**All Space Missions from 1957**

**(Data Analysis and Visualization)**

Here is a detail on what the columns represent:

**Company Name**: The organization responsible for launching the mission (e.g., NASA, SpaceX, Roscosmos).

**Location**: The launch site or base (e.g., Kennedy Space Center, Baikonur Cosmodrome).

**Datum**: Likely the date of the mission (in German, "Datum" means "date"). It may be formatted as YYYY-MM-DD or another date format.

**Detail**: Rocket Name

**Status Rocket**: The condition or outcome of the rocket itself (StatusRetired, StatusActive).

**Cost (Million $)**: The cost of the mission.

**Status Mission**: The overall mission outcome (e.g., Success, Failure).

Cleaning Steps:

Since the dataset was in good condition, basic checking was done to figure out a few things.

First checked for null columns. Since only the Cost column has missing columns which totally makes sense, and that too approximately 77%, I’ve decided to leave it that way. Still leaving the code for calculating the medium.

df['Cost (Million $)'].fillna(df['Cost (Million $)'].median(), inplace=True)

**ipynb File Code**

import pandas as pd

# Load the dataset

file\_path = 'Space\_Corrected.csv'

df = pd.read\_csv(file\_path)

#Inspecting the Data

print(df.info())

print(df.head())

print(df.describe())

#Handling Mission values

print(df.isnull().sum())

Now I’ve realized there is a problem with the “Datum” Column. It doesn’t comply with the Power BI Terms of converting to date/time/timezone format. Below are the details.

The date format "Fri Aug 07, 2020 05:12 UTC" is not directly recognized as a standard date format in Power BI when converting to Date/Time/Timezone. This happens because Power BI expects a more structured date format.

**Why Does This Error Occur?**

* The format includes **weekday names** (e.g., Fri) and **timezone information** (UTC), which Power BI does not automatically parse in Date/Time/Timezone type.
* Power BI expects formats like:
  + 2020-08-07 05:12:00
  + 08/07/2020 05:12 AM

**Solution: Steps to Fix It in Power BI**

Here’s how you can clean and convert the date correctly:

**Step 1: Removed Unnecessary Parts**

* Use **Power Query** to clean the data.

1. Go to **Power Query Editor**.
2. Select the date column.
3. Navigate to **Add Column** → **Custom Column**.
4. Use this formula to remove the weekday and timezone:

powerquery

CopyEdit

= Text.Middle([Datum], 4, 20)

✅ This keeps only "Aug 07, 2020 05:12".

**Step 2: Converted to Date/Time**

1. Select the cleaned column.
2. Go to **Transform** → **Data Type** → **Date/Time**.

After some type conversions and removing of unnecessary columns, the dataset was ready for visualization.

Now there are various questions that can be asked from this data.

**Mission Trends Over Time**

1. **How has the number of space missions evolved over the years?**  
   📊 *Line plot or bar chart showing mission count by year.*
2. **Which year had the highest number of successful missions?**  
   📈 *Line plot of successful missions over time.*

**Company and Organization Analysis**

1. **Which companies have conducted the most missions?**  
   📊 *Bar chart showing mission count by company.*
2. **Which companies have the highest success rates?**  
   📊 *Bar chart showing success rate (%) by company.*
3. **How do different companies compare in terms of mission costs?**  
   📊 *Box plot showing cost distribution by company.*

**Location Insights**

1. **Which launch sites are most frequently used?**  
   📍 *Bar chart or map showing mission count by location.*
2. **Are certain locations more prone to mission failures?**  
   ⚠️ *Bar chart comparing successful vs failed missions by location.*

**Cost Analysis**

1. **What is the trend of mission costs over time?**  
   📈 *Line plot showing average cost per year.*
2. **Which missions had the highest costs, and were they successful?**  
   📊 *Bar chart of top 10 most expensive missions with status indicators.*

**Mission Outcome Patterns**

1. **What is the overall success rate of space missions?**  
   🟢🔴 *Pie chart showing success, failure, and partial success rates.*
2. **Are certain years more prone to failures than others?**  
   📈 *Stacked bar chart showing successful vs failed missions by year.*
3. **Is there a correlation between mission cost and success?**  
   📉 *Scatter plot of mission cost vs success status.*

**Detailed Analysis**

1. **What types of missions were conducted most frequently?**  
   📊 *Bar chart showing mission types from the 'Detail' column.*
2. **Which companies are investing more in high-cost missions?**  
   💰 *Bar chart comparing total expenditure by company.*

And after a bit of playing with chatgpt, I got a background image (wireframe) created using DALL-E tech.