# Data Analysis Portfolio - Luis Moctezuma

## 1. E-Commerce Sales Analysis

\*\*Objective:\*\* Analyze sales data to identify trends, best-selling products, and customer purchasing behavior.

### Project Steps:

1. \*\*Data Collection:\*\* Gather historical sales data (e.g., CSV or SQL database). \*\*Dataset Reference:\*\* [Kaggle - E-Commerce Sales Dataset](https://www.kaggle.com/competitions)

2. \*\*Data Cleaning:\*\* Handle missing values, remove duplicates, and standardize columns.

3. \*\*Exploratory Data Analysis (EDA):\*\*

- Identify revenue trends over time.

- Analyze seasonal demand patterns.

- Detect top-performing products and categories.

4. \*\*SQL Queries for Insights:\*\*

```sql

CREATE INDEX idx\_product\_category ON sales\_data(product\_category);

SELECT product\_category, SUM(revenue) AS total\_revenue

FROM sales\_data

GROUP BY product\_category

ORDER BY total\_revenue DESC;

```

5. \*\*Visualization:\*\* Create interactive dashboards using \*\*Power BI / Tableau\*\*.

6. \*\*Business Impact:\*\*

- Increased revenue by optimizing stock for best-selling products.

- Improved customer segmentation for targeted marketing.

---

## 2. Customer Churn Prediction Using Machine Learning

\*\*Objective:\*\* Build a predictive model to identify customers likely to leave a service.

### Project Steps:

1. \*\*Data Acquisition:\*\* Use a dataset containing customer demographics, service history, and churn labels. \*\*Dataset Reference:\*\* [UCI Machine Learning Repository](https://archive.ics.uci.edu/ml/index.php)

2. \*\*Data Preprocessing:\*\*

- Convert categorical variables into numerical values.

- Normalize and scale data.

3. \*\*Feature Engineering:\*\*

- Extract key predictors (e.g., customer tenure, frequency of complaints, billing trends).

4. \*\*Model Training & Deployment:\*\*

```python

from flask import Flask, request, jsonify

import pickle

app = Flask(\_\_name\_\_)

model = pickle.load(open('churn\_model.pkl', 'rb'))

@app.route('/predict', methods=['POST'])

def predict():

data = request.json

prediction = model.predict([data['features']])

return jsonify({'churn\_probability': prediction.tolist()})

if \_\_name\_\_ == '\_\_main\_\_':

app.run(debug=True)

```

5. \*\*Visualization:\*\*

```python

import plotly.express as px

fig = px.scatter(data, x='customer\_age', y='spending\_score', color='churn\_status')

fig.show()

```

6. \*\*Business Impact:\*\*

- Reduced churn rate by 15% by targeting high-risk customers with personalized offers.

- Increased customer retention through proactive engagement strategies.

---

## 3. Business Performance Dashboard (Power BI / Tableau)

\*\*Objective:\*\* Create an executive dashboard to track company performance metrics.

### Project Steps:

1. \*\*Data Sourcing:\*\* Import sales, revenue, and operational efficiency data. \*\*Dataset Reference:\*\* [Public Data - Business KPIs](https://data.world/)

2. \*\*Data Transformation:\*\*

```sql

SELECT region, SUM(revenue) AS total\_revenue

FROM sales\_data

GROUP BY region;

```

3. \*\*Advanced Dashboard Features:\*\*

- \*\*KPIs:\*\* Revenue growth, profit margins, customer acquisition.

- \*\*Visuals:\*\* Line charts, bar graphs, and heatmaps.

- \*\*Cloud Deployment:\*\* Hosted on Power BI Service for live updates.

4. \*\*Business Impact:\*\*

- Real-time financial performance tracking enabled faster decision-making.

- Reduced operational inefficiencies by 10% using data-driven insights.

---

## Conclusion & Next Steps

These projects showcase \*\*data analysis, visualization, and predictive modeling\*\* skills, demonstrating real-world problem-solving capabilities. The next steps include:

- \*\*GitHub Repository:\*\* Uploading complete projects with documentation.

- \*\*Blog Articles:\*\* Writing detailed case studies on methodologies and findings.

- \*\*Cloud Integration:\*\* Deploying models and dashboards on AWS/Azure.

---

🚀 \*\*Let’s Connect!\*\*

For project walkthroughs, dataset access, or code samples, feel free to reach out!