

# Presentacion Proyecto Ciencia de Datos - SpaceX

## Resumen Ejecutivo

Objetivo:

Realizar un análisis completo de los lanzamientos espaciales de SpaceX, incluyendo limpieza y manejo de datos, análisis exploratorio, consultas SQL, visualizaciones interactivas y análisis predictivo.

Resultados clave:

- Se analizaron más de 60 lanzamientos desde 2010.
- Se aplicaron técnicas de codificación, limpieza y visualización.
- Se creó un mapa interactivo con Folium para mostrar los sitios de lanzamiento.
- Se desarrolló un dashboard con Plotly Dash para exploración dinámica de datos.
- El modelo predictivo SVM con kernel RBF alcanzó una precisión del 83.33%.

## Introducción

Contexto:

SpaceX ha revolucionado la exploración espacial con sus lanzamientos de cohetes Falcon 9 y Falcon Heavy. Entender los datos históricos de estos lanzamientos puede ayudar a predecir futuros éxitos y mejorar la planificación.

Propósito del proyecto:

Analizar datos de lanzamientos para identificar patrones, evaluar el desempeño y construir modelos predictivos para el éxito de la misión.

## Metodología de Recopilación y Manejo de Datos

- Se obtuvo la información de la API pública de SpaceX y datos adicionales de Wikipedia.
- Datos principales: fecha de lanzamiento, masa de carga útil, tipo de órbita, sitio de lanzamiento, resultado y detalles del cohete.
- Se limpió el conjunto de datos eliminando valores faltantes y codificando variables categóricas mediante One-Hot Encoding.
- Se utilizaron pandas para la manipulación y normalización de datos.

# Presentacion Proyecto Ciencia de Datos - SpaceX

## Metodología de Análisis Exploratorio de Datos (EDA) y Visualización Interactiva

- Análisis estadístico básico para detectar valores atípicos y distribución de variables.
- Visualización de conteos de lanzamientos por año, sitios y resultados mediante gráficos de barras y líneas.
- Uso de seaborn para gráficos categóricos y scatter plots para relaciones entre variables.
- Creación de mapas con Folium para la ubicación geográfica de sitios de lanzamiento.
- Desarrollo de dashboards interactivos con Plotly Dash para filtrar y explorar los datos dinámicamente.

## Metodología de Análisis Predictivo

- Se entrenaron modelos de clasificación para predecir el éxito o fracaso del lanzamiento.
- Modelos evaluados: Support Vector Machine (SVM), Árbol de Decisión, Regresión Logística.
- Selección del mejor modelo con base en validación cruzada y métricas como precisión y  $R^2$ .
- El SVM con kernel RBF tuvo el mejor rendimiento en el conjunto de validación.

## Resultados EDA y Visualización

- El sitio CCAFS SLC 40 tuvo la mayor cantidad de lanzamientos (~41).
- La tasa de éxito general fue aproximadamente 80%.
- Las órbitas más comunes fueron LEO y GTO.
- La mayoría de los aterrizajes exitosos ocurrieron en naves no tripuladas y plataformas oceánicas.
- Gráficos muestran la evolución anual y el tipo de carga útil por tipo de cohete.

(Aquí incluir imágenes de gráficos de barras, pie charts, scatter plots, etc.)

## Resultados de Consultas SQL

- Se utilizaron consultas para obtener masa mínima y máxima de carga útil, conteos por sitio de lanzamiento y estado de la misión.
- Ejemplo: `SELECT MIN(payload__mass__kg_) FROM spacextbl;`
- Conteo de lanzamientos exitosos por sitio, filtrado de datos para análisis específico.
- Se identificaron las ubicaciones exactas de los sitios y se relacionaron con la tasa de éxito.

# Presentacion Proyecto Ciencia de Datos - SpaceX

## Mapa Interactivo con Folium

- Se generó un mapa con marcadores para cada sitio de lanzamiento usando Folium.
- Los marcadores contienen información relevante como número de lanzamientos y coordenadas.
- Uso de clusters para manejar múltiples lanzamientos en ubicaciones similares.
- Visualización geoespacial que permite la interacción y exploración del histórico de lanzamientos.

## Dashboard Plotly Dash

- Creación de un dashboard que permite filtrar lanzamientos por sitio, año y tipo de órbita.
- Visualización dinámica de gráficos que actualizan según selección del usuario.
- Integración de dropdowns, sliders y gráficos interactivos para análisis profundo.
- Facilita la comprensión y comunicación de patrones en los datos.

## Resultados de Análisis Predictivo

- El modelo SVM con kernel RBF logró una precisión del 83.33% en datos de prueba.
- Evaluación con métricas: Accuracy, F1 Score, y Matriz de Confusión.
- El análisis mostró la importancia de ciertas características como masa de la carga útil y tipo de órbita para la predicción.
- Los modelos de árbol y regresión tuvieron menor desempeño en comparación.

## Conclusiones

- El análisis mostró patrones claros en los lanzamientos exitosos de SpaceX.
- Herramientas de visualización y mapas son esenciales para explorar datos complejos.
- Los modelos predictivos son prometedores para anticipar resultados futuros con buena precisión.
- Se recomienda seguir mejorando la calidad y cantidad de datos para futuros análisis.

## Creatividad e Innovación

- Integración de múltiples herramientas: API, SQL, Python, Folium y Dash.
- Uso de visualizaciones interactivas para facilitar la exploración.
- Desarrollo de una narrativa clara para comunicar resultados de ciencia de datos.

# **Presentacion Proyecto Ciencia de Datos - SpaceX**

- Presentación estructurada que conecta análisis descriptivo y predictivo.