

A photograph of a SpaceX Falcon 9 rocket launching vertically against a clear blue sky. The rocket is white with black markings, and a large, bright orange and white plume of fire and smoke trails behind it. In the bottom left corner, the top of the launch pad service structure is visible.

Ciencia de Datos

Análisis y predicción de aterrizajes de Falcon 9 de SpaceX

**Humberto Abel Fernández
Rodríguez**

20-07-2024



Esquema

- Resumen ejecutivo
- Introducción
- Metodología
- Resultados
- Discussion
- Conclusiones
- Apendice

Resumen ejecutivo

Resumen de metodologías:

- Obtención de datos a través de la API de SpaceX y de Wikipedia con técnicas de Webscrapping.
- Gestión de los datos (wrangling): obtención de datos de lanzamiento con éxitos y fracasos (1,0).
- Exploración y análisis de los datos con SQL y técnicas de visualización.
- Visualización interactiva de los lugares de lanzamiento con Folium y Dashboards.
- Predicción con machine learning usando Regresión Logística, SVM, Árbol de decisiones y KNN.

Resumen de resultados:

- Resultados del análisis exploratorio de datos.
- Manipulación de los datos para obtener un conjunto de lanzamiento con éxitos y fracasos.
- Análisis visual e interactivo de resultados.
- Resultados del análisis predictivo con machine learning.

Introducción

La era espacial comercial ha comenzado. Empresas como Virgin Galactic (vuelos espaciales suborbitales), Rocket Lab (pequeños proveedores de satélites), Blue Origin (cohetes reutilizables suborbitales y orbitales) y SpaceX compiten por un mercado en auge

- Concretamente, SpaceX ofrece:
 - Envíos a la Estación Espacial Internacional.
 - Starlink: Internet satelital.
 - Misiones tripuladas.
 - Costos bajos: \$62 millones por lanzamiento del Falcon 9 frente al resto del mercado que sube a casi el triple. Esto lo consigue reutilizando la primera etapa.

Objetivo del Proyecto:

- Determinar si la primera etapa del Falcon 9 de SpaceX aterrizará, con lo cual se podría reutilizar y abaratar el siguiente lanzamiento de un modo notable ya que es la parte más grande y costosa.

Dificultades:

- El aterrizaje de la primera etapa no siempre es exitoso y, además, SpaceX puede sacrificar la primera etapa dependiendo de la misión.

Cometido del Científico de Datos:

- Determinar el precio de cada lanzamiento.
- Recopilar información sobre SpaceX y crear paneles de datos para identificar los factores que influyen en el resultado del aterrizaje.
- Predecir si SpaceX reutilizará la primera etapa usando modelos de aprendizaje automático y datos públicos.



Metodología

Resumen de la metodología

Obtención de datos	A través de la API de SpaceX obtenemos un conjunto de datos. Con técnicas de Webscrapping obtenemos otro conjunto de datos. Ambos los pasamos a dataframes de Pandas.
Tratamiento de datos (wrangling)	Limpieza y preparación de los datos
Análisis exploratorio de datos (EDA)	Técnicas de visualización y SQL
Visualización interactiva de los lugares de lanzamiento	Folium
Predicción con machine learning	Modelos de clasificación supervisados: Regresión Logística SVM Árbol de decisiones KNN

De la API de SpaceX

```
response = requests.get(spacex_url)
```

Información masiva sin formato

IBM Developer

Pandas Dataframe

SKILLS NETWORK 

Obtención de datos

Con Web Scraping desde Wikipedia

[illegible]

Tabla de lanzamientos en Wikipedia

[illegible]

```
df= pd.DataFrame({ key:pd.Series(value) for key, value in launch_dict.items() })
```

HTTP GET para obtener el HTML de la página con los lanzamientos del Falcon 9 para recibir una respuesta HTTP.

Creamos un objeto BeautifulSoup desde la HTML "response".

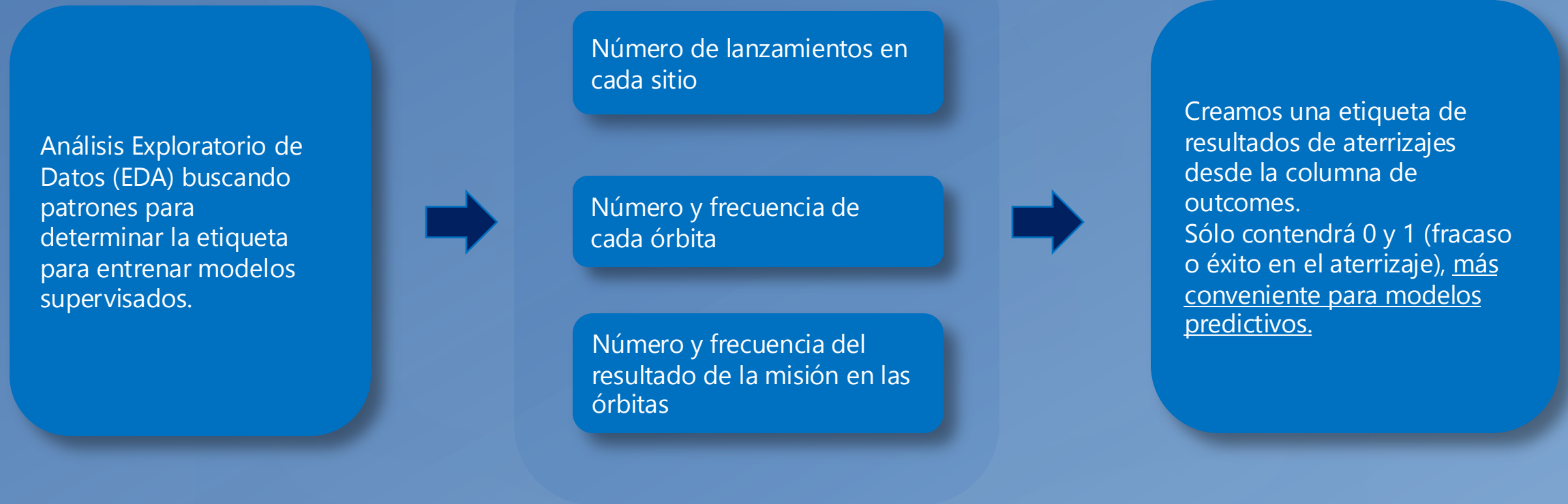
Extraemos los nombres de las columnas relevantes desde las cabeceras de las tablas HTML.

Creamos un diccionario con los nombres de las columnas y rellenamos los datos mediante un bucle de python. Y creamos un dataframe con Pandas a partir del diccionario.

Iteramos a través de la los elementos "<th>" y para obtener los nombres de las columnas uno por uno aplicamos la función:

```
extract column from header()
```


Manejo de los datos (Wrangling)



<https://github.com/Perlujum/Space-X/blob/main/labs-jupyter-spacex-Data%20wrangling.ipynb>

EDA con SQL

Después de convertir nuestro archivo csv a una tabla SQL llamada SPACEXTBL se hicieron las siguientes consultas para mostrar lo siguiente:

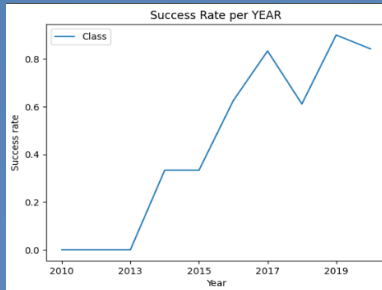


1. Los distintos nombres de los lugares de lanzamiento.
2. Los 5 primeros registros donde el nombre del lugar de lanzamiento empieza por 'CCA'.
3. La masa total de la carga útil transportada por los cohetes lanzados por la NASA (CRS).
4. La masa promedio de la carga útil transportada por la versión F9 v1.1 del cohete.
5. La fecha en que se logró el primer aterrizaje exitoso en una plataforma terrestre.
6. Los nombres de los cohetes que han aterrizado con éxito en una nave dron y tienen una masa de carga útil entre 4000 y 6000.
7. El número total de resultados de misiones exitosas y fallidas.
8. Los nombres de las versiones de los cohetes que han transportado la masa máxima de carga útil.
9. Registros que muestren los nombres de los meses, los resultados de aterrizajes fallidos en naves dron, las versiones de cohetes y el sitio de lanzamiento para los meses del año 2015.
10. Recuento de resultados de aterrizaje (como Falla (nave dron) o Éxito (plataforma terrestre)) entre las fechas 2010-06-04 y 2017-03-20, en orden descendente.

https://github.com/Perlujum/Space-X/blob/main/jupyter-labs-eda-sql-coursera_sqlite.ipynb

EDA con técnicas de visualización

Usando matplotlib y seaborn haremos las siguientes visualizaciones:



1. Relación entre Flight Number y Payload.
2. Relación entre Flight Number y Launch Site.
3. Relación entre Payload y Launch Site.
4. Porcentaje de éxito de cada órbita.
5. Relación entre FlightNumber y Orbit type.
6. Relación entre Payload y Orbit type
7. Tendencia anual de lanzamientos con éxito.

<https://github.com/Perlujum/Space-X/blob/main/edadataviz.ipynb>

Análisis visual interactivo con Folium



Usando matplotlib y seaborn haremos las siguientes visualizaciones:



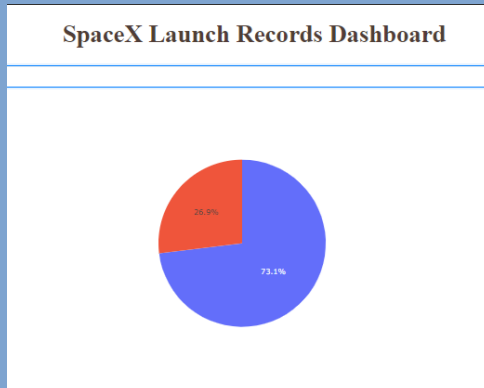
1. Señalar en el mapa los sitios de lanzamiento.
2. Señalar en el mapa los lanzamientos exitosos y fallidos para cada sitio.
3. Calcular las distancias del lugar de lanzamiento a sus proximidades.

https://github.com/Perlujum/Space-X/blob/main/lab_jupyter_launch_site_location.ipynb

Análisis visual interactivo con Dash



Usando Dash de un modo interactivo responderemos a las siguientes preguntas:



1. ¿Qué sitio tiene más lanzamientos exitosos.
2. ¿Qué sitio tiene la proporción más alta de lanzamientos exitosos?
3. ¿Qué rango de carga útil tiene la mayor proporción de lanzamientos exitosos?
4. ¿Qué rango de carga útil tiene la proporción más baja de lanzamientos exitosos?
5. Which F9 Booster version (v1.0, v1.1, FT, B4, B5, etc.) has the highest launch success rate?
6. ¿Qué versión del cohete F9 (v1.0, v1.1, FT, B4, B5, etc.) tiene la mayor tasa de éxito en lanzamientos?

https://github.com/Perlujum/Space-X/blob/main/spacex_dash_app.py

Análisis Predictivo

Creamos un pipeline de aprendizaje automático para predecir si la primera etapa aterrizará, dado los datos de los laboratorios anteriores:

1. Importamos las librerías y funciones auxiliares
2. Cargamos el dataframe
3. Creamos un numpy array para la columna class (éxito=1, fracaso=0) como variable **Y**.
4. Estandarizamos los datos para **X** usando la función transform().
5. Separamos los datos en datos de entrenamiento y datos de prueba con la función train_test_split.
6. Usamos varios modelos de machine learning y ajustamos los hiperparámetros usando GridsearchCV.
7. Calculamos la precisión para todos los modelos con score y así poder encontrar cual es **el más preciso**.

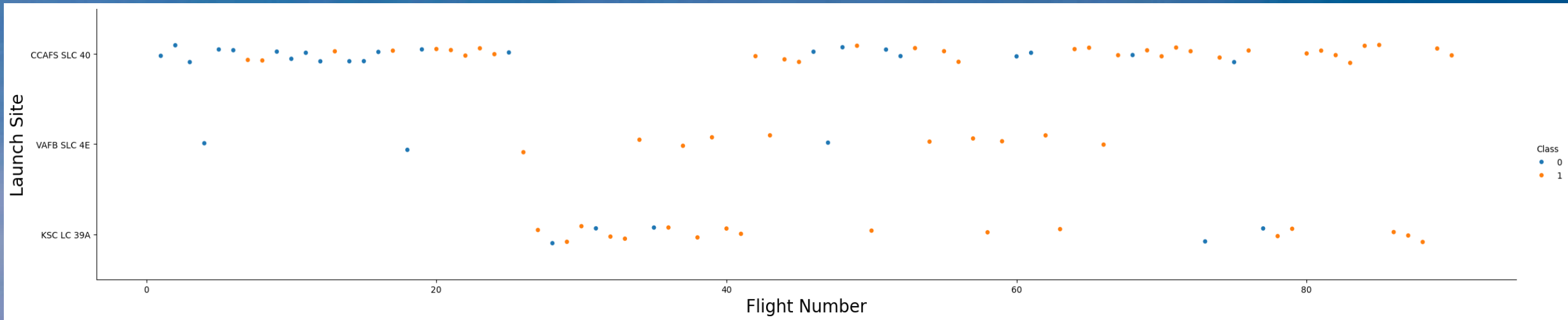
https://github.com/Perlujum/Space-X/blob/main/SpaceX_Machine%20Learning%20Prediction_Part_5.ipynb



Resultados

Resultados visualizaciones

Flight Number y LaunchSite

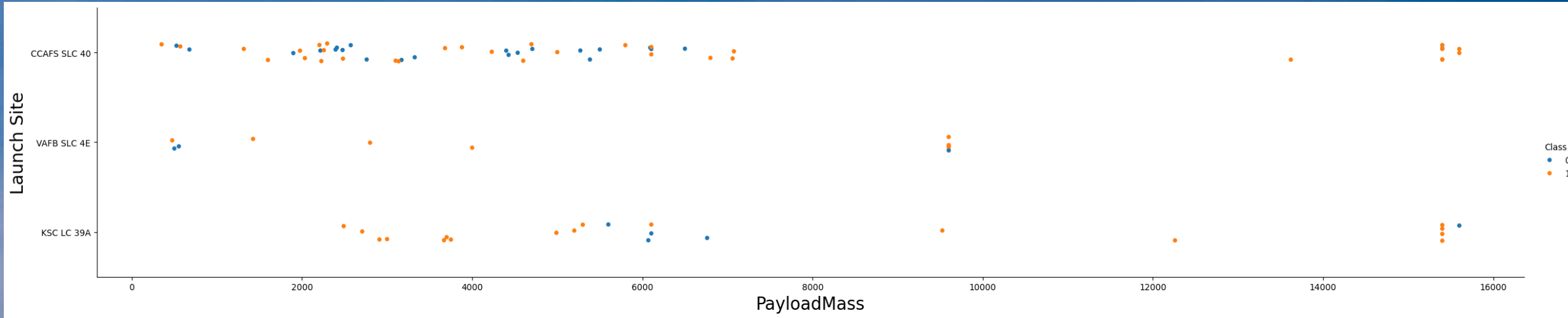


Se puede observar que:

- A medida que aumenta el número de lanzamientos, la proporción de aterrizajes con éxito aumenta para todos los lugares de lanzamiento.
- De los tres sitios, KSC es el que más tarde se ha incorporado la actividad.
- De los tres lugares, CCAFS es el que más cantidad de lanzamientos tiene.

Resultado visualizaciones

PayloadMass y LaunchSite

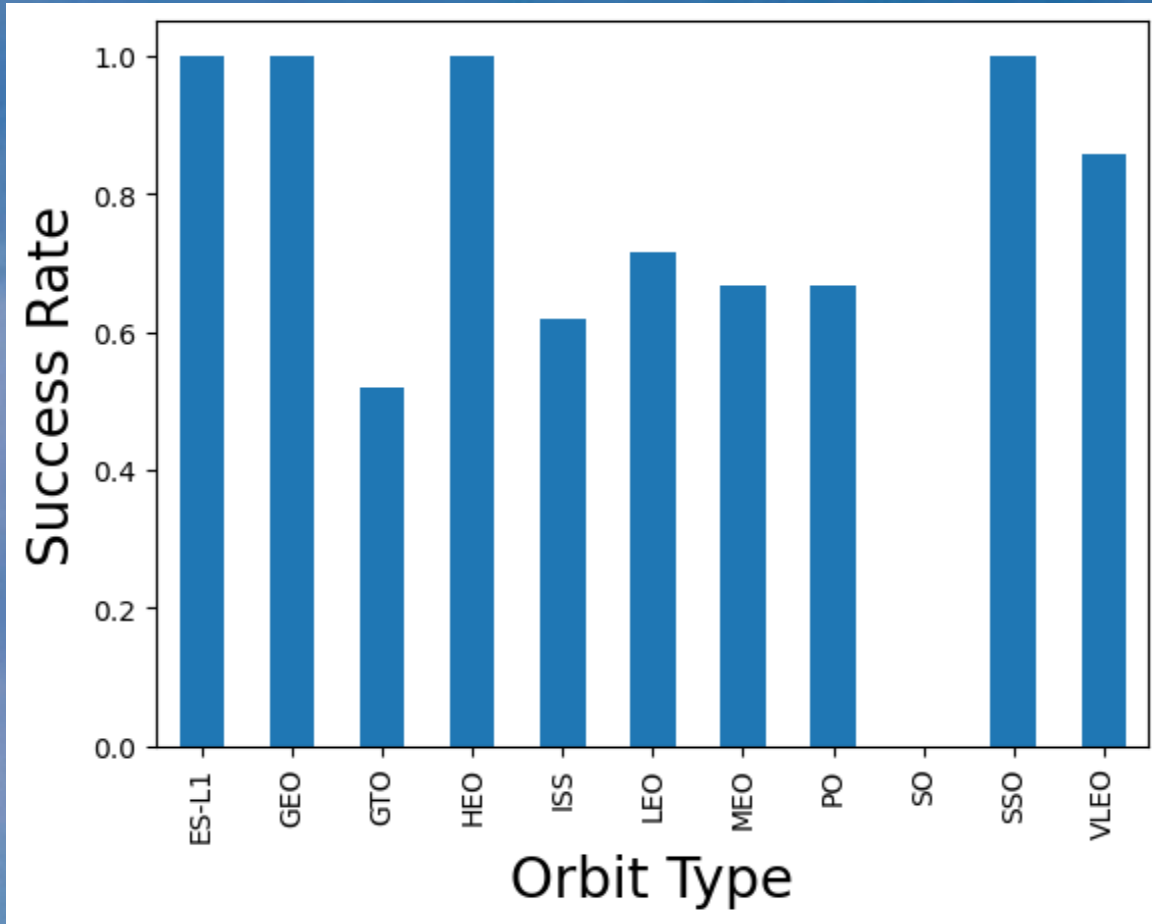


Se puede observar que:

- CCAFS es quien tiene el mayor nº de lanzamientos y casi todos tienen una carga útil entre 300 y 7000. Tiene la menor proporción de éxito.
- VAFB no ha lanzado ningún cohete cuya carga útil fuera mayor de 10.000.
- KSC tiene una tasa de éxito del 100% con cargas entre 0 y 5.500.
- La tasa de éxito parece aumentar entre 9.500 y 16.000.

Resultado visualizaciones

Tipo Órbita y Éxito

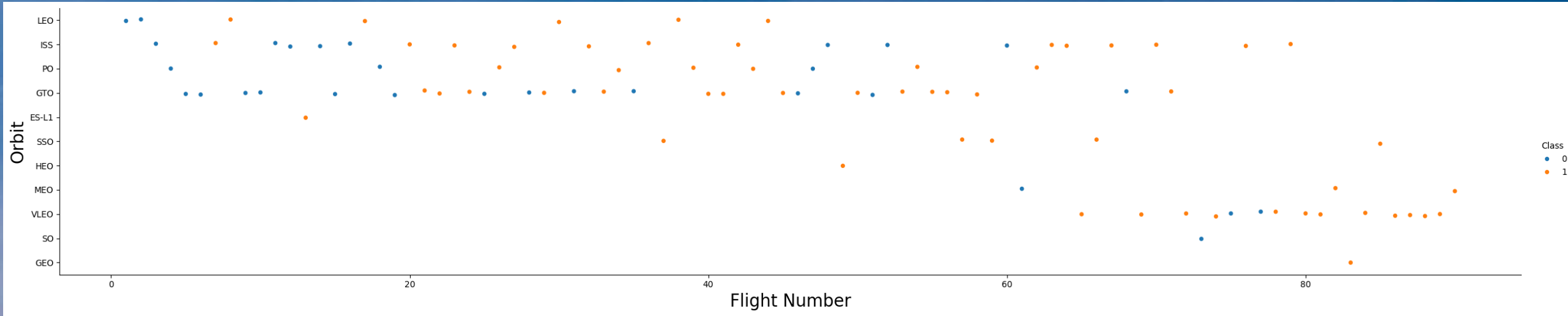


Se puede observar que:

- Las órbitas ES-L1, GEO, HEO y SSO tienen una tasa de éxito del 100%.
- Le siguen GTO, ISS, LEO, MEO y PO con una tasa de éxito entre 70 y 50%
- SO tiene 0% de éxito.

Resultado visualizaciones

FlightNumber y Orbit type

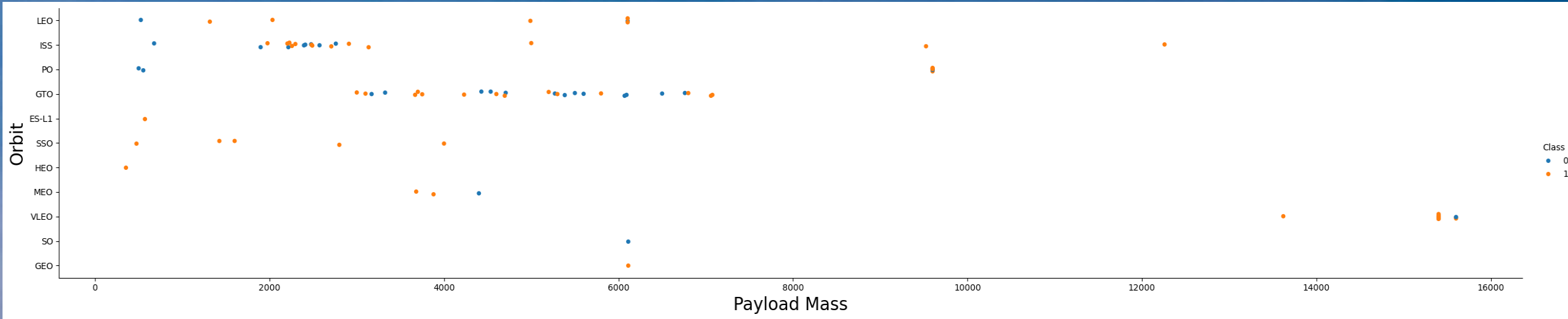


Se puede observar que:

- Las últimas órbitas abordadas han sido SO y GEO, que tienen un solo lanzamiento cada una con resultados dispares.
- A partir del lanzamiento nº 77 hay una tasa de éxito del 100% en todas las órbitas.
- En casi todas las órbitas, a medida que aumentan sus lanzamientos, la tasa de éxito tiende al 100% salvo GTO.

Resultado visualizaciones

Payload y Orbit

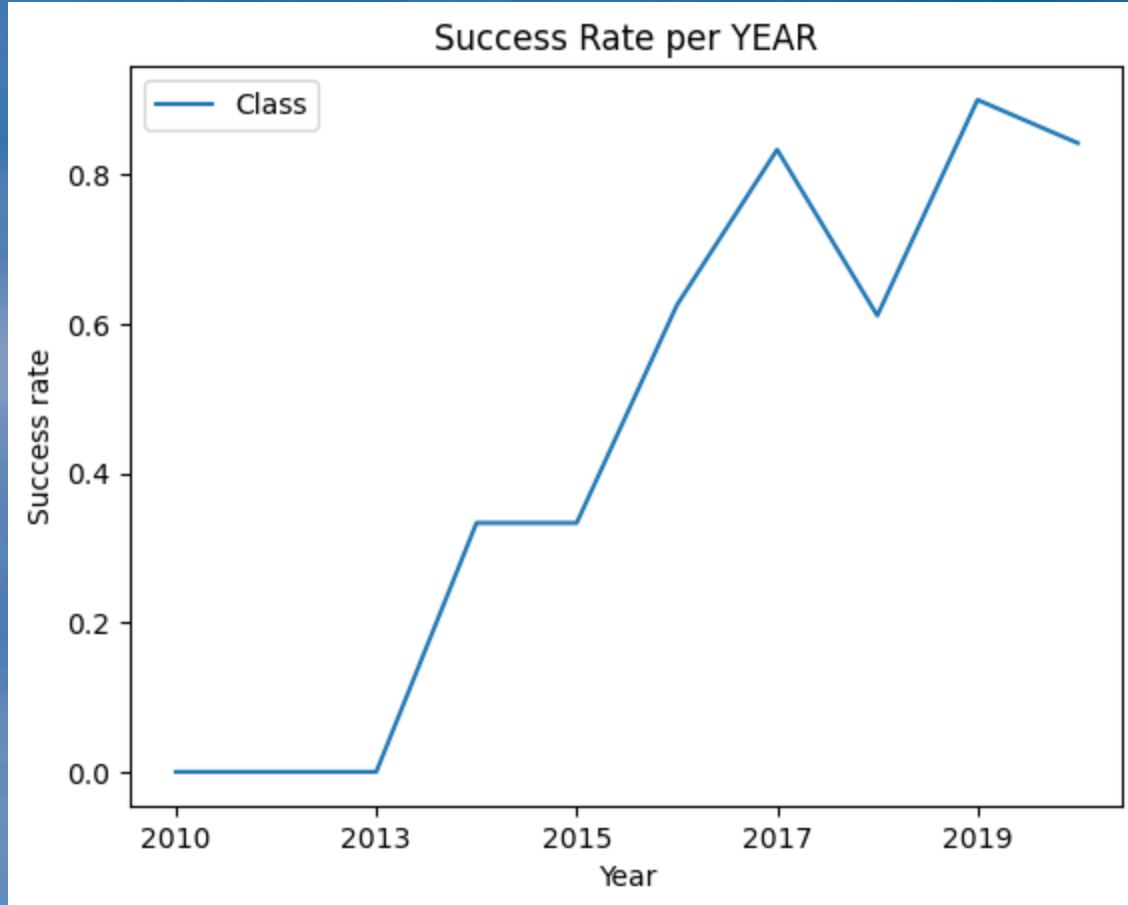


Se puede observar que:

- VLEO es la órbita donde se han lanzado las mayores cargas y casi todas exitosas.
- En los lanzamientos de LEO, ISS y PO se aprecia que el éxito está relacionado con las cargas mayores
- La órbita SSO tiene un 100% de éxito en los lanzamientos realizados, estando estos entre 500 y 4.000 aprox.
- En GTO se alternan éxitos y fracasos en todos sus lanzamientos sin que parezca influir la carga útil.

Resultado visualizaciones

Tendencia anual: éxito en los lanzamientos.



- Hay un aumento claro de la tasa de éxitos con el paso del tiempo.
- El aumento empieza en el año 2013 hasta 2017, donde llega a alcanzar más del 80%.
- Hay una bajada de la tasa de éxito hasta el 60% aprox. en el año 2018 que habría que investigar a qué se debe.
- Finalmente, se logra una tasa de éxito máxima en 2019, en torno al 90%.

Resultados consultas SQL

Consulta general

```
%sql select * from SPACEXTBL limit 3
```

```
* sqlite:///my\_data1.db
```

```
Done.
```

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

Resultados consultas SQL

Nombres de los lugares únicos de lanzamiento

```
%sql select distinct Launch_Site FROM SPACEXTBL
[16]
... * sqlite:///my\_data1.db
Done.
...
Launch_Site
CCAFS LC-40
VAFB SLC-4E
KSC LC-39A
CCAFS SLC-40
```


Resultados consultas SQL

Mostrar 5 registros donde el lugar de lanzamiento comienza con la cadena 'CCA'

```
%sql select * from SPACEXTBL WHERE LAUNCH_SITE LIKE 'CCA%' LIMIT 5
```

[18]

... * [sqlite:///my_data1.db](#)

Done.

...

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

Resultados consultas SQL

Mostrar la carga útil total transportada por cohetes lanzados por NASA (CRS)

```
▶ [38] %sql SELECT SUM(PAYLOAD_MASS_KG_) FROM SPACEXTBL WHERE UPPER(CUSTOMER) ='NASA (CRS)' --Tengo dudas... Con LIKE 'NASA (CRS)%' se añade otro registro (como mínimo) que es NASA (CRS) Pacific
... * sqlite:///my\_data1.db
... Done.
... SUM(PAYLOAD_MASS_KG_)
... 45596
```

+ Code + Markdown

Resultados consultas SQL

Mostrar la media de la carga útil transportada por los cohetes versión F9 v1.1

```
▶ [37] %sql select AVG(PAYLOAD_MASS_KG_) from SPACEXTBL WHERE Booster_Version = 'F9 v1.1'

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

...
AVG(PAYLOAD_MASS_KG_)
2928.4
```

Resultados consultas SQL

Mostrar la fecha cuando se consiguió aterrizar exitosamente en una plataforma terrestre.

```
[44] %sql select date from SPACEXTBL WHERE LANDING_OUTCOME LIKE '%ground pad%' ORDER BY DATE limit 1
... * sqlite:///my\_data1.db
Done.
...
      Date
2015-12-22
```

Resultados consultas SQL

Mostrar los nombres de los cohetes que aterrizaron con éxito en drone ship y tenían una carga útil entre 4000 y 6000

```
[54] %sql select Booster_Version from SPACEXTBL WHERE LANDING_OUTCOME = 'Success (drone ship)' AND PAYLOAD_MASS_KG_ BETWEEN 4000 AND 6000
... * sqlite:///my\_data1.db
Done.
...
Booster_Version
F9 FT B1022
F9 FT B1026
F9 FT B1021.2
F9 FT B1031.2
```


Resultados consultas SQL

Mostrar el número total de resultados exitosos y no exitosos

```
▶ %sql select Mission_Outcome, COUNT(Mission_Outcome) from SPACEXTBL GROUP BY Mission_Outcome
[68]
... * sqlite:///my\_data1.db
Done.
...


| Mission_Outcome                  | COUNT(Mission_Outcome) |
|----------------------------------|------------------------|
| Failure (in flight)              | 1                      |
| Success                          | 98                     |
| Success                          | 1                      |
| Success (payload status unclear) | 1                      |


```

Resultados consultas SQL

Mostrar los nombres de las versiones de los cohetes que han transportado la máxima carga útil.

```
%sql select Booster_Version, payload_mass_kg_ from SPACEXTBL WHERE PAYLOAD_MASS_KG_ = (SELECT MAX(PAYLOAD_MASS_KG_) FROM SPACEXTBL)
```

[66]

... * [sqlite:///my_data1.db](#)

Done.

...

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

Resultados consultas SQL

Listar los registros que muestren los meses, resultados fallidos en drone ship, versión del cohete y sitio de lanzamiento en el año 2015.

```
%sql SELECT SUBSTR(Date, 6, 2) AS MONTH, landing_outcome, Booster_Version, Launch_Site FROM SPACEXTBL WHERE substr(Date,0,5)='2015' and landing_outcome = 'Failure (drone ship)'
```

... * [sqlite:///my_data1.db](#)
Done.

...

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

Resultados consultas SQL

Ordena la cantidad de resultados de aterrizaje tales como fallos en drone ship o éxitos en plataforma terrestre entre 2010-06-04 y 2017-03-20 en orden descendente.

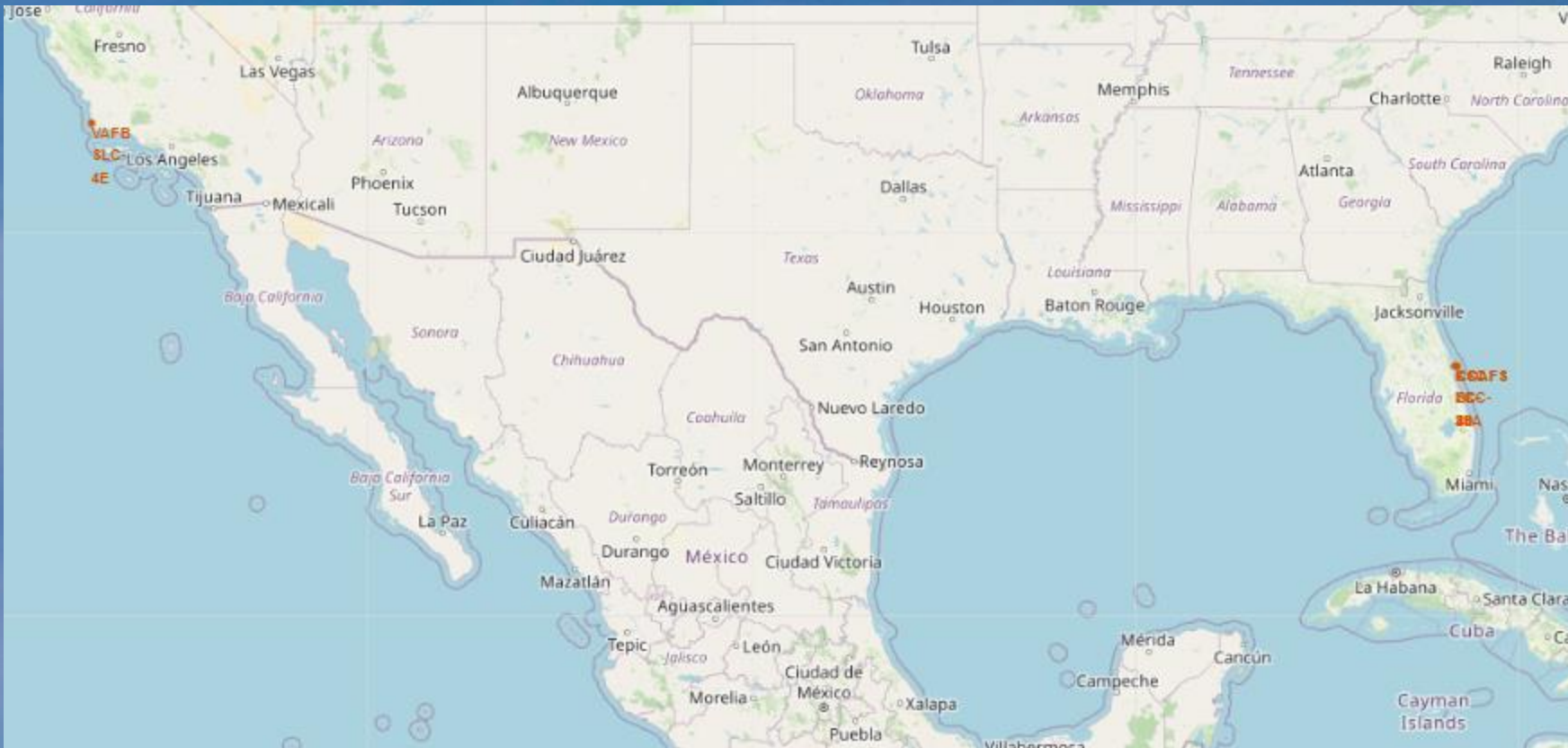
```
%sql select count(landing_outcome), landing_outcome from SPACEXTBL where DATE between '2010-06-04' and '2017-03-20' group by landing_outcome order by count(landing_outcome) desc
```

```
... * sqlite:///my\_data1.db  
Done.
```

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

Resultados mapas interactivos con Folium

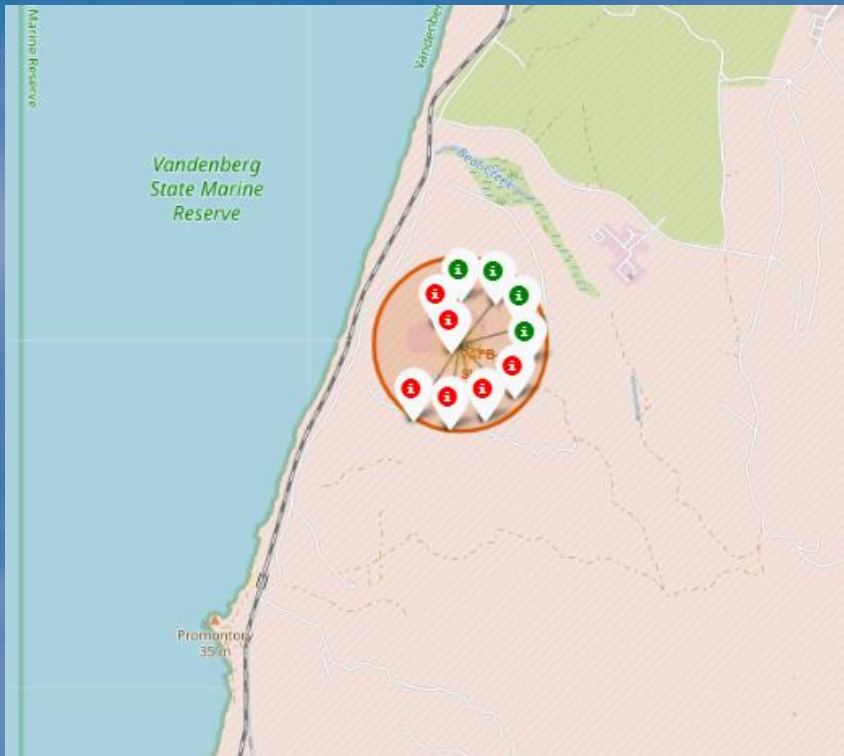
Mostrar todos los lugares de lanzamiento



	Launch Site	Lat	Long
0	CCAFS LC-40	28.562302	-80.577356
1	CCAFS SLC-40	28.563197	-80.576820
2	KSC LC-39A	28.573255	-80.646895
3	VAFB SLC-4E	34.632834	-120.610745

Resultados mapas interactivos con Folium

Posición de los lanzamientos exitosos y no exitosos en cada lugar de lanzamiento.

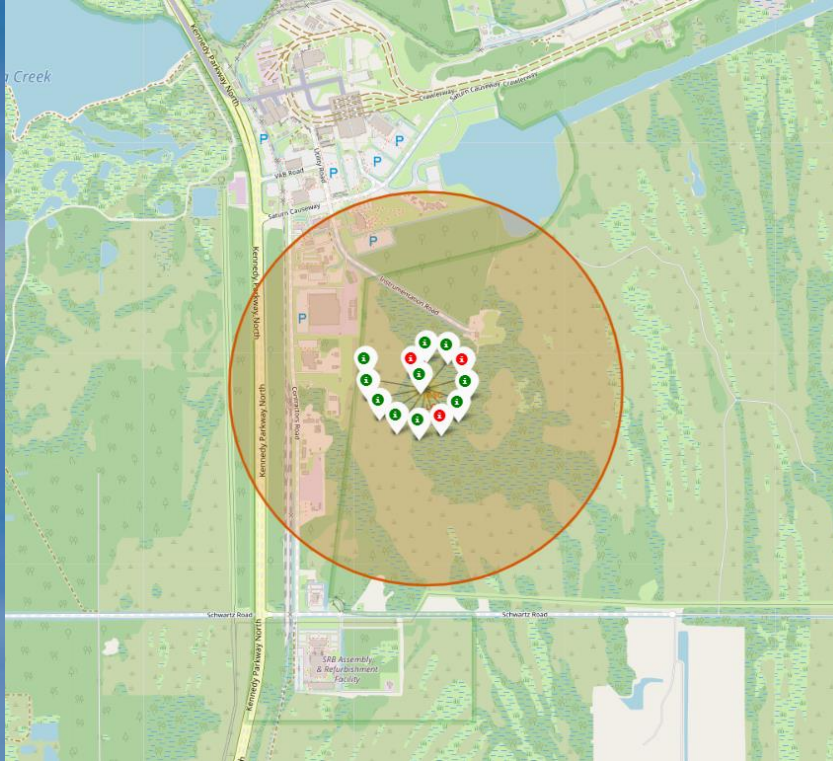


Lugar: VAFB SLC-4E en costa oeste

- 10 lanzamientos totales. 4 exitosos y 6 fallidos.
- Es el lugar menos usado para los lanzamientos

Resultados mapas interactivos con Folium

Posición de los lanzamientos exitosos y no exitosos en cada lugar de lanzamiento.

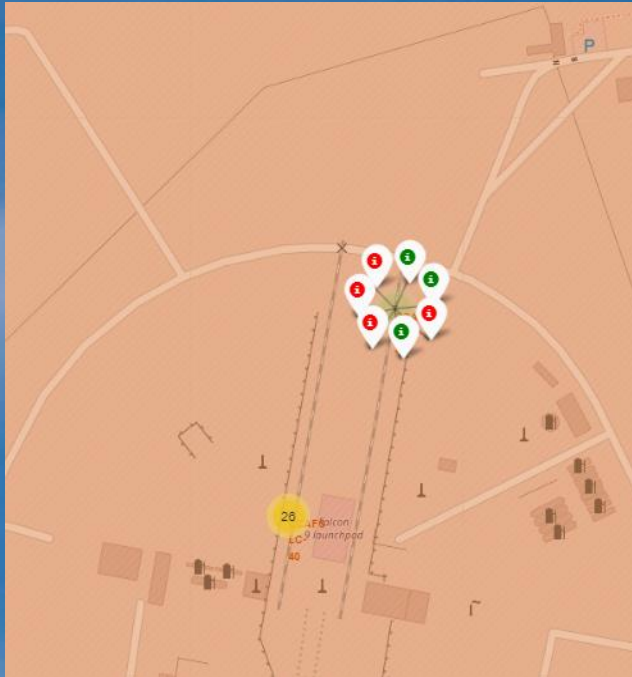


Lugar: KSC LC-39A en costa este

- 13 lanzamientos totales. 10 exitosos y 3 fallidos.

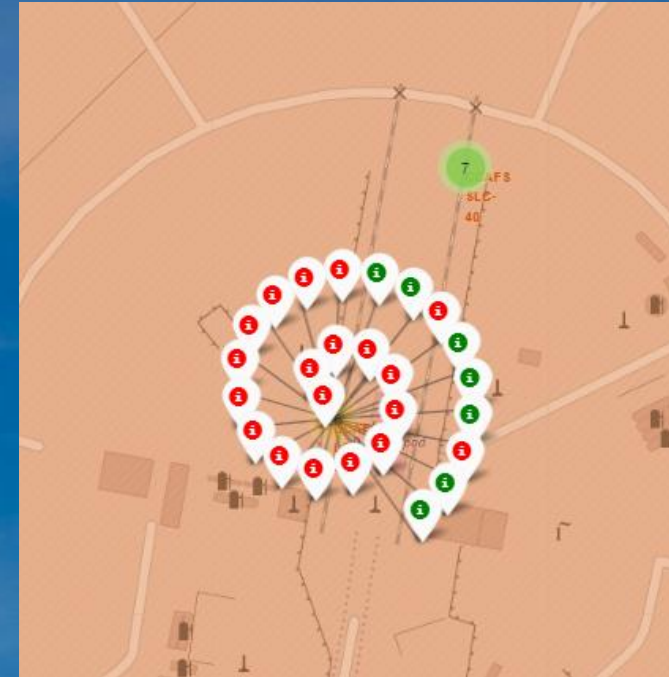
Resultados mapas interactivos con Folium

Posición de los lanzamientos exitosos y no exitosos en cada lugar de lanzamiento.



Lugar: CCAFS SLC -40

- 7 lanzamientos totales. 3 exitosos y 4 fallidos.

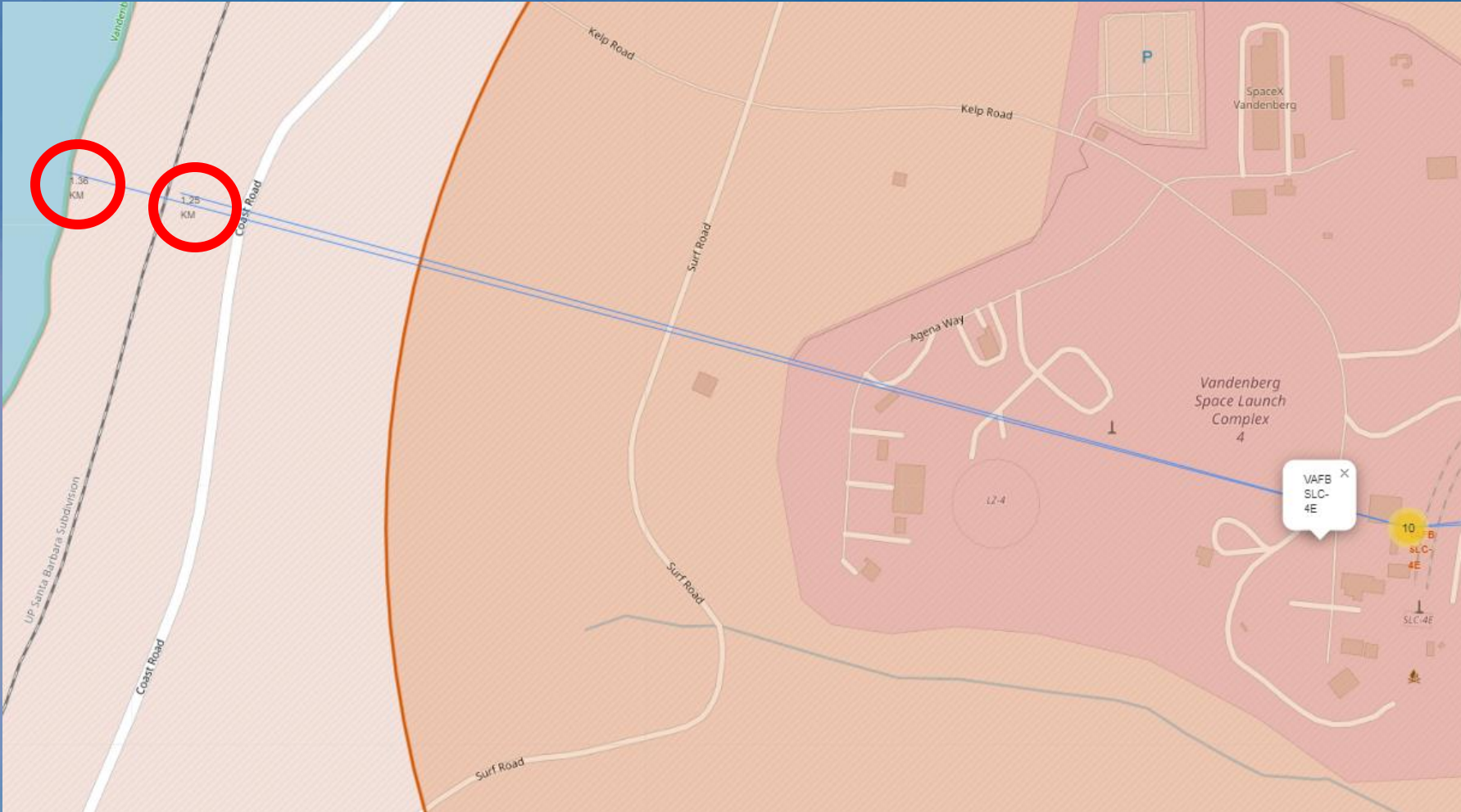


Lugar: CCAFS LC-40 en costa este

- 26 lanzamientos totales. 7 exitosos y 19 fallidos.

Resultados mapas interactivos con Folium

Mostrar las distancias a la costa, ciudad, vías de tren y autopista más cercanas al punto de lanzamiento VAFB SLC-4E en costa oeste.



Distancias:

- A la costa:

1.3554890309623102 km

- A la vía de tren más cercana.

1.2450966556116494 km

Resultados mapas interactivos con Folium

Mostrar las distancias a la costa, ciudad, vías de tren y autopista más cercanas al punto de lanzamiento VAFB SLC-4E en costa oeste.



Resultados Plotly Dash Dashboard

Distribución de éxito por lugares de lanzamiento.

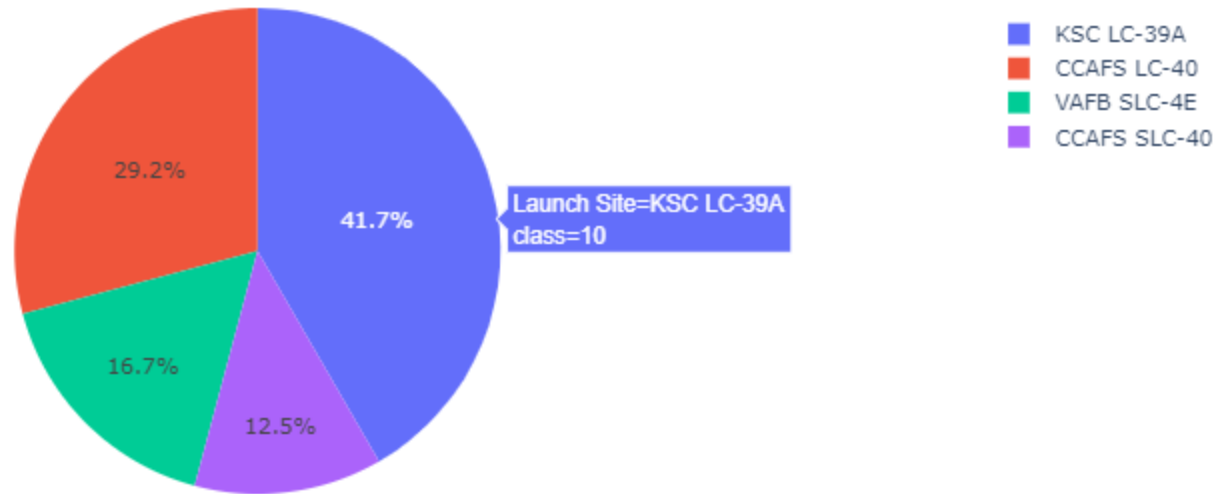
SpaceX Launch Records Dashboard

All Sites



Succes count for all launch Sites

Testing

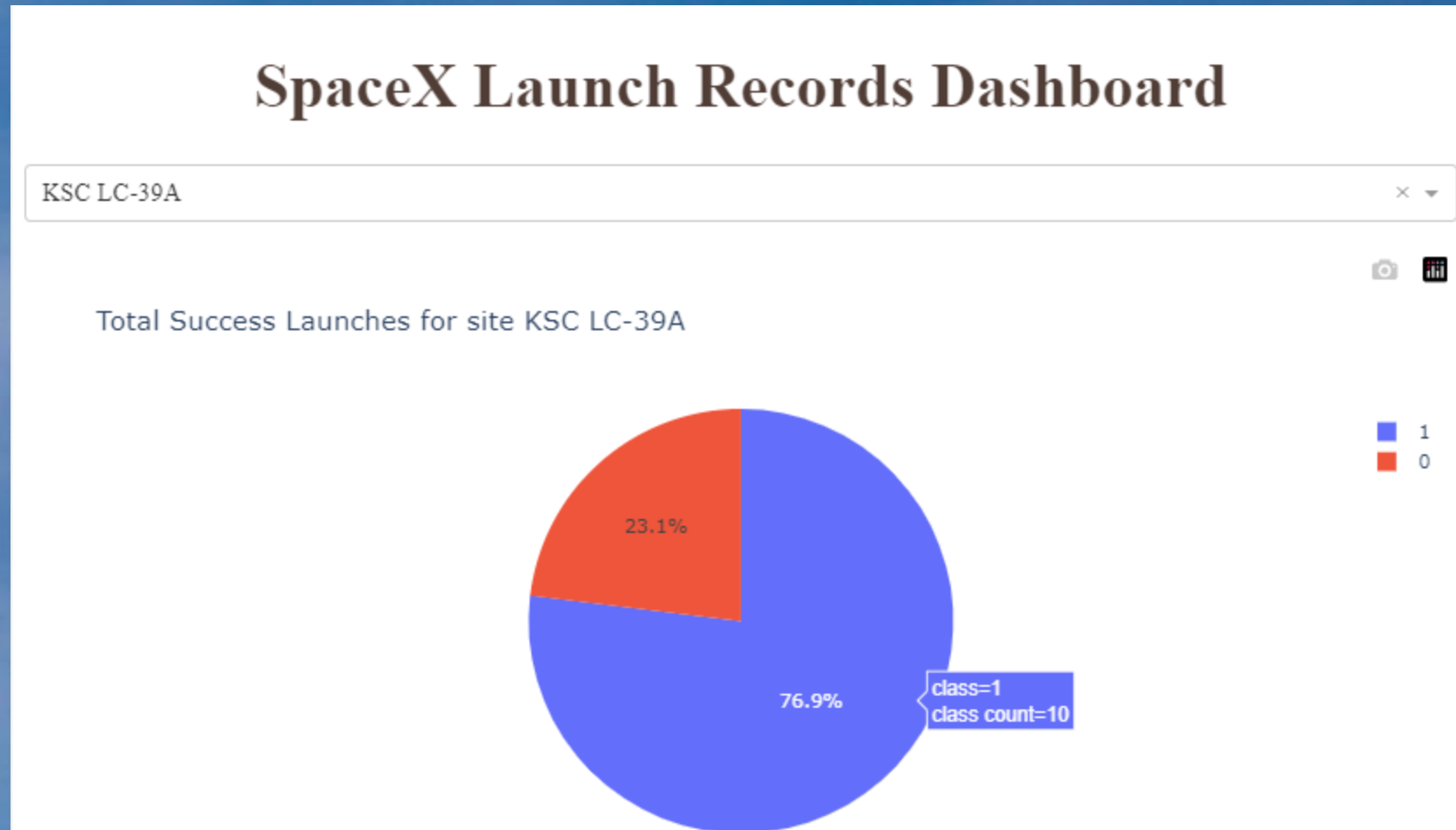


KSC LC-39A tiene la mejor tasa de lanzamientos exitosos.

La estación **CCAFS SLC-40** tiene la tasa más baja de éxito.

Resultados Plotly Dash Dashboard

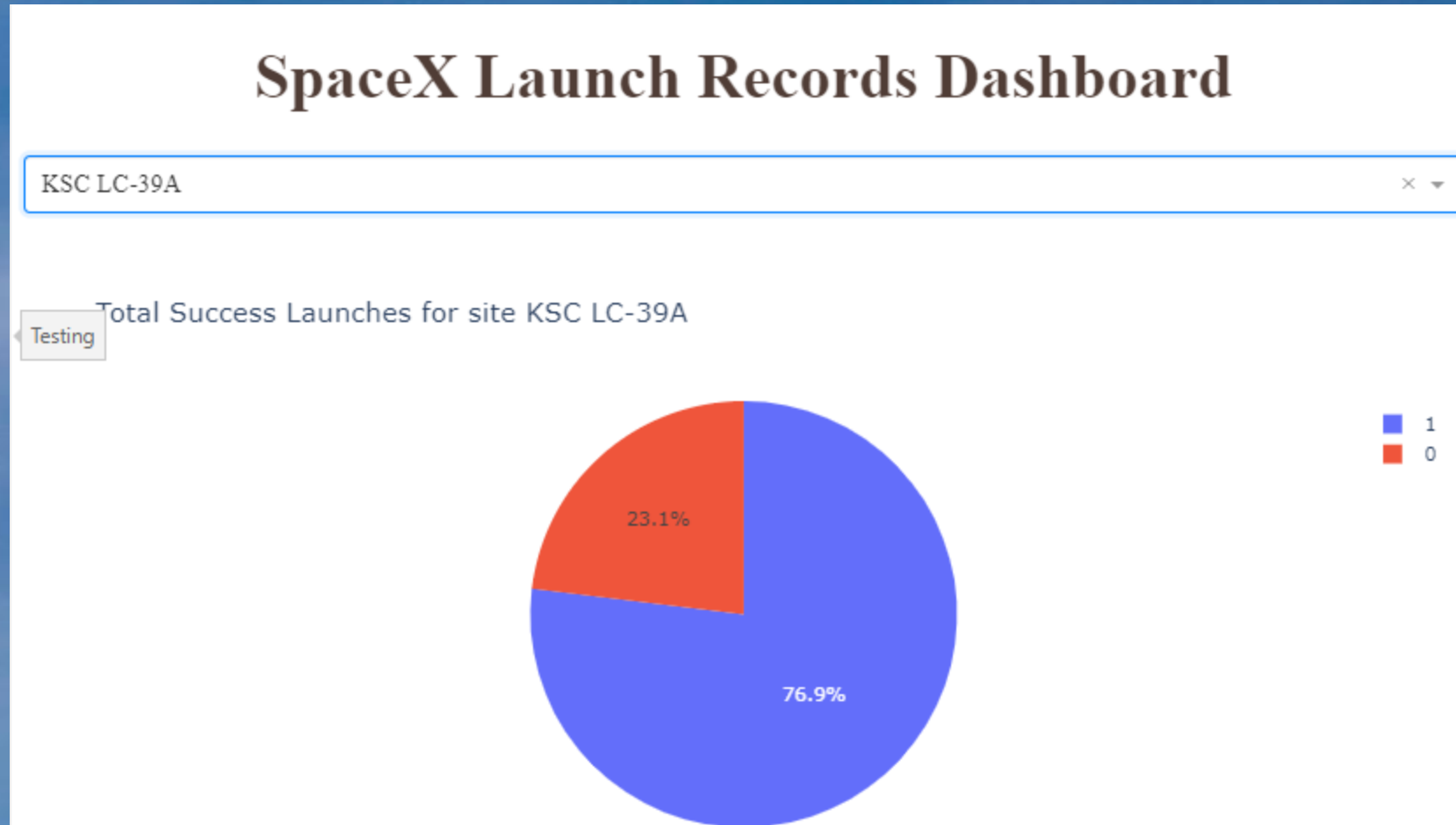
¿Cuál es el lugar con mayor número de lanzamientos exitosos?



KSC LC-39A es el sitio con mayor número de lanzamientos exitosos (10).

Resultados Plotly Dash Dashboard

¿Cuál es el lugar con la tasa más alta de éxito en los lanzamientos?



KSC LC-39A tiene la mejor tasa de lanzamientos exitosos con un 76,9% de éxito.

Resultados Plotly Dash Dashboard

¿Qué rango de carga útil tiene el mejor ratio de lanzamientos con éxito?



Según mi criterio, el rango que tiene la mayor proporción de lanzamientos exitosos está entre 2.000 y 4000.

Resultados Plotly Dash Dashboard

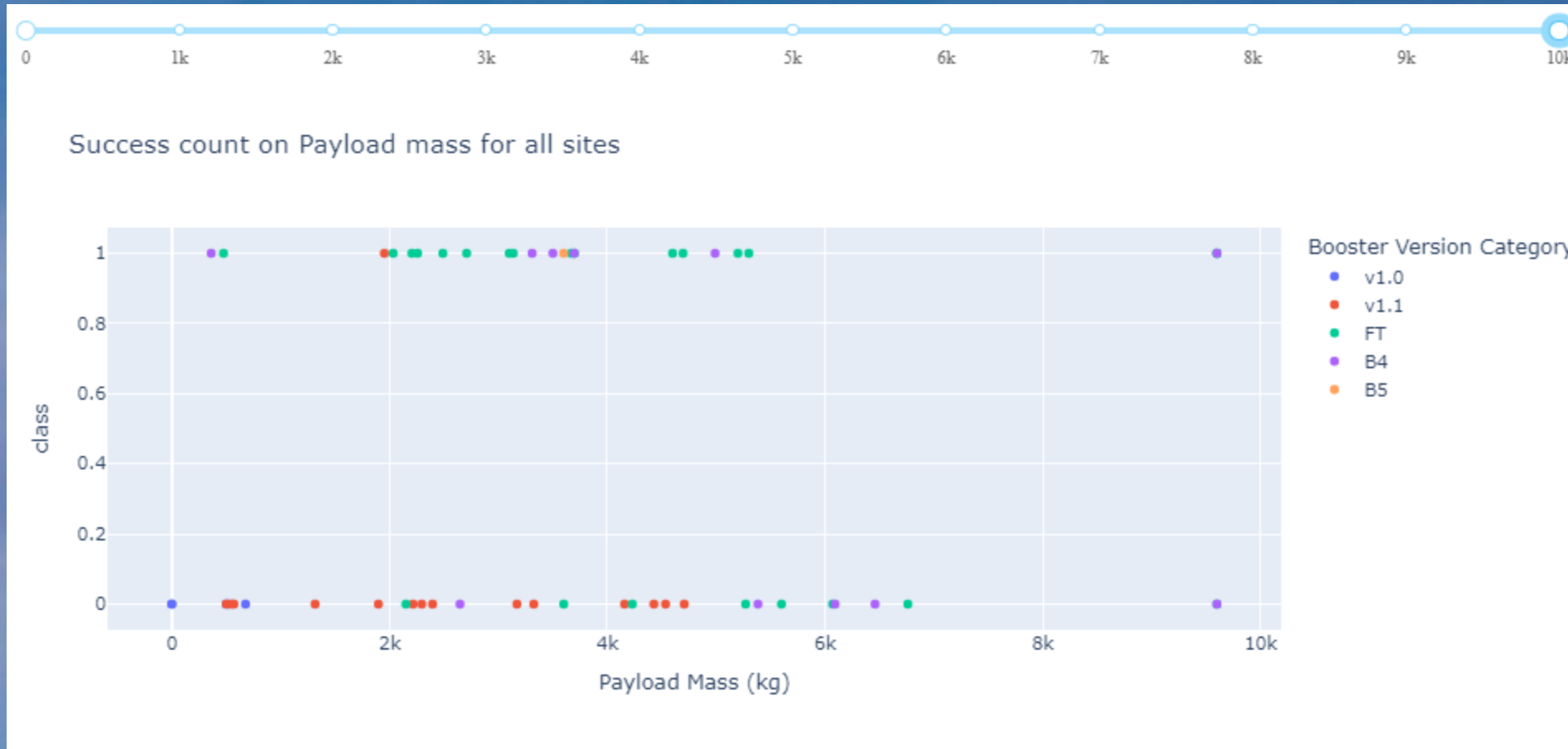
¿Qué rango o rangos de carga útil tiene la proporción más baja de lanzamientos con éxito?



El rango 6.000 y 9.000 tiene la tasa más baja de éxito ya que todos los lanzamientos han sido fallidos.

Resultados Plotly Dash Dashboard

¿Qué versión del cohete F9 Booster (v1.0, v1.1, FT, B4, B5, etc.) tiene la tasa de éxito de lanzamiento más alta?



Se puede apreciar que **FT** tiene un número considerable de lanzamientos exitosos frente a los fallidos, teniendo el ratio más alto (13/6). Le sigue **B4** (6/5).

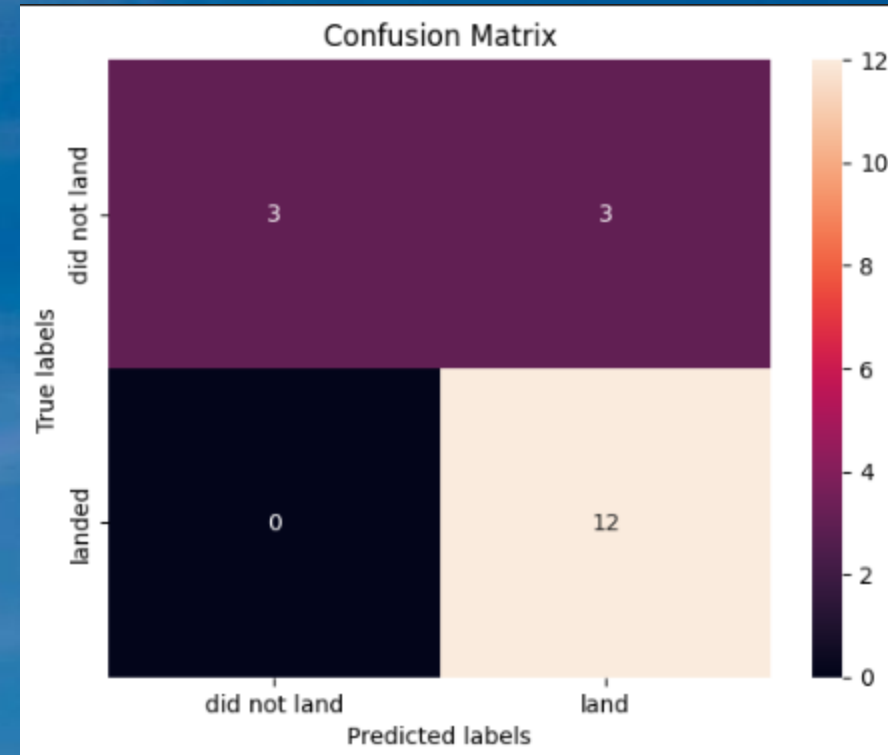
Resultados Análisis Predictivo - Clasificación

Regresión logística

```
tuned hpyerparameters :(best parameters) {'C': 0.01, 'penalty': 'l2', 'solver': 'lbfgs'}  
accuracy : 0.8464285714285713
```

```
logreg_cv.score(X_test, Y_test)
```

```
0.8333333333333334
```



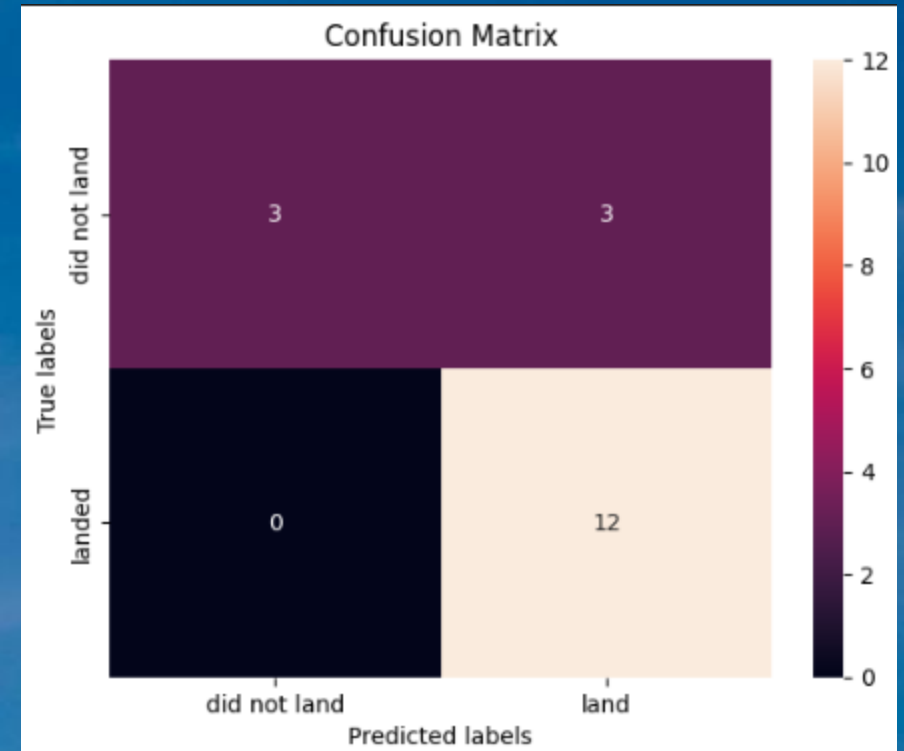
Resultados Análisis Predictivo - Clasificación

Support Vector Machine

```
tuned hpyerparameters :(best parameters) {'C': 1.0, 'gamma': 0.03162277660168379, 'kernel': 'sigmoid'}  
accuracy : 0.8482142857142856
```

```
svm_cv.score(X_test, Y_test)
```

```
0.8333333333333334
```



Resultados Análisis Predictivo - Clasificación

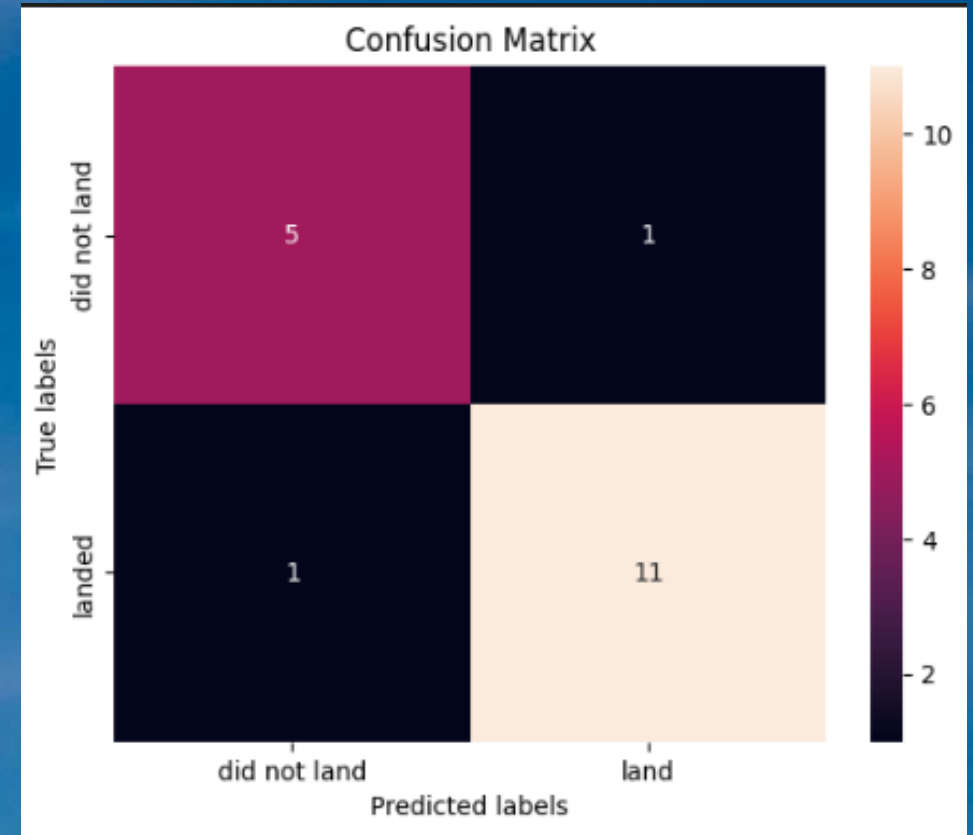
Árbol de Decisiones

```
tuned hpyerparameters :(best parameters) {'criterion': 'gini', 'max_depth': 16, 'max_features': 'sqrt', 'min_samples_leaf': 4, 'min_samples_split': 10, 'splitter': 'random'}  
accuracy : 0.8875
```

[+ Code](#)[+ Markdown](#)

```
tree_cv.score(X_test, Y_test)
```

```
0.8888888888888888
```



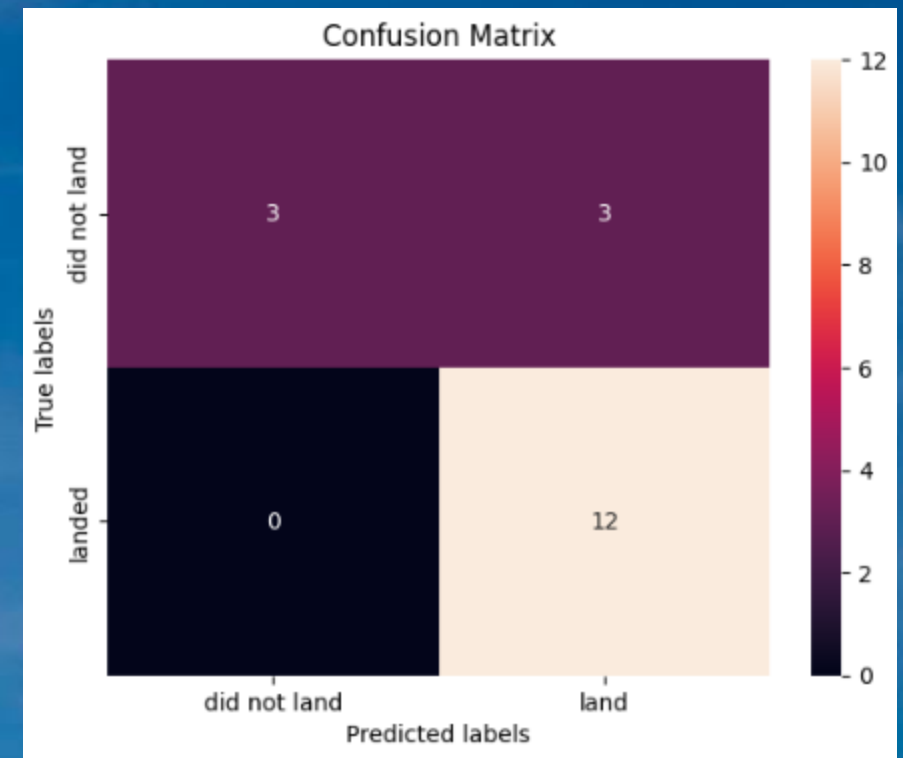
Resultados Análisis Predictivo - Clasificación

K Vecinos más Cercanos

```
tuned hpyerparameters :(best parameters) {'algorithm': 'auto', 'n_neighbors': 10, 'p': 1}  
accuracy : 0.8482142857142858
```

```
knn_cv.score(X_test, Y_test)
```

```
0.8333333333333334
```



Resultados Análisis Predictivo - Clasificación

Buscando el modelo más preciso con el método score()

```
print('=' * 76)
print('Podemos observar que la precisión es SUPERIOR para el árbol de decisiones.')
print('=' * 76)
print(f'Precisión para Regresión Logística: {logreg_cv.score(X_test, Y_test)}')
print(f'Precisión para SVM: {svm_cv.score(X_test, Y_test)}')
print(f'Precisión para Árbol de Decisiones: {tree_cv.score(X_test, Y_test)}')
print(f'Precisión para KNN: {knn_cv.score(X_test, Y_test)}')
```

```
=====
Podemos observar que la precisión es SUPERIOR para el árbol de decisiones.
=====
```

```
Precisión para Regresión Logística: 0.8333333333333334
Precisión para SVM: 0.8333333333333334
Precisión para Árbol de Decisiones: 0.8888888888888888
Precisión para KNN: 0.8333333333333334
```

Observamos que el árbol de decisiones es el más preciso de los cuatro modelos con una puntuación de 0.89 usando score().

Resultados Análisis Predictivo - Clasificación

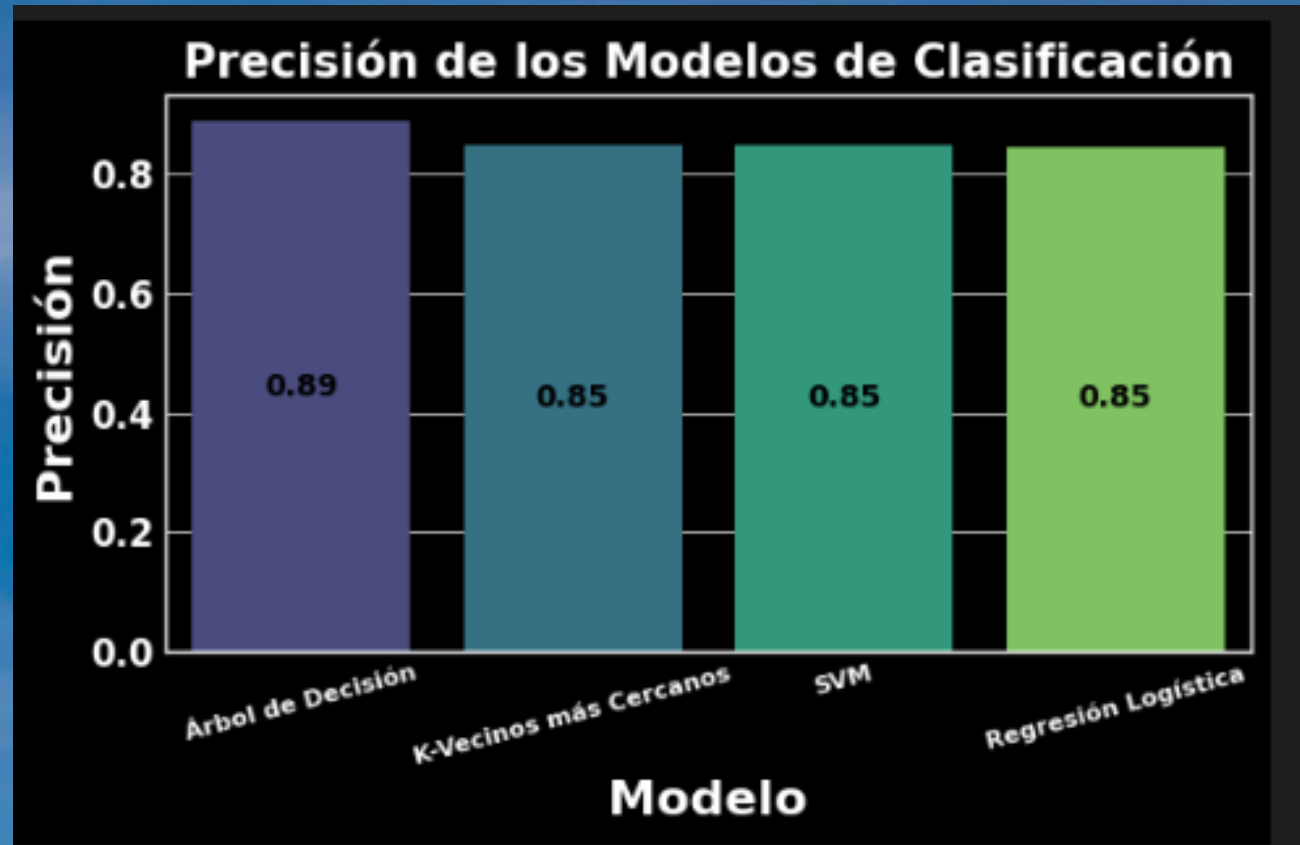
Ordenando los modelos de mejor a peor exactitud con `best_score`

	Model	Accuracy (Best Scores)
2	Árbol de Decisión	0.887500
3	K-Vecinos más Cercanos	0.848214
1	SVM	0.848214
0	Regresión Logística	0.846429

Observamos que también nos sale el árbol de decisiones como modelo más exacto usando el atributo `best_score` de `GridSearchCV`

Resultados Análisis Predictivo - Clasificación

Visualización de la precisión de los cuatro modelos usados



CONCLUSIONES

- Con el paso del tiempo, la tasas de éxito ha aumentado de un modo notable en general. En el año 2018 bajó al desde 80 al 60% para alcanzar el 90% en 2019. Esta bajada de 2018 sería interesante averiguar si de debe a algo en concreto.
- CCAFS (costa este) es el centro que más lanzamientos ha efectuado hasta el momento.
- Las cargas entre 9.500 y 16.000 kg. están asociadas a una tasa de éxito superior.
- Hay determinadas órbitas que tienen una tasa de éxito del 100%. Estas son ES-L1, GEO, HEO y SSO.
- La media de las cargas transportadas por los cohetes de SpaceX es de 2928.4 y han transportado hasta el momento 45596 kg.
- Los lugares de lanzamiento de SpaceX se sitúan convenientemente lejos de ciudades, autopistas y vías de tren, y muy cerca de la costa.
- KSC LC-39A tiene la tasa de lanzamientos exitosos más alta y CCAFS SLC-40 la más baja.
- KSC LC-39A es el sitio con mayor número de lanzamientos exitosos.
- Las cargas transportadas con mayor éxito están en el rango 2000-4000 y las que menos, 6000-8000
- La versión FT del cohete F) es la que mejor tasa de éxito tiene en sus lanzamientos.
- Los cuatro modelos predictivos usados tienen una precisión similar, pero el **árbol de decisión es apreciablemente mejor**, tanto en sus valores score como best_score y en su matriz de confusión. ya que ha producido 11 verdaderos positivos y 5 verdaderos negativos (16 aciertos de 18) frente a los 12 verdaderos positivos y sólo 3 verdaderos negativos (16 aciertos de 18) de los otros tres modelos, aunque tiene un falso negativo que los otros modelos no tienen.