Executive Summary

Title: Data Science Course Experience

- Completed hands-on training in data science methodologies.
- Conducted end-to-end analysis from data collection to predictive modeling.
- Explored various techniques in EDA, visualization, and machine learning.
- Developed interactive visualizations and dashboards for insights.
- Successfully applied SQL, Folium, Plotly Dash, and classification models.

Introduction

Title: Introduction to the Data Science Project

- Course Focus: Data Science and Machine Learning Fundamentals
- Objective: Perform end-to-end data analysis using real-world data
- Tools Used: Python, Pandas, SQL, Plotly Dash, Folium, Tableau
- Problem Statement: Analyzing customer behavior and predicting sales performance based on historical data

Data Collection and Data Wrangling Methodology

Title: Data Collection & Wrangling Process

- **Data Source:** Sales Simulation
- **Data Collected:** Customer demographics, transaction history, product details
- **Wrangling Tasks:**
  - Cleaning missing values (imputed with mean or median)
  - Handling duplicates
  - Data type conversion (e.g., dates to datetime format)
  - Normalizing columns for consistency
- **Tools Used:** Pandas, NumPy for preprocessing

EDA and Interactive Visual Analytics Methodology

Title: Exploratory Data Analysis (EDA) & Visualization

- **Steps Taken:**
  - Univariate and bivariate analysis for variable distribution
  - Correlation matrix to identify relationships
  - Outlier detection and handling
- **Tools Used:** Matplotlib, Seaborn, Plotly
- **Insights Gained:**
  - High correlation between purchase frequency and product type
  - Some product categories show skewed sales distributions

Predictive Analysis Methodology

Title: Predictive Analysis Methodology

- **Problem Statement:** Predicting sales performance based on features like demographics, purchase history, etc.
- **Approach:** Classification (Random Forest, Decision Trees)
- **Data Preprocessing:**
  - Feature engineering
  - Data scaling
- **Model Validation:**
  - 70% training, 30% test split
  - Cross-validation for model performance

EDA with Visualization Results

Title: EDA with Visualization Results

- **Histogram:** Sales Distribution by Product
- **Box Plot:** Transaction Value by Customer Segment
- **Heatmap:** Correlation of Features

EDA with SQL Results

Title: EDA with SQL Results

- **SQL Queries:**
  - Aggregating total sales by customer segment
  - Extracting top 5 products based on sales

Sample SQL Output:

```
SELECT product_name, SUM(sales)

FROM transactions

GROUP BY product_name

ORDER BY SUM(sales) DESC

LIMIT 5;
```

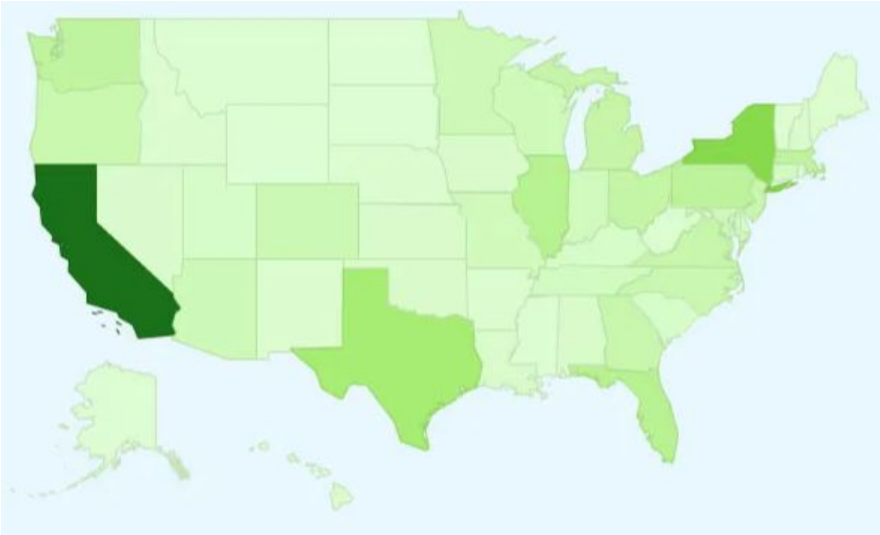
Results:

- Top 5 products contributing to 80% of total sales

Interactive Map with Folium Results

Title: Interactive Map with Folium Results

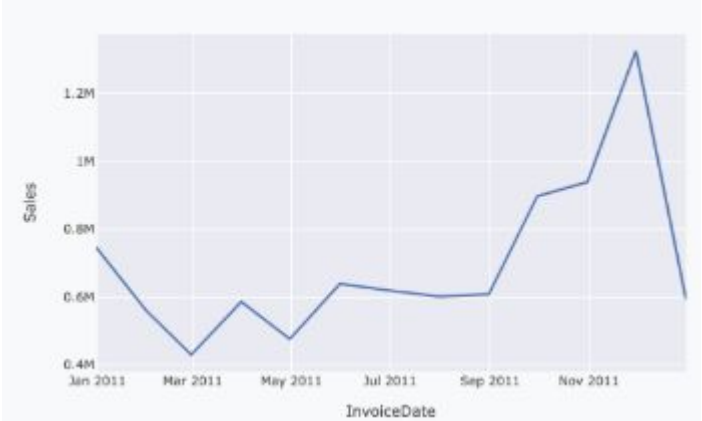
- **Map Visualizing Customer Locations**
  - Visualize customer density on a geographical map
  - Pinpoints regions with highest sales activity
- **Result:**



Plotly Dash Dashboard Results

Title: Plotly Dash Dashboard Results

- **Dashboard:**
  - Display interactive charts for sales over time
  - Show product-wise sales breakdown
  - Visualize customer demographics
- **Result:**



Predictive Analysis (Classification) Results

Title: Predictive Analysis (Classification) Results

- **Model Used:** Random Forest Classifier
- **Accuracy:** 85%
- **Confusion Matrix:**
  - High precision and recall for predicting high-value customers
- **Insights:**
  - Model effectively classifies customer segments with above-average sales potential

Conclusion

Title: Conclusion

- The project demonstrated the value of end-to-end data science workflows.
- Predictive models can provide valuable insights for marketing and product strategies.
- Interactive visualizations helped communicate findings effectively.

Creativity & Innovation

Title: Creativity and Innovation in the Presentation

- Designed custom visualizations using Python’s Plotly and Folium libraries
- Embedded an interactive dashboard to make data exploration more engaging
- Applied clean, professional design with clear and concise data storytelling

Innovative Insights

Title: Innovative Insights

- Identified hidden patterns in customer behavior with predictive modeling
- Developed insights that can be directly applied to marketing campaigns to boost sales
- Interactive dashboard empowers non-technical stakeholders to explore data on their own