Note:

This is the original blank project template from Angela Yu's 100 Days course, to reference and compare with `main.py` and see which questions/details are hers and which solutions/notes are mine.

Some of the formatting, etc..., I've changed for my organizational needs, so you can easily see what's different when looking back and forth between the two .ipynb files.

Introduction

This dataset was scraped from nextspaceflight.com and includes all the space missions since the beginning of Space Race between the USA and the Soviet Union in 1957!

Install Package with Country Codes

In [ ]:
```
%pip install iso3166
```

Upgrade Plotly

Run the cell below if you are working with Google Colab.

In [ ]:
```
%pip install --upgrade plotly
```

Import Statements

In [ ]:
```
import numpy as np
import pandas as pd
import plotly.express as px
```

```
import matplotlib.pyplot as plt
import seaborn as sns

# These might be helpful:
from iso3166 import countries
from datetime import datetime, timedelta
```

Notebook Presentation

In [ ]:
```
pd.options.display.float_format = '{:,.2f}'.format
```

Load the Data

In [ ]:
```
df_data = pd.read_csv('mission_launches.csv')
```

Preliminary Data Exploration

- What is the shape of `df_data` ?
- How many rows and columns does it have?
- What are the column names?
- Are there any NaN values or duplicates?

In [ ]:

In [ ]:

Data Cleaning - Check for Missing Values and Duplicates

Consider removing columns containing junk data.

In [ ]:

In [ ]:

Descriptive Statistics

In [ ]:

In [ ]:

Number of Launches per Company

Create a chart that shows the number of space mission launches by organisation.

In [ ]:

In [ ]:

Number of Active versus Retired Rockets

How many rockets are active compared to those that are decomissioned?

In [ ]:

In [ ]:

Distribution of Mission Status

How many missions were successful? How many missions failed?

In [ ]:

In [ ]:

How Expensive are the Launches?

Create a histogram and visualise the distribution. The price column is given in USD millions (careful of missing values).

In [ ]:

In [ ]:

Use a Choropleth Map to Show the Number of Launches by Country

- Create a choropleth map using the plotly documentation
- Experiment with plotly's available colours. I quite like the sequential colour `matter` on this map.
- You'll need to extract a `country` feature as well as change the country names that no longer exist.

Wrangle the Country Names

You'll need to use a 3 letter country code for each country. You might have to change some country names.

- Russia is the Russian Federation
- New Mexico should be USA
- Yellow Sea refers to China
- Shahrud Missile Test Site should be Iran
- Pacific Missile Range Facility should be USA
- Barents Sea should be Russian Federation
- Gran Canaria should be USA

You can use the iso3166 package to convert the country names to Alpha3 format.

In [ ]:

In [ ]:

Use a Choropleth Map to Show the Number of Failures by Country

In [ ]:

In [ ]:

Create a Plotly Sunburst Chart of the countries, organisations, and mission status.

In [ ]:

In [ ]:

In [ ]:

Analyse the Total Amount of Money Spent by Organisation on Space Missions

In [ ]:

In [ ]:

In [ ]:

Analyse the Amount of Money Spent by Organisation per Launch

In [ ]:

In [ ]:

In [ ]:

Chart the Number of Launches per Year

In [ ]:

In [ ]:

Chart the Number of Launches Month-on-Month until the Present

Which month has seen the highest number of launches in all time? Superimpose a rolling average on the month on month time series chart.

In [ ]:

In [ ]:

Launches per Month: Which months are most popular and least popular for launches?

Some months have better weather than others. Which time of year seems to be best for space missions?

In [ ]:

In [ ]:

How has the Launch Price varied Over Time?

Create a line chart that shows the average price of rocket launches over time.

In [ ]:

In [ ]:

Chart the Number of Launches over Time by the Top 10 Organisations.

How has the dominance of launches changed over time between the different players?

In [ ]:

In [ ]:

In [ ]:

Cold War Space Race: USA vs USSR

The cold war lasted from the start of the dataset up until 1991.

In [ ]:

In [ ]:

Create a Plotly Pie Chart comparing the total number of launches of the USSR and the USA

Hint: Remember to include former Soviet Republics like Kazakhstan when analysing the total number of launches.

In [ ]:

In [ ]:

Create a Chart that Shows the Total Number of Launches Year-On-Year by the Two Superpowers

In [ ]:

In [ ]:

Chart the Total Number of Mission Failures Year on Year.

In [ ]:

In [ ]:

Chart the Percentage of Failures over Time

Did failures go up or down over time? Did the countries get better at minimising risk and improving their chances of success over time?

In [ ]:

In [ ]:

In [ ]:

For Every Year Show which Country was in the Lead in terms of Total Number of Launches (up to and including including 2020)

Do the results change if we only look at the number of successful launches?

In [ ]:

In [ ]:

Create a Year-on-Year Chart Showing the Organisation Doing the Most Number of Launches

Which organisation was dominant in the 1970s and 1980s? Which organisation was dominant in 2018, 2019 and 2020?

In [ ]:

In [ ]:

In [ ]: