Executive SUMMARY

- This project examines global rocket launches to uncover insights into success rates, payload trends, and potential predictive outcomes.
- Data was collected, cleaned, and analyzed using Python, SQL, and interactive visualization tools to explore patterns and performance over time.
- A classification model was developed to predict the likelihood of success for future rocket launches based on historical data.

Introduction

- The space industry has undergone rapid evolution, with the number of launches increasing steadily each year.
- This analysis focuses on evaluating the performance of different space companies, launch sites, and payload characteristics.
- The aim is to extract key insights and develop a predictive model to estimate the probability of launch success.

Data Collection & Wrangling Methodology

- Data sources: Historical rocket launch datasets, including CSV files and online repositories.
- Tools: Python, Pandas, NumPy, and SQL were used for data extraction and preprocessing.
- Data cleaning: Missing values were addressed, company names standardized, and date fields converted.
- Output: A structured dataset prepared for exploratory analysis and predictive modeling.

EDA & Interactive Visual Analytics – Methodology

• Exploratory data analysis: Conducted using Matplotlib, Seaborn, and Plotly to investigate launch patterns.

• Data visualization: Trends in launches, payload distributions, and company performance metrics were effectively visualized.

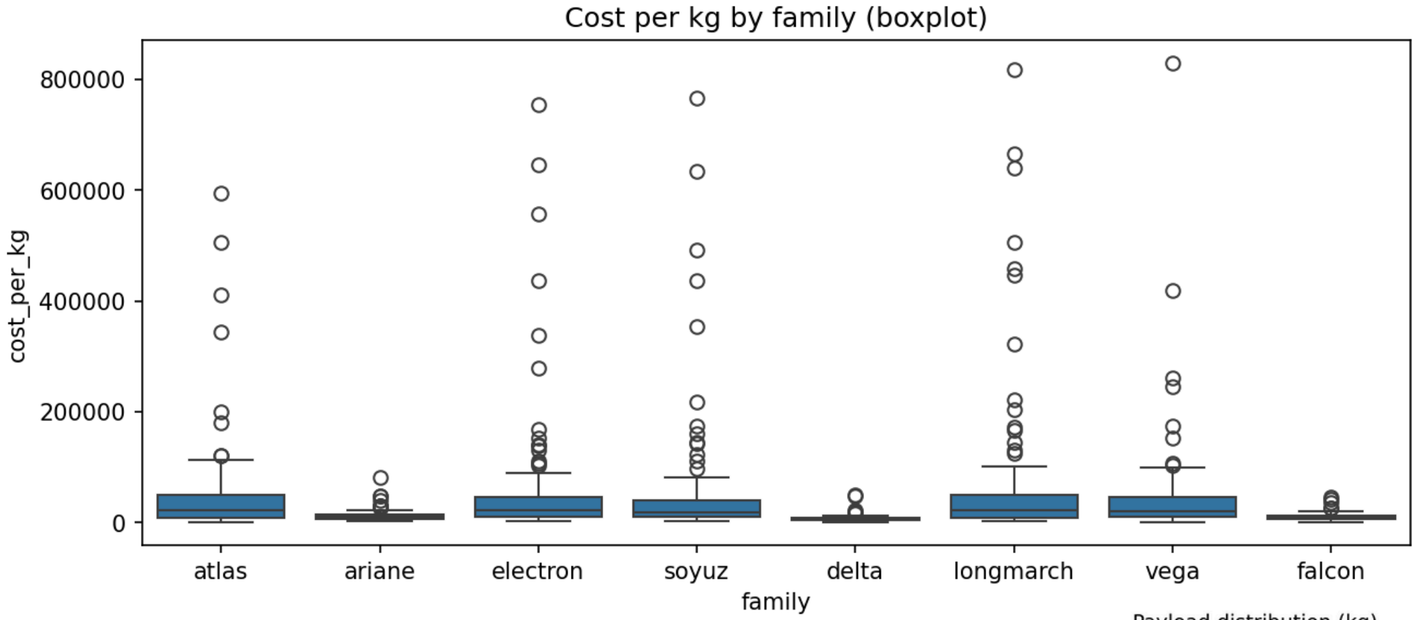
• Interactive dashboards: Built with Plotly Dash to enable deeper insights and facilitate dynamic exploration of the data.

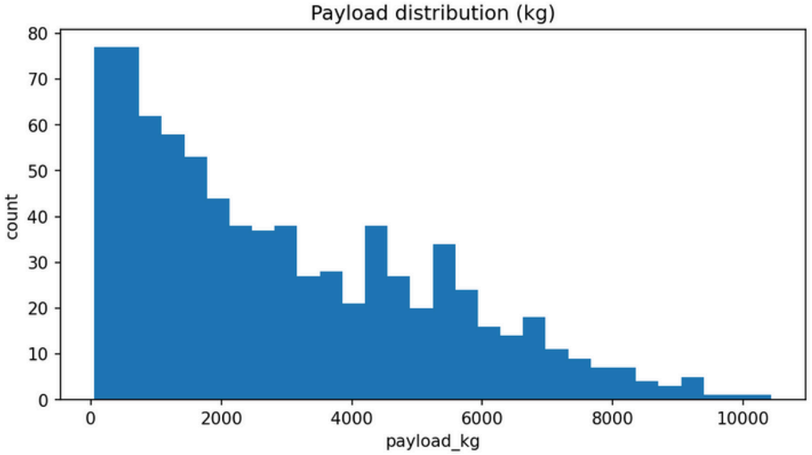
PREDICTIVE ANALYSIS – METHODOLOGY

- Model implementation: A supervised classification model was implemented to predict rocket launch success.
- Models tested: Logistic Regression, Random Forest, and Support Vector Machines were evaluated.
- Performance evaluation: Model performance was assessed using the confusion matrix, ROC curve, and accuracy score.

EDA RESULTS – VISUALIZATIONS

- Visual exploration: Analysis revealed notable insights into launch frequency and company success rates.
- Key results: The main findings from the visual analysis are highlighted for further interpretation.

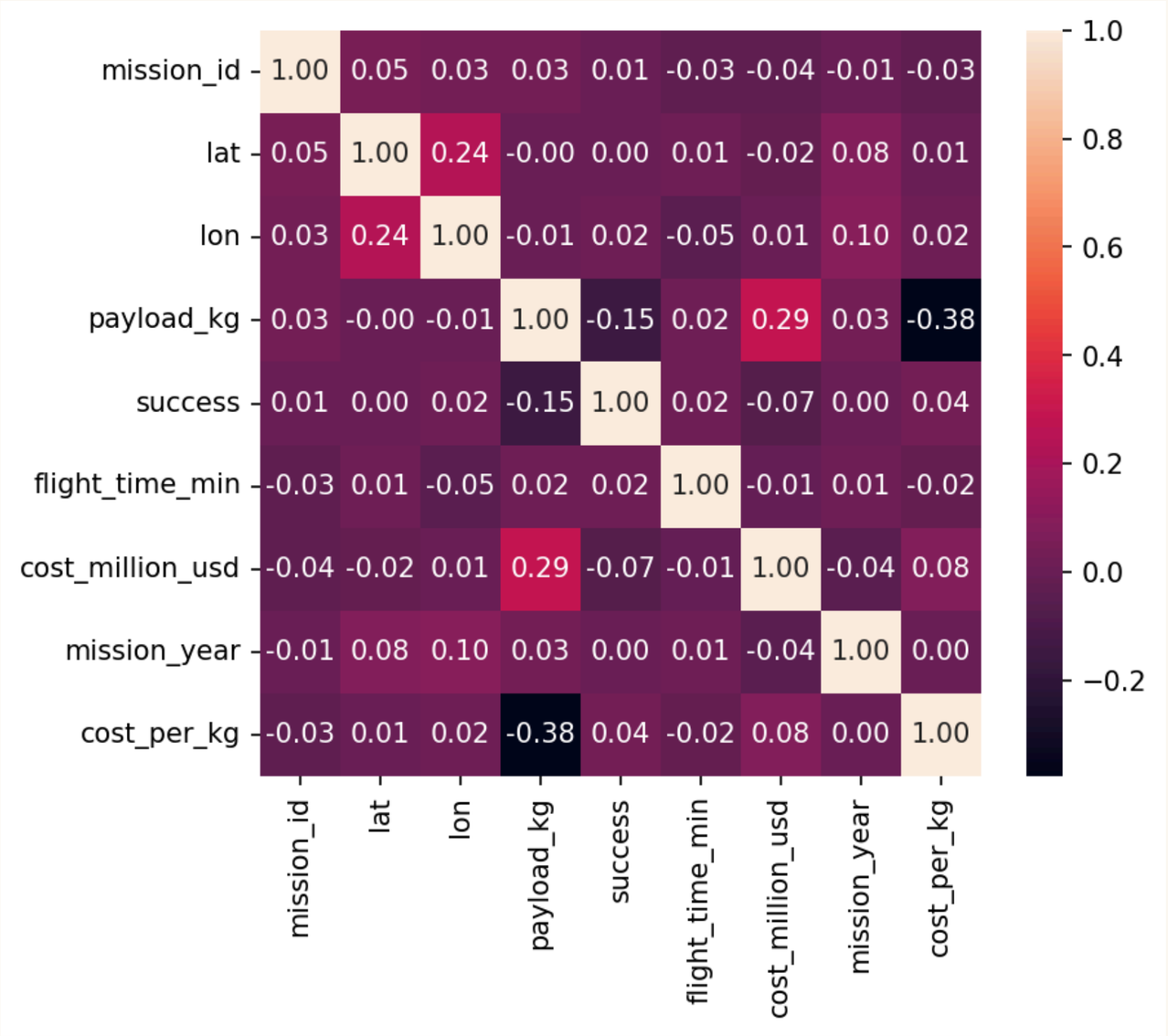




mission_id	rocket_name	family	launch_date	launch_site	lat	lon	payload_kg	orbit	success	flight_time_s	cost_million_u	mission_yea	payload_clas	cost_per_kg
1	rocket_13	atlas	03 00:00:00	[illegible]	[illegible]	[illegible]	1493	gto	1	22	34.78	2012	[illegible]	[illegible]
2	rocket_53	ariane	18 00:00:00	[illegible]	[illegible]	[illegible]	3838	sso	1	124	21.3	2024	large	[illegible]
3	rocket_17	electron	03 00:00:00	[illegible]	[illegible]	[illegible]	1525	gto	1	30	5.0	2001	medium	[illegible]
4	rocket_259	soyuz	10 00:00:00	[illegible]	[illegible]	[illegible]	5618	transfer	1	96	35.64	2001	heavy	[illegible]
5	rocket_113	electron	24 00:00:00	[illegible]	[illegible]	[illegible]	703	leo	1	4	29.92	2020	medium	42500.0
6	rocket_217	ariane	06 00:00:00	[illegible]	[illegible]	[illegible]	2469	gto	1	18	52.95	2015	large	[illegible]
7	rocket_173	soyuz	01 00:00:00	[illegible]	[illegible]	[illegible]	268	heo	1	135	21.93	2004	small	[illegible]
8	rocket_184	atlas	06 00:00:00	[illegible]	[illegible]	[illegible]	1761	leo	0	1	39.18	2015	medium	[illegible]

EDA RESULTS – SQL QUERIES

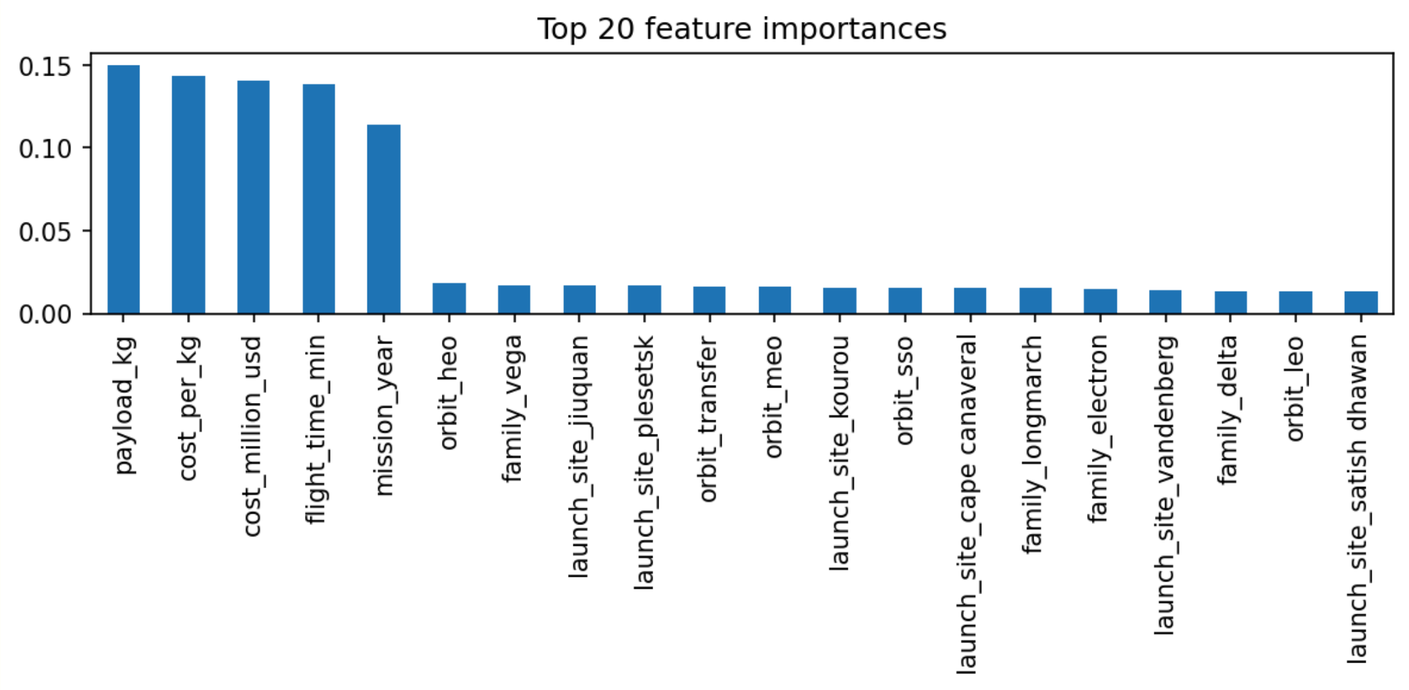
- SQL queries: Used to summarize and aggregate key metrics from the dataset.
- Key outputs: Highlight the top-performing companies and yearly averages derived from the SQL queries.



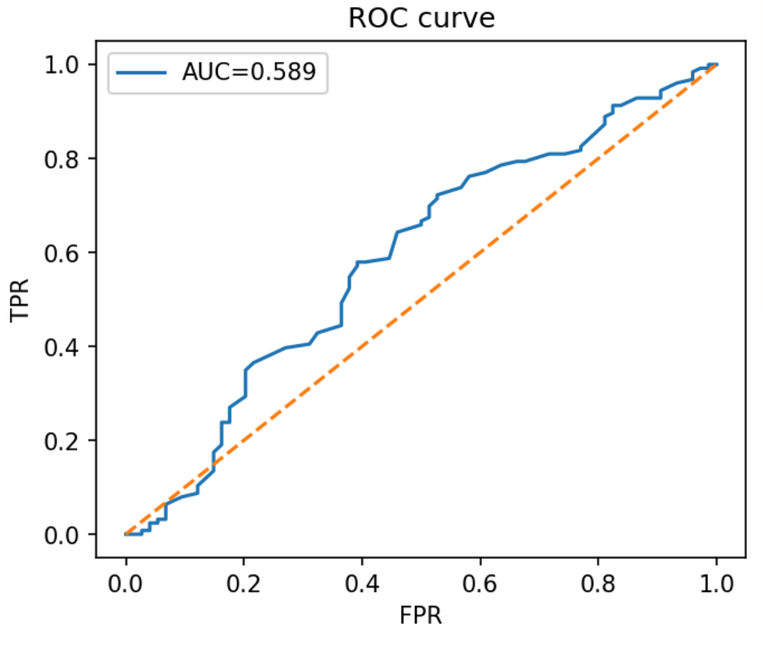
INTERACTIVE MAP – FOLIUM

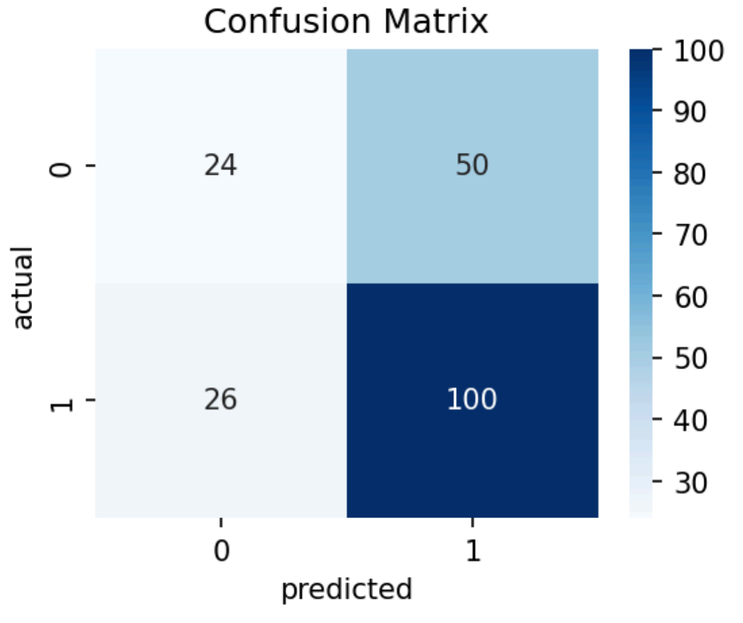
• Global launch visualization: A Folium map was created to display rocket launch locations worldwide.

• Markers and metadata: Each marker represents a launch site and includes relevant information about the launches.



PLOTLY DASH DASHBOARD





- Interactive dashboard: Developed using Plotly Dash to facilitate dynamic exploration of the data.
- User functionality: Allows users to filter and visualize rocket launches interactively through charts and controls.

Predictive Analysis Results

- Predictive model performance: The model achieved promising results, highlighting the key features influencing launch success.
  - Core results: The main findings from the classification analysis are summarized below.

family	cnt	avg_payload	avg_cost
vega	117	1767.8	30.1
ariane	106	4824.3	47.39
delta	105	5034.7	28.21
electron	98	1579.0	29.46
longmarch	97	1708.2	30.62
soyuz	95	1578.1	28.07
falcon	83	5463.1	51.23
atlas	83	1600.9	29.22
	16	3226.3	37.26

CREATIVITY & INNOVATIVE INSIGHTS

- Project integration: Combined interactive tools, visual storytelling, and predictive modeling to create an engaging presentation.
- Presentation quality: Visual design, structured narrative, and insightful conclusions enhanced the clarity and impact of the findings.

CONCLUSION

• Insights from analysis: This study provided meaningful insights into the dynamics and patterns of rocket launches.

• Predictive model performance: The model demonstrated the ability to estimate launch success probabilities with reasonable accuracy.

• Future work: Potential improvements include refining the models and integrating real-time data for enhanced predictions.