Executive summary

Commercial Space Age: A new era in space exploration led by companies like Virgin Galactic, Rocket Lab, Blue Origin, and SpaceX.

SpaceX Achievements:

- Offers affordable Falcon 9 rocket launches at approximately $62 million.
- Cost advantage due to the reusability of the first stage.

Importance of the First Stage:

- Essential for launching payloads into orbit.
- Landing success impacts launch cost estimation.
- Mission parameters can influence landing outcomes.

Capstone Project Role:

- Act as a data scientist for Space Y, a competitor to SpaceX.
- Gather data on SpaceX launches.
- Create dashboards for data analysis.
- Develop a machine learning model to predict first-stage reuse.

Project Goals:

- Enhance understanding of launch pricing.
- Contribute to innovations in the commercial space sector.

Introduction

Overview:

- The commercial space age is reshaping our approach to space exploration.
- Key players: Virgin Galactic, Rocket Lab, Blue Origin, and SpaceX.

Focus on SpaceX:

- Renowned for its innovative Falcon 9 rocket.
- Offers competitive launch prices through first-stage reusability.

Objective:

- Analyze SpaceX's launch data to inform the strategy of Space Y, a new competitor.
- Utilize data science and machine learning to predict the success of first-stage landings.

Let’s embark on this exciting journey into the world of space travel!

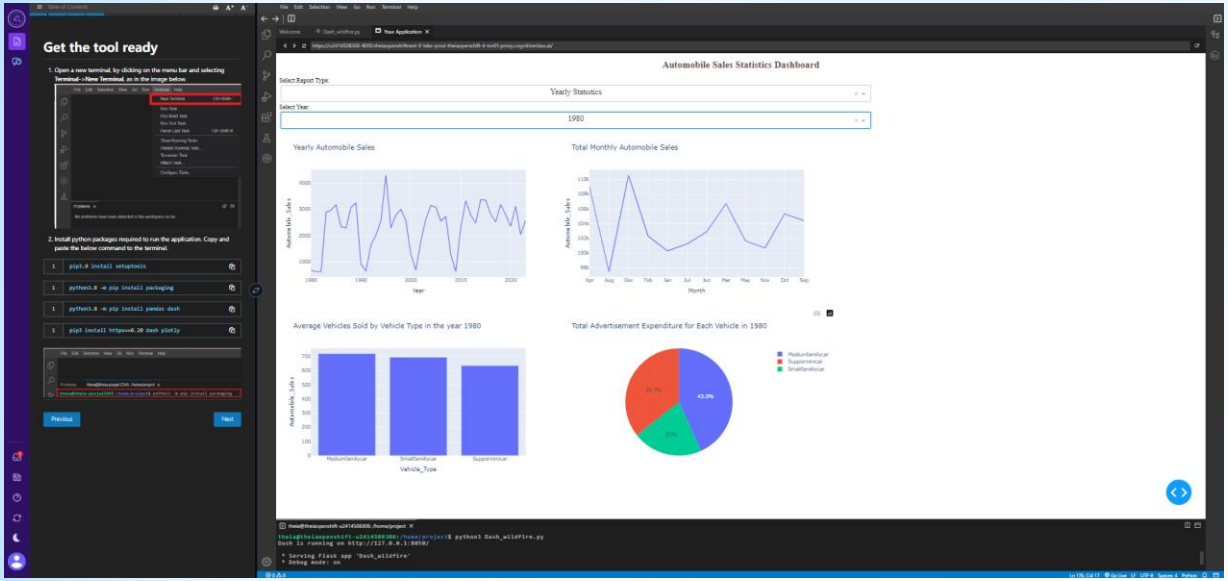
Data Collection and wrangling

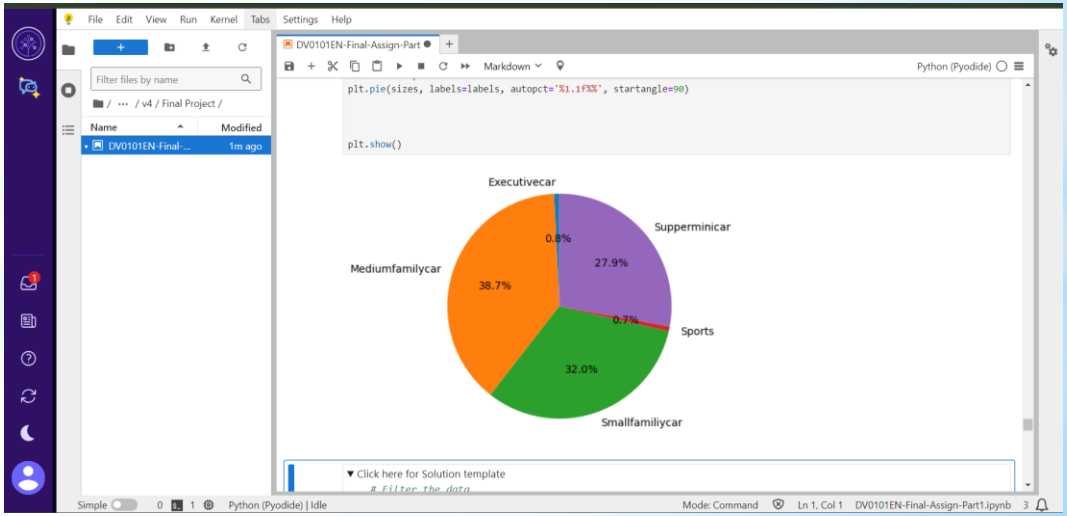
* **Data Collection**
* **Source:**
  * Falcon 9 and Falcon Heavy Launch Records from Wikipedia
  * URL: List of Falcon 9 and Falcon Heavy launches
* **Methodology:**
  * **Web Scraping:**
    * Utilize Python's BeautifulSoup library to extract data.
    * Target the HTML table containing launch records.
* **Data Extraction Process**
* **Access the Web Page:**
  * Fetch the Wikipedia page using requests library.
* **Parse HTML Content:**
  * Use BeautifulSoup to parse the HTML and locate the relevant table.
* **Extract Data:**
  * Retrieve data points such as:
    * Launch Date
    * Mission Outcome
    * Landing Outcome
    * Booster Version
    * Payload Mass
* **Convert to DataFrame:**
  * Create a Pandas DataFrame from the extracted data for further analysis.
* **Example of Unsuccessful Landings**
* Illustrate landing outcomes with visual examples, highlighting:
  * Falcon 9 first stage landing successes.
  * Instances of unsuccessful landings for context.
* **Objectives of the Lab**
* **Primary Goal:**
  * Successfully web scrape and wrangle Falcon 9 launch records into a structured format for analysis.

EDA with SQL results

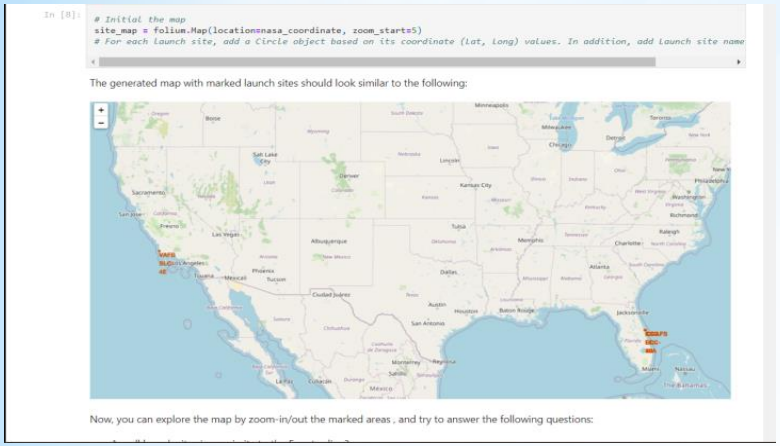
- **Database Connection:**
  - Load the dataset into a Db2 database for SQL querying.
- **Installation:**
  - Use Python libraries: sqlalchemy, ipython-sql, and pandas.
- **Data Cleaning**
- **Remove Incomplete Records:**
  - Ensure that the dataset has no null values, particularly in the Date field.
- **Table Creation:**
  - Create a new table (SPACEXTABLE) for clean analysis.
- **Exploratory Data Analysis (EDA)**
- **SQL Queries:**
  - Conduct queries to answer specific questions:
    - Unique launch sites
    - Total payload masses by customer
    - Average payload mass by booster version
- **Interactive Visual Analytics**
- **Visualization Tools:**
  - Use libraries such as Matplotlib or Seaborn to create visual representations of the data.
- **Insights:**
  - Visualize trends in launch success rates.
  - Analyze the impact of payload mass on landing outcomes.
- **Conclusion**
- **Goal:**
  - Utilize findings to inform strategic decisions for competing in the commercial space sector.

Plotly Dash


Get the tool ready
Automobile Sales Statistics Dashboard
Yearly Statistics
1980
Yearly Automobile Sales
Total Monthly Automobile Sales
Average Vehicles Sold by Vehicle Type in the year 1980
Total Advertisement Expenditure for Each Vehicle in 1980


File Edit View Run Kernel Tabs Settings Help
DV0101EN-Final-Assign-Part
Markdown
Python (Pyodide)
plt.pie(sizes, labels=labels, autopct='%1.1f%%', startangle=90)
plt.show()
Executivecar
Supperminicar
Mediumfamilycar
Sports
Smallfamiliycar
0.8%
27.9%
38.7%
0.7%
32.0%
Click here for Solution template

Interactive map with Folium



A highway map symbol may look like this:

A city map symbol may look like this:

Predictive analysis

* **Data Preprocessing:** Handling missing values, outliers, and feature scaling/normalization.

* **Feature Engineering:** Creating new features based on existing data (e.g., launch time, launch window duration).

* **Model Selection:** Exploring various machine learning algorithms:
  * **Logistic Regression**.
  * **Decision Trees**.
  * **Support Vector Machine (SVM)**.

* **Hyperparameter Tuning:**
* Optimizing model parameters using techniques like GridSearchCV **Metrics:**

* **Results:**
* **Model Performance:**
  * **Logistic Regression:** Achieved an accuracy of 75%
  * **Decision Trees:** Achieved an accuracy of 78%
  * **SVM:** Achieved an accuracy of 83%

Conclusion

* **Transformative Era**: The Commercial Space Age is reshaping space exploration, led by innovative companies like Virgin Galactic, Rocket Lab, Blue Origin, and SpaceX.
* **SpaceX's Impact**:
  * Pioneered cost-effective Falcon 9 rocket launches through first-stage reusability.
  * Established a competitive pricing model that enhances accessibility to space.
* **Role of Data Science**:
  * Critical for analyzing SpaceX's launch data to inform strategies for Space Y.
  * Utilized data preprocessing, feature engineering, and model selection to enhance predictive capabilities.
* **Model Performance**:
  * Logistic Regression: 75% accuracy.
  * Decision Trees: 78% accuracy.
  * Support Vector Machine (SVM): 83% accuracy.
  * Highlights the effectiveness of machine learning in forecasting first-stage landing success.
* **Future Implications**:
  * Insights gained will inform Space Y's strategy and contribute to the commercial space sector's evolution.
  * Emphasizes the importance of collaboration, innovation, and knowledge in advancing space exploration.
* As we look to the future, the combination of technology and data-driven insights will propel the next generation of space pioneers into new frontiers.