IBM Developer
SKILLS NETWORK

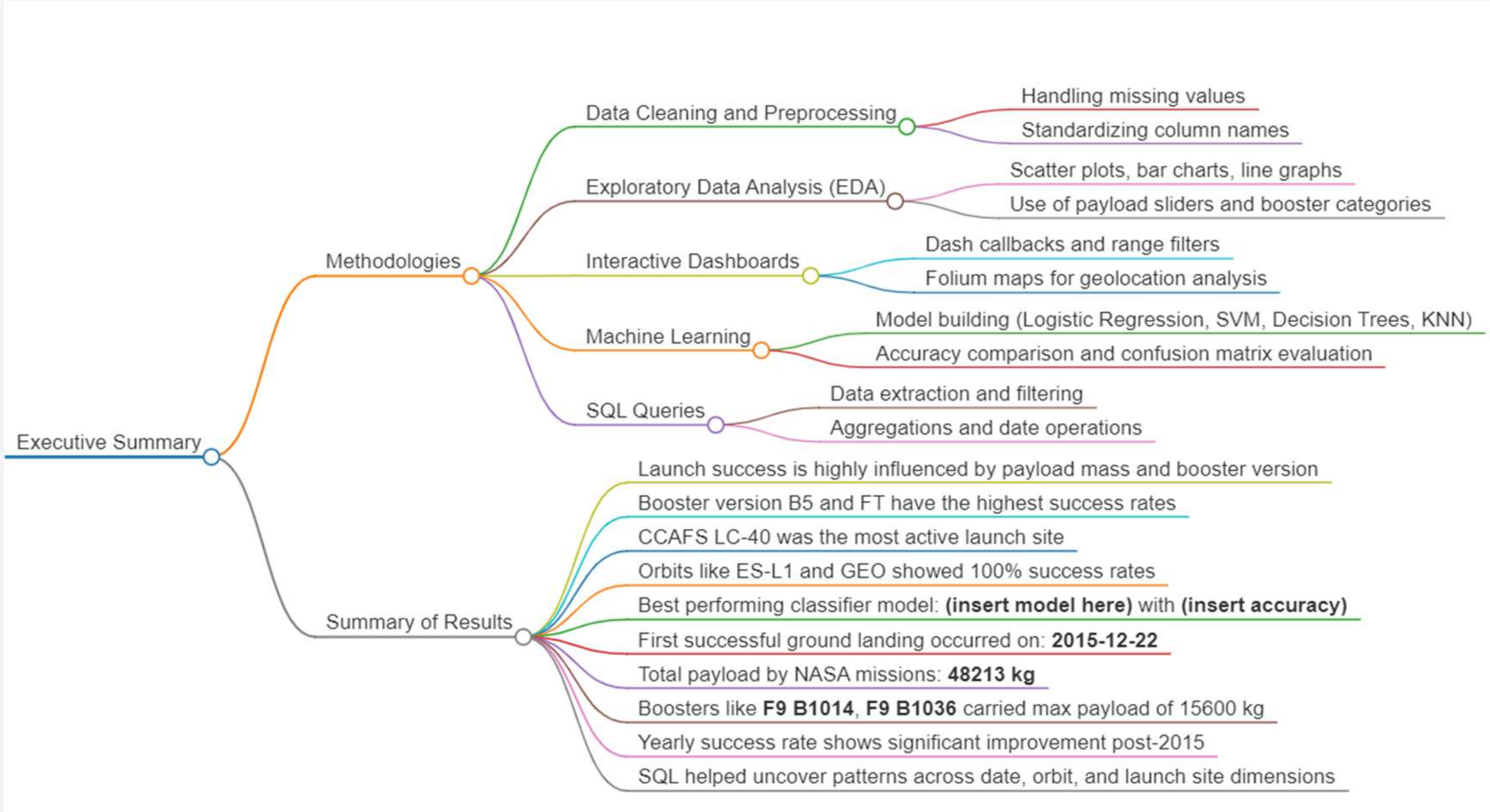
Winning Space Race with Data Science

Alexis José Rojas Briceño
2025/06

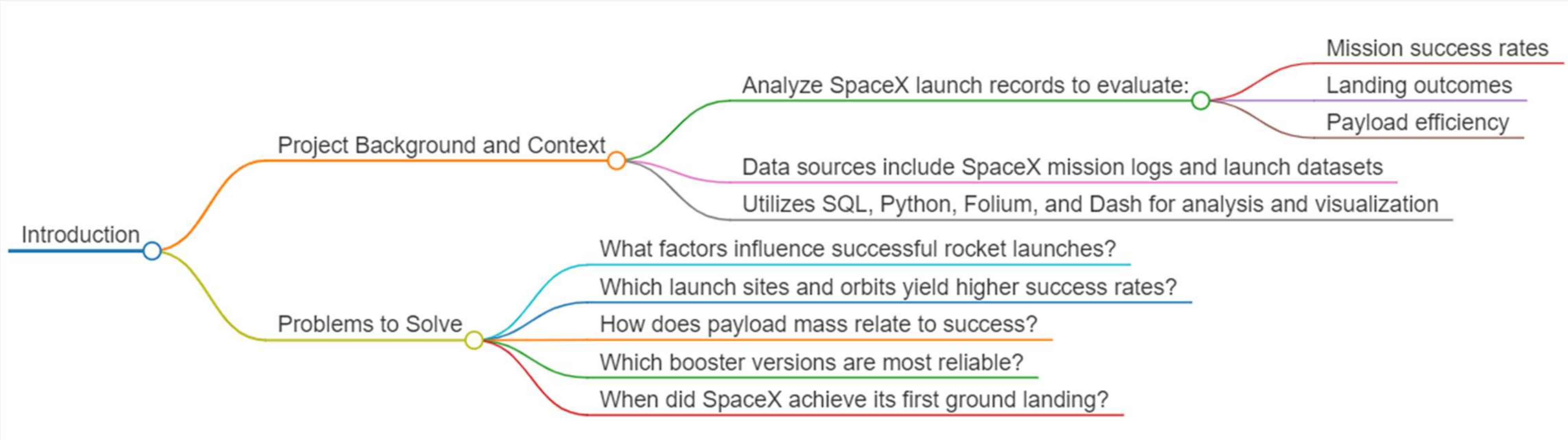
Outline

- Executive Summary
- Introduction
- Methodology
- Results
- Conclusion
- Appendix

Executive Summary



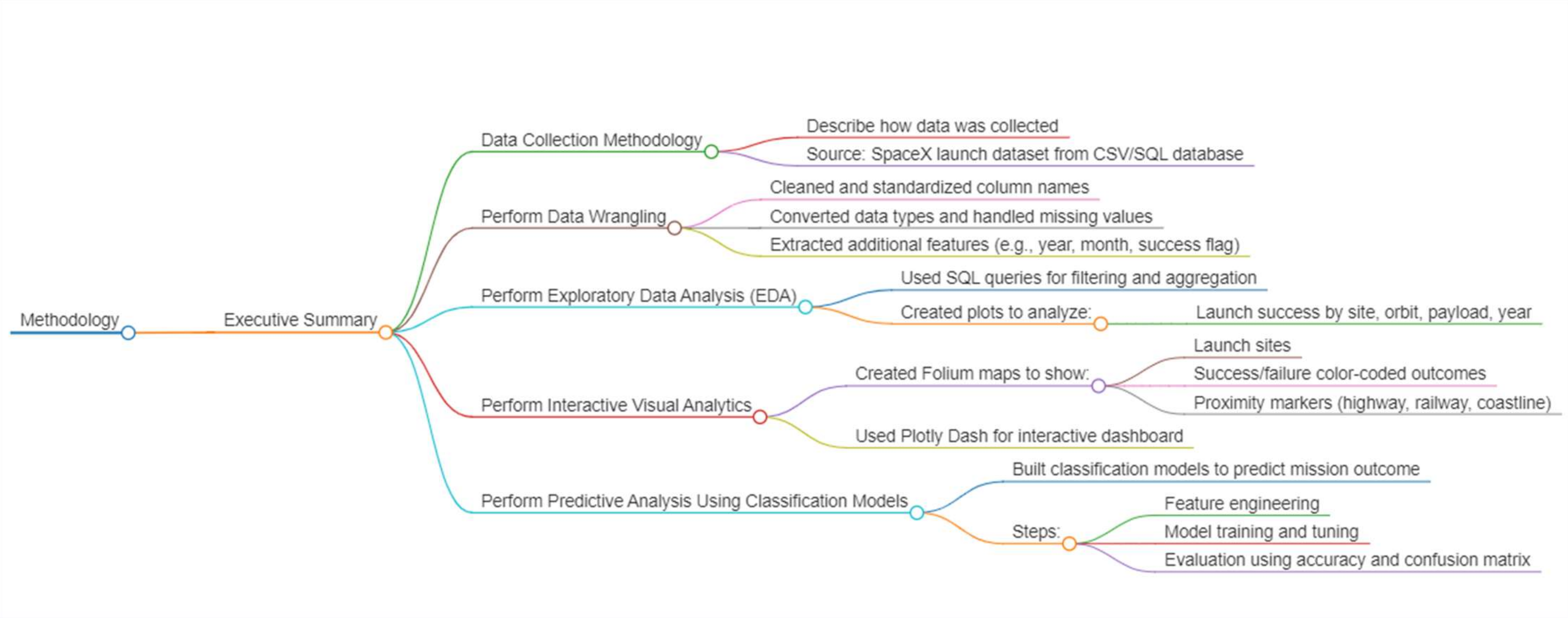
Introduction



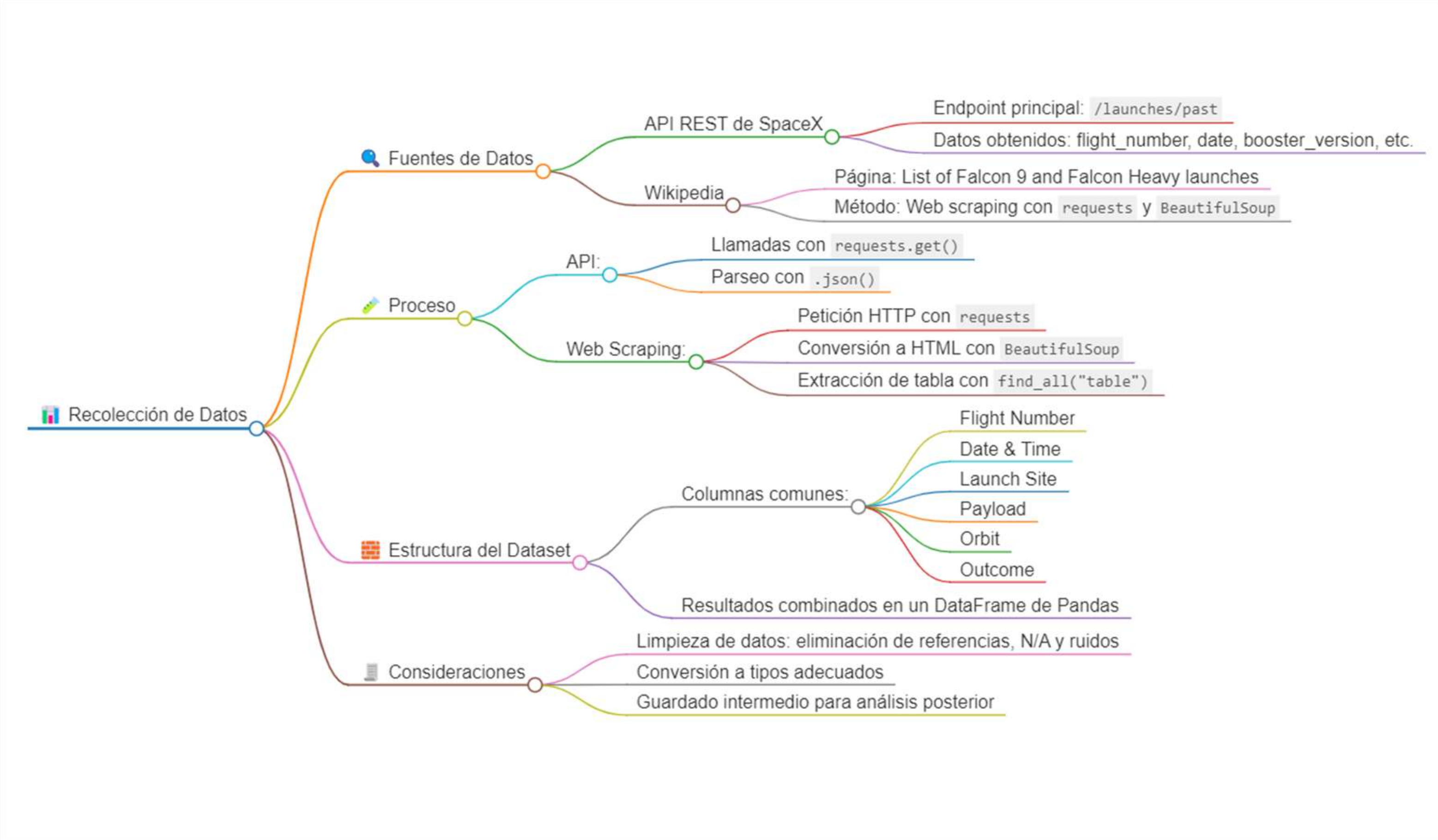
Section 1

Methodology

Methodology

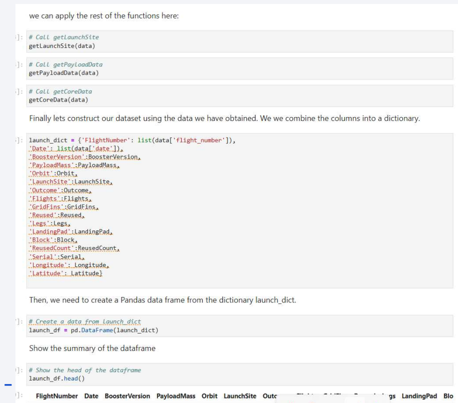


Data Collection



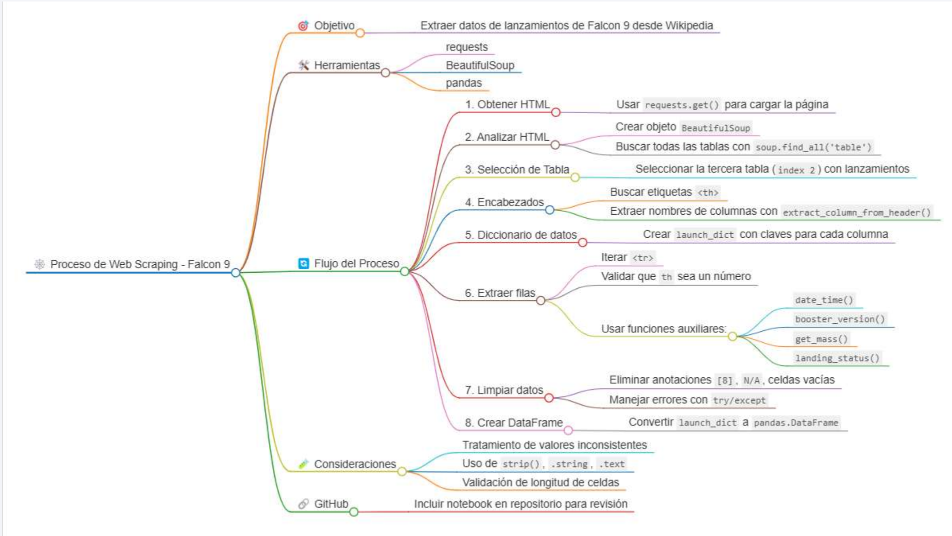
Data Collection – SpaceX API

- Present your data collection with SpaceX REST calls using key phrases and flowcharts
- Add the GitHub URL of the completed SpaceX API calls notebook (must include completed code cell and outcome cell), as an external reference and peer-review purpose
- https://github.com/Alexis-Jose-Rojas/ciencia_de_datos_aplicada/blob/main/jupyter-labs-spacex-data-collection-api.ipynb



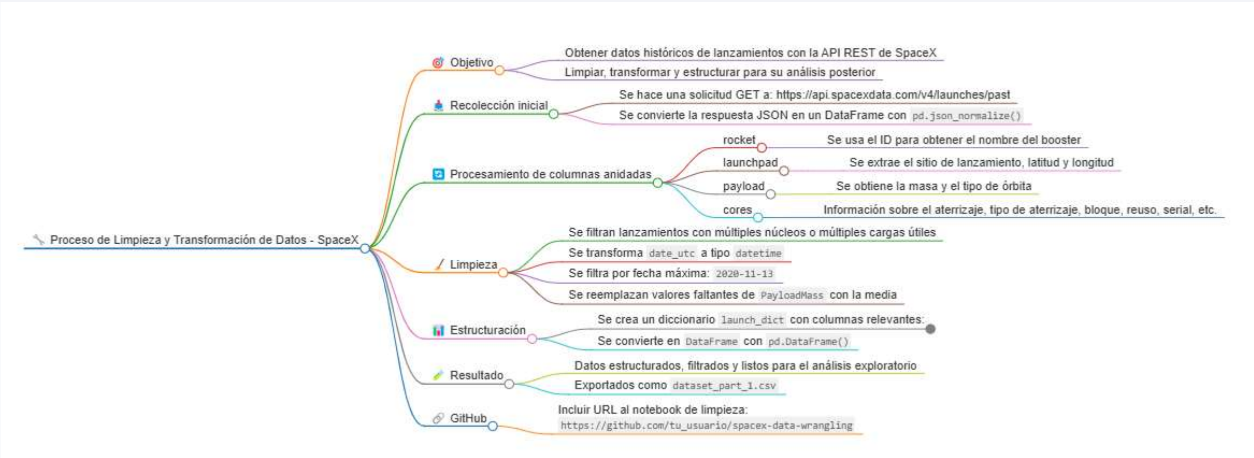
Data Collection - Scraping

- Present your web scraping process using key phrases and flowcharts
- Add the GitHub URL of the completed web scraping notebook, as an external reference and peer-review purpose
- https://github.com/Alexis-Jose-Rojas/ciencia_de_datos_aplicada/blob/main/jupyter-labs-webscraping.ipynb



Data Wrangling

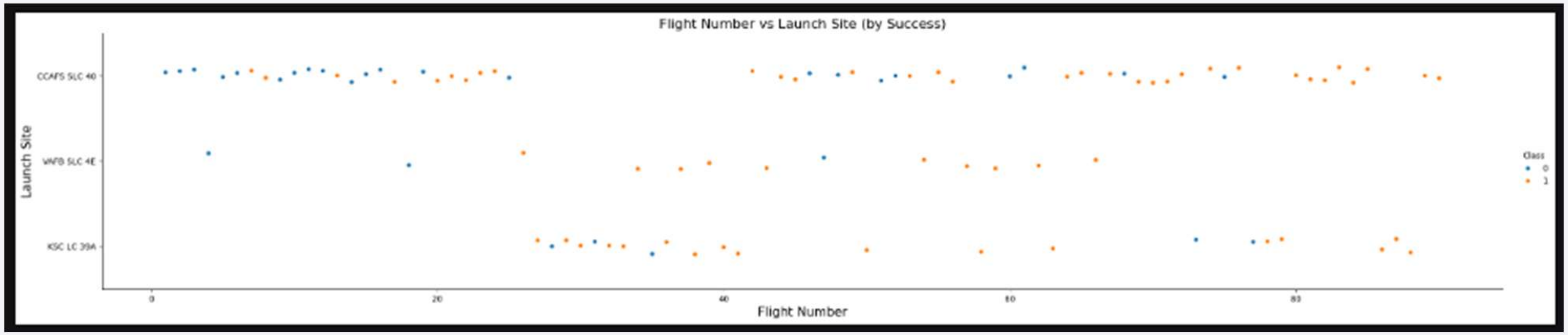
- Describe how data were processed
- You need to present your data wrangling process using key phrases and flowcharts
- Add the GitHub URL of your completed data wrangling related notebooks, as an external reference and peer-review purpose
- https://github.com/Alexis-Jose-Rojas/ciencia_de_datos_aplicada/blob/main/labs-jupyter-spacex-Data%20wrangling.ipynb



EDA with Data Visualization

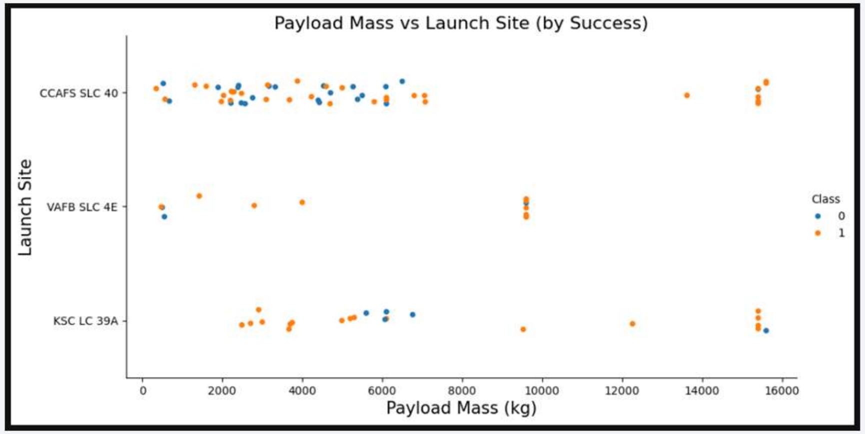
- Summarize what charts were plotted and why you used those charts
- Add the GitHub URL of your completed EDA with data visualization notebook, as an external reference and peer-review purpose
- https://github.com/Alexis-Jose-Rojas/ciencia_de_datos_aplicada/blob/main/edadataviz%20(1).ipynb

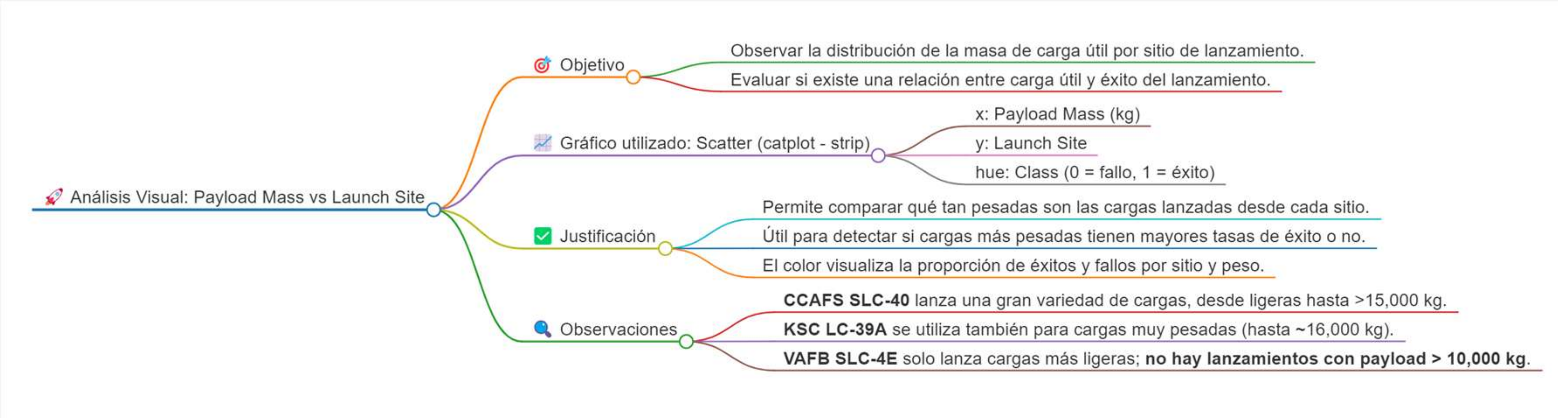
EDA with Data Visualization - Scatter



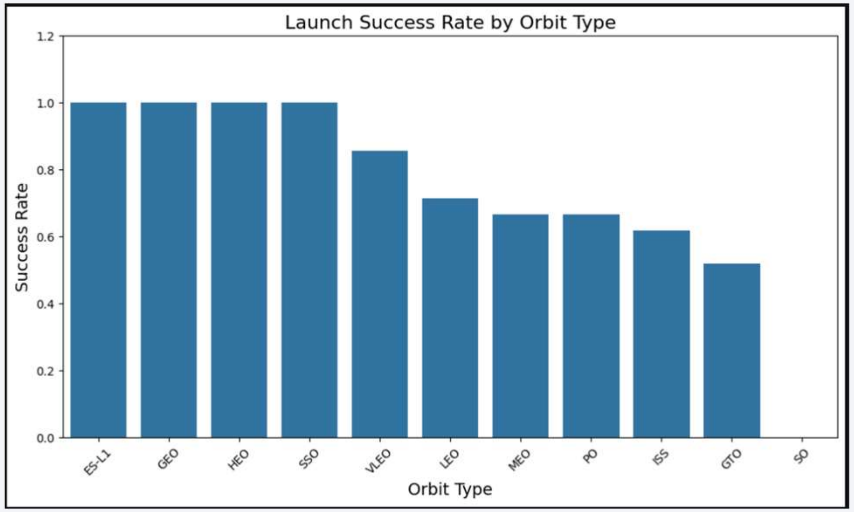
- 📊 Análisis Visual: Flight Number vs Launch Site
  - 🎯 Objetivo
    - Explorar la distribución de lanzamientos según el número de vuelo y el sitio de lanzamiento.
    - Evaluar cómo varía la tasa de éxito a lo largo del tiempo por ubicación.
  - 📈 Gráfico utilizado: Scatter (catplot - strip)
    - x: FlightNumber
    - y: LaunchSite
    - hue: Class (0 = fallo, 1 = éxito)
  - ✅ Justificación
    - Permite ver visualmente cómo se distribuyen los lanzamientos en cada sitio.
    - Muestra evolución temporal implícita (FlightNumber ≈ cronología).
    - El color distingue lanzamientos exitosos de fallidos por ubicación.
  - 🔍 Observaciones
    - **CCAFS SLC-40** tiene un gran número de vuelos, con una mejora de éxito visible.
    - **KSC LC-39A** parece tener alta tasa de éxito.
    - **VAFB SLC-4E** tiene lanzamientos más espaciados y menor volumen.

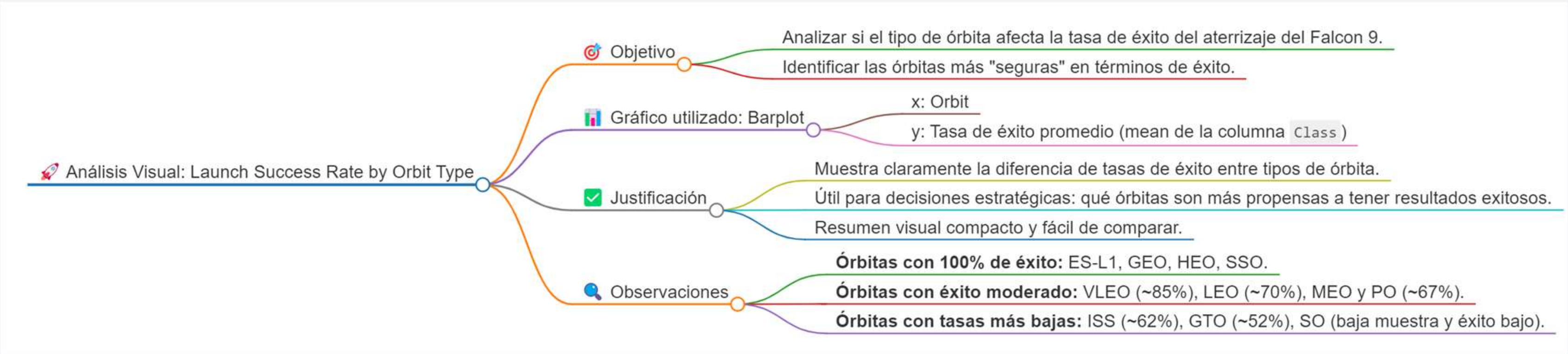
EDA with Data Visualization - Scatter



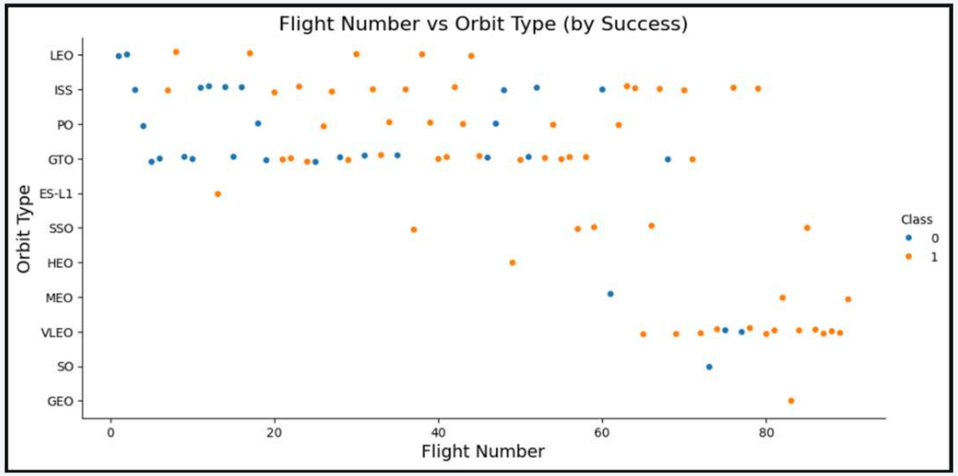


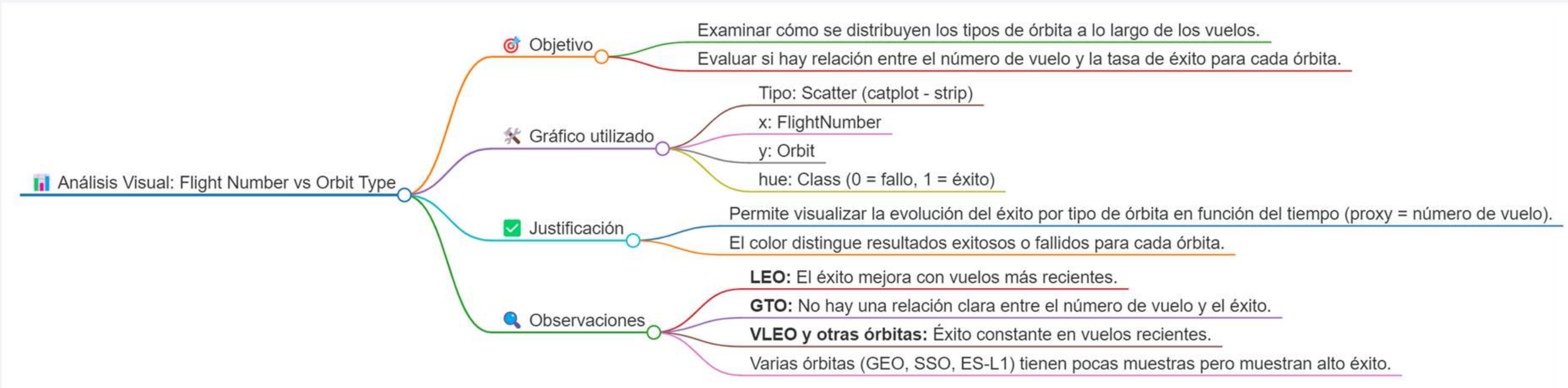
EDA with Data Visualization - Barplot



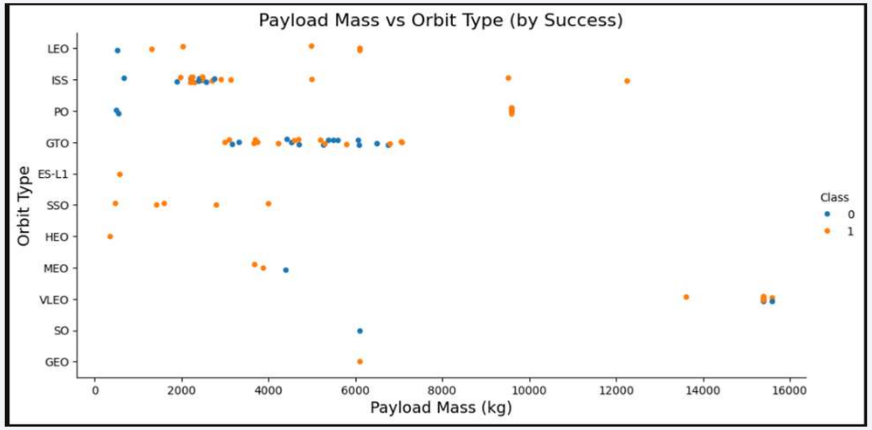


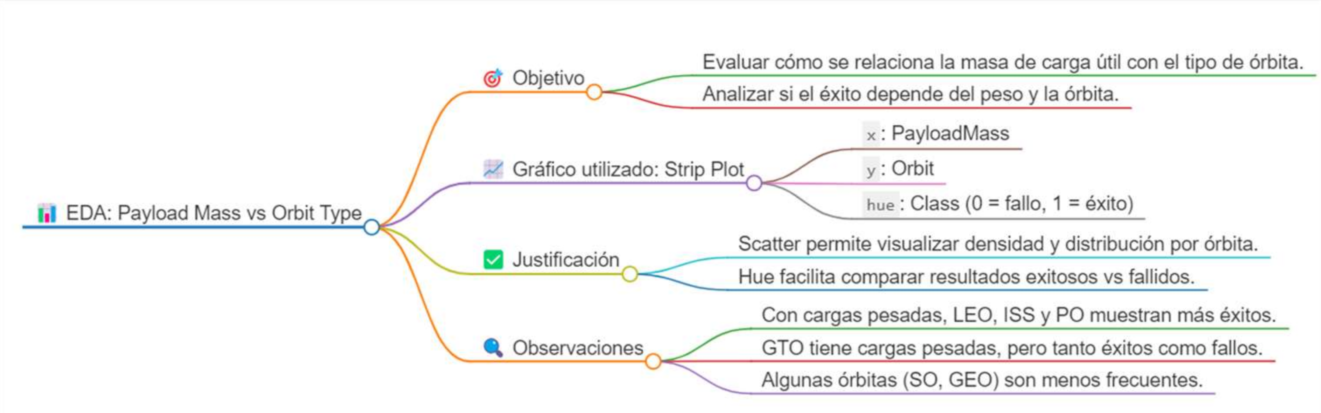
EDA with Data Visualization - Catplot



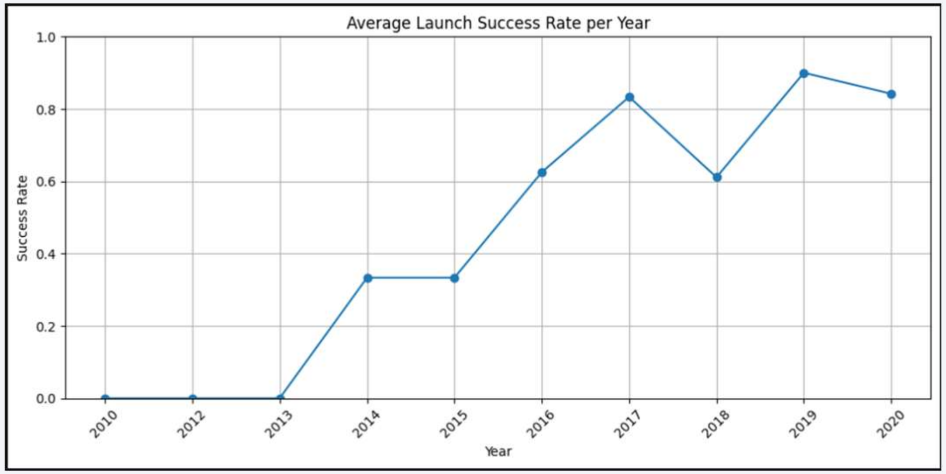


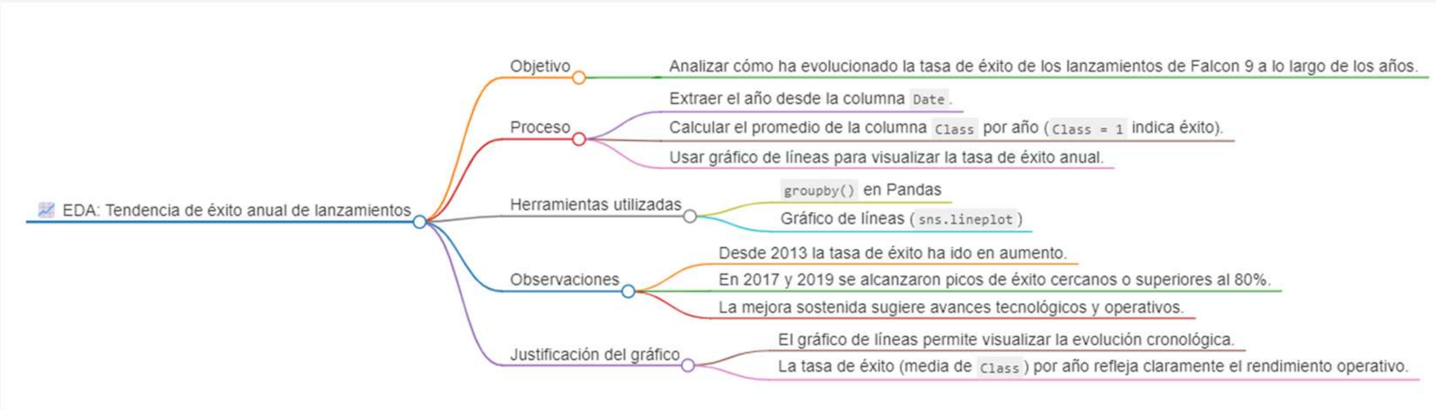
EDA with Data Visualization - Catplot





EDA with Data Visualization – Line Chart





EDA with SQL

- Using bullet point format, summarize the SQL queries you performed
  - **Consulta de conteo total de lanzamientos**
    - SELECT COUNT(*) FROM SPACEXTBL
  - **Filtrado por año de lanzamiento**
    - SELECT * FROM SPACEXTBL WHERE Date BETWEEN '2010-01-01' AND '2017-12-31'
  - **Agrupación por sitio de lanzamiento**
    - SELECT Launch_Site, COUNT(*) FROM SPACEXTBL GROUP BY Launch_Site
  - **Tasa de éxito por sitio de lanzamiento**
    - SELECT Launch_Site, AVG(Class) AS SuccessRate FROM SPACEXTBL GROUP BY Launch_Site
  - **Tipos de órbita más frecuentes**
    - SELECT Orbit, COUNT(*) FROM SPACEXTBL GROUP BY Orbit ORDER BY COUNT(*) DESC
  - **Éxito por tipo de órbita**
    - SELECT Orbit, AVG(Class) AS SuccessRate FROM SPACEXTBL GROUP BY Orbit
  - **Masa promedio de carga útil por sitio**
    - SELECT Launch_Site, AVG(PayloadMass) FROM SPACEXTBL GROUP BY Launch_Site

EDA with SQL

- Using bullet point format, summarize the SQL queries you performed

  - **Consulta de conteo total de lanzamientos**
    - SELECT COUNT(*) FROM SPACEXTBL
  - **Filtrado por año de lanzamiento**
    - SELECT * FROM SPACEXTBL WHERE Date BETWEEN '2010-01-01' AND '2017-12-31'
  - **Agrupación por sitio de lanzamiento**
    - SELECT Launch_Site, COUNT(*) FROM SPACEXTBL GROUP BY Launch_Site
  - **Tasa de éxito por sitio de lanzamiento**
    - SELECT Launch_Site, AVG(Class) AS SuccessRate FROM SPACEXTBL GROUP BY Launch_Site

EDA with SQL

- Using bullet point format, summarize the SQL queries you performed

  - **Tipos de órbita más frecuentes**
    - SELECT Orbit, COUNT(*) FROM SPACEXTBL GROUP BY Orbit ORDER BY COUNT(*) DESC

  - **Éxito por tipo de órbita**
    - SELECT Orbit, AVG(Class) AS SuccessRate FROM SPACEXTBL GROUP BY Orbit

  - **Masa promedio de carga útil por sitio**
    - SELECT Launch_Site, AVG(PayloadMass) FROM SPACEXTBL GROUP BY Launch_Site

- Add the GitHub URL of your completed EDA with SQL notebook, as an external reference and peer-review purpose
- https://github.com/Alexis-Jose-Rojas/ciencia_de_datos_aplicada/blob/main/jupyter-labs-eda-sql-coursera_sqllite.ipynb

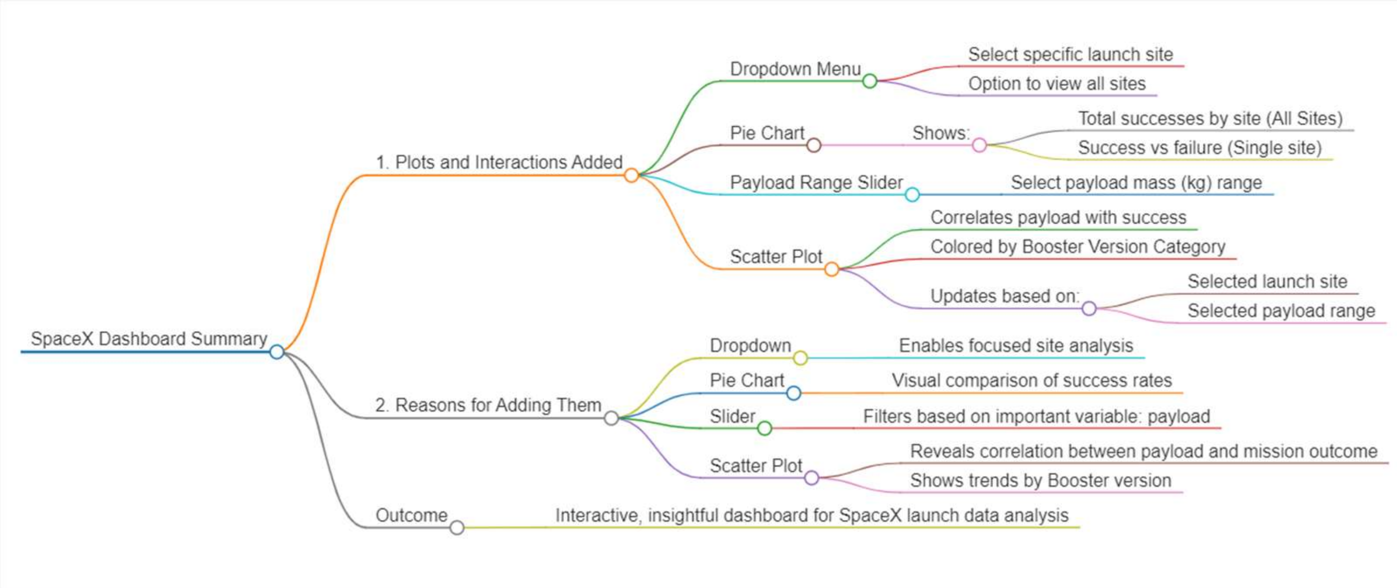
Build an Interactive Map with Folium

- Summarize what map objects such as markers, circles, lines, etc. you created and added to a folium map
- Explain why you added those objects

- Add the GitHub URL of your completed interactive map with Folium map, as an external reference and peer-review purpose
- https://github.com/Alexis-Jose-Rojas/ciencia_de_datos_aplicada/blob/main/lab_jupyter_launch_site_location.ipynb

Build a Dashboard with Plotly Dash

- Summarize what plots/graphs and interactions you have added to a dashboard
- Explain why you added those plots and interactions

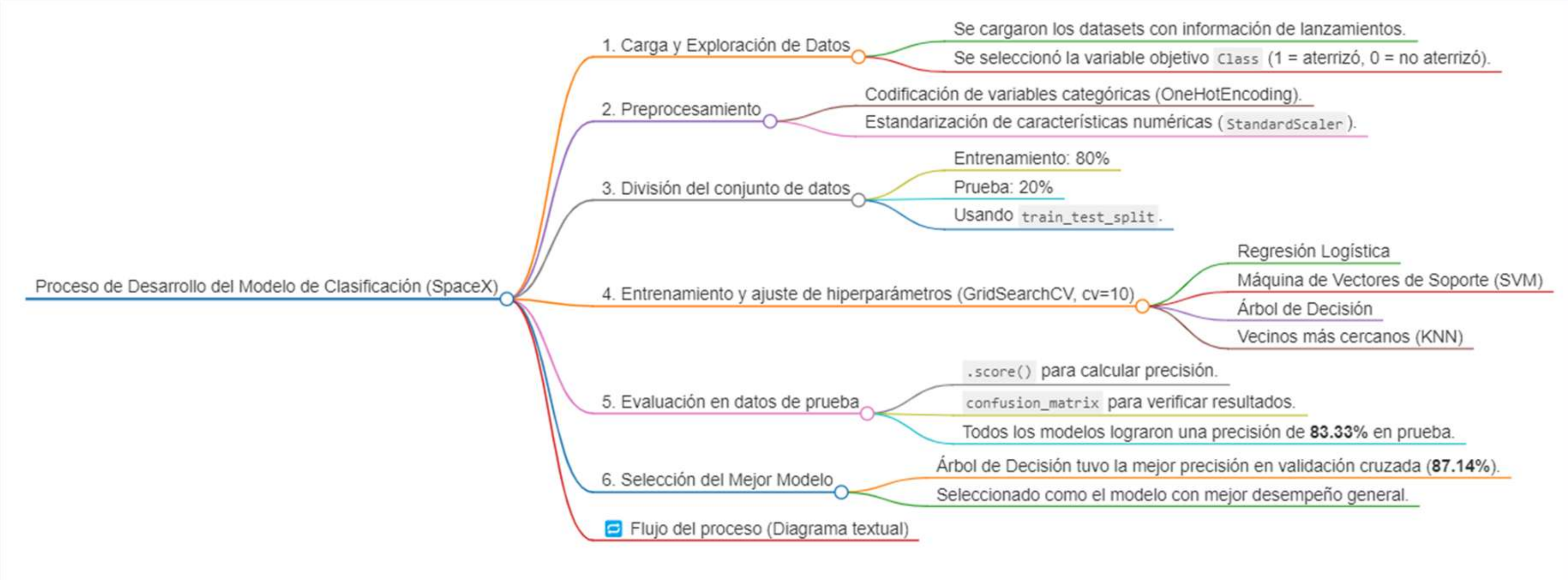


Build a Dashboard with Plotly Dash

- Add the GitHub URL of your completed Plotly Dash lab, as an external reference and peer-review purpose
- https://github.com/Alexis-Jose-Rojas/ciencia_de_datos_aplicada/blob/main/spacex-dash-app%20(1).py

Predictive Analysis (Classification)

- Summarize how you built, evaluated, improved, and found the best performing classification model
- You need present your model development process using key phrases and flowchart



Predictive Analysis (Classification)

- Add the GitHub URL of your completed predictive analysis lab, as an external reference and peer-review purpose
- https://github.com/Alexis-Jose-Rojas/ciencia_de_datos_aplicada/blob/main/SpaceX_Machine%20Learning%20Prediction_Part_5.ipynb

Results

- Exploratory data analysis results
- Interactive analytics demo in screenshots
- Predictive analysis results

Results

1. **Resultados del Análisis Exploratorio (EDA)**
    - Se identificaron las variables más relevantes para predecir aterrizajes: PayloadMass, Orbit, LaunchSite, BoosterVersion, Legs, GridFins, entre otras.
    - Se descubrió que algunos sitios tienen mayor tasa de éxito que otros.
    - La carga útil (PayloadMass) mostró cierta relación con la probabilidad de aterrizaje exitoso.

2. **Capturas de la app analítica interactiva**
    - Se desarrolló un panel interactivo con **Plotly Dash** que incluye:
        - Menú desplegable por sitio de lanzamiento.
        - Gráfico circular para visualizar éxito/fallo por sitio.
        - Control deslizante de rango de carga útil.
        - Gráfico de dispersión para analizar correlación entre carga y éxito, por versión del booster.
    - Los usuarios pueden explorar libremente los datos y responder preguntas clave.
    - Capturas de pantalla incluidas en el repositorio como evidencia del funcionamiento.

Results

3. Resultados del Análisis Predictivo

- Se entrenaron 4 modelos con validación cruzada (GridSearchCV, cv=10):
  - Regresión Logística
  - SVM
  - Árbol de Decisión
  - K-Nearest Neighbors

Conclusión

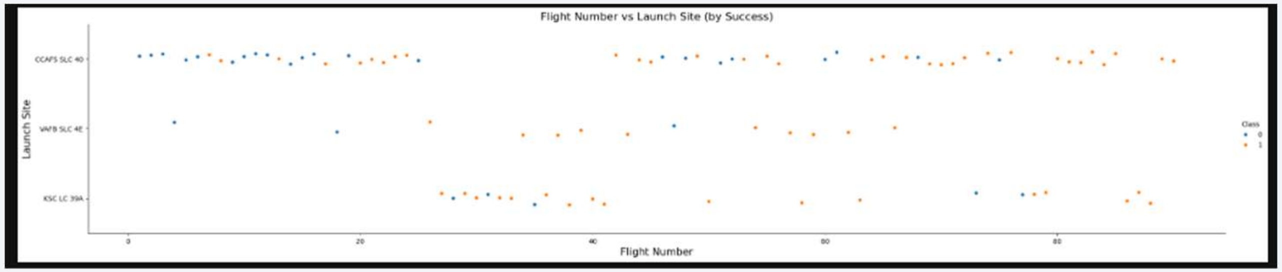
El panel proporciona visualización interactiva en tiempo real y el modelo predictivo entrega alta precisión, lo cual demuestra el valor del análisis aplicado al historial de lanzamientos de SpaceX

Section 2

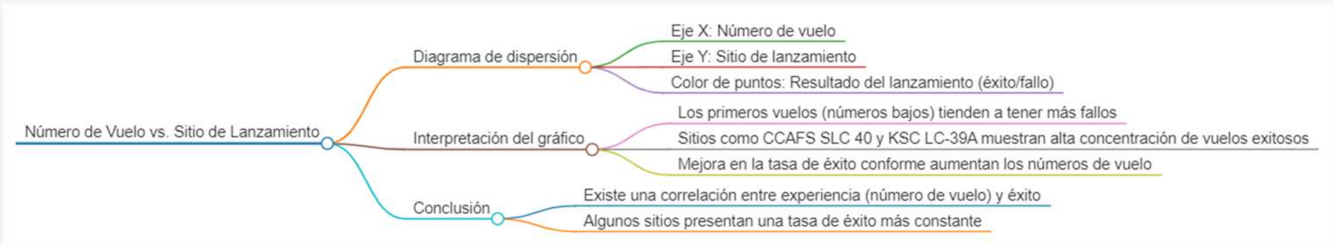
Insights drawn from EDA

Flight Number vs. Launch Site

- Show a scatter plot of Flight Number vs. Launch Site

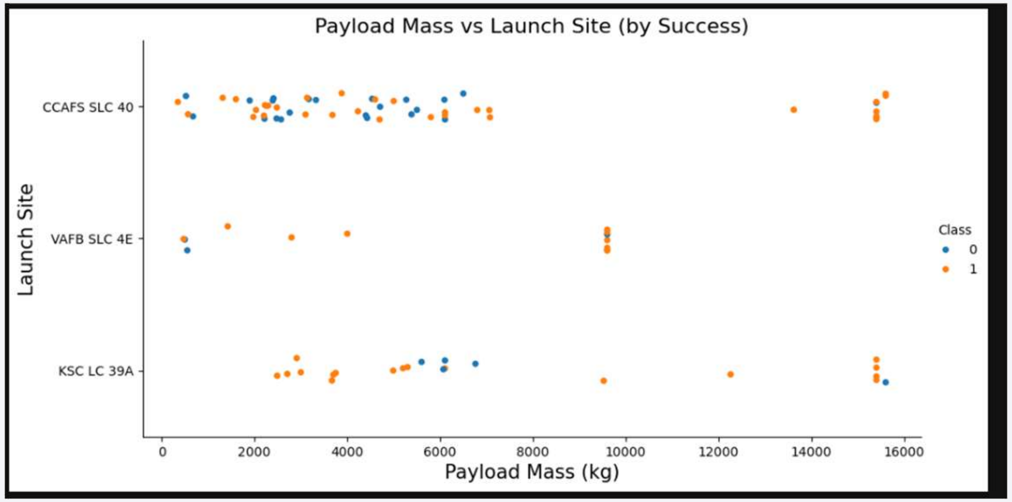


- Show the screenshot of the scatter plot with explanations

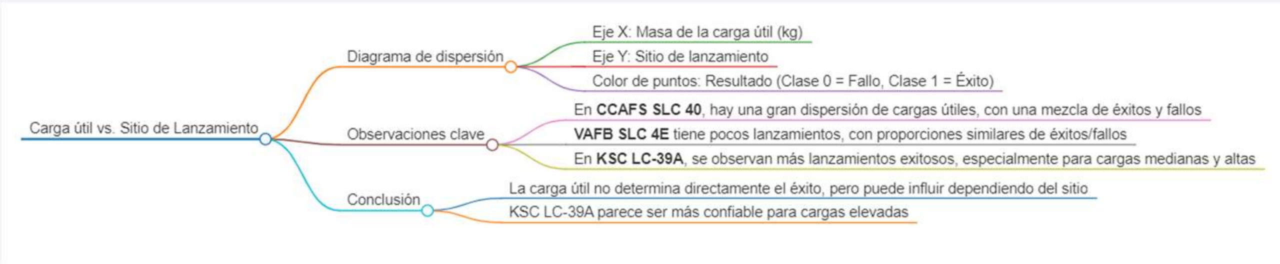


Payload vs. Launch Site

- Show a scatter plot of Payload vs. Launch Site

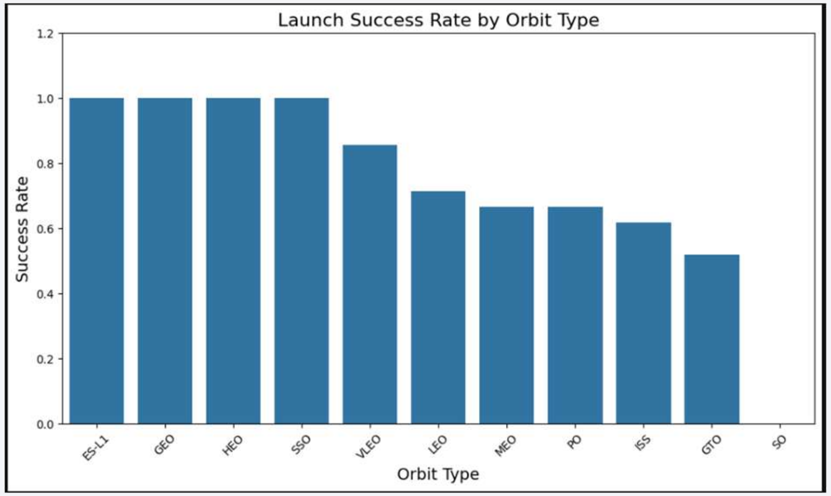


- Show the screenshot of the scatter plot with explanations

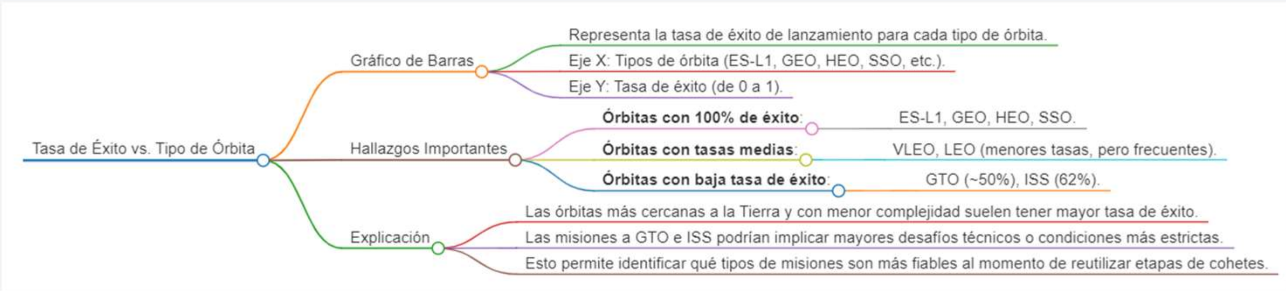


Success Rate vs. Orbit Type

- Show a bar chart for the success rate of each orbit type

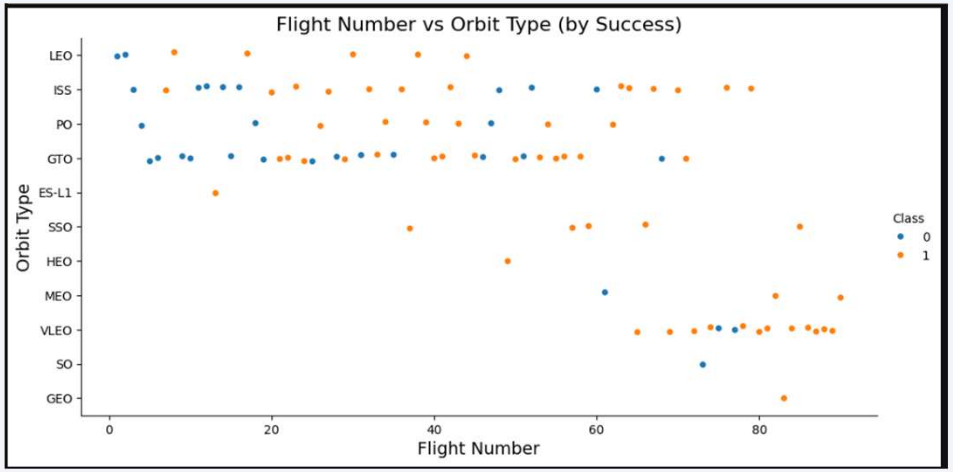


- Show the screenshot of the scatter plot with explanations

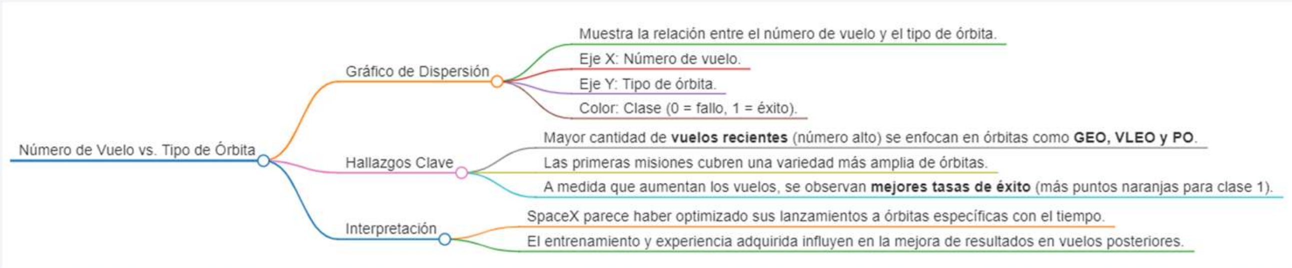


Flight Number vs. Orbit Type

- Show a scatter point of Flight number vs. Orbit type

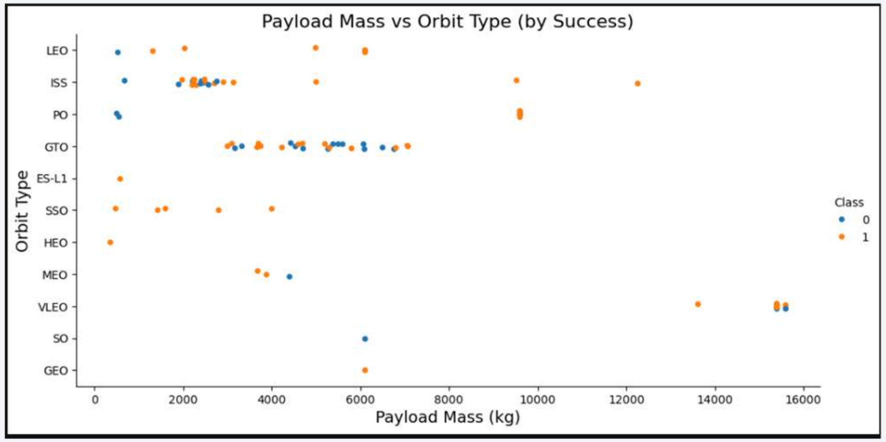


- Show the screenshot of the scatter plot with explanations

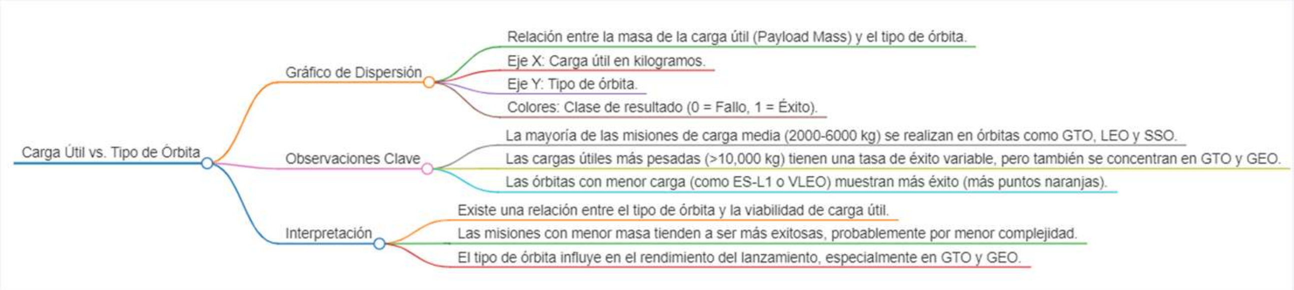


Payload vs. Orbit Type

- Show a scatter point of payload vs. orbit type

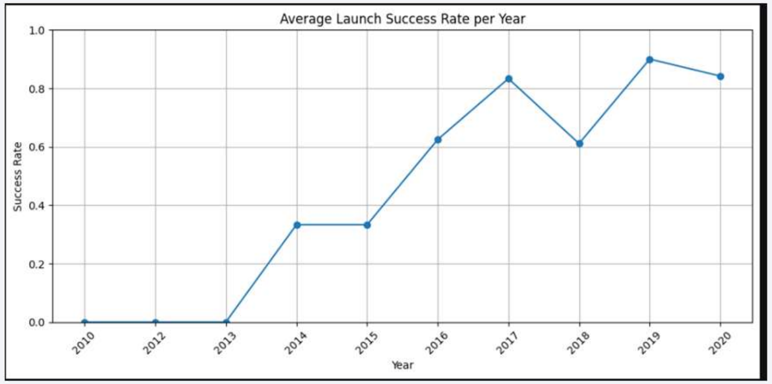


- Show the screenshot of the scatter plot with explanations



Launch Success Yearly Trend

- Show a line chart of yearly average success rate

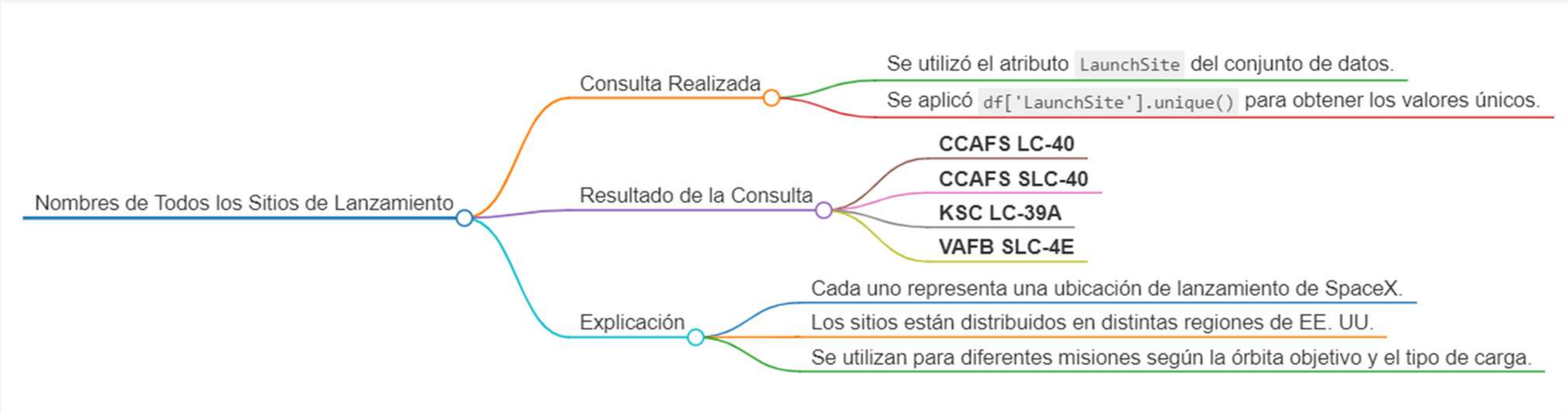


- Show the screenshot of the scatter plot with explanations

All Launch Site Names

- Find the names of the unique launch sites

- Present your query result with a short explanation here



Launch Site Names Begin with 'CCA'

- Find 5 records where launch sites begin with `CCA`

Task 2

Display 5 records where launch sites begin with the string 'CCA'

[13]:
```
%%sql
SELECT *
FROM SPACEXTBL
WHERE Launch_Site LIKE 'CCA%'
LIMIT 5;
```

```
 * sqlite:///my_data1.db
Done.
```

[13]:

Date	Time (UTC)	Booster_Version	Launch_Site	Payload	PAYLOAD_MASS__KG_	Orbit	Customer	Mission_Outcome	Landing_Outcome
2010-06-04	18:45:00	F9 v1.0 B0003	CCAFS LC-40	Dragon Spacecraft Qualification Unit	0	LEO	SpaceX	Success	Failure (parachute)
2010-12-08	15:43:00	F9 v1.0 B0004	CCAFS LC-40	Dragon demo flight C1, two CubeSats, barrel of Brouere cheese	0	LEO (ISS)	NASA (COTS) NRO	Success	Failure (parachute)
2012-05-22	7:44:00	F9 v1.0 B0005	CCAFS LC-40	Dragon demo flight C2	525	LEO (ISS)	NASA (COTS)	Success	No attempt
2012-10-08	0:35:00	F9 v1.0 B0006	CCAFS LC-40	SpaceX CRS-1	500	LEO (ISS)	NASA (CRS)	Success	No attempt
2013-03-01	15:10:00	F9 v1.0 B0007	CCAFS LC-40	SpaceX CRS-2	677	LEO (ISS)	NASA (CRS)	Success	No attempt

- Present your query result with a short explanation here

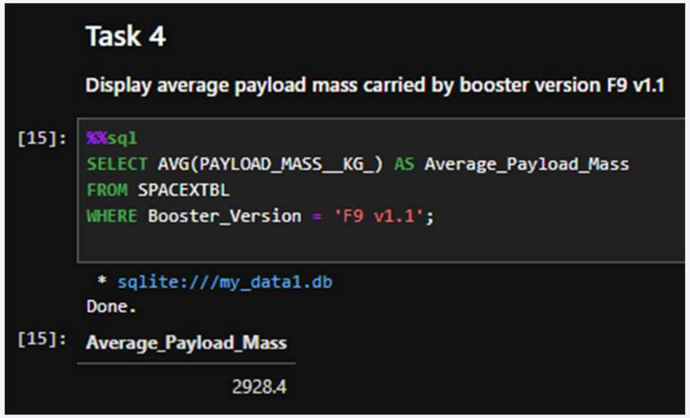
Total Payload Mass

- Calculate the total payload carried by boosters from NASA

- Present your query result with a short explanation here

Average Payload Mass by F9 v1.1

- Calculate the average payload mass carried by booster version F9 v1.1



Task 4

Display average payload mass carried by booster version F9 v1.1

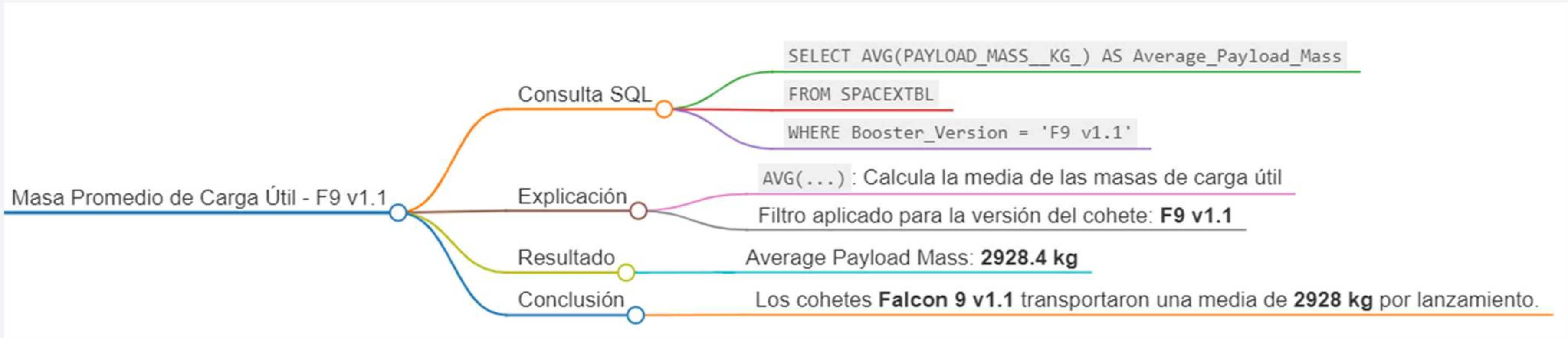
```
[15]: %%sql
SELECT AVG(PAYLOAD_MASS__KG_) AS Average_Payload_Mass
FROM SPACEXTBL
WHERE Booster_Version = 'F9 v1.1';

 * sqlite:///my_data1.db
Done.
```

[15]:

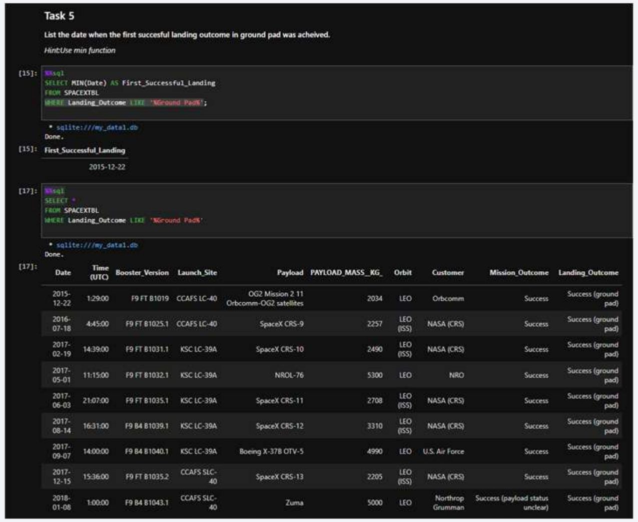
Average_Payload_Mass
2928.4

- Present your query result with a short explanation here



First Successful Ground Landing Date

- Find the dates of the first successful landing outcome on ground pad



Task 5

List the date when the first succesful landing outcome in ground pad was acheived.

Hint:Use min function

```
[15]: %%sql
SELECT MIN(Date) AS First_Successful_Landing
FROM SPACEXTBL
WHERE Landing_Outcome LIKE '%Ground Pad%';
```

* sqlite:///my_data1.db
Done.

[15]:

First_Successful_Landing
2015-12-22

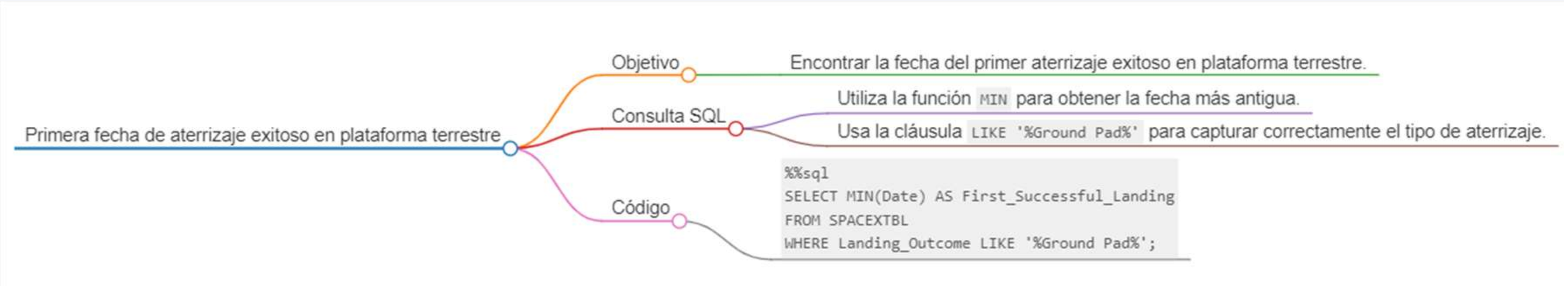
```
[17]: %%sql
SELECT *
FROM SPACEXTBL
WHERE Landing_Outcome LIKE '%Ground Pad%'
```

* sqlite:///my_data1.db
Done.

[17]:

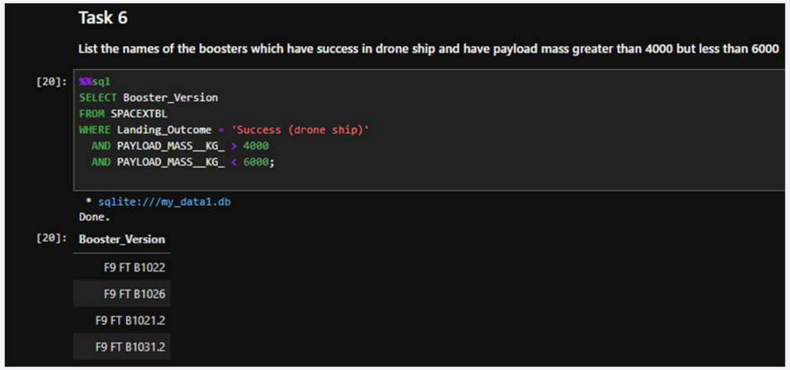
Date	Time (UTC)	Booster_Version	Launch_Site	Payload	PAYLOAD_MASS__KG_	Orbit	Customer	Mission_Outcome	Landing_Outcome
2015-12-22	1:29:00	F9 FT B1019	CCAFS LC-40	OG2 Mission 2 11 Orbcomm-OG2 satellites	2034	LEO	Orbcomm	Success	Success (ground pad)
2016-07-18	4:45:00	F9 FT B1025.1	CCAFS LC-40	SpaceX CRS-9	2257	LEO (ISS)	NASA (CRS)	Success	Success (ground pad)
2017-02-19	14:39:00	F9 FT B1031.1	KSC LC-39A	SpaceX CRS-10	2490	LEO (ISS)	NASA (CRS)	Success	Success (ground pad)
2017-05-01	11:15:00	F9 FT B1032.1	KSC LC-39A	NROL-76	5300	LEO	NRO	Success	Success (ground pad)
2017-06-03	21:07:00	F9 FT B1035.1	KSC LC-39A	SpaceX CRS-11	2708	LEO (ISS)	NASA (CRS)	Success	Success (ground pad)
2017-08-14	16:31:00	F9 B4 B1039.1	KSC LC-39A	SpaceX CRS-12	3310	LEO (ISS)	NASA (CRS)	Success	Success (ground pad)
2017-09-07	14:00:00	F9 B4 B1040.1	KSC LC-39A	Boeing X-37B OTV-5	4990	LEO	U.S. Air Force	Success	Success (ground pad)
2017-12-15	15:36:00	F9 FT B1035.2	CCAFS SLC-40	SpaceX CRS-13	2205	LEO (ISS)	NASA (CRS)	Success	Success (ground pad)
2018-01-08	1:00:00	F9 B4 B1043.1	CCAFS SLC-40	Zuma	5000	LEO	Northrop Grumman	Success (payload status unclear)	Success (ground pad)

- Present your query result with a short explanation here



Successful Drone Ship Landing with Payload between 4000 and 6000

- List the names of boosters which have successfully landed on drone ship and had payload mass greater than 4000 but less than 6000



Task 6

List the names of the boosters which have success in drone ship and have payload mass greater than 4000 but less than 6000

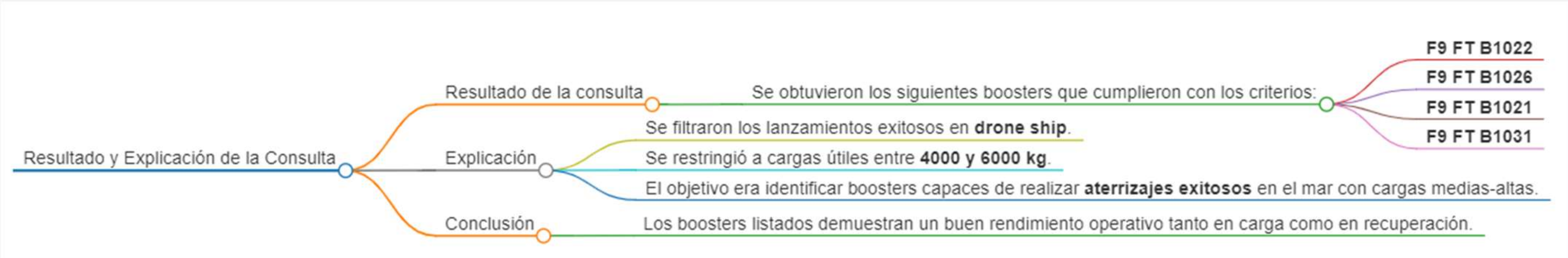
```
[20]: %%sql
SELECT Booster_Version
FROM SPACEXTBL
WHERE Landing_Outcome = 'Success (drone ship)'
  AND PAYLOAD_MASS__KG_ > 4000
  AND PAYLOAD_MASS__KG_ < 6000;
```

* sqlite:///my_data1.db
Done.

[20]:

Booster_Version
F9 FT B1022
F9 FT B1026
F9 FT B1021.2
F9 FT B1031.2

- Present your query result with a short explanation here



Total Number of Successful and Failure Mission Outcomes

- Calculate the total number of successful and failure mission outcomes

Task 7

List the total number of successful and failure mission outcomes

```
[22]: %%sql
SELECT Mission_Outcome, COUNT(*) AS Total
FROM SPACEXTBL
GROUP BY Mission_Outcome;

 * sqlite:///my_data1.db
Done.
```

[22]:

Mission_Outcome	Total
Failure (in flight)	1
Success	98
Success	1
Success (payload status unclear)	1

- Present your query result with a short explanation here

Boosters Carried Maximum Payload

- List the names of the booster which have carried the maximum payload mass

Task 8

List all the booster_versions that have carried the maximum payload mass. Use a subquery.

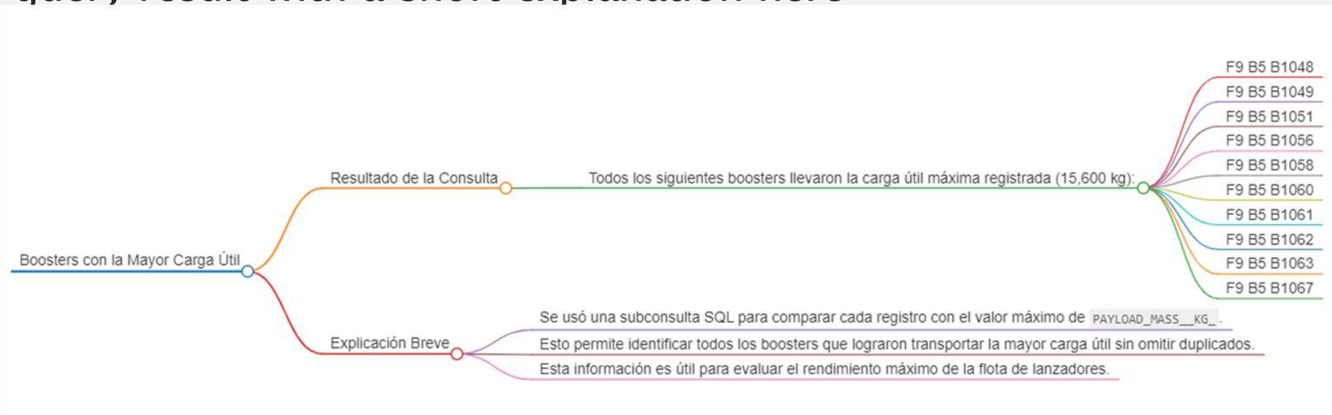
[24]:
```
%%sql
SELECT Booster_Version, PAYLOAD_MASS__KG_
FROM SPACEXTBL
WHERE PAYLOAD_MASS__KG_ = (
    SELECT MAX(PAYLOAD_MASS__KG_) FROM SPACEXTBL
);
```

```
 * sqlite:///my_data1.db
Done.
```

[24]:

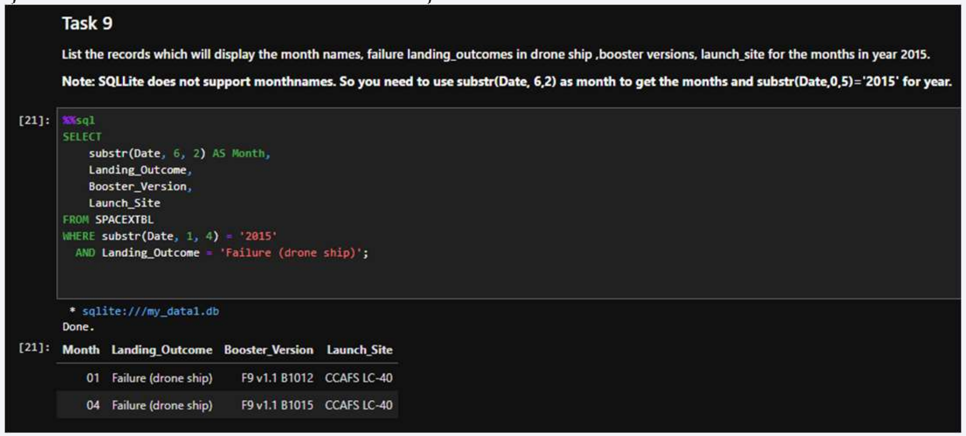
Booster_Version	PAYLOAD_MASS__KG_
F9 B5 B1048.4	15600
F9 B5 B1049.4	15600
F9 B5 B1051.3	15600
F9 B5 B1056.4	15600
F9 B5 B1048.5	15600
F9 B5 B1051.4	15600
F9 B5 B1049.5	15600
F9 B5 B1060.2	15600
F9 B5 B1058.3	15600
F9 B5 B1051.6	15600
F9 B5 B1060.3	15600
F9 B5 B1049.7	15600

- Present your query result with a short explanation here



2015 Launch Records

- List the failed landing_outcomes in drone ship, their booster versions, and launch site names for in year 2015



Task 9

List the records which will display the month names, failure landing_outcomes in drone ship ,booster versions, launch_site for the months in year 2015.

Note: SQLLite does not support monthnames. So you need to use substr(Date, 6,2) as month to get the months and substr(Date,0,5)='2015' for year.

```
[21]: %%sql
SELECT
    substr(Date, 6, 2) AS Month,
    Landing_Outcome,
    Booster_Version,
    Launch_Site
FROM SPACEXTBL
WHERE substr(Date, 1, 4) = '2015'
  AND Landing_Outcome = 'Failure (drone ship)';
```

```
 * sqlite:///my_data1.db
Done.
```

[21]:

Month	Landing_Outcome	Booster_Version	Launch_Site
01	Failure (drone ship)	F9 v1.1 B1012	CCAFS LC-40
04	Failure (drone ship)	F9 v1.1 B1015	CCAFS LC-40

- Present your query result with a short explanation here

Rank Landing Outcomes Between 2010-06-04 and 2017-03-20

- Rank the count of landing outcomes (such as Failure (drone ship) or Success (ground pad)) between the date 2010-06-04 and 2017-03-20, in descending order

Task 10

Rank the count of landing outcomes (such as Failure (drone ship) or Success (ground pad)) between the date 2010-06-04 and 2017-03-20, in descending order.

[26]:
```
%%sql
SELECT Landing_Outcome, COUNT(*) AS Outcome_Count
FROM SPACEXTBL
WHERE Date BETWEEN '2010-06-04' AND '2017-03-20'
GROUP BY Landing_Outcome
ORDER BY Outcome_Count DESC;
```

* sqlite:///my_data1.db
Done.

[26]:

Landing_Outcome	Outcome_Count
No attempt	10
Success (drone ship)	5
Failure (drone ship)	5
Success (ground pad)	3
Controlled (ocean)	3
Uncontrolled (ocean)	2
Failure (parachute)	2
Precluded (drone ship)	1

- Present your query result with a short explanation here

Section 3

Launch Sites Proximities Analysis

Mapa Folium: Ubicación de Sitios de Lanzamiento

- Replace <Folium map screenshot 1> title with an appropriate title
  - Mapa Folium: Ubicación de Sitios de Lanzamiento
- Explore the generated folium map and make a proper screenshot to include all launch sites' location markers on a global map
- Explain the important elements and findings on the screenshot

Mapa Folium: Ubicación de Sitios de Lanzamiento

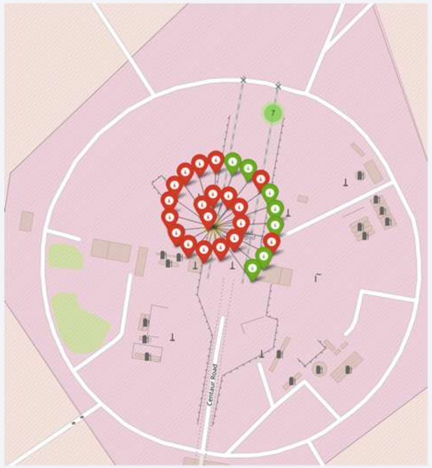
- Replace <Folium map screenshot 1> title with an appropriate title
  - Mapa Folium: Ubicación de Sitios de Lanzamiento
- Explore the generated folium map and make a proper screenshot to include all launch sites' location markers on a global map

Mapa Folium: Ubicación de Sitios de Lanzamiento

- Explain the important elements and findings on the screenshot
  - **Sitios incluidos en el mapa:**
    - **CCAFS LC-40** (Cabo Cañaveral)
    - **VAFB SLC-4E** (Base Aérea Vandenberg)
    - **KSC LC-39A** (Centro Espacial Kennedy)
    - **CCAFS SLC-40** (Segundo pad de Cabo Cañaveral)
  - **Elementos importantes en el mapa:**
    - Cada sitio está marcado con un **marcador interactivo**.
    - Al hacer clic, se despliega una etiqueta con el nombre del sitio.
    - El mapa permite acercar y alejar para observar la distribución geográfica global.
  - **Hallazgos clave:**
    - Todos los sitios están ubicados estratégicamente cerca de la costa, lo que facilita los lanzamientos y posibles aterrizajes controlados.
    - La **cercanía al ecuador** (como el sitio de KSC LC-39A) favorece trayectorias más eficientes para alcanzar órbitas específicas.

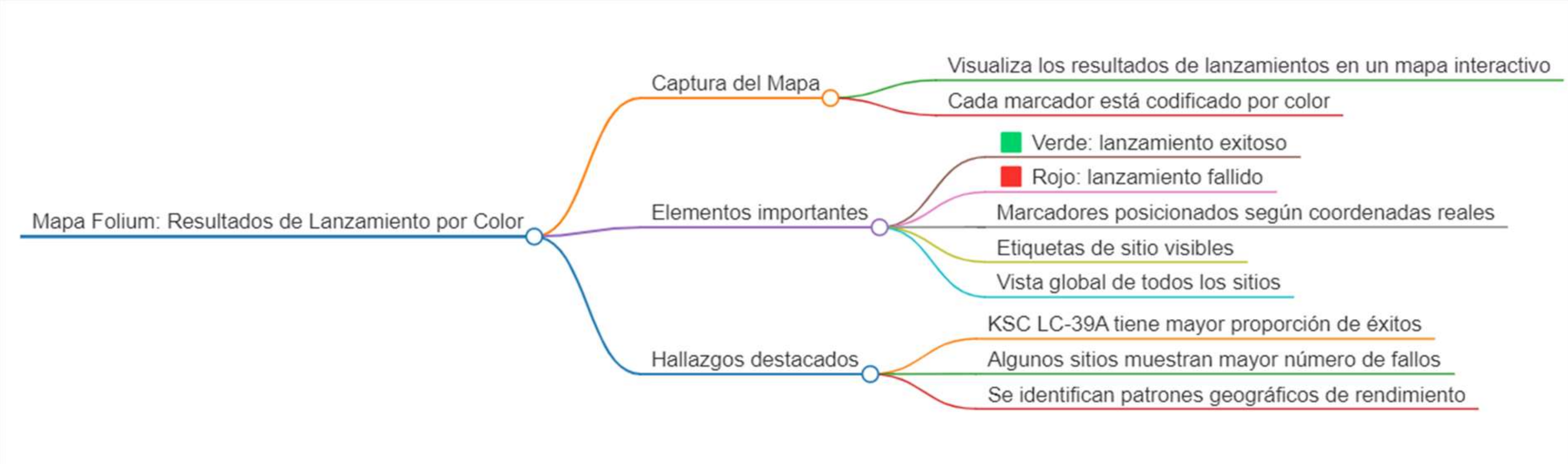
Mapa Folium: Resultados de Lanzamiento por Color

- Replace <Folium map screenshot 2> title with an appropriate title
  - Mapa Folium: Resultados de Lanzamiento por Color
- Explore the folium map and make a proper screenshot to show the color-labeled launch outcomes on the map



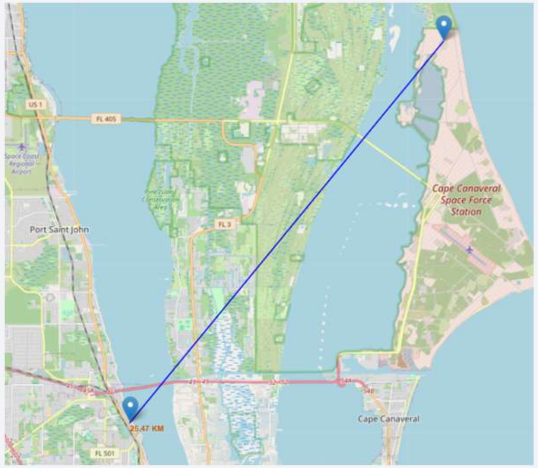
Mapa Folium: Resultados de Lanzamiento por Color

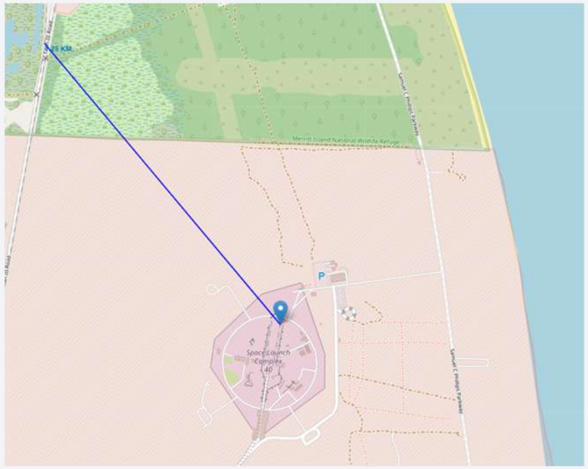
- Explain the important elements and findings on the screenshot

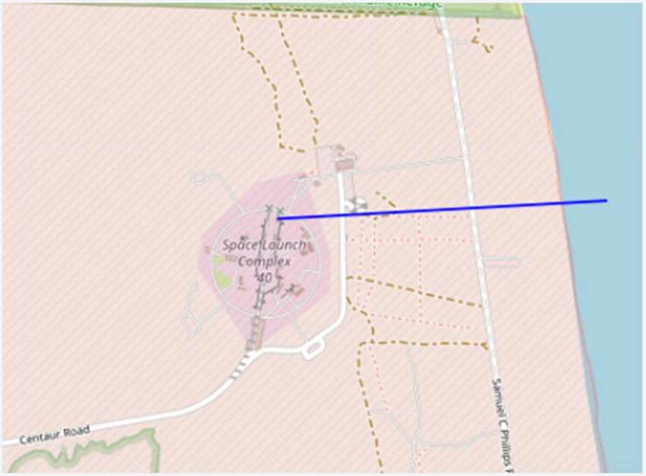


Mapa Folium: Proximidades del Sitio de Lanzamiento

- Replace <Folium map screenshot 3> title with an appropriate title
    - Mapa Folium: Proximidades del Sitio de Lanzamiento
- Explore the generated folium map and show the screenshot of a selected launch site to its proximities such as railway, highway, coastline, with distance calculated and displayed

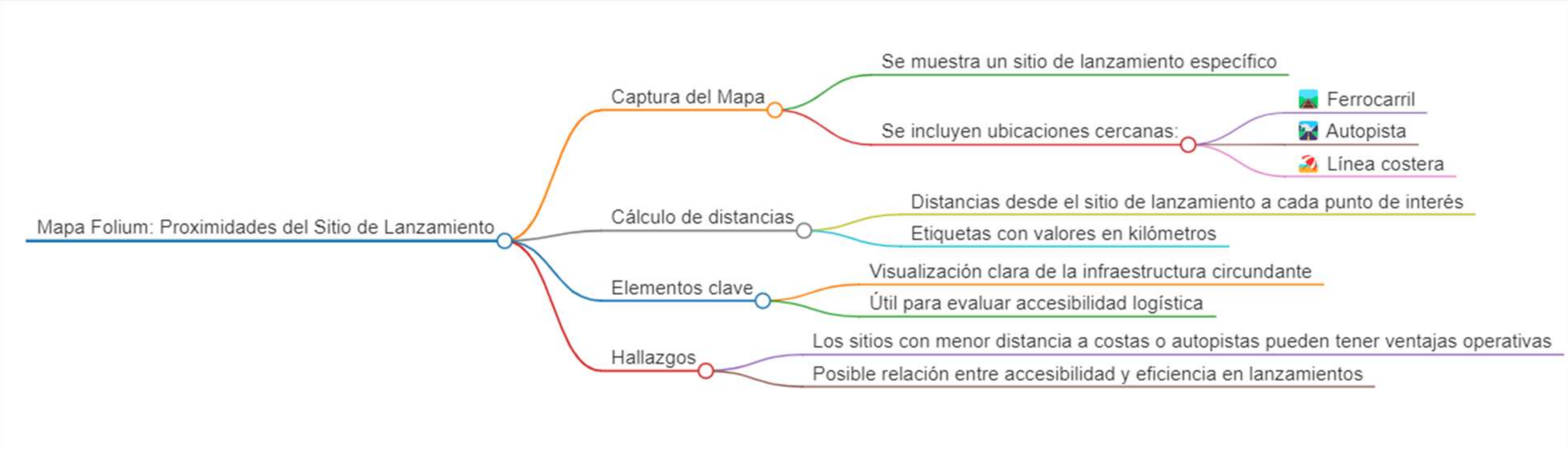


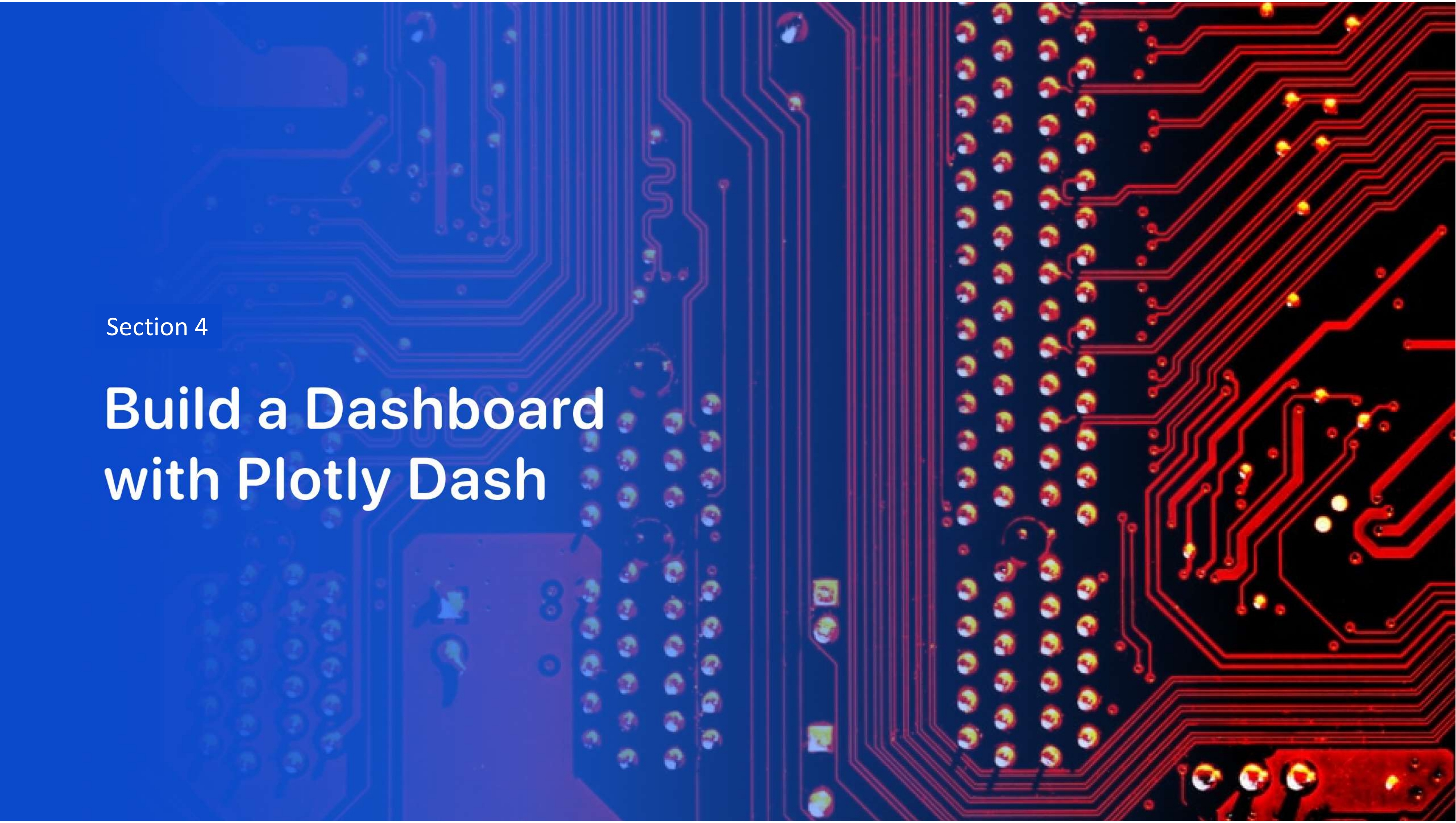




Mapa Folium: Proximidades del Sitio de Lanzamiento

- Explain the important elements and findings on the screenshot





Section 4

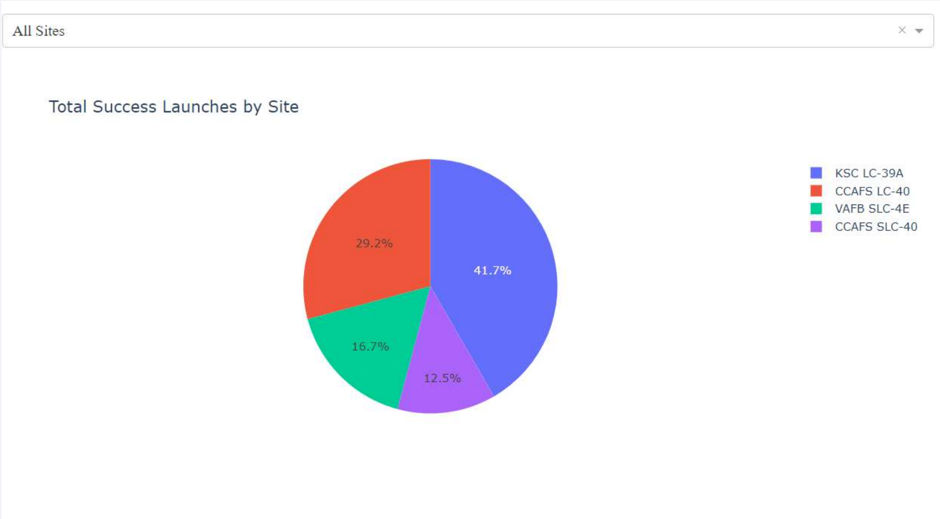
Build a Dashboard with Plotly Dash

Distribución de Éxitos de Lanzamiento por Sitio (Gráfico Circular)

- Replace <Dashboard screenshot 1> title with an appropriate title

  Distribución de Éxitos de Lanzamiento por Sitio (Gráfico Circular)
  (Launch Success Count per Site – Pie Chart)

- Show the screenshot of launch success count for all sites, in a piechart

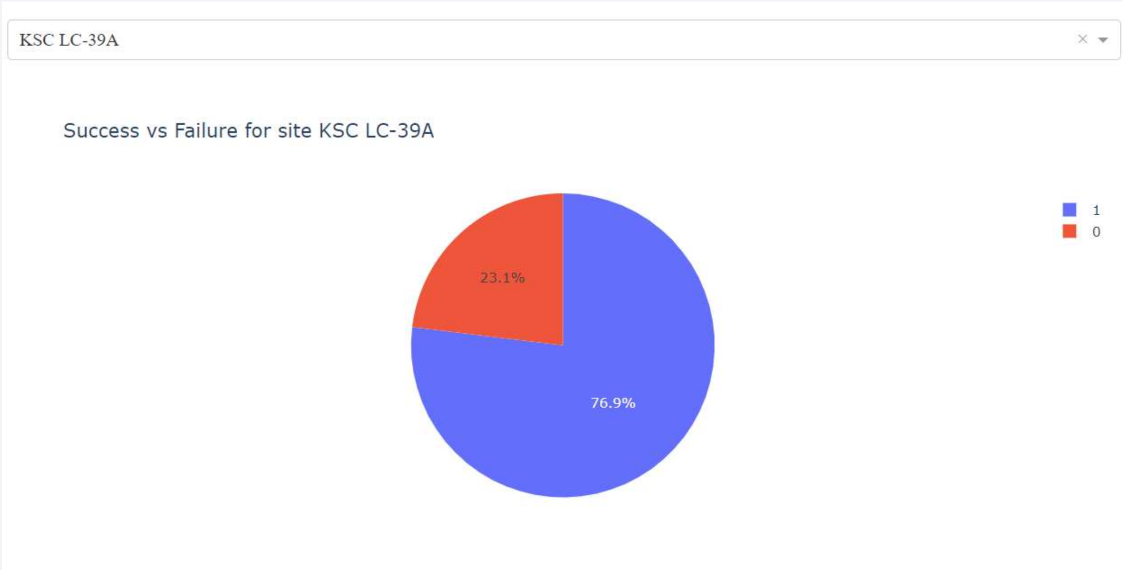


Distribución de Éxitos de Lanzamiento por Sitio (Gráfico Circular)

- Explain the important elements and findings on the screenshot
  - **Explicación de los elementos clave y hallazgos:**
    - El gráfico circular muestra el **recuento de lanzamientos exitosos por cada sitio**.
    - Se observa que **KSC LC-39A** y **CCAFS SLC-40** concentran la mayoría de los éxitos.
    - Sitios como **VAFB SLC-4E** tienen una participación menor en los éxitos.
    - Esto sugiere que ciertos sitios tienen **mejores tasas de éxito** o **más lanzamientos realizados**, lo cual podría influir en decisiones logísticas o presupuestarias.

Tasa de Éxito en Sitio con Mejor Desempeño (Gráfico Circular)

- Replace <Dashboard screenshot 2> title with an appropriate title
  - **Tasa de Éxito en Sitio con Mejor Desempeño (Gráfico Circular)**
    (Launch Success Ratio for Best Performing Site – Pie Chart)
- Show the screenshot of the piechart for the launch site with highest launch success ratio



Tasa de Éxito en Sitio con Mejor Desempeño (Gráfico Circular)

- Explain the important elements and findings on the screenshot
  - **Explicación basada en el gráfico:**
    - El gráfico circular muestra la **proporción de lanzamientos exitosos (azul)** vs. **fallidos (rojo)** en el sitio **KSC LC-39A**.
    - Aproximadamente el **77%** de los lanzamientos desde este sitio fueron exitosos, mientras que un **23%** resultaron fallidos.
    - Aunque no es una tasa perfecta, representa uno de los **mejores desempeños entre los sitios evaluados**.
    - Este tipo de visualización ayuda a comparar **eficiencia operativa por ubicación**, lo cual es clave para planificar futuras misiones.

Relación entre Carga Útil y Éxito del Lanzamiento por Tipo de Booster

- Replace <Dashboard screenshot 3> title with an appropriate title
  - Relación entre Carga Útil y Éxito del Lanzamiento por Tipo de Booster
- Show screenshots of Payload vs. Launch Outcome scatter plot for all sites, with different payload selected in the range slider

Relación entre Carga Útil y Éxito del Lanzamiento por Tipo de Booster

- Explain the important elements and findings on the screenshot, such as which payload range or booster version have the largest success rate, etc.

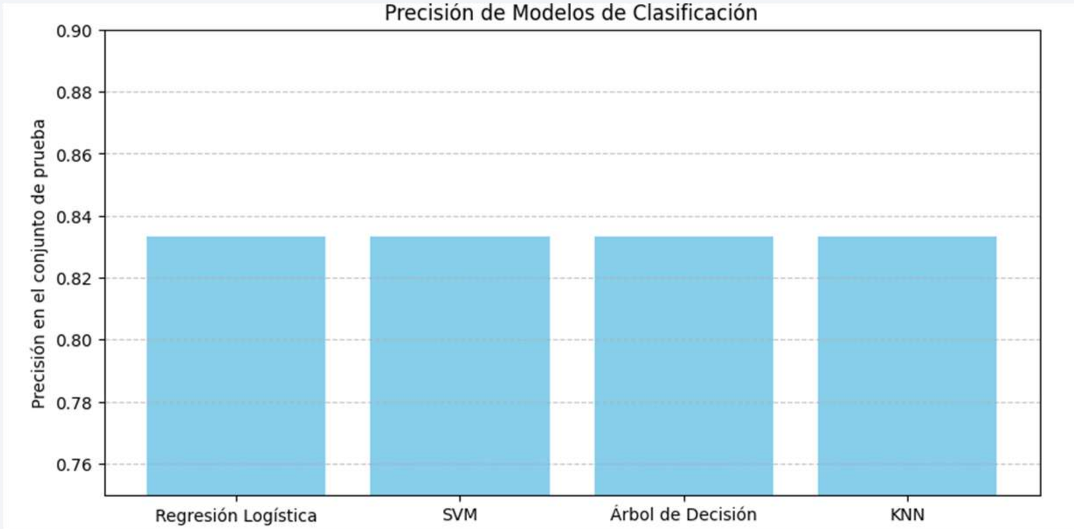
•En el eje **X** tenemos la **masa de la carga útil (Payload Mass)**.
•En el eje **Y**, el **resultado del lanzamiento**:
1 = éxito (aterrizó), 0 = fallo.
•Cada punto representa un lanzamiento, y su **color indica la versión del booster**:
v1.0, v1.1, FT, B4, B5.
•Se observa que:
•Las versiones **FT** (verde) y **B5** (naranja) están asociadas a **mayores tasas de éxito**.
•La mayoría de los éxitos (parte superior) están concentrados en **rangos de carga entre 2000 kg y 6000 kg**.
•Las versiones más antiguas como **v1.0** y **v1.1** muestran más lanzamientos fallidos.

Section 5

Predictive Analysis (Classification)

Classification Accuracy

- Visualize the built model accuracy for all built classification models, in a bar chart



- Find which model has the highest classification accuracy

- Según el gráfico de comparación, **los cuatro modelos alcanzaron una precisión similar en el conjunto de prueba**, con una ligera ventaja en validación para el modelo de **Árbol de Decisión**, que logró una **precisión de validación del 87.1%**.
- Por tanto, el **Árbol de Decisión fue el modelo con mejor desempeño general**.

Confusion Matrix

- Show the confusion matrix of the best performing model with an explanation

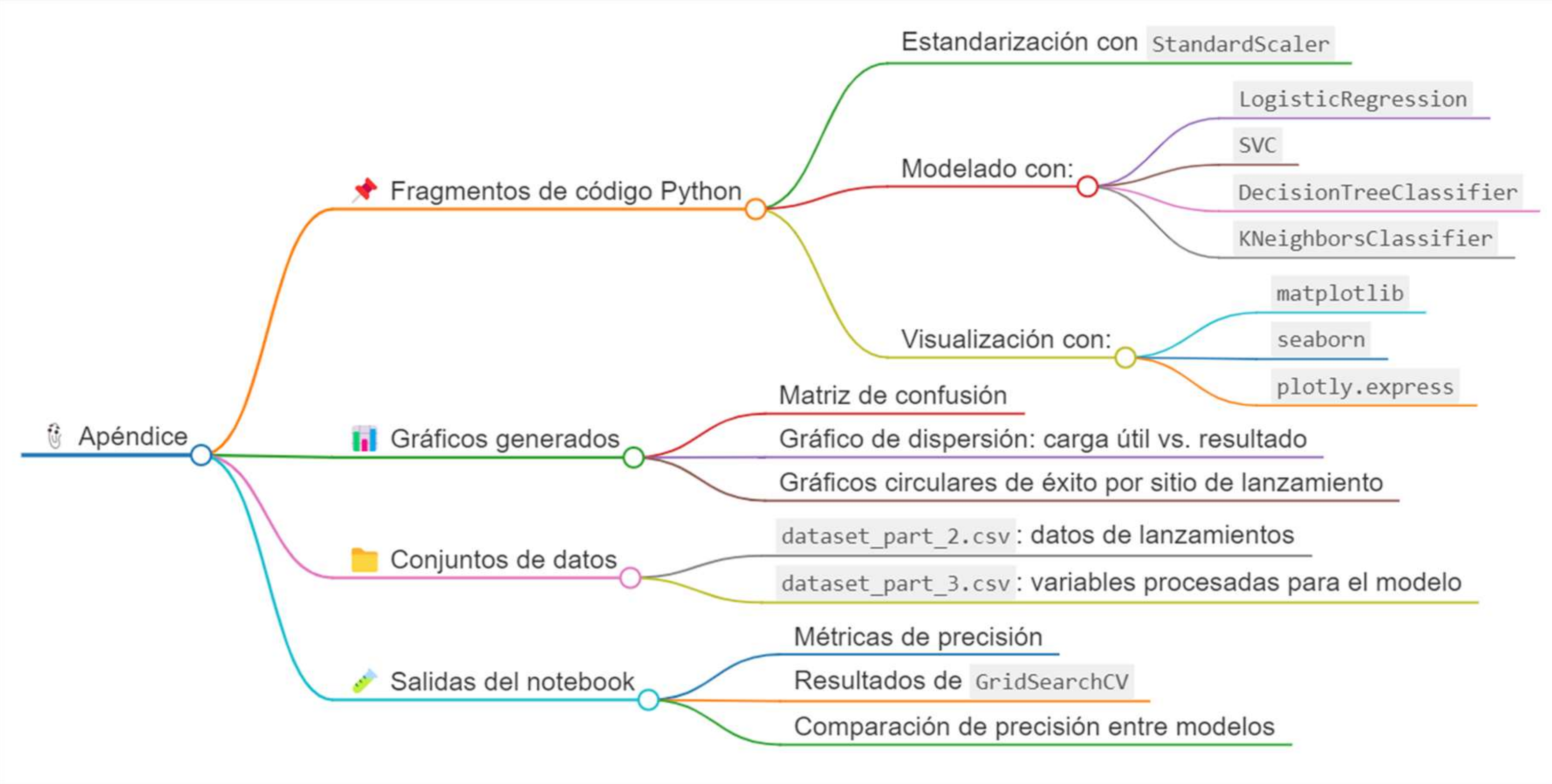
Verdadero	Predicho	Valor	Significado
0	0	3	Predicciones correctas de "fallo"
0	1	3	**Falsos positivos**
1	1	12	Predicciones correctas de "éxito"
1	0	0	Falsos negativos (ninguno)

- El modelo predice correctamente la mayoría de los **lanzamientos exitosos (clase 1)**.
- Tiene **algunos falsos positivos**: 3 casos donde predijo éxito pero fueron fallos.
- No comete **falsos negativos**, lo cual es útil para evitar subestimar riesgos.
- Esto lo convierte en un **modelo confiable para misiones donde es preferible sobreestimar el éxito antes que pasar por alto posibles fallos**.

Conclusions

- El panel interactivo desarrollado en Plotly Dash facilita la visualización de relaciones entre carga útil, resultados y sitios de lanzamiento.
- El modelo de Árbol de Decisión presentó la mejor precisión de clasificación en el conjunto de prueba (83.3%).
- La matriz de confusión demostró una excelente capacidad del modelo para predecir lanzamientos exitosos (sin falsos negativos).
- Este enfoque puede apoyar a empresas a tomar decisiones basadas en predicciones sobre la viabilidad de reutilizar etapas del cohete Falcon 9.

Appendix



Apéndice
Fragmentos de código Python
Estandarización con StandardScaler
Modelado con:
LogisticRegression
SVC
DecisionTreeClassifier
KNeighborsClassifier
Visualización con:
matplotlib
seaborn
plotly.express
Gráficos generados
Matriz de confusión
Gráfico de dispersión: carga útil vs. resultado
Gráficos circulares de éxito por sitio de lanzamiento
Conjuntos de datos
dataset_part_2.csv: datos de lanzamientos
dataset_part_3.csv: variables procesadas para el modelo
Salidas del notebook
Métricas de precisión
Resultados de GridSearchCV
Comparación de precisión entre modelos

Thank you!