Data Analysis on the Foreign Satellites Launched by ISRO

Internship Report Submitted in partial fulfillment of the
requirement for undergraduate degree of
Bachelor of Technology
In
COMPUTER SCIENCE AND ENGINEERING

By
Rudhramyna Amshu
HU22CSEN0101640

Under the Guidance of
Dr. Jhansi Rani T
Assistant Professor

Department Of Computer Science and Engineering
GITAM School of Technology
GITAM (Deemed to be University)
Hyderabad-502329
March 2025

DECLARATION

I hereby declare that the summer internship report entitled "**Data Analysis on the Foreign Satellites Launched by ISRO**" is an original work done in the Department of Computer Science and Engineering, GITAM School of Technology, GITAM (Deemed to be University), submitted in partial fulfillment of the requirements for the award of the degree of "Bachelor of Technology" in Computer Science and Engineering. The work had not been submitted to any other college or university for the award of any degree or diploma. I declare that it was carried out independently by me under the guidance of **Dr Jhansi Rani T,** Asst. Professor, GITAM (Deemed To Be University), Hyderabad, India.

.

Place: HYDERABAD

Date: 10-03-2025

Name: Rudhramyna Amshu

Student Roll No: HU22CSEN0101640

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

GITAM SCHOOL OF TECHNOLOGY

GITAM

(DEEMED TO BE UNIVERSITY)
HYDERABAD CAMPUS

CERTIFICATE

This is to certify that the Internship Report entitled **"Data Analysis on the Foreign Satellites Launched by ISRO"** is being submitted by **Rudhramyna Amshu** (HU22CSEN0101640) in partial fulfillment of the requirement for the award of **Bachelor of Technology** in Computer Science and Engineering at
GITAM (Deemed to Be University), Hyderabad during the academic year 2024-2025.

Dr Jhansi Rani T
Assistant Professor
Department of CSE

Dr Shaik Mahaboob Basha
Professor and HOD
Department of CSE

CERTIFICATE OF COMPLETION

Networking Academy

Certificate of Course Completion

Amshu Rudhramyna

has successfully achieved student level credential for completing the Data Analytics Essentials course.

The student was able to proficiently:

- *Explain how the data analytics process creates value from data.*
- *Explain the characteristics of data, including formats, availability and methods to acquire.*
- *Transform data using analytics tools.*
- *Analyze data using basic statistical and data preparation techniques.*
- *Complete hands-on lab using Excel, SQL, Tableau and other tools.*
- *Evaluate and share project portfolio.*

Networking Academy Verified Data Analytics Essentials

Scan to Verify

Lynn Bloomer

Lynn Bloomer
Director, Cisco Networking Academy

Issued on: Mar 11, 2025

ACKNOWLEDGEMENT

I sincerely thank **Dr. D. Sambasiva Rao, Pro Vice-Chancellor, GITAM Hyderabad**, and **Dr. N. Seetharamaiah, Principal, GITAM Hyderabad**, for providing the opportunity to undertake this internship.

My heartfelt gratitude to **Dr Shaik Mahaboob Basha, Head of the Department of CSE**, for her guidance and support in expanding my knowledge and preparing this report.

Thanks to my mentor, **Dr Jhansi Rani T**, for her invaluable support and insights throughout this internship.

I would also like to thank my friends who helped me to make my work more organized and well-stacked till the end.

Rudhramyna Amshu

HU22CSEN0101640

Abstract

India has launched 431 satellites for 34 countries as of 30 July 2023. This project analyzes foreign satellites launched by ISRO, focusing on launch trends, payload distributions, and global collaborations. Data was collected through web scraping from Wikipedia, processed using Pandas, and visualized with **Matplotlib, Seaborn, and Power BI**.

The analysis highlights the steady growth of ISRO's foreign satellite launches, reinforcing its position as a leading global space service provider. **PSLV emerged as the most frequently used launch vehicle**, known for its reliability and cost-effectiveness. Additionally, the study found a diverse distribution of satellite masses, reflecting varying mission requirements.

Through **exploratory data analysis (EDA)**, the project uncovers ISRO's increasing role in commercial satellite launches and its expanding global partnerships. The findings provide a foundation for **predictive modeling**, which could help forecast future launch trends based on historical data.

Beyond descriptive analytics, this project sets the stage for advanced machine learning applications, offering strategic insights for ISRO and global space organizations. The study contributes to a broader understanding of **India's influence in the space industry**, helping policymakers, researchers, and stakeholders optimize future satellite missions. This research serves as a baseline for further exploration into space technology trends and commercial launch strategies.

TABLE OF CONTENTS

Introduction

Data analysis is essential for uncovering valuable insights from extensive datasets. The main objective of traditional data analysis is to convey information in a manner that is accessible to non-technical decision-makers.

In this project, we analyze foreign satellites launched by the Indian Space Research Organisation (ISRO). The dataset is sourced from Wikipedia (List of foreign satellites launched by India) through web scraping. The project follows key data analysis steps: **data collection, data cleaning, exploratory data analysis (EDA), and visualization using Power BI**.

This report also covers essential concepts of data analysis, including structured and unstructured data, types of statistics, data visualization methods, SQL for data management, and ethical considerations in data analysis as covered in the Cisco Networking Academy course 'Data Analytics Essentials' which serves as a prerequisite.

Types of Data Analysis

Data analysis is broadly categorized into:

1. **Descriptive Analysis** - Summarizing past data through mean, median, mode, etc.
   It asks, "What happened?"
2. **Diagnostic Analysis** - Identifying reasons for trends.
   It asks, "Why did this happen?"
3. **Predictive Analysis** - Using historical data to make future predictions.
   It asks, "What might happen in the future?"
4. **Prescriptive Analysis** - Suggesting actions based on predictions.
   It asks, "What should be done next?"

Process

Step 1: Ask a question

Write down your question. What is the benefit of determining the answer to this question? In most data analytic projects, the potential benefit is the reason for the project.

Review the four different types of data analysis. What fits the best?

Step 2: Determine the Data Needed

a. Make a list of the data elements that you identify.

a. Open a web browser and use any search engine to search for the data elements that you have identified.
b. List sources that you found that may have data relating to the elements you identified.
c. You may also want to investigate sources, such as Kaggle, that provide sample data for analysis.

Reflection:

1. Why is it important to identify the question that needs to be answered by the analysis before beginning the project?
2. Name some sources of open data for analysis that you found while searching for your data elements.

Project:

Step 1: Asking the Question

Question:

How has the launch of foreign satellites by ISRO evolved over time, and what trends can be identified in terms of launch frequency, mass, and country of origin?

Understanding the answer to this question sheds light on ISRO's position in the global satellite launch market. It offers valuable insights into the increasing demand for foreign satellite launches and highlights India's role as a cost-effective and dependable launch service provider. This analysis can be beneficial for governments, businesses, and researchers in their strategic planning and policy formulation.

For this project, a **Descriptive Analysis** is the most suitable approach, as it aims to summarize the dataset, identify trends in satellite launches, and present them in a clear and effective manner.

Step 2: Determining the Data Needed

Identified Data Elements:

- Launch Date
- Satellite Name
- Country of Origin
- Launch Mass (kg)
- Launch Vehicle Used
- Additional Remarks (if any)

Data Sources:

The primary source of data is Wikipedia:

- Wikipedia: List of Foreign Satellites Launched by India

Search Approach:

A web search was conducted using search terms like:

- “ISRO foreign satellite launches dataset”
- “India satellite launch statistics”
- “Foreign satellites launched by India historical data”

Reflection:

Why is it important to identify the question before beginning the project?

Formulating a well defined prominent question guarantees that the analysis does not stray off the topic. It is useful when deciding which dataset to select, the method of analysis to use, and the insights to be drawn. In the absence of focus, the analysis may turn out to be sloppy and useless.

Data

A collection of data is referred to as a dataset.

Datasets often contain multiple related files stored in different formats. Information about a dataset, including a description of what it contains and how it is formatted, is called **metadata**.

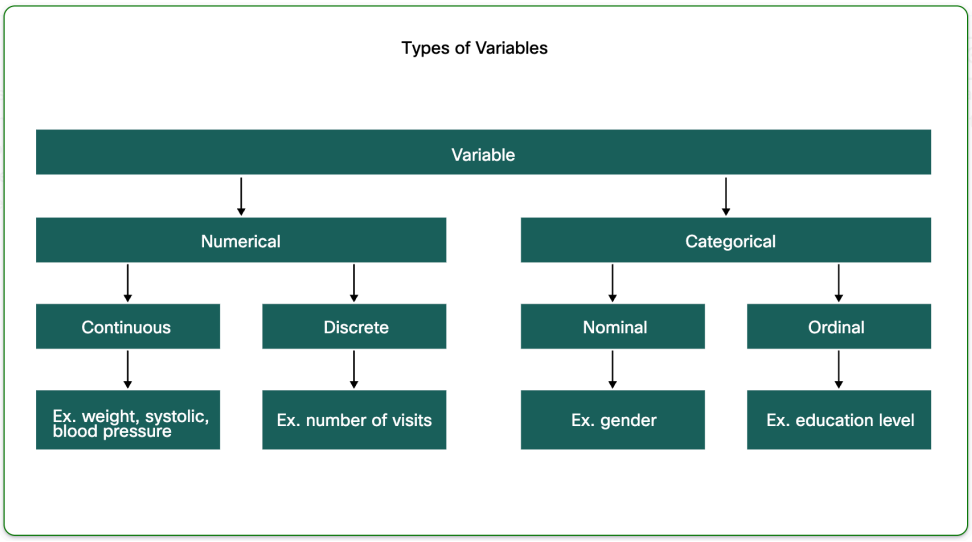
- **Structured Data**: Well-organized, stored in tabular formats like Excel, SQL databases (e.g., CSV, JSON, databases).
- **Unstructured Data**: Raw data in formats like images, videos, PDFs, and free-text logs.

The dataset in this project is structured, stored in CSV files, and preprocessed using **Pandas** in Python.

Structured Query Language, is a powerful database management tool that allows data analysts to retrieve and interact with selections of data that are stored in relational databases. Relational databases have a defined structure and contain multiple interrelated data tables that need to be queried.

When performing any kind of data experiment or analysis, it is critical to define the key characteristics that need to be measured or observed. These characteristics to be studied are called variables. A **variable** in this context is anything that varies from one instance to another, that can be measured, and whose value can be manipulated or controlled in theoretical scenarios.

The recordings of the values, patterns, and occurrences for a set of variables are **observations**. The value or set of values for a specific observation is called a **data point**.



Some questions that we should ask when selecting a data source:

1. What data points are necessary to inform your analysis?
2. Do I already have access to this data, or must I find a dataset from another source?
3. Where are reliable and verifiable sources of this data?
4. How often is the relevant data collected and updated?
5. How is the data licensed for use, and is there a cost?
6. Is the data in a format that I can use, or convert to use, with my tools?

Data Collection (Web Scraping)

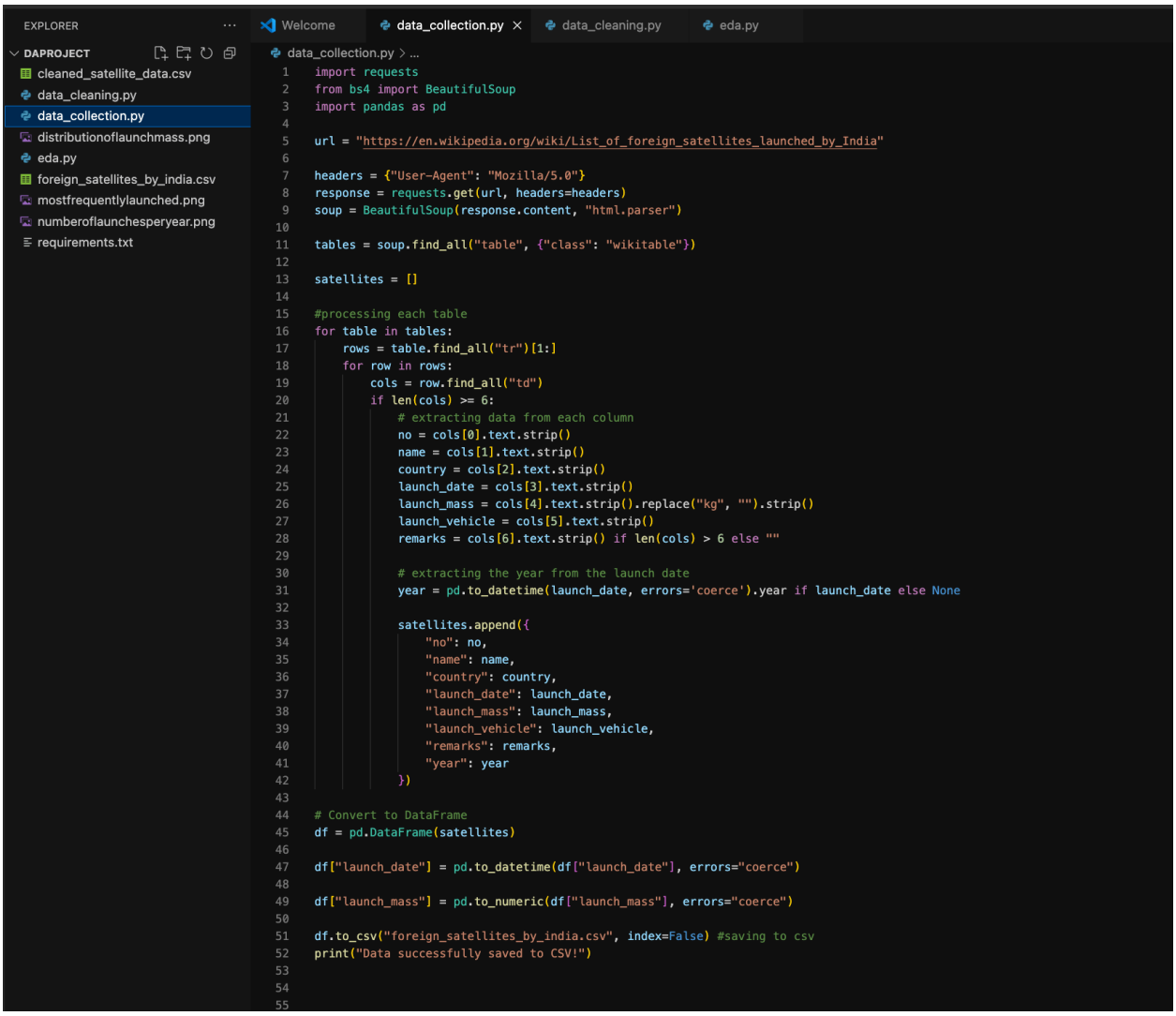
The data_collection.py script scrapes the Wikipedia page using:

- **Requests** library to fetch webpage data.
- **BeautifulSoup** for parsing HTML tables.
- Extracts columns: Satellite No., Name, Country, Launch Date, Launch Mass, Launch Vehicle, and Remarks.
- Saving the dataset as foreign_satellites_by_india.csv.

This project was composed and implemented on Visual Studio Code, which is an integrated development environment developed by Microsoft for Windows, Linux, macOS and web browsers.

Features include support for debugging, syntax highlighting, intelligent code completion, snippets, code refactoring, and embedded version control with Git. Users can change the theme, keyboard shortcuts, preferences, and install extensions that add functionality.

INPUT:


EXPLORER
DAPROJECT
cleaned_satellite_data.csv
data_cleaning.py
data_collection.py
distributionoflaunchmass.png
eda.py
foreign_satellites_by_india.csv
mostfrequentlylaunched.png
numberoflaunchesperyear.png
requirements.txt
Welcome
data_collection.py
data_cleaning.py
eda.py
data_collection.py > ...
import requests
from bs4 import BeautifulSoup
import pandas as pd
url = "https://en.wikipedia.org/wiki/List_of_foreign_satellites_launched_by_India"
headers = {"User-Agent": "Mozilla/5.0"}
response = requests.get(url, headers=headers)
soup = BeautifulSoup(response.content, "html.parser")
tables = soup.find_all("table", {"class": "wikitable"})
satellites = []
#processing each table
for table in tables:
rows = table.find_all("tr")[1:]
for row in rows:
cols = row.find_all("td")
if len(cols) >= 6:
# extracting data from each column
no = cols[0].text.strip()
name = cols[1].text.strip()
country = cols[2].text.strip()
launch_date = cols[3].text.strip()
launch_mass = cols[4].text.strip().replace("kg", "").strip()
launch_vehicle = cols[5].text.strip()
remarks = cols[6].text.strip() if len(cols) > 6 else ""
# extracting the year from the launch date
year = pd.to_datetime(launch_date, errors='coerce').year if launch_date else None
satellites.append({
"no": no,
"name": name,
"country": country,
"launch_date": launch_date,
"launch_mass": launch_mass,
"launch_vehicle": launch_vehicle,
"remarks": remarks,
"year": year
})
# Convert to DataFrame
df = pd.DataFrame(satellites)
df["launch_date"] = pd.to_datetime(df["launch_date"], errors="coerce")
df["launch_mass"] = pd.to_numeric(df["launch_mass"], errors="coerce")
df.to_csv("foreign_satellites_by_india.csv", index=False) #saving to csv
print("Data successfully saved to CSV!")

OUTPUT:


PROBLEMS
OUTPUT
DEBUG CONSOLE
TERMINAL
Filter
Code
[Running] python -u "/Applications/folders/DAproject/data_collection.py"
Data successfully saved to CSV!
[Done] exited with code=0 in 1.265 seconds

cleaned_satellite_data

no	name	country	launch_date	launch_mass	launch_vehicle	remarks	year
1	DLR-Tubsat	Germany	1999-05-26	45.0	PSLV-C2	ISRO's 1st commercial launch with foreign satellites as payload. India's Oceansat-1 was also launched. This was PSLV's 3rd launch overall.	1999
3	BIRD	Germany	2001-10-22	92.0	PSLV-C3	ISRO's 2nd commercial launch.	2001
5	Lapan-TUBsat	Indonesia	2007-01-10	56.0	PSLV-C7	ISRO's 2nd commercial launch.	2007
7	AGILE	Italy	2007-04-23	352.0	PSLV-C8	PSLV's 11th flight.	2007
8	TecSAR	Israel	2008-01-21	295.0	PSLV-C10	PSLV's 12th launch.	2008
9	CAN-X2	Canada	2008-04-28	3.5	PSLV-C9	ISRO launched 10 satellites, of which 8 were foreign.[7]	2008
17	UWE-2	Germany	2009-09-23	1.0	PSLV-C14	ISRO launched 7 satellites, of which 6 were foreign.[8]	2009
23	Alsat-2A	Algeria	2010-07-12	116.0	PSLV-C15	ISRO launched 5 satellites, of which 3 were foreign.[9]	2010
26	VESSELSAT-1	Luxembourg	2011-01-12	28.7	PSLV-C18	ISRO launched 4 satellites, of which 1 was foreign.[10]	2011
27	X-SAT	Singapore	2011-04-20	106.0	PSLV-C16	ISRO launched 3 satellites, of which 1 was foreign.[11]	2011
28	SPOT-6	France	2012-09-09	712.0	PSLV-C21	PSLV's 22nd flight.	2012
30	Sapphire	Canada	2013-02-25	148.0	PSLV-C20	ISRO launched 7 satellites, of which 6 were foreign.[12]	2013
36	SPOT-7	France	2014-06-30	714.0	PSLV-C23	PSLV's 10th flight in 'core-alone' configuration (i.e. without the use of solid strap-on motors).	2014
41	UK-DMC 3A	United Kingdom	2015-07-10	447.0	PSLV-XL C28[13]	India's first exclusive foreign satellites launch, all the 5 payloads were from United Kingdom. At the time it was the heaviest commercial mission (1439 kg) successfully accomplished using a launch vehicle assembled by ISR	2015
46	LAPAN-A2	Indonesia	2015-09-28	76.0	PSLV-C30	Commercial satellites from United States were launched on an Indian rocket for the first time. Astrosat, India's first dedicated astronomy satellite, was also launched on this flight.[14]	2015
52	TeLEOS-1	Singapore	2015-12-16	400.0	PSLV-C29	Exclusive commercial launch of 6 Singaporean satellites.	2015
58	LAPAN A3	Indonesia	2016-06-22	120.0	PSLV-XL C34	ISRO launched 20 satellites (including 3 Indian satellites) aboard PSLV-C34, the highest number of satellites that the agency has launched aboard a single flight.[15][16]	2016
75	AISAT-1N	Algeria	2016-09-26	7.0	PSLV-G C35	ISRO launches 8 satellites in its 15th flight of the 'XL' version of the PSLV - 5 foreign satellites and 3 Indian satellites (SCATSAT-1, PRATHAM and PISAT).[17]	2016
80-167	88 x Flock-3p	United States	2017-02-15	7.0	PSLV-XL 37	ISRO launched 104 satellites, of which 3 were Indian satellites. It was[18] the largest number of satellites launched on a single flight by any space agency.[19]	2017
181	Pegasus	Austria	2017-06-23	2.0	PSLV-C38	ISRO launched 31 satellites, of which 29 were foreign.[20]	2017
210	Telesat Phase-1 LEO	Canada	2018-01-12	168.0	PSLV-XL C40	ISRO Launched 31 satellites, of which 28 were foreign.[22]	2018
238	NovaSAR	United Kingdom	2018-09-16	445.0	PSLV-CA C42	Exclusive commercial launch of two foreign satellites belonged to Surrey Satellite Technologies Ltd (SSTL), United Kingdom. The satellites were put into Sun-synchronous orbit under a commercial arrangement with Antrix C	2018
240	Centauri -1	Australia	2018-11-29	10.0	PSLV-CA C43	-	2018
270	Bluewalker 1	Lithuania	2019-04-01	10.0	PSLV-QL C45	-	2019
298	Meshbed	United States	2019-04-01	4.5	PSLV-XL C47	-	2019
311	Izanagi (QPS-SAR)	Japan	2019-04-01	4.5	PSLV-QL C48	-	2019
320	R2	Lithuania	2019-04-01	4.5	PSLV-DL C49	Technology demonstration satellite.	2019
329	Amazônia-1	Brazil	2019-04-01	637.0	PSLV-DL C51	First Earth observation satellite entirely developed by Brazil.	2019
343	DS-EO	Singapore	2019-04-01	365.0	PSLV-CA C53	First Earth observation satellite entirely developed by Brazil.	2019
346-381	36 × OneWeb	United Kingdom	2019-04-01	365.0	LVM 3 M2	First commercial launch of LVM 3.	2019
382-385	4 × Astrocast	United States	2019-04-01	17.92	PSLV-XL C54	Satellites Developed by Spaceflight,United States For Astrocast,Switzerland	2019
387	Janus-1	United States	2019-04-01	11.5	SSLV-D2	First Successful launch of SSLV	2019
388-423	36 × OneWeb	United Kingdom	2019-04-01	11.5	LVM 3 M3	Second commercial launch of LVM 3.It is the heaviest payload that is launched by a LVM 3 and ISRO to date.	2019
424	TeLEOS-2	Singapore	2019-04-01	741.0	PSLV-CA C55	57th Mission of PSLV	2019
426	DS-SAR	Singapore	2019-04-01	352.0	PSLV-CA C56	58th PSLV Mission. Commercial Launch for Singapore's DS-SAR Satellite and 6 Co-Passenger satellites from Singapore.	2019
433	PROBA-3	European Union	2019-04-01	352.0	PSLV-XL C59	58th PSLV Mission. Commercial Launch for Singapore's DS-SAR Satellite and 6 Co-Passenger satellites from Singapore.	2019

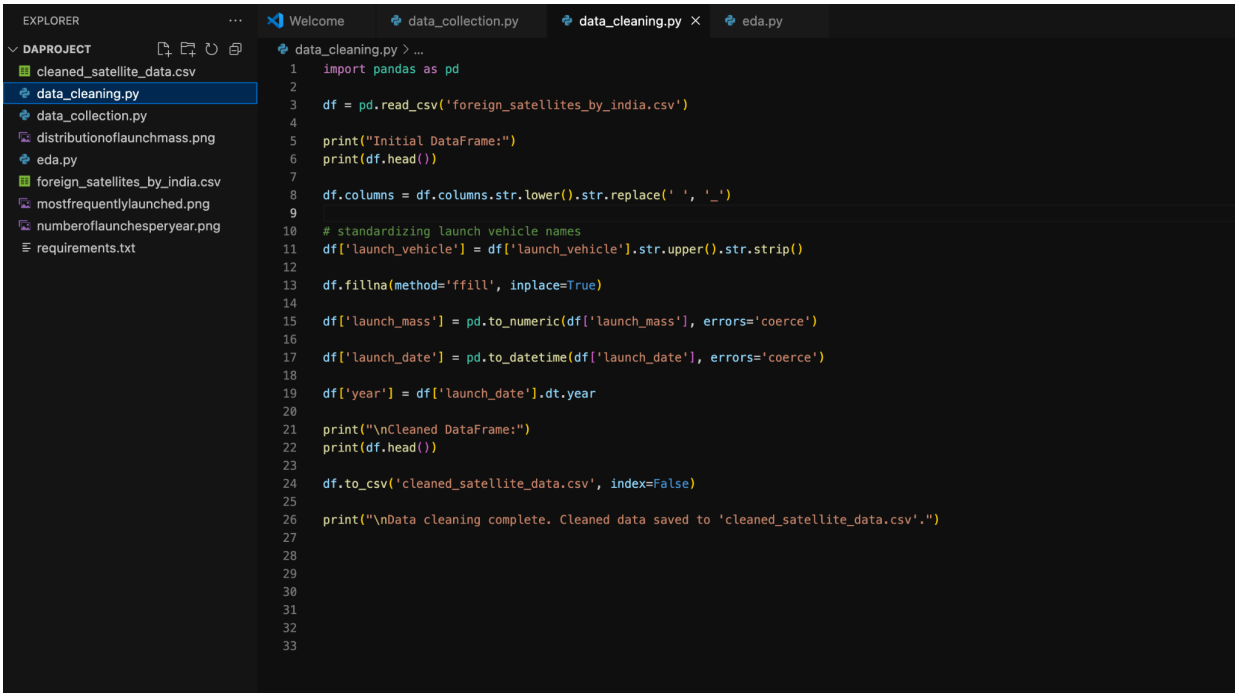
CSV file of the scraped data

Data Cleaning

The data_cleaning.py script performs:

- **Standardization**: Converts column names to lowercase and replaces spaces with underscores.
- **Missing Value Handling**: Uses forward fill (fillna(method='ffill')) to propagate missing values.
- **Data Type Conversions**:
    - Converts launch_date to datetime format.
    - Converts launch_mass to numeric values.
    - Extracts the launch year.
- Saves cleaned data as cleaned_satellite_data.csv.

INPUT:



OUTPUT:

```
[Running] python -u "/Applications/folders/DAproject/data_cleaning.py"
Initial DataFrame:
  no          name  ...                                        remarks  year
0  1    DLR-Tubsat  ...  ISRO's 1st commercial launch with foreign sate...  1999
1  3          BIRD  ...                     ISRO's 2nd commercial launch.  2001
2  5  Lapan-TUBsat  ...                                               NaN  2007
3  7         AGILE  ...                              PSLV's 11th flight.  2007
4  8        TecSAR  ...                              PSLV's 12th launch.  2008

[5 rows x 8 columns]
/Applications/folders/DAproject/data_cleaning.py:21: FutureWarning: DataFrame.fillna with 'method' is deprecated and will raise in a future version.
Use obj.ffill() or obj.bfill() instead.
  df.fillna(method='ffill', inplace=True)

Cleaned DataFrame:
  no          name  ...                                        remarks  year
0  1    DLR-Tubsat  ...  ISRO's 1st commercial launch with foreign sate...  1999
1  3          BIRD  ...                     ISRO's 2nd commercial launch.  2001
2  5  Lapan-TUBsat  ...                     ISRO's 2nd commercial launch.  2007
3  7         AGILE  ...                              PSLV's 11th flight.  2007
4  8        TecSAR  ...                              PSLV's 12th launch.  2008

[5 rows x 8 columns]

Data cleaning complete. Cleaned data saved to 'cleaned_satellite_data.csv'.

[Done] exited with code=0 in 0.736 seconds
```

Exploratory Data Analysis (EDA)

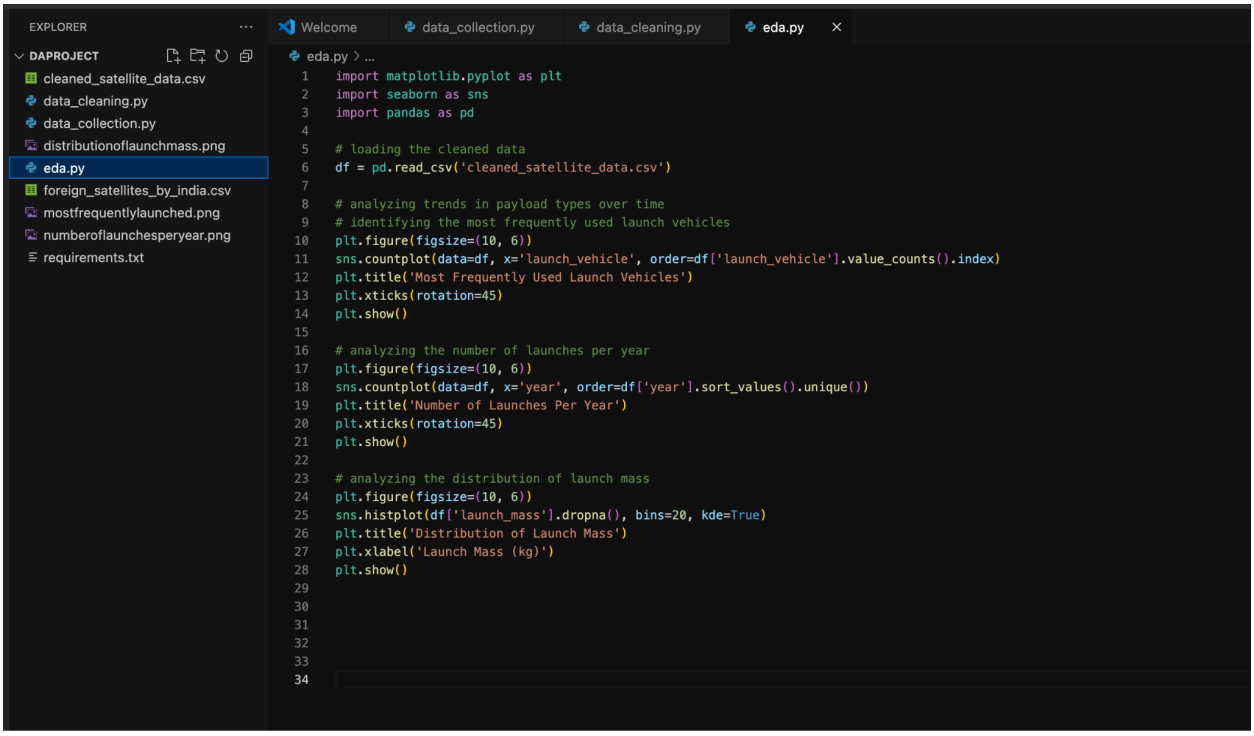
The eda.py script visualizes data using **Matplotlib and Seaborn**:

1. **Most Frequently Used Launch Vehicles** - Count plot to show the frequency of launch vehicles.
2. **Number of Launches Per Year** - Trend analysis using a count plot.
3. **Distribution of Launch Mass** - Histogram to visualize launch mass variations.

Exploratory Data Analysis (EDA) is a systematic and iterative approach used to analyze and understand the characteristics of a dataset. It involves summarizing and visualizing data to identify patterns, relationships, and anomalies without preconceived notions or hypotheses:

- Enhances understanding of data structure and relationships.
- Ensures data quality by identifying errors and outliers.
- Facilitates informed decision-making for further analysis or modeling.

INPUT:



OUTPUT:

```
[Running] python -u "/Applications/folders/DAproject/eda.py"
2025-03-10 19:46:47.613 python[56969:3833663] +[IMKClient subclass]: chose IMKClient_Modern
2025-03-10 19:46:47.613 python[56969:3833663] +[IMKInputSession subclass]: chose IMKInputSession_Modern
2025-03-10 19:47:31.741 python[56969:3833663] The class 'NSSavePanel' overrides the method identifier.  This method is implemented by class 'NSWindow'

[Done] exited with code=0 in 845.233 seconds
```

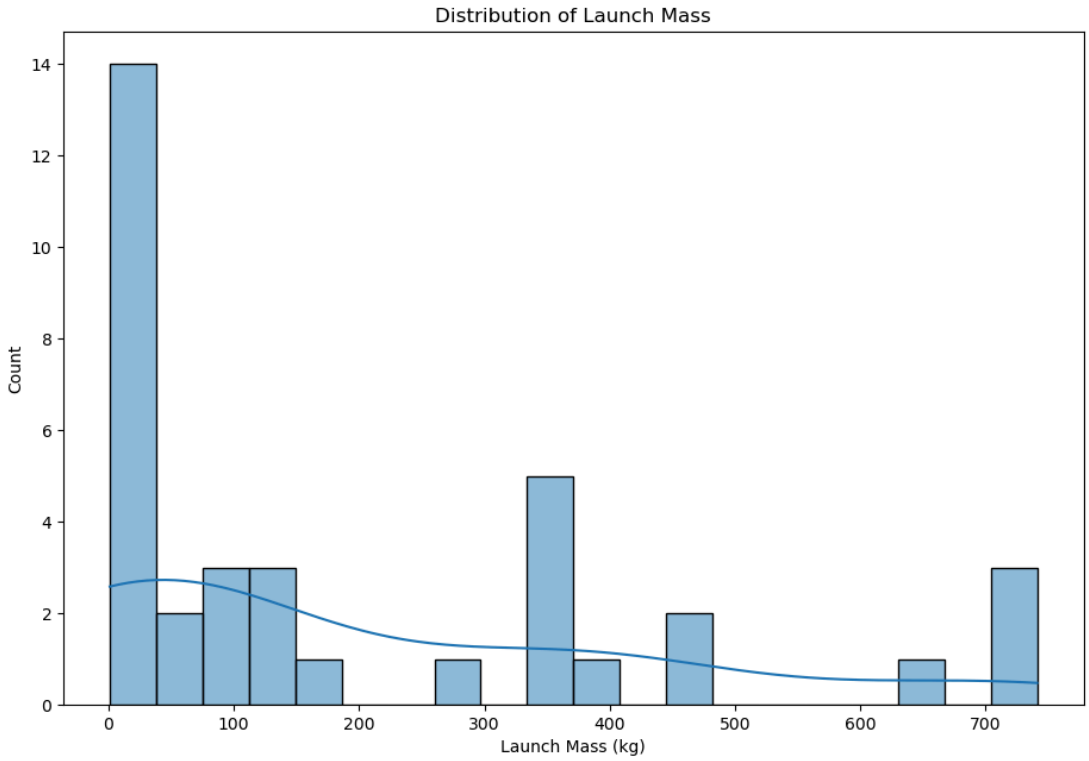
Data Visualization

Choosing the right visualizations improves insights:

- **Line Charts**: Time series trends (e.g., number of launches per year).
- **Bar/Column Charts**: Categorical comparisons (e.g., frequently used launch vehicles).
- **Histograms**: Distribution of numerical values (e.g., launch mass distribution).
- **Pie Charts**: Proportions (e.g., country-wise satellite launches, best done in Power BI/Tableau).

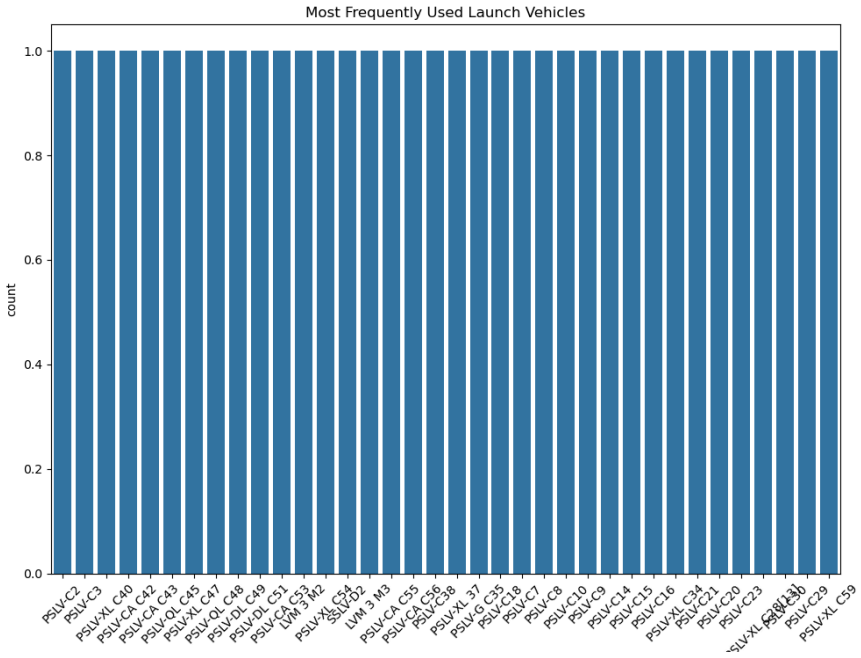
The visualizations for the project were done via Matplotlib and Power BI.

Microsoft Power BI is a business intelligence platform designed to transform disparate data sources into interactive and visually immersive insights. It is part of the Microsoft Power Platform and offers a range of tools and services, including Power BI Desktop for report creation, the Power BI Service for cloud-based analytics, and Power BI Mobile apps for on-the-go access. Power BI supports various data sources such as Excel, SQL Server, and web files, allowing users to create customized dashboards and reports. Its user-friendly interface makes it accessible to both non-technical business users and BI developers, facilitating data-driven decision-making through real-time analytics and data visualization.

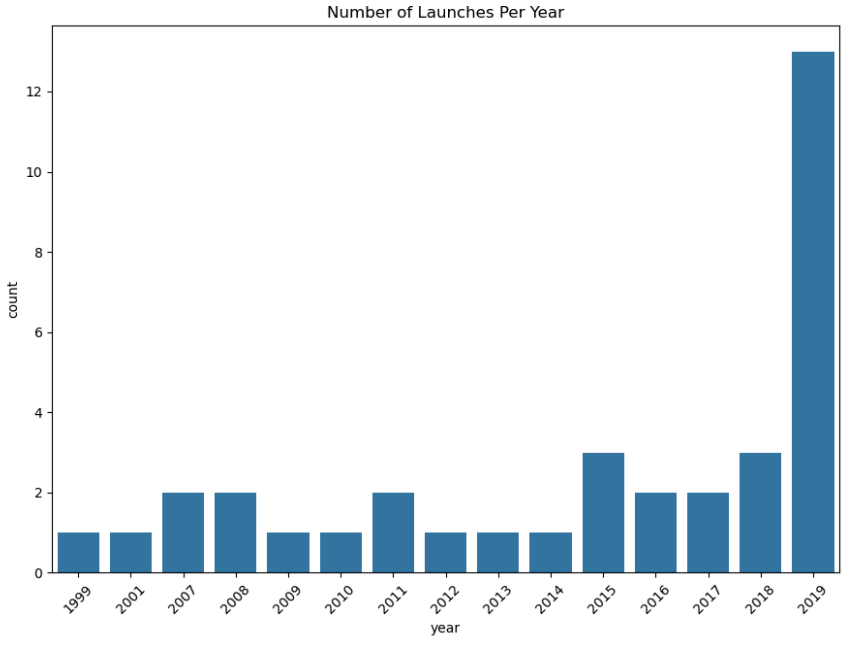


Distribution of Launch Mass

The histogram above represents the **distribution of launch mass** (in kilograms) for foreign satellites launched by India. The data shows a **high concentration of satellites with relatively low masses**, particularly in the **0-100 kg range**, which dominates the distribution. A smaller number of satellites fall into higher mass categories, with a few peaks around **400 kg, 500 kg, and 700 kg**. The kernel density estimate (KDE) curve provides a smooth approximation of the distribution, indicating a right-skewed pattern where **most satellites are lightweight, and only a few are heavier.** This trend aligns with the growing preference for **small and microsatellites**, particularly for commercial and research purposes, which require lower launch costs and greater deployment flexibility.

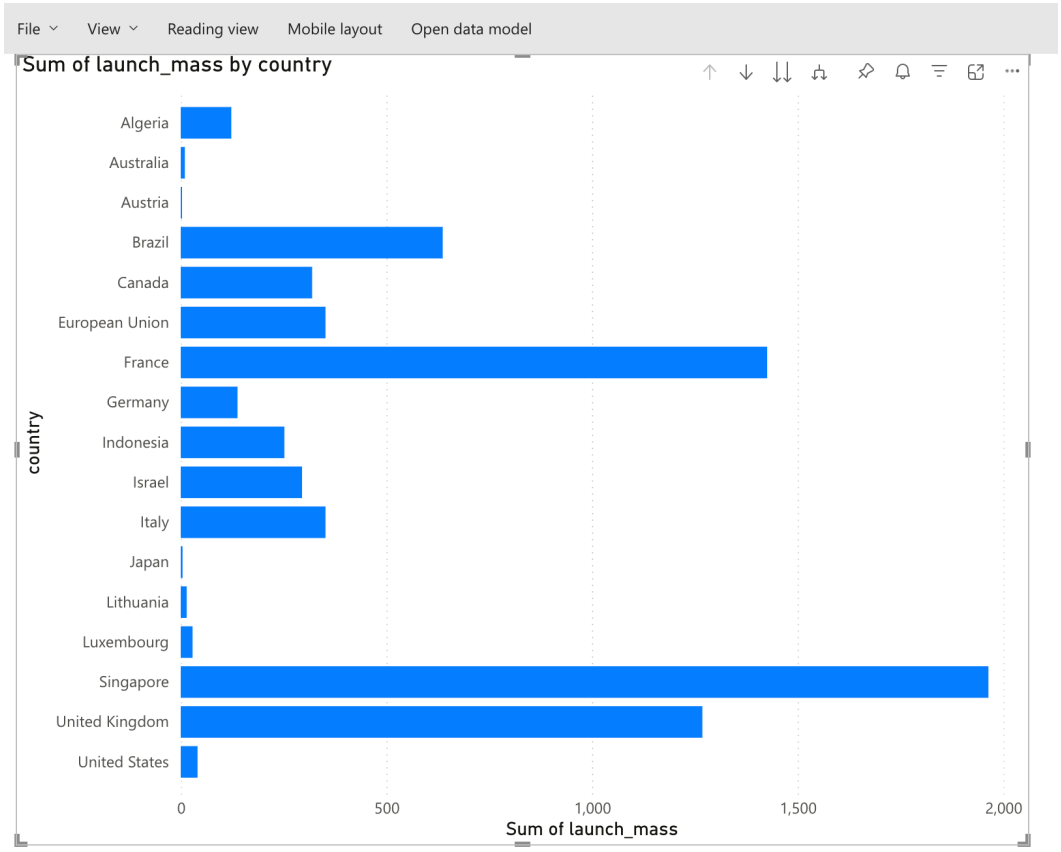


Most frequently launched vehicles

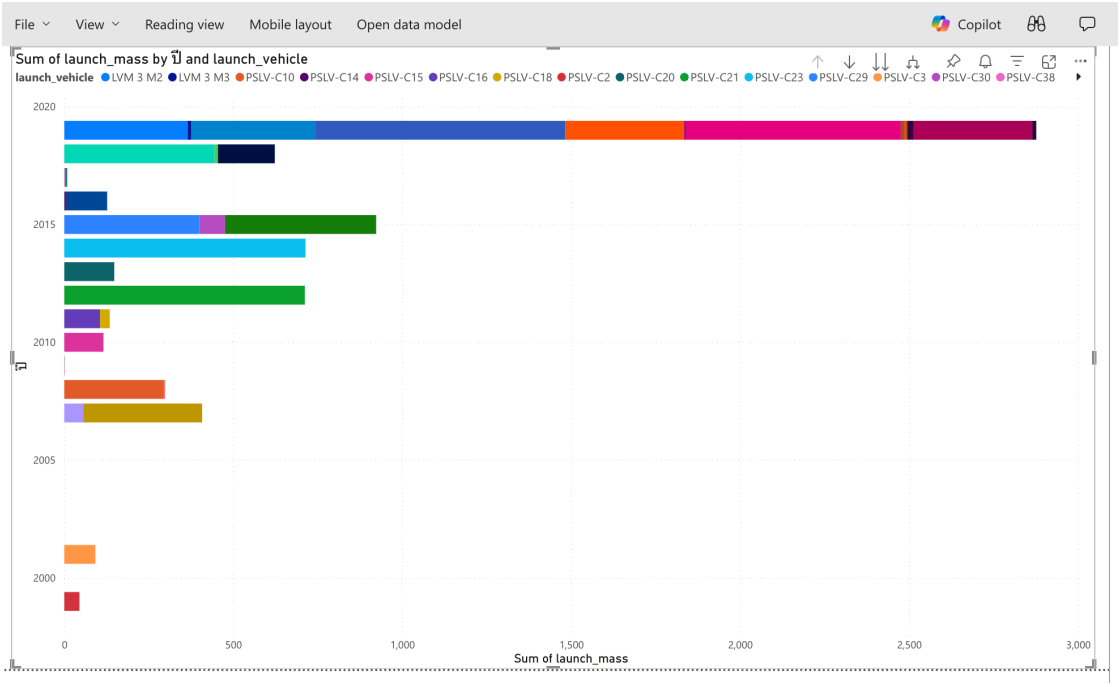


Total number of Launches per year

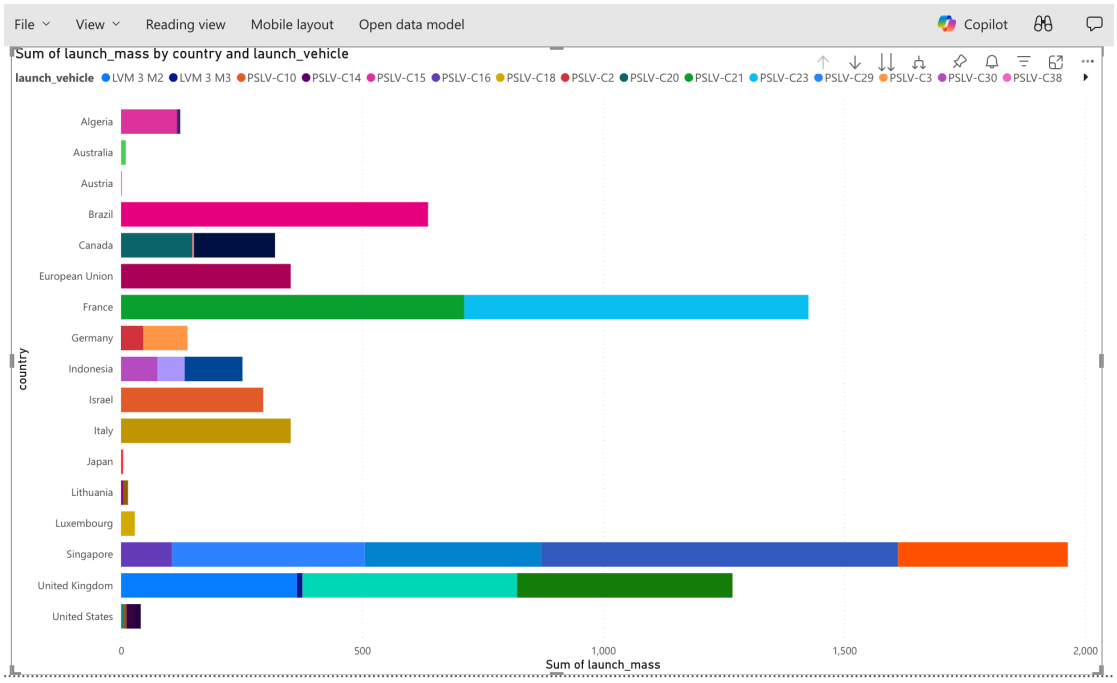
Power BI visualizations:



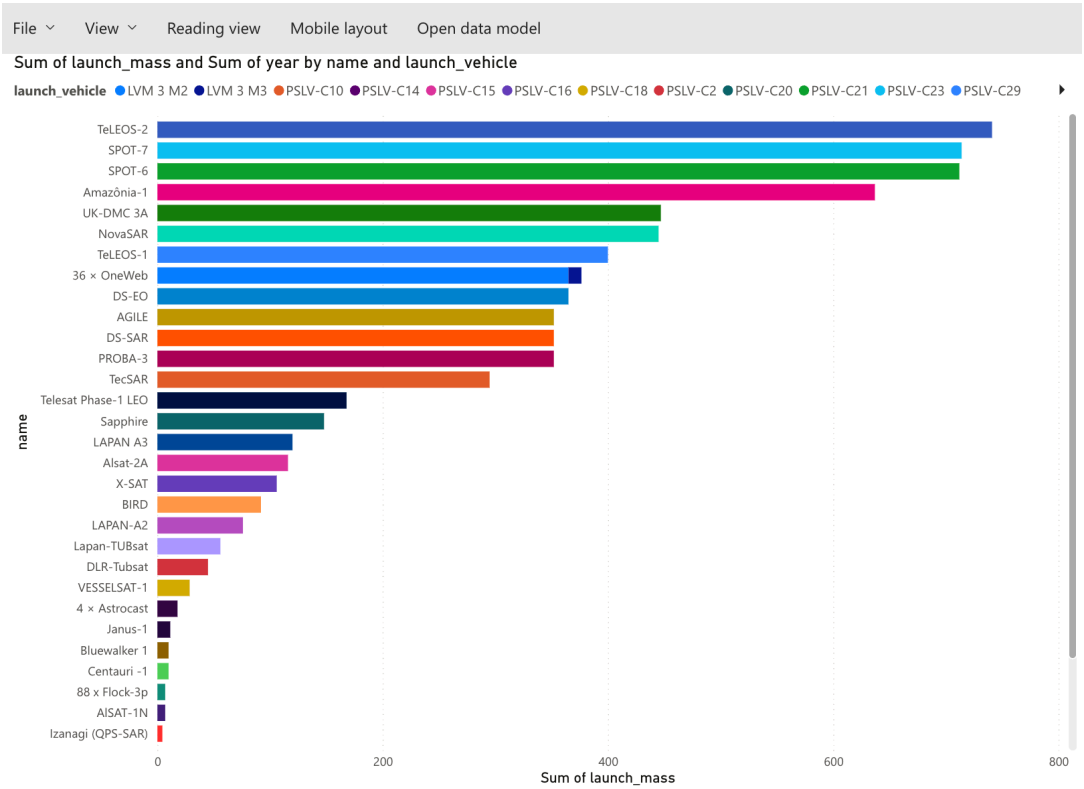
Sum of launch mass by country



Sum of launch mass by year and launch vehicle



Sum of launch mass by country and launch vehicle



Sum of launch mass and year by name and launch vehicle

Role of Excel in Data Analysis

Excel plays a pivotal role in data analysis by offering a wide range of tools and features that facilitate data organization, cleaning, analysis, and visualization. Here's an expanded view of Excel's role in these processes:

Storing and Structuring Data

- Data Organization: Excel allows users to store data in a structured format using tables, which can be easily manipulated and analyzed.
- Data Import: Excel supports importing data from various sources, such as CSV files, databases, and other spreadsheets.
- Data Management: Excel's features like sorting and filtering enable users to manage large datasets efficiently.

Performing Basic Data Cleaning

- Sorting and Filtering: These functions help in organizing data based on specific criteria, making it easier to identify patterns or anomalies.
- Conditional Formatting: This feature highlights cells based on conditions, such as values above or below a certain threshold, aiding in quick visual inspection of data quality.
- Removing Duplicates: Excel allows users to remove duplicate entries, ensuring data consistency and reducing redundancy.
- Error Handling: Functions like IFERROR help in identifying and managing errors within formulas, ensuring data integrity.

SQL for Data Management

SQL (Structured Query Language) is an exceptional tool for managing large datasets efficiently. It provides a structured approach to storing, querying, and manipulating data in relational databases.

Key SQL Commands for Data Management

1. **SELECT**: This command is used to query very specific data from a database. It allows users to retrieve only the necessary columns and rows, reducing data redundancy and also improving query efficiency.
2. **GROUP BY**: This clause is essential for aggregating statistics. It groups data based on one or more columns and applies aggregate functions like SUM, AVG, or COUNT to each group.
3. **JOIN**: SQL JOINs merge datasets from multiple tables based on related columns. There are several types of JOINs:
    - INNER JOIN: Returns records that have matching values in both tables.
    - LEFT JOIN: Returns all records from the left table and the matched records from the right table.
    - RIGHT JOIN: Similar to LEFT JOIN but returns all records from the right table.
    - FULL OUTER JOIN: Returns all records when there is a match in either left or right records.

4. **ORDER BY**: This clause sorts the query results in ascending or descending order based on one or more columns.

Additional SQL Features for Data Management

- INSERT: Adds new data into a table.
- UPDATE: Modifies existing data in a table.
- DELETE: Removes data from a table.
- CREATE TABLE: Creates a new table.
- DROP TABLE: Deletes a table.
- ALTER TABLE: Modifies the structure of a table.

Ethics and Bias in Data Analysis

Ethics and Bias in Data Analysis are essential factors to consider for making sure that data-driven decisions are fair, reliable, and respectful of individual rights:

Data Accuracy

- Correct Web Scraping: It's important to ensure that web scraping techniques are precise and do not introduce errors into the dataset. This means validating the data against known sources and checking for any inconsistencies.

- Proper Data Processing: Implementing strong data processing methods is crucial for managing missing values, outliers, and inconsistencies. This includes using suitable statistical techniques to clean and transform the data.

- Data Validation: Regularly validating data against external benchmarks or datasets is necessary to maintain accuracy and reliability.

Bias Consideration

- Avoiding Misrepresentation: It's vital to ensure that data analysis does not misrepresent the underlying reality. This requires being aware of potential biases in data collection, processing, and interpretation.

- Algorithmic Bias: It's important to recognize that algorithms can inherit biases from the data they are trained on, which can lead to unfair outcomes. Techniques such as data preprocessing and model auditing can help reduce these biases.
- Ensuring fairness in analysis involves adopting strategies that promote equity in data-driven decision-making. This can include the use of fairness metrics and tools designed to detect bias.

Summary

This project analyzed **foreign satellites launched by ISRO**, following a structured approach: data collection, cleaning, exploratory analysis, and visualization. The goal was to uncover trends in **launch frequency, payload distribution,** and **ISRO's** role in **global satellite launches**.

Key Findings

- ISRO's foreign satellite launches have grown steadily, establishing it as a major global launch provider.
- **PSLV** is the most frequently used launch vehicle, reinforcing its reliability and cost-effectiveness.
- Satellite mass varies significantly, reflecting diverse mission needs.
- Global collaborations in space technology are increasing, with ISRO playing a key role.

What I've Learned

Through this project, I gained valuable insights into data analysis, web scraping, and exploratory data visualization. I learned how to efficiently extract structured data from the web using BeautifulSoup, clean and preprocess it with Pandas, and create meaningful visualizations with **Matplotlib, Seaborn, and Power BI**. Analyzing ISRO's foreign satellite launches helped me understand trends in global space collaborations, the significance of different launch vehicles, and the impact of payload variations on mission planning.

Additionally, I realized the importance of data cleaning in ensuring accuracy and consistency, as raw datasets often contain inconsistencies and missing values. The project also reinforced the value of exploratory data analysis (EDA) in uncovering patterns that may not be immediately obvious. Most importantly, I learned how **data-driven insights can guide decision-making**, such as predicting future launch trends and optimizing launch strategies.

This project provided a strong foundation for further exploration into predictive analytics and machine learning applications in space technology. It also deepened my appreciation for ISRO's growing role in the **commercial satellite launch market** and India's increasing influence in the global space industry.

Conclusion

This study highlights ISRO's growing influence in commercial satellite launches. It provides valuable insights into how India has positioned itself as a dependable and economical space launch service provider. The findings also pave the way for predictive modeling to anticipate future trends.

With further **machine learning and advanced analytics**, this research could help ISRO optimize its launch strategies and partnerships, shaping the future of global space exploration.

Further Exploration and Future Prospects

This project serves as a strong foundation for deeper analysis into **ISRO's role in global space collaborations** and the evolving trends in commercial satellite launches. While we focused on past launch patterns, the next step could involve **predictive analytics** to forecast future trends. Machine learning models, such as time series forecasting, could be applied to predict the number of satellite launches ISRO might undertake in the coming years based on historical data.

Another area worth exploring is **comparative analysis** between ISRO and other major space organizations like NASA, SpaceX, Roscosmos, and ESA. Studying differences in launch frequency, cost-effectiveness, payload capacities, and mission success rates could provide valuable insights into India's competitive position in the space industry. This could also help identify potential areas of improvement for ISRO's commercial space operations.

Furthermore, **geospatial analysis** using GIS tools could help visualize launch site distributions and trajectory patterns of foreign satellites launched by ISRO. This would offer insights into how different orbits are utilized for various mission types.

On a broader scale, this research could be expanded to study the economic impact of ISRO's satellite launch services, assessing how foreign collaborations contribute to India's space economy. Additionally, analyzing policy frameworks and international agreements could provide a clearer picture of how geopolitical factors influence ISRO's commercial satellite launches.

By integrating big data analytics, AI-driven insights, and real-time satellite tracking, future research can help optimize **ISRO's strategic planning and decision-making**, ultimately strengthening India's role in global space exploration.

Key Insights

- ISRO has positioned itself as a leading commercial launch service provider.
- Launch trends indicate a rise in small satellite launches, reflecting global demand.
- Polar Satellite Launch Vehicle (PSLV) remains the most reliable and frequently used launch vehicle.
- ISRO's commercial partnerships are expanding, reinforcing India's space economy.

References

- Wikipedia: List of Foreign Satellites Launched by India
- **Pandas Documentation**: https://pandas.pydata.org/
- **Matplotlib Documentation**: https://matplotlib.org/
- **Seaborn Documentation**: https://seaborn.pydata.org/
- **Power BI Official Guide**: https://powerbi.microsoft.com/en-us/
- **SQL Tutorials**: https://www.w3schools.com/sql/