IBM Capstone Project

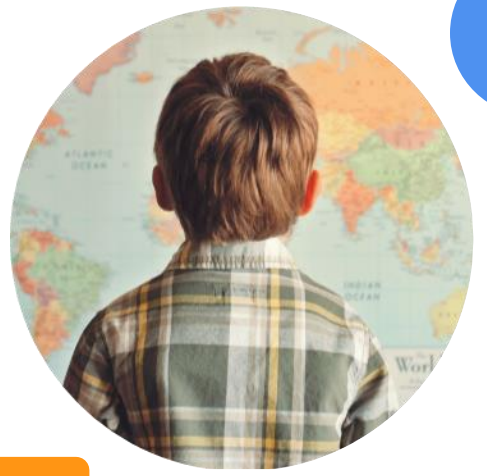
Olcay-09 June 2023

Executive Summary

This executive summary provides an overview IBM Capstone Project conducted by XYZ Company. The project aimed to analyze The Spacex's attempts to decrease the expenses and help to develop an actionable insights to optimize strategies and enhance the experiences.

The methodology employed a combination of exploratory data analysis, clustering techniques, and predictive modeling. A comprehensive dataset was utilized to uncover meaningful patterns and segments.

Table of Contents



Introduction

The IBM Data Science Capstone project serves as the culmination of the Data Science Professional Certificate program, offering an opportunity to apply the acquired knowledge and skills to a real-world scenario. This project focuses on solving a practical problem using data science techniques and methodologies.

The objective of this capstone project is to analyze a given dataset and extract meaningful insights that can drive decision-making and provide value to stakeholders by leveraging data science tools and techniques, participants will explore, preprocess, analyze, and visualize the data to derive actionable insights.

Methodology

1.Data Understanding: Data Experts will familiarize themselves with the dataset, its structure, and the variables involved.

2.Data Preparation: This step involves cleaning and preprocessing the data to address missing values, outliers, and inconsistencies. Data Experts will also perform feature engineering and data transformation to enhance the dataset's suitability for analysis.

3.Exploratory Data Analysis: Through visualizations and statistical analysis, Data experts will uncover patterns, relationships, and trends within the data. This stage aims to gain a deeper understanding of the dataset and identify potential variables of interest.

4.Model Development: Data Specialists will apply appropriate machine learning algorithms to develop predictive or descriptive models, depending on the project's goals. This step involves training, evaluating, and fine-tuning the models to achieve optimal performance.

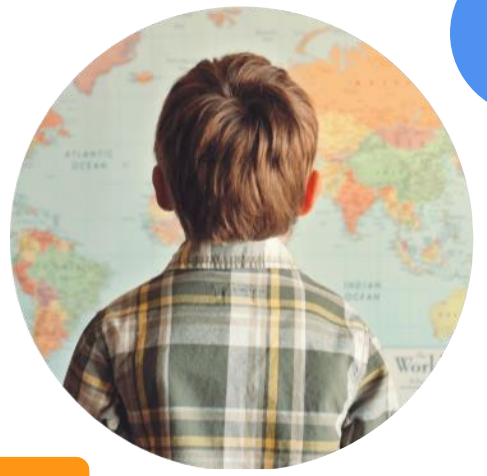
5.Results Evaluation: The performance of the developed models will be assessed using relevant metrics and techniques. Participants will interpret the results and evaluate the models' effectiveness in addressing the problem statement.

6.Conclusion and Recommendations: Based on the findings, Specialists , summarize the insights gained, and propose actionable recommendations for stakeholders. These recommendations aim to drive informed decision-making and potentially provide solutions to the problem at hand.

EDA & Interactive Visual Analytics Methodology

For the project, I conducted a comprehensive Exploratory Data Analysis (EDA) using various techniques and employed interactive visual analytics methodologies to gain insights into the dataset. Checking for missing values and performing data cleaning and preprocessing steps to ensure the dataset's integrity.

I employed interactive visual analytics tools and techniques to explore the relationships and patterns within the data. Interactive charts, such as scatter plots, bar charts, and heatmaps, to visualize the distribution of variables and identify potential trends or clusters. These interactive visualizations allowed us to dynamically explore different aspects of the dataset and uncover hidden patterns that would have been difficult to identify through traditional static plots.



Predictive Analysis Methodology

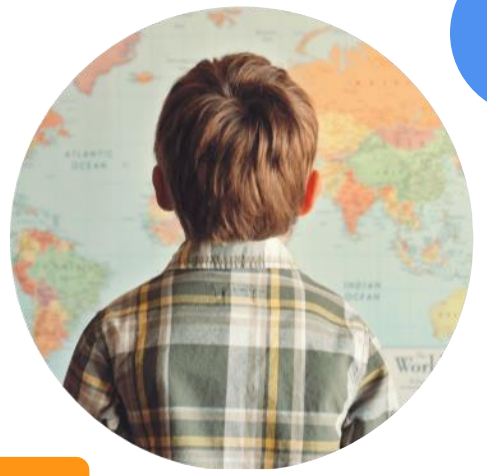
1.Model Training and Evaluation: Train the selected model using the training dataset. Adjust the model's parameters and hyperparameters to optimize its performance. Evaluate the trained model using the testing dataset and calculate relevant evaluation metrics such as accuracy, precision, recall, and F1-score.

2.Model Optimization: Fine-tune the model by iterating on the model selection, parameter tuning, and feature engineering steps to improve its performance. Consider techniques such as cross-validation and grid search to identify the best combination of model parameters.

3.Predictive Analysis Results: Use the optimized model to make predictions on new, unseen data. Analyze the results and assess the model's predictive power. Evaluate its ability to accurately predict the success of SpaceX rocket launches based on the chosen factors.

4.Interpretation and Insights: Interpret the results of the predictive analysis. Identify the most influential factors in predicting launch success and gain insights into the underlying patterns and relationships. Communicate the findings to stakeholders, providing actionable recommendations based on the predictive analysis.

5.Model Deployment: If applicable, deploy the predictive model in a production environment to make real-time predictions. Ensure the model is scalable, efficient, and continuously monitored for performance.



SQL Results

It is obvious from the results, the expenses can be decresed.

Landing_Outcome	Count
Success	20

Avg(PAYLOAD_MASS__KG_)
2928.4

Done.

t[47]:

Sum(PAYLOAD_MASS__KG_)
45596.0

Folium Maps Results

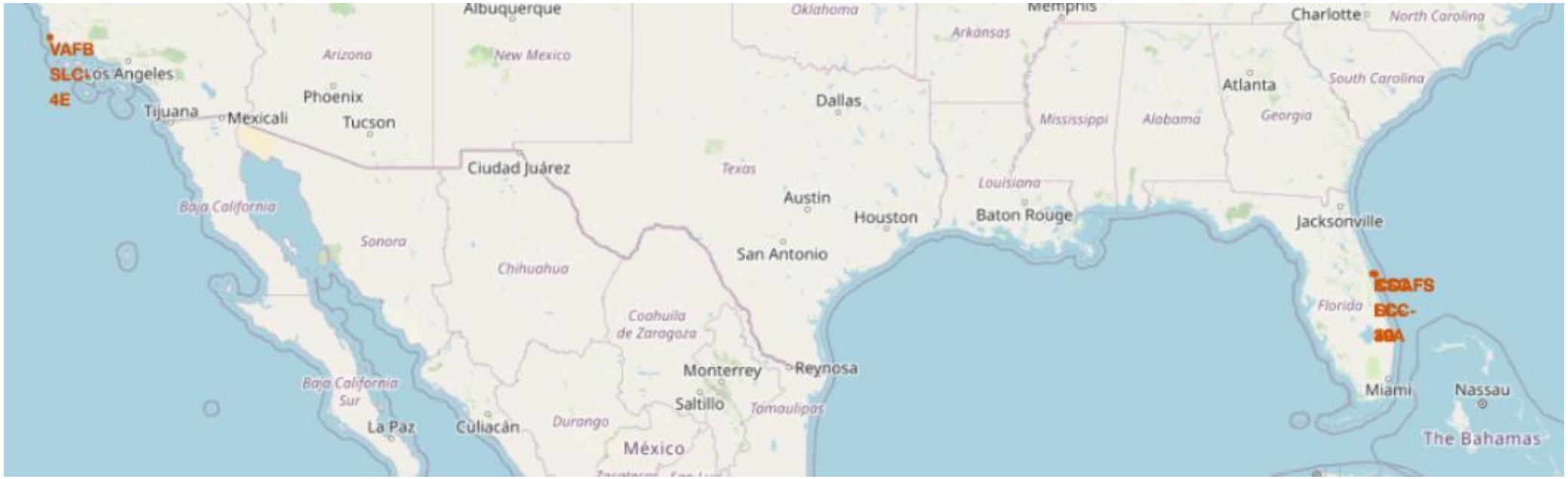
The Longtitude and and the Latitude of the SpaceX locations

7]:

	FlightNumber	Date	BoosterVersion	PayloadMass	Orbit	LaunchSite	Outcome	Flights	GridFins	Reused	Legs	LandingPad	Block	ReusedCount	Serial	Longitude	Latitude	Class
0	1	2010-06-04	Falcon 9	6104.959412	LEO	CCAFS SLC 40	None None	1	False	False	False	NaN	1.0	0	B0003	-80.577366	28.561857	0
1	2	2012-05-22	Falcon 9	525.000000	LEO	CCAFS SLC 40	None None	1	False	False	False	NaN	1.0	0	B0005	-80.577366	28.561857	0
2	3	2013-03-01	Falcon 9	677.000000	ISS	CCAFS SLC 40	None None	1	False	False	False	NaN	1.0	0	B0007	-80.577366	28.561857	0
3	4	2013-09-29	Falcon 9	500.000000	PO	VAFB SLC 4E	False Ocean	1	False	False	False	NaN	1.0	0	B1003	-120.610829	34.632093	0
4	5	2013-12-03	Falcon 9	3170.000000	GTO	CCAFS SLC 40	None None	1	False	False	False	NaN	1.0	0	B1004	-80.577366	28.561857	0

Folium Maps Results

The SpaceX locations

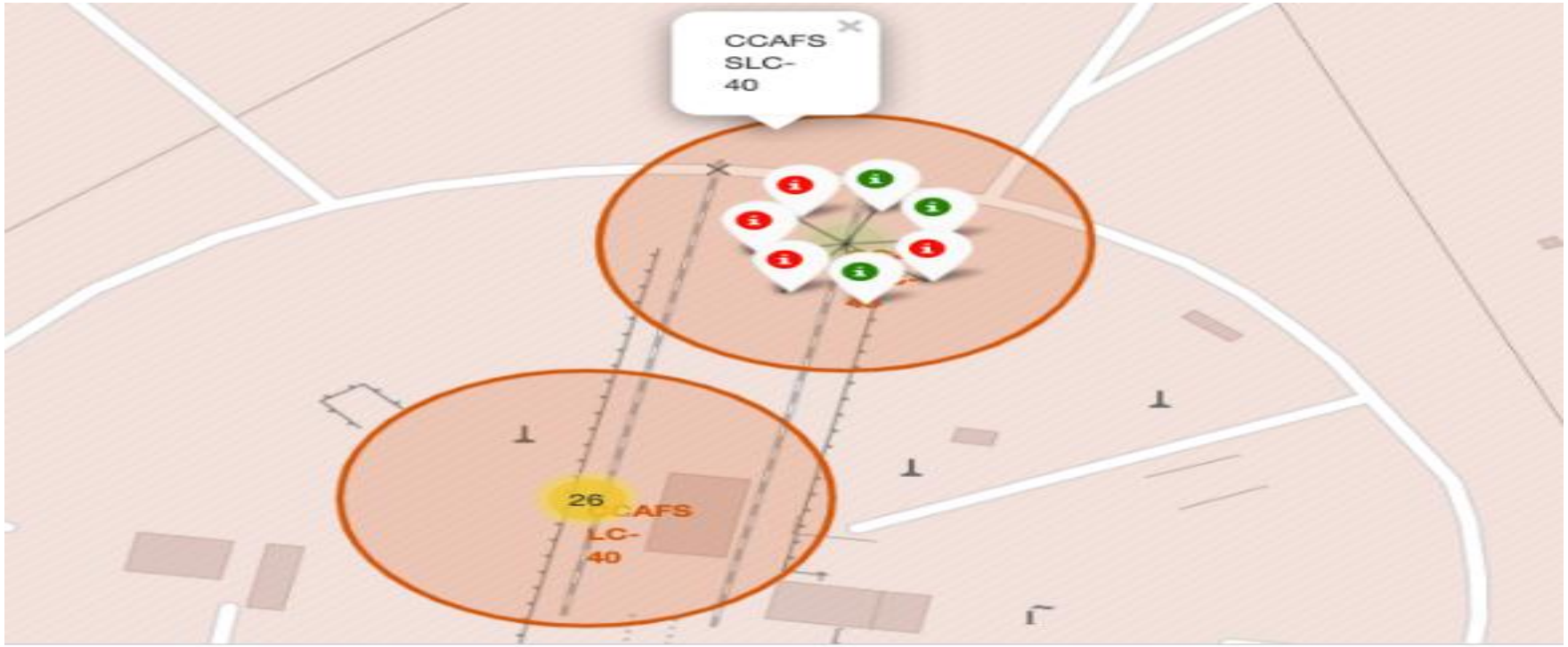


Folium Maps Results

The Most sucessfull SpaceX location

Folium Maps Results

Details

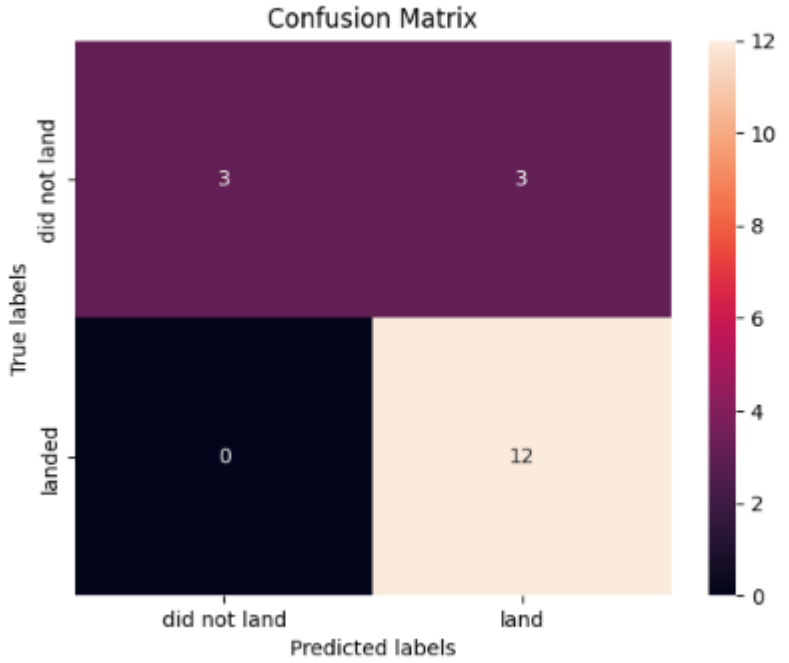


Dash Board Analysis

Predictive Analysis

The predictive Analysis also show it will be beneficial to continue the SpaceX Project.

```
]: logreg_cv.score(X_test, Y_test)
]: 0.8333333333333334
```



Conclusion

In conclusion, the SpaceX Capstone project has provided a comprehensive analysis of SpaceX's missions and launch data. Through the exploration and visualization of the dataset, we gained valuable insights into the trends, patterns, and performance of SpaceX.

The project started with data wrangling, where we cleaned and prepared the dataset for further analysis. Exploratory data analysis (EDA) allowed us to understand the distribution, relationships, and characteristics of the variables in the dataset. Interactive visual analytics enhanced our understanding by providing interactive and visually appealing representations of the data, enabling us to explore different aspects of SpaceX's missions and launches.

Furthermore, predictive modeling techniques were applied to forecast future mission success based on historical data. This allowed us to assess the performance of SpaceX and identify factors that contribute to mission success or failure. The predictive models provided valuable insights into potential areas of improvement and helped in making data-driven decisions.

Overall, the project demonstrated the importance of data analysis and visualization in gaining insights and assisting informed decisions. The findings and recommendations derived from this project can be valuable for SpaceX and other stakeholders in optimizing mission planning, resource allocation, and overall operational efficiency.

Thank you

Olcay-11.06.2023

The way to get started is to quit talking and begin doing.

Walt Disney