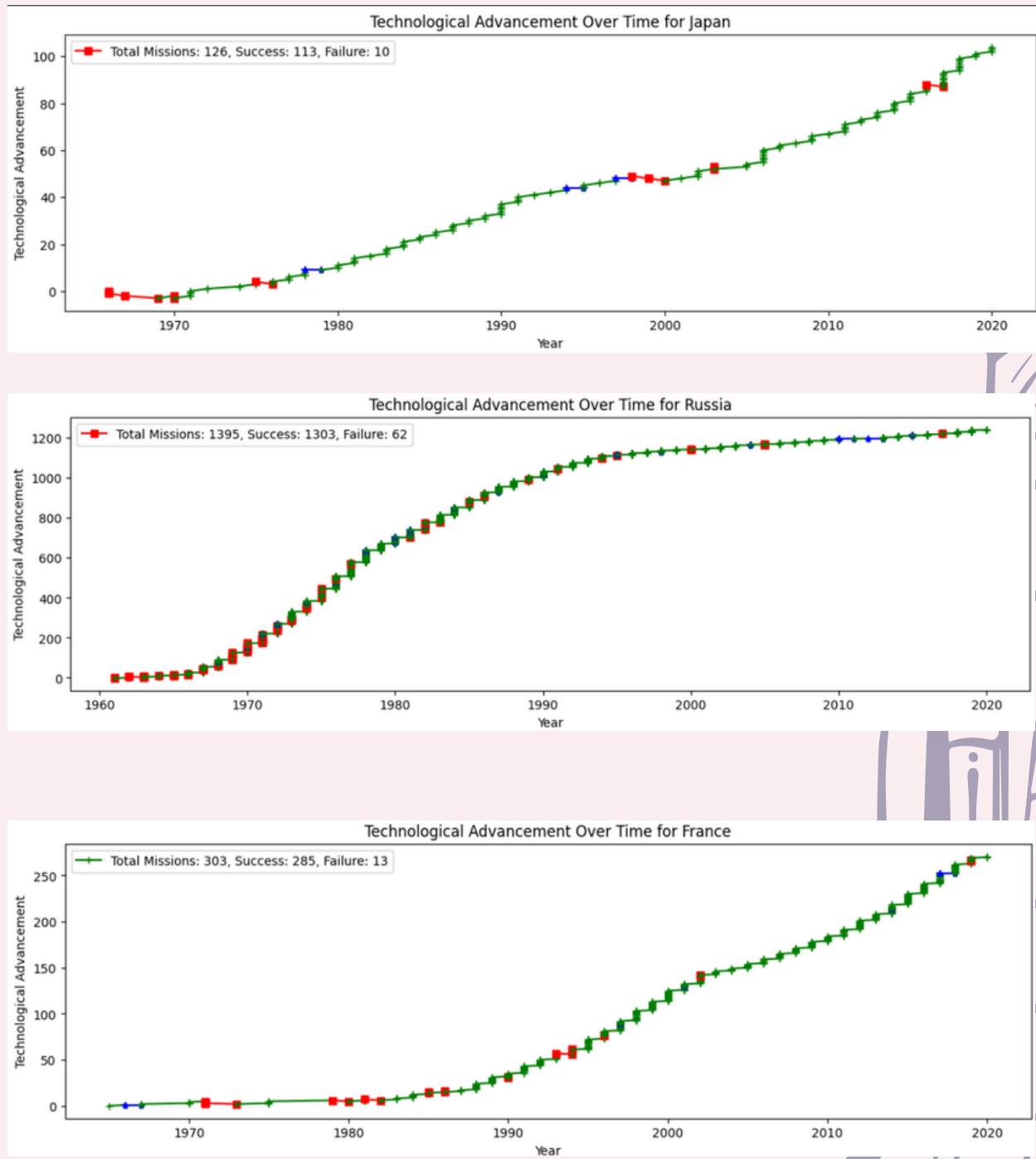
STEPS

- **Data Filtering:** This method restricts the missions from a given country by filtering the Data Frame (df1).
- **Date Processing:** This method extracts the year by converting the 'date' column to datetime format.
- **Sorting:** To guarantee chronological charting, the data is sorted by date.
- The 'tech_advancement' column is initialized and then incremented by 1 for each successful mission and decremented by 1 for each unsuccessful mission. This calculates the technological advancement.
- **Plotting:** Creates a line plot where the x-axis represents the year, and the y-axis represents the cumulative technological advancement. Different colors indicate success, failure, and other statuses.
- **Legend and Conclusion:** Adds a legend showing the total number of missions, successes, and failures. It also prints whether the country is advancing technologically based on the net trend.
- **Usage:** Loops through each country in df1 and generates the respective plots.

Technological Development by Country

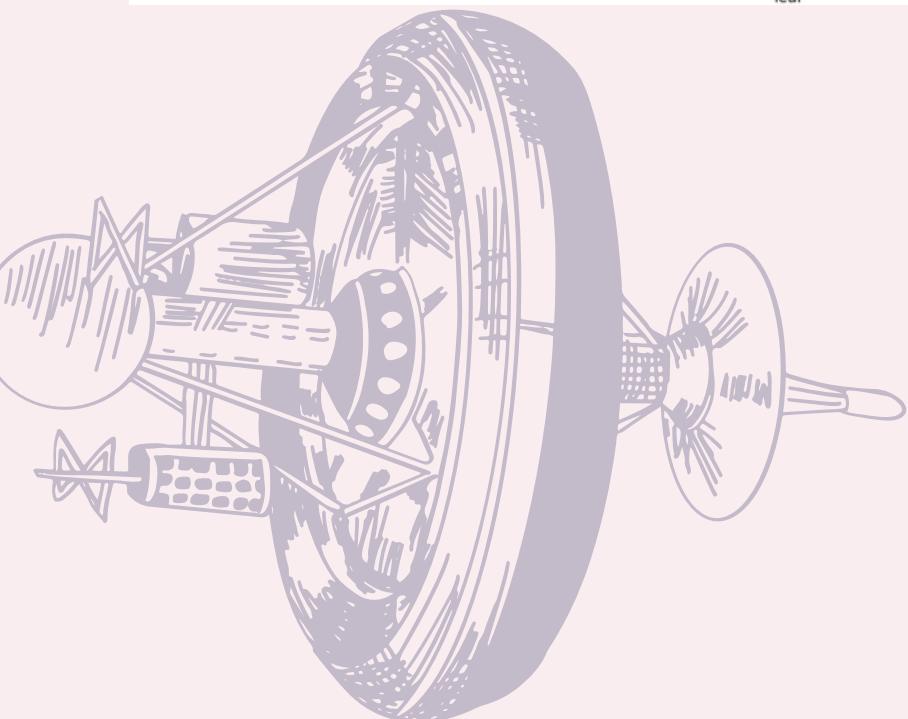
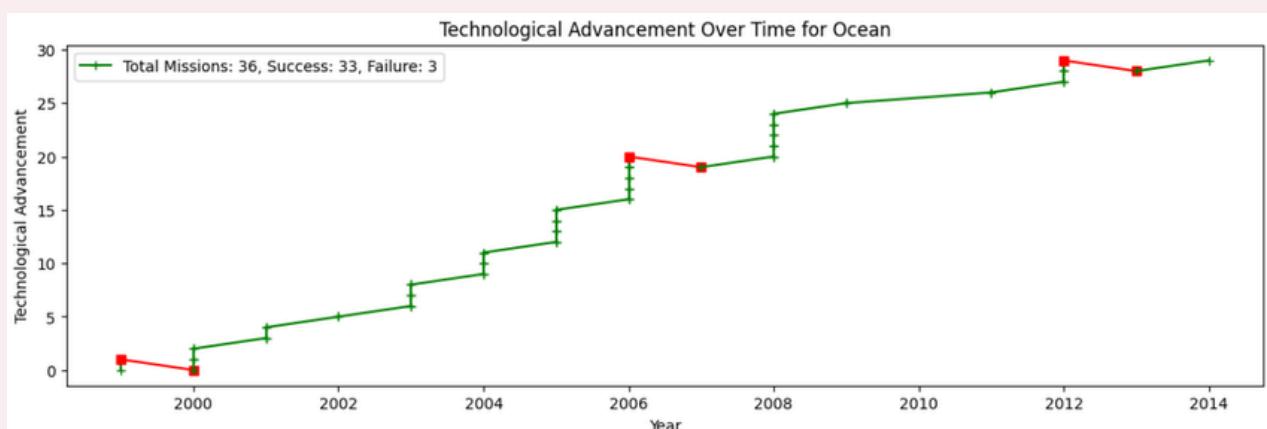
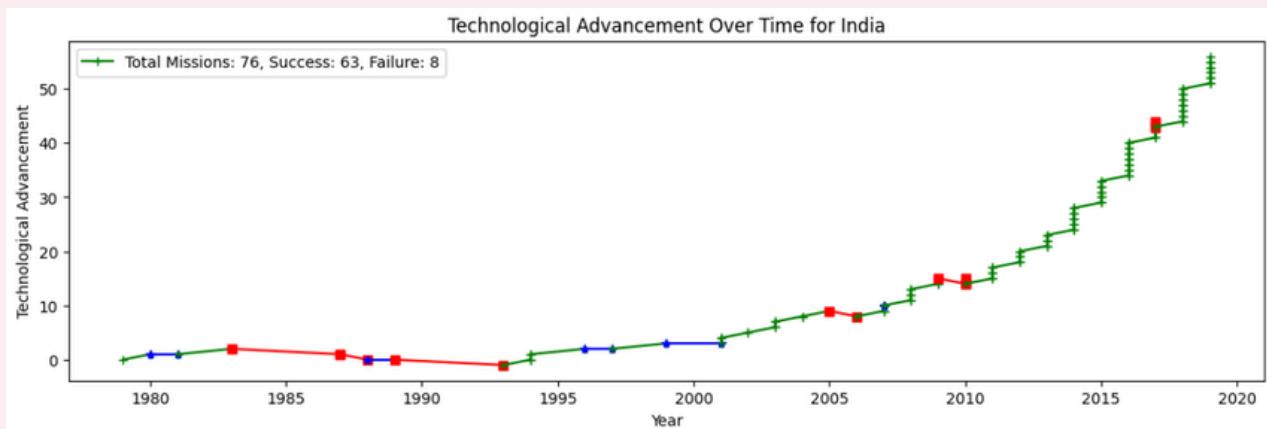
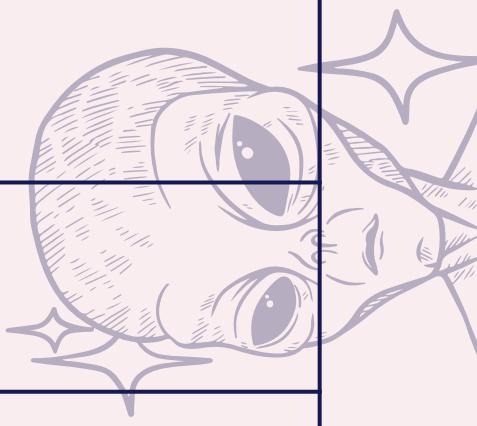


Technological Development by Country

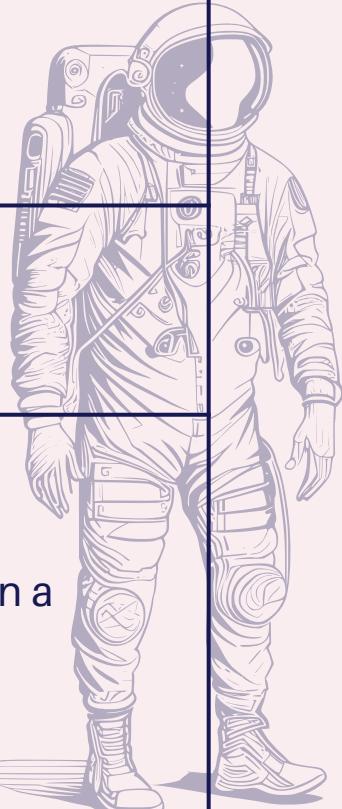


SpaceItem from 1957 Data Analysis

TECHNOLOGICAL DEVELOPMENT BY COUNTRY



CONTRIBUTIONS OF COMPANIES WITHIN EACH COUNTRY



Function: plot_company_missions(country_name)

Purpose: To visualize the contributions of companies within a specific country to space missions.

STEPS

Data Filtering: Filters the Data Frame to include only the missions from the specified country.

Grouping and Counting: Groups the data by 'Company Name' to count the total number of missions, successes, and failures.

Merging Data Frames: Merges the data to create a single Data Frame containing total missions, successes, and failures for each company.

Plotting: Creates horizontal bar charts where each bar represents a company, coloured segments show successes (green) and failures (red), and the text inside the bars shows the counts.

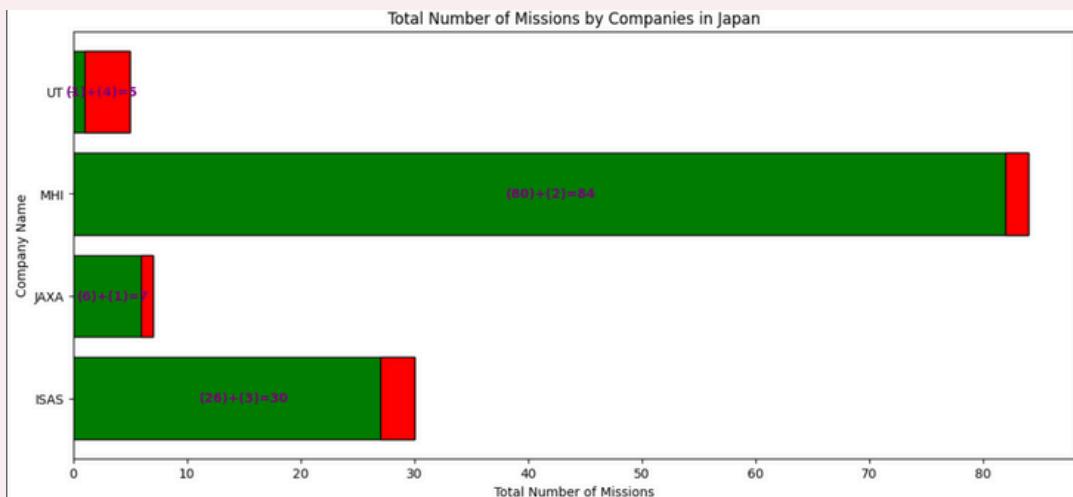
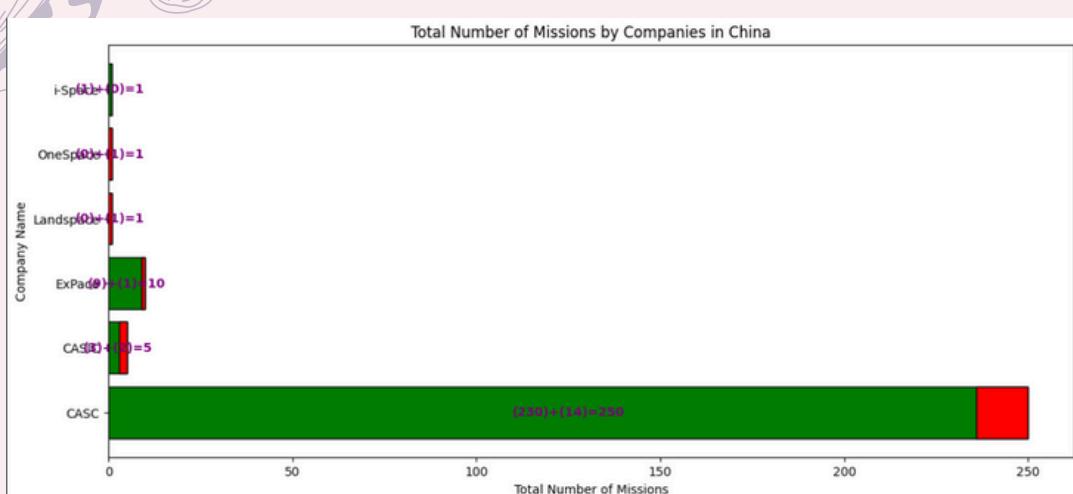
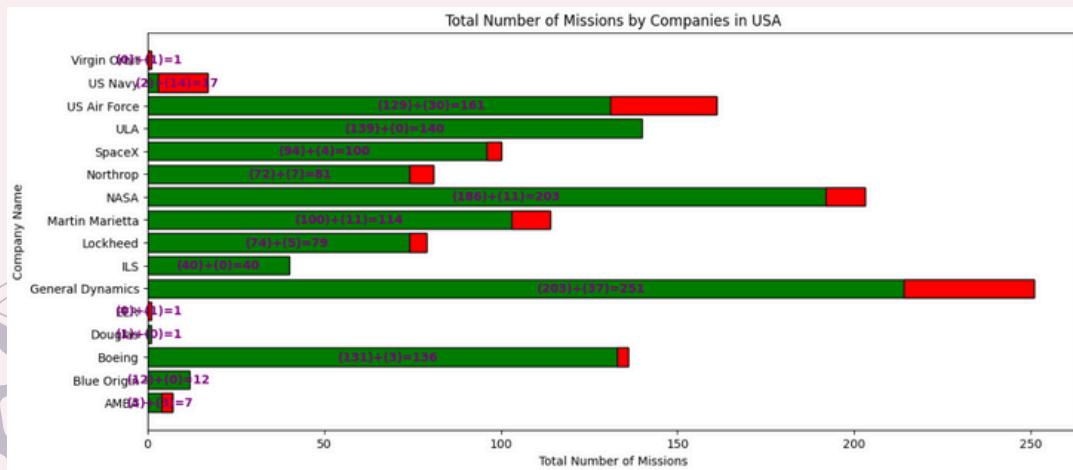
Details Display: Displays detailed text annotations within the bars for better clarity.

Usage: Loops through each country and generates the respective plots.



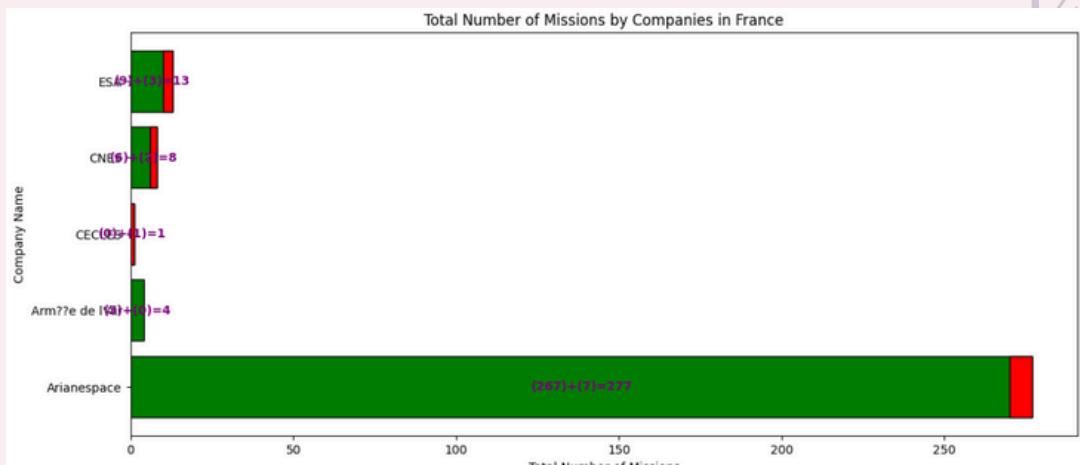
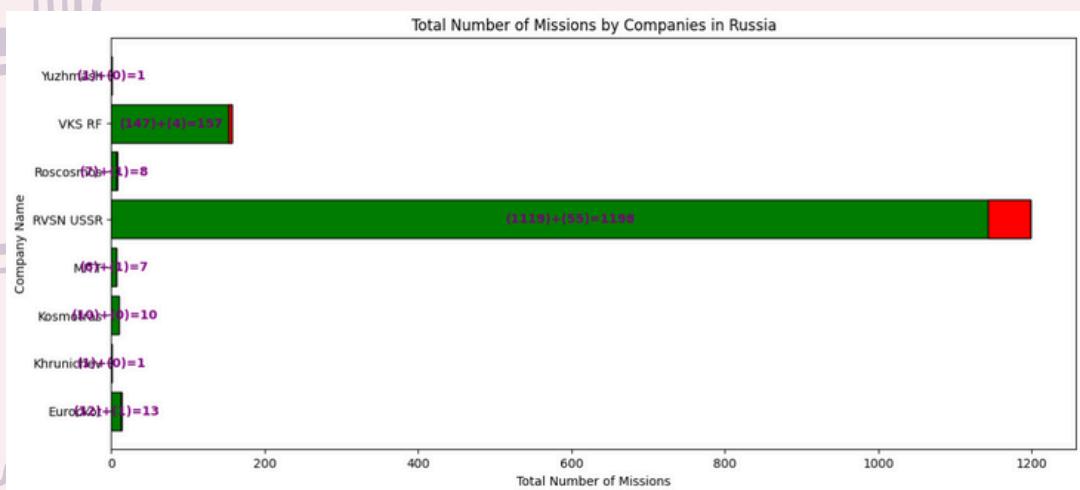
SpaceItem FROM 1957 DATA ANALYSIS

CONTRIBUTIONS OF COMPANIES WITHIN EACH COUNTRY



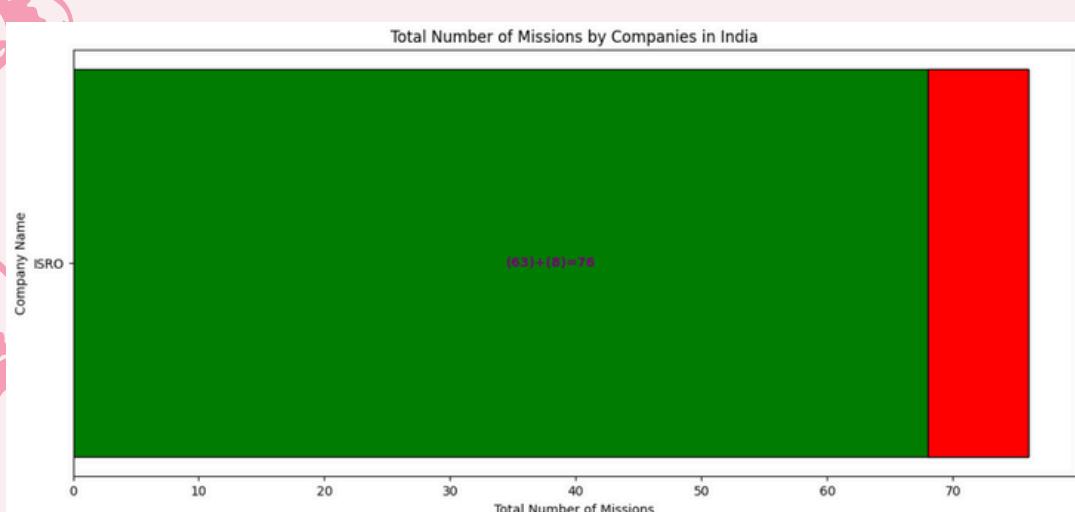
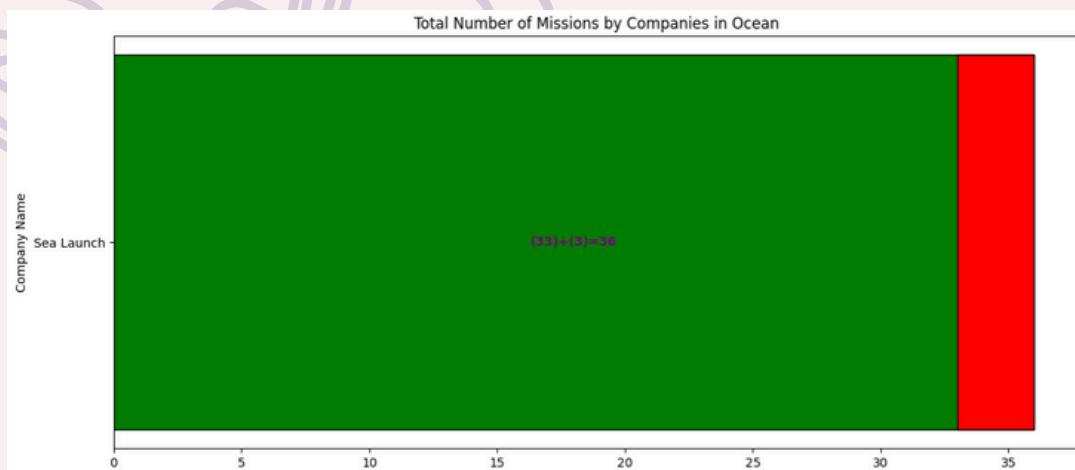
SPACE MISSION FROM 1957 DATA ANALYSIS

CONTRIBUTIONS OF COMPANIES WITHIN EACH COUNTRY



*SPACE MISSION FROM 1957 DATA ANALYSIS

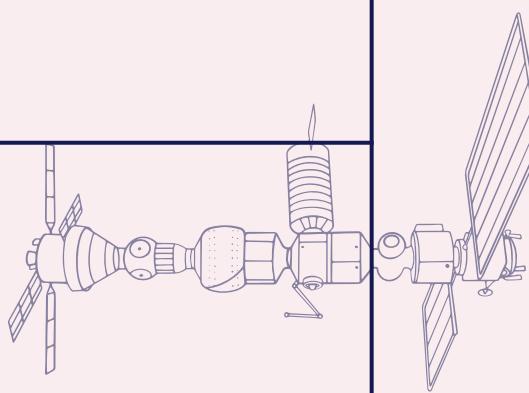
CONTRIBUTIONS OF COMPANIES WITHIN EACH COUNTRY



CONTRIBUTIONS OF COMPANIES WITHIN EACH COUNTRY

Function: plot_company_missions(country_name)

STEPS



Purpose: To visualize the contributions of companies within a specific country to space missions)

Data Filtering: Filters the Data Frame to include only the missions from the specified country. **Grouping and Counting:** Groups the data by 'Company Name' to count the total number of missions, successes, and failures.

Merging Data Frames: Merges the data to create a single Data Frame containing total missions, successes, and failures for each company.

Plotting: Creates horizontal bar charts where each bar represents a company, coloured segments show successes (green) and failures (red), and the text inside the bars shows the counts.

Details Display: Displays detailed text annotations within the bars for better clarity. **Usage:** Loops through each country and generates the respective plots.

COMPANY MISSIONS OVERVIEW

Function: plot_company_missions(company_data, title) and generate_plots(df1)

STEPS

Purpose: To provide a comprehensive overview of the contributions of companies to space missions, categorized by the number of missions they have accomplished.

Steps:

Grouping and Counting: Similar to the previous function, it groups the data by 'Company Name' and counts total missions, successes, and failures.

Filtering: Filters companies into different categories based on the number of missions:

More than 290 missions

Between 90 and 290 missions

Between 20 and 80 missions

Fewer than 20 missions

Plotting: Creates separate horizontal bar charts for each category, with detailed annotations inside the bars.

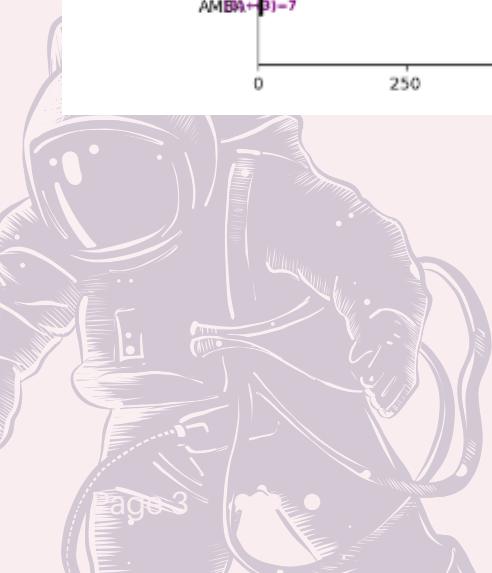
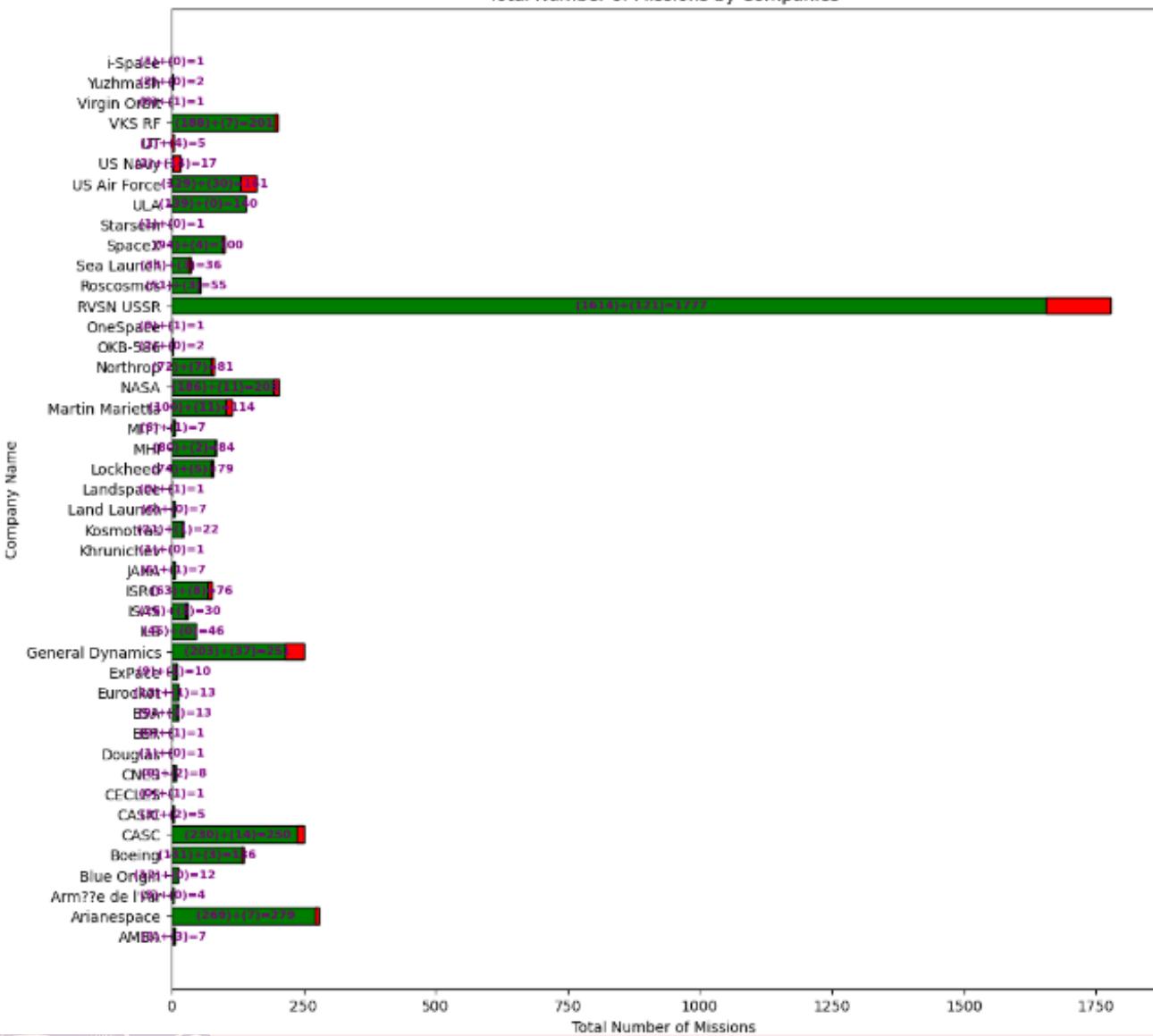
Customization: Each plot is titled according to the category of mission counts.

Usage: Calls the generate_plots function with df1 to generate the plots for all categories.



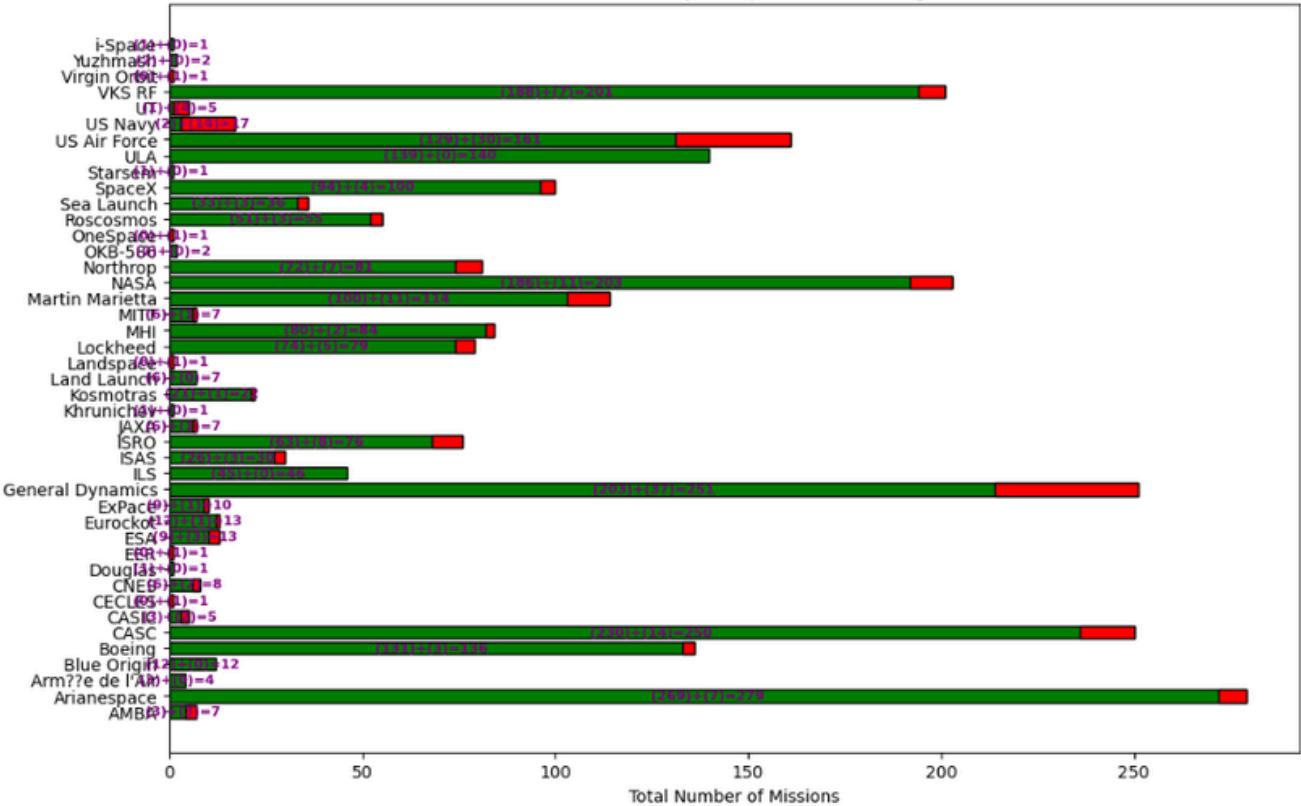
COMPANY MISSIONS OVERVIEW

Total Number of Missions by Companies

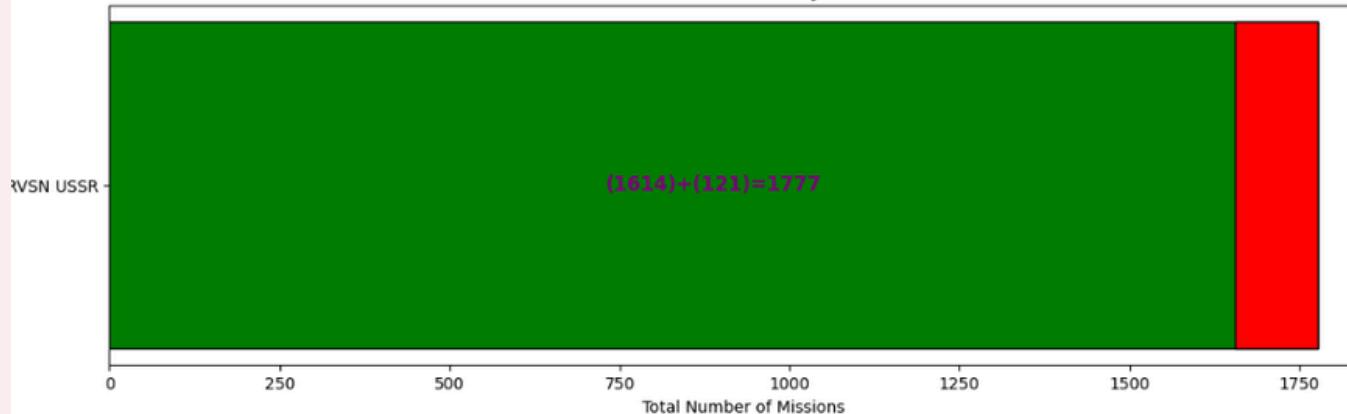


COMPANY MISSIONS OVERVIEW

Total Number of Missions by Companies (Excluding RVSN USSR)

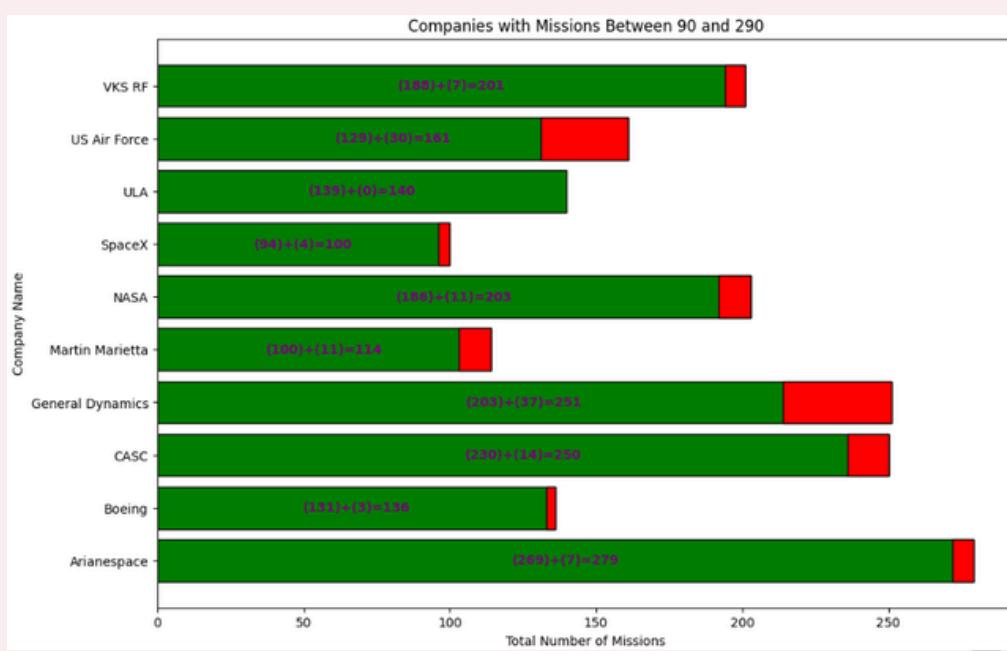
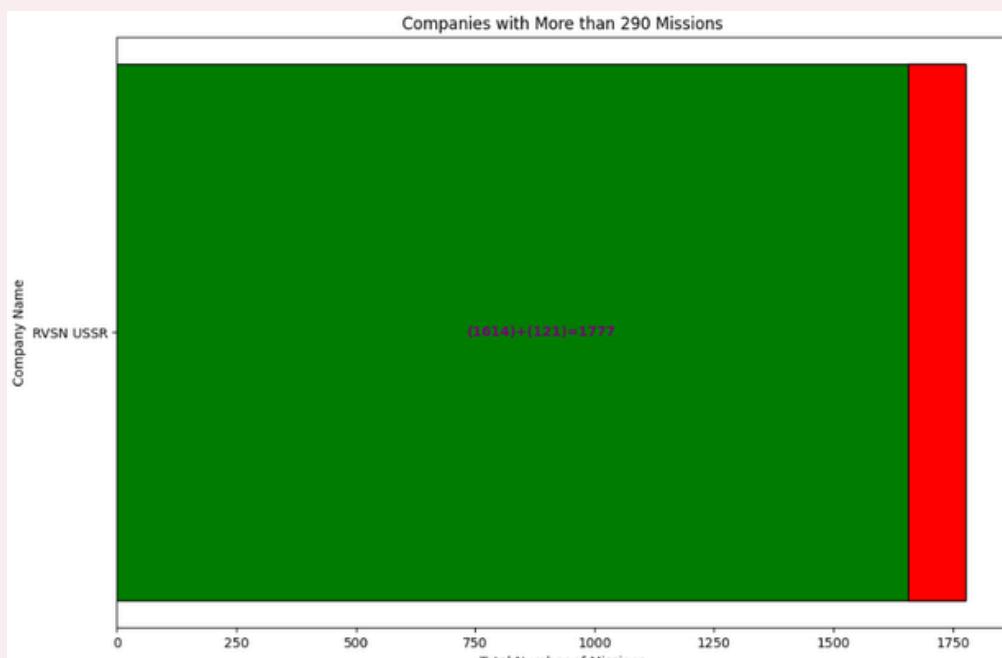


Total Number of Missions by RVSN USSR

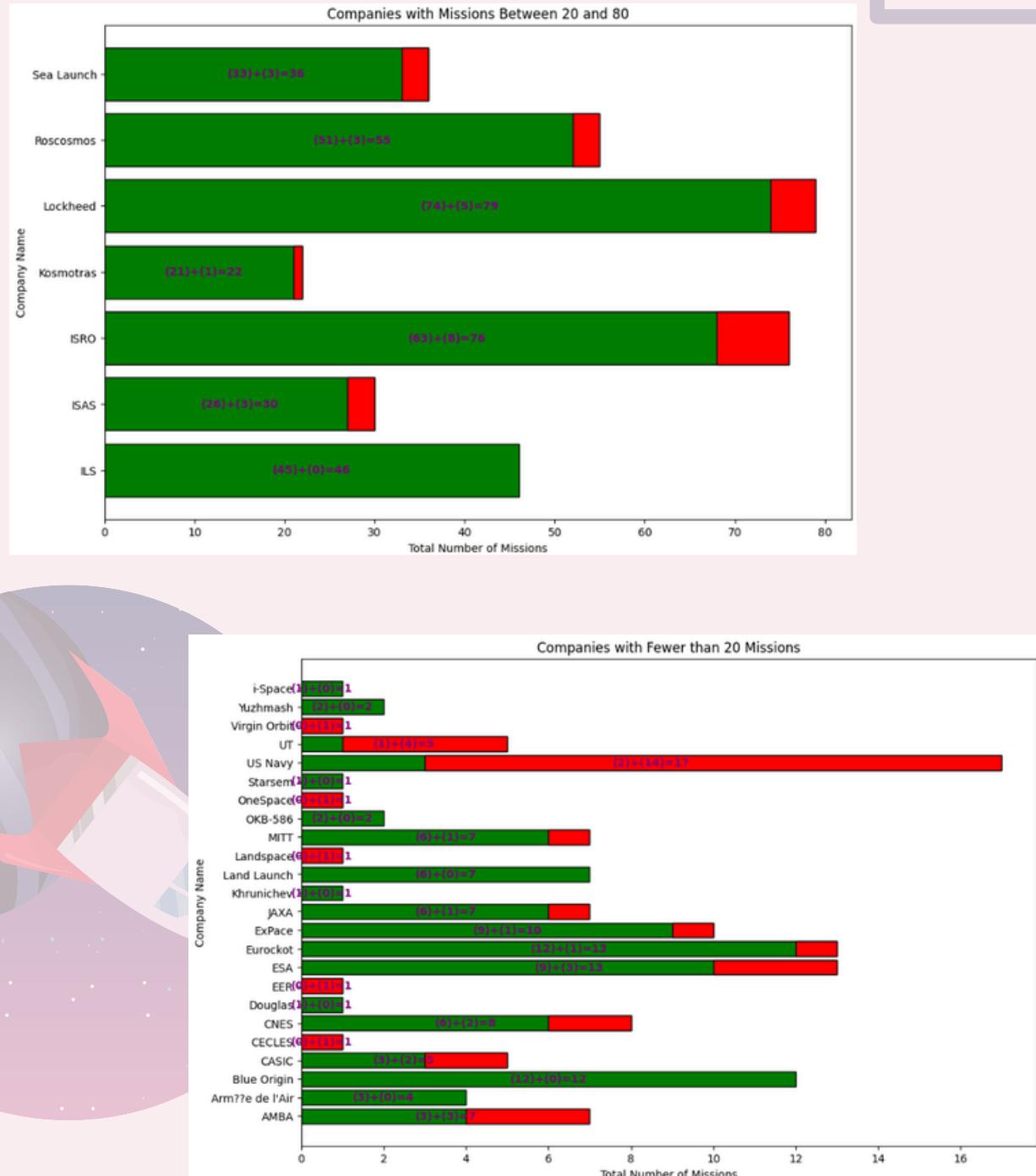


COMPANY MISSIONS OVERVIEW

FILTERING



COMPANY MISSIONS OVERVIEW



*SPACE MISSION FROM 1957 DATA ANALYSIS

SUMMARY OF TECHNOLOGICAL ADVANCEMENTS



ROCKET
PROPULSION

SPACECRAFT
DESIGN

NAVIGATION
AND CONTROL
SYSTEMS

COMMUNICATION
TECHNOLOGIES

SCIENTIFIC
INSTRUMENTS
AND
EXPERIMENTATION

HUMAN
SPACEFLIGHT

UPGRADE



COMPARATIVE ANALYSIS:



The comparative analysis revealed that the USSR led the early Space Race with milestones like Sputnik 1 and Yuri Gagarin's flight, while the US quickly advanced with significant investments, culminating in the Apollo 11 Moon landing. The USSR focused on incremental advancements, whereas the US pursued ambitious goals, driving rapid innovation and significant progress in space exploration for both nations.

DATA FILTERING:



The initial step involves filtering the primary Data Frame df1 to segregate the data based on the countries aligned with the USSR and the US. The arrays ussr_aligned and us_aligned list the countries in each group: USSR-aligned: Russia, Kazakhstan. US-aligned: USA, Japan, France. Using these arrays, the data is filtered into two separate DataFrames: ussr_data for USSR-aligned countries and us_data for US-aligned countries

```
> # Arrays containing countries aligned with the USSR and the US
  ussr_aligned = ['Russia', 'Kazakhstan']
  us_aligned = ['USA', 'Japan', 'France']

  # Step 1: Filter the Data
  ussr_data = df1[df1['country'].isin(ussr_aligned)]
  us_data = df1[df1['country'].isin(us_aligned)]
```



DATA AGGREGATION

The next step involves aggregating the filtered data to summarize the total missions, successes, and failures for each country within the two groups: The `ussr_summary` DataFrame is created by grouping `ussr_data` by country and applying aggregation functions to count the total missions and calculate the number of successes and failures. Similarly, the `us_summary` DataFrame is created for the US-aligned country

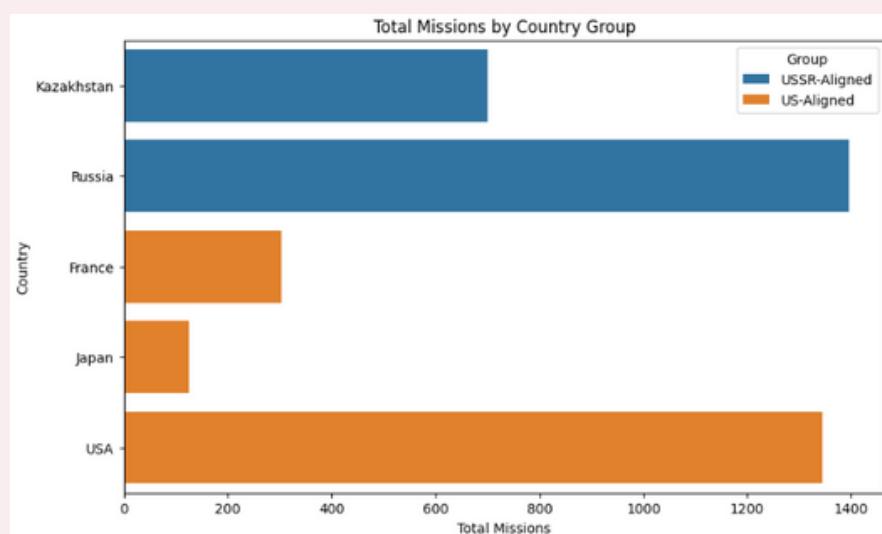
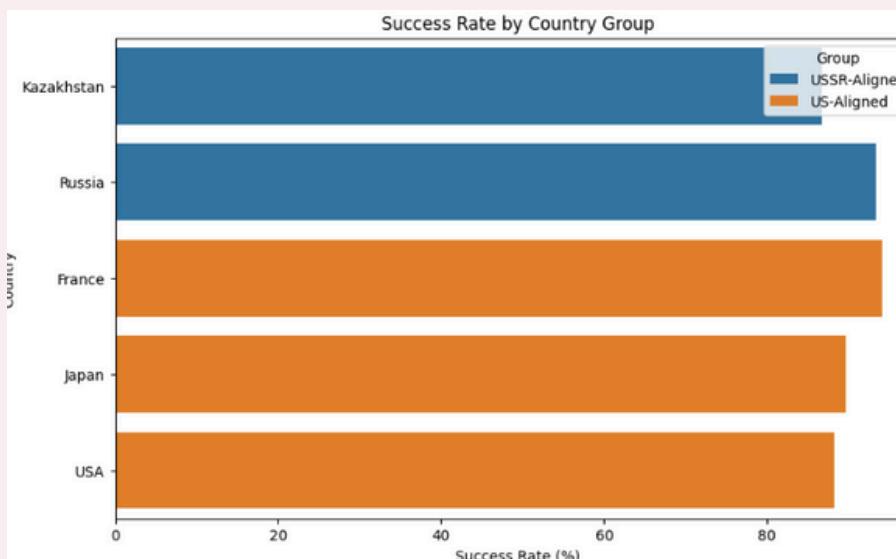
COMBINING SUMMARIES AND CALCULATING SUCCESS RATES

The aggregated summaries are combined into a single DataFrame `comparison_summary` using the `pd.concat()` function. The combined Data Frame includes columns for the group (USSR-aligned or US-aligned), country, total missions, successes, and failures. Additionally, a new column, Success Rate (%), is calculated to represent the success rate as a percentage of total missions.



VISUALIZATION

The aggregated summaries are combined into a single DataFrame `comparison_summary` using the `pd.concat()` function. The combined Data Frame includes columns for the group (USSR-aligned or US-aligned), country, total missions, successes, and failures. Additionally, a new column, Success Rate (%), is calculated to represent the success rate as a percentage of total missions.



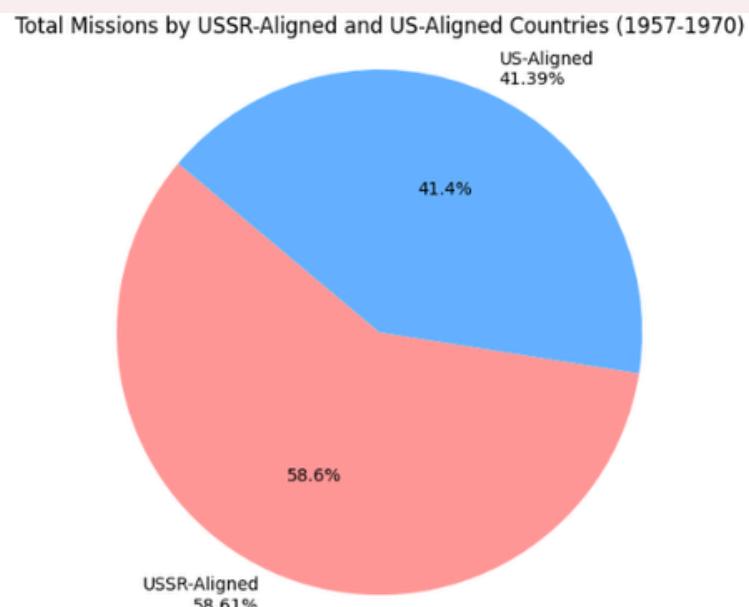
ADDITIONAL INSIGHTS

For each country, detailed insights are printed, showing the group, country name, total missions, successes, failures, and success rate percentage.

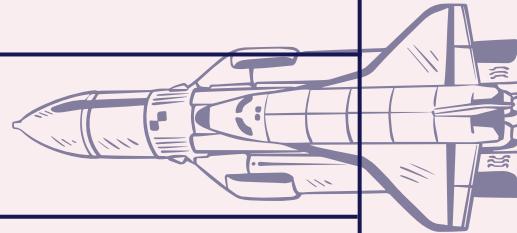
```
Group: USSR-Aligned, Country: Kazakhstan  
Total Missions: 701, Successes: 608, Failures: 72, Success Rate: 86.73%  
  
Group: USSR-Aligned, Country: Russia  
Total Missions: 1395, Successes: 1303, Failures: 62, Success Rate: 93.41%  
  
Group: US-Aligned, Country: France  
Total Missions: 303, Successes: 285, Failures: 13, Success Rate: 94.06%  
  
Group: US-Aligned, Country: Japan  
Total Missions: 126, Successes: 113, Failures: 10, Success Rate: 89.68%  
  
Group: US-Aligned, Country: USA  
Total Missions: 1344, Successes: 1186, Failures: 127, Success Rate: 88.24%
```

PIE CHART VISUALIZATIONS

To further enhance the comparative analysis, two additional pie charts are generated to show the distribution of missions and the success-failure ratio for both USSR-aligned and US-aligned countries within the specified date range (1957-1970). The function plot_pie_chart_ussr_vs_us creates a pie chart displaying the total number of missions conducted by USSR aligned and US-aligned countries. The chart includes percentage labels for each group to highlight their relative contributions.



ADDITIONAL INSIGHTS



For each country, detailed insights are printed, showing the group, country name, total missions, successes, failures, and success rate percentage.

```
Group: USSR-Aligned, Country: Kazakhstan
Total Missions: 701, Successes: 608, Failures: 72, Success Rate: 86.73%

Group: USSR-Aligned, Country: Russia
Total Missions: 1395, Successes: 1303, Failures: 62, Success Rate: 93.41%

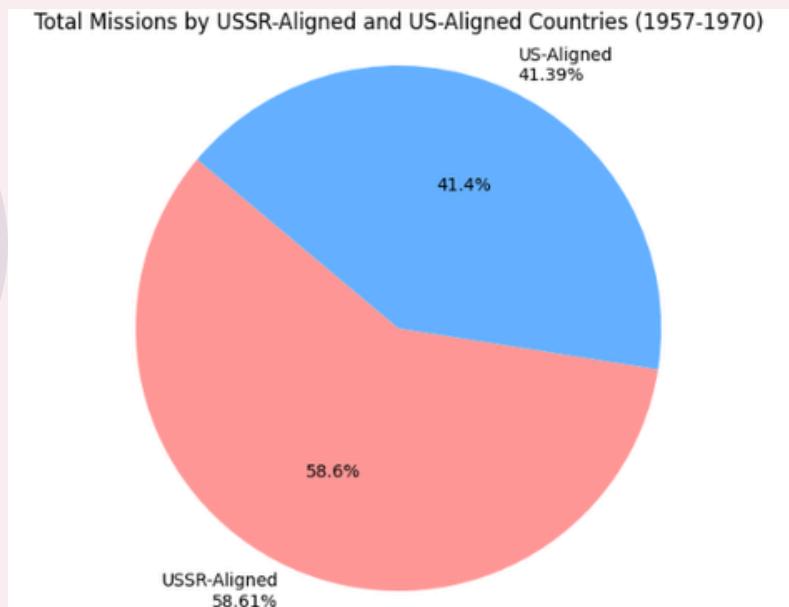
Group: US-Aligned, Country: France
Total Missions: 303, Successes: 285, Failures: 13, Success Rate: 94.06%

Group: US-Aligned, Country: Japan
Total Missions: 126, Successes: 113, Failures: 10, Success Rate: 89.68%

Group: US-Aligned, Country: USA
Total Missions: 1344, Successes: 1186, Failures: 127, Success Rate: 88.24%
```

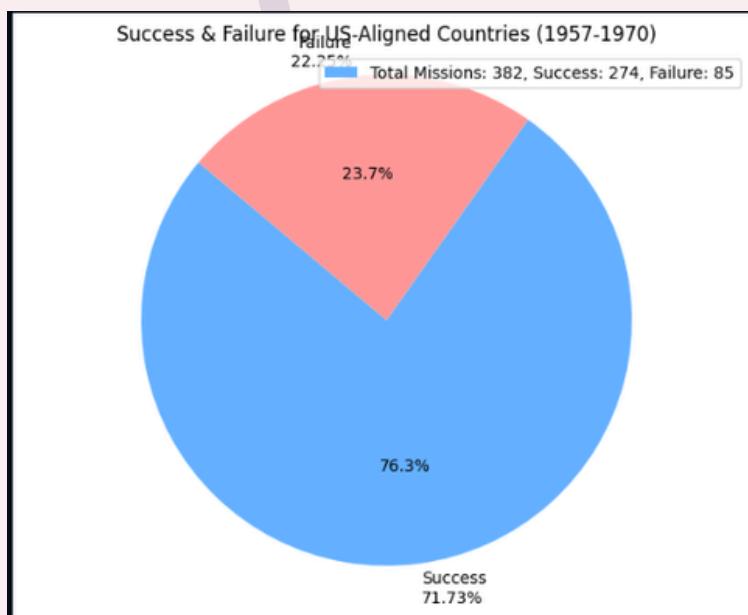
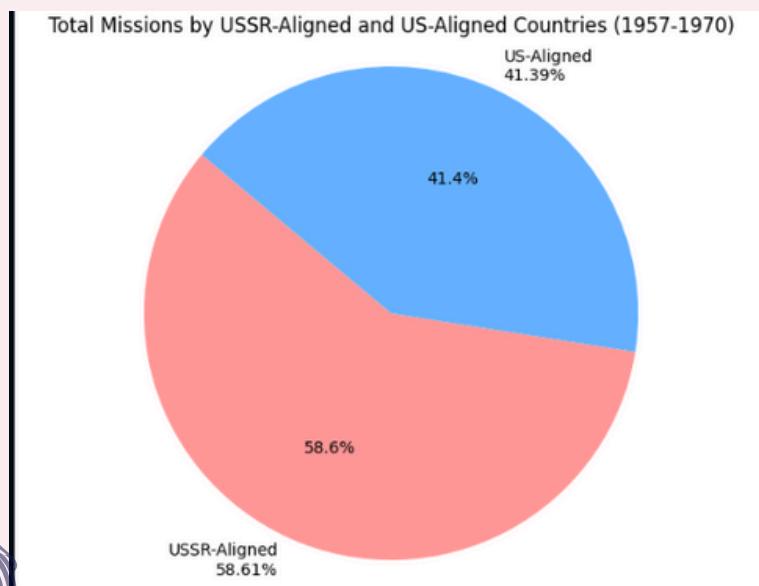
PIE CHART VISUALIZATIONS

To further enhance the comparative analysis, two additional pie charts are generated to show the distribution of missions and the success-failure ratio for both USSR-aligned and US-aligned countries within the specified date range (1957-1970).



PIE CHART FOR SUCCESS AND FAILURE RATES

Two additional functions, `plot_pie_success_failure_ussr` and `plot_pie_success_failure_us`, generate pie charts to display the success and failure rates for USSR-aligned and US-aligned countries, respectively. These charts provide a clear visual representation of the success and failure distribution within each group.





VISUALIZATION AND INTERPRETATION:

Visualizations, including graphs and bar charts, highlighted key trends and milestones in the Space Race. They showed the number of launches, success rates, and mission types, revealing the competitive dynamics between the USSR and the US. Line graphs depicted significant achievements like the first satellite, human spaceflight, and lunar landing. These visual tools provided clear insights into the rapid progress and pivotal moments of space exploration.

DATE CONVERSION AND YEAR EXTRACTION

The function begins by converting the date column in the DataFrame df1 to datetime format, which facilitates accurate date manipulations. It then extracts the year from the date and stores it in a new column named year.

```
def plot_space_exploration_trends(df1):
    # Convert date column to datetime format
    df1['date'] = pd.to_datetime(df1['date'])

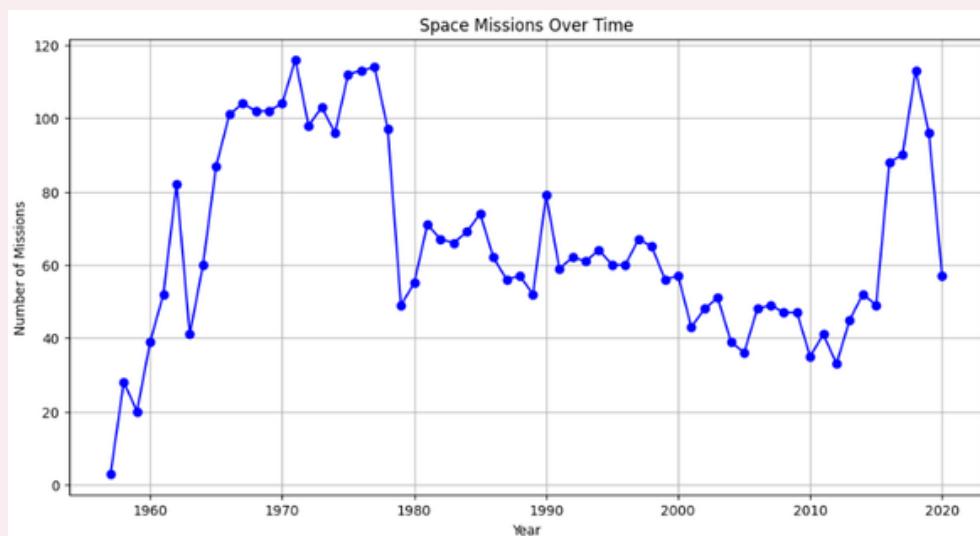
    # Extract year from date
    df1['year'] = df1['date'].dt.year

    # Group by year and count missions
    missions_by_year = df1.groupby('year').size()
```



MISSION TRENDS OVER TIME

The first visualization focuses on the overall trend of space missions over time. By grouping the data by year and counting the number of missions each year, a line plot is created to display the trend of space missions from the earliest recorded mission to the latest. This plot highlights the increase or decrease in space exploration activities over the years, marked by notable peaks and troughs.



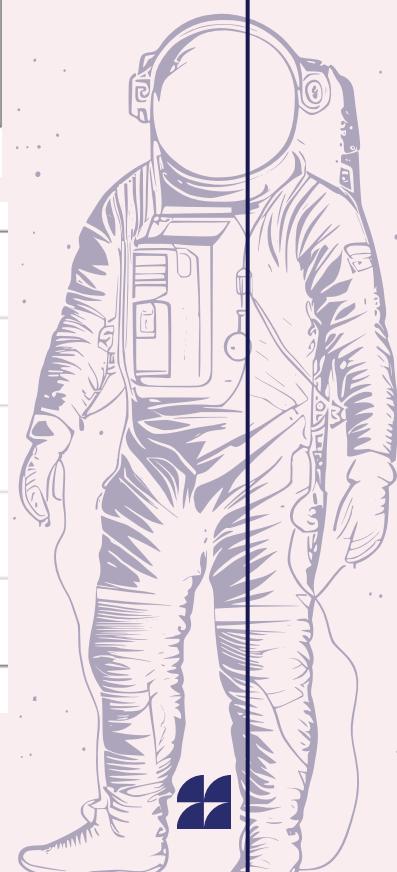
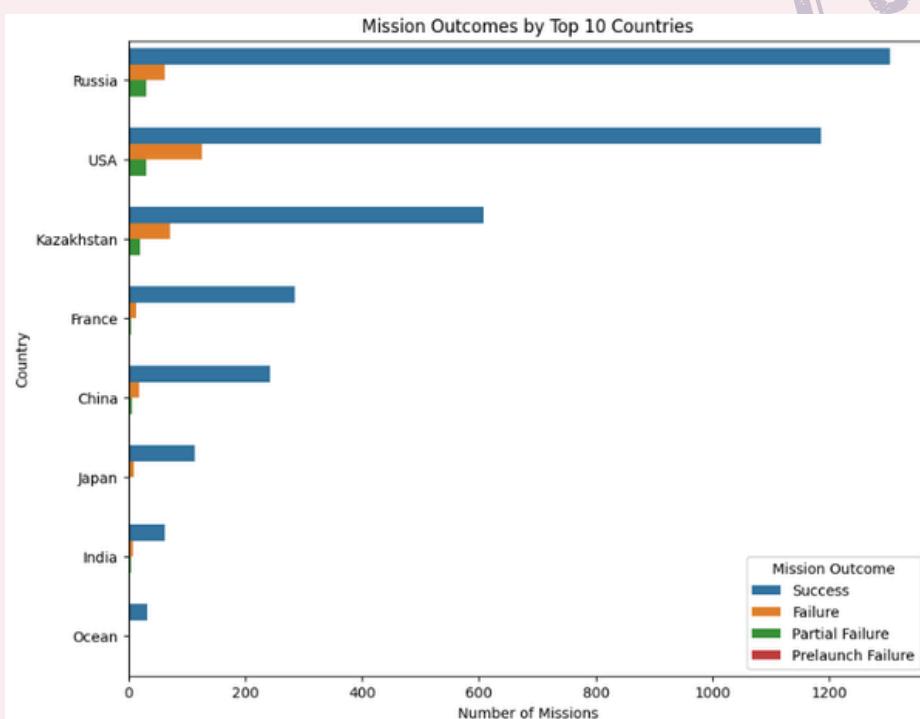
SUCCESS RATE OVER TIME

The second visualization shifts the focus to the success rate of these missions. By creating a new column `success`, which indicates whether a mission was successful, the function calculates the annual success rate as a percentage. A line plot is then generated to display the success rate over time, providing insights into the effectiveness and reliability of space missions across different periods.



MISSION OUTCOMES BY COUNTRY

The third visualization delves into the competitive dynamics among countries in space exploration. It identifies the top 10 countries with the highest number of space missions and filters the Data Frame accordingly. A count plot is then created to show the distribution of mission outcomes (Success, Failure, etc.) for these top countries. This plot effectively highlights the comparative performance of different nations in the realm of space exploration.



LESSONS LEARNED



CHALLENGES ENCOUNTERED

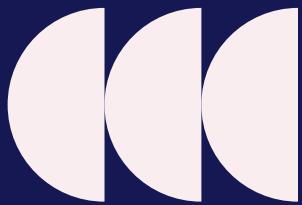
The analysis of space missions from 1957 onwards highlighted several crucial lessons. Firstly, technological innovation proved to be the driving force behind the rapid progress in space exploration. Advances in rocket propulsion, spacecraft design, and communication systems were essential for mission success. The competitive dynamic between the USSR and the US during the Space Race significantly accelerated these advancements, demonstrating how competition can spur excellence. Additionally, the contrasting approaches of the two nations—incremental advancements by the USSR and ambitious goals by the US—both uniquely contributed to the overall progress, emphasizing the value of diverse strategic methodologies.

Secondly, the importance of collaboration became evident. Successful space missions required teamwork among scientists, engineers, and international partners, showcasing the collective effort needed to overcome significant challenges. Data-driven decision-making was another critical factor, as detailed analysis and visualization of mission data provided valuable insights, guiding future missions and technological investments. Lastly, the ability to adapt to challenges and learn from failures was vital. Both nations demonstrated resilience, leading to improved designs and successful missions over time, underscoring the necessity of flexibility and continuous improvement in space exploration.



*SPACE MISSION FROM 1957 DATA ANALYSIS

THANK YOU



CONTACT INFORMATION FOR FURTHER INQUIRIES

📞 +92 333 4471066

✉️ info@itsolera.com

📍 Javeed Arcade, Office #2A St 3, G-12/1 G-12,
Islamabad, Islamabad Capital Territory 46000

🌐 www.itsolera.com