Analysis of Electricity Consumption in Maryland

Unraveling Sales, Revenue, and Prices (2010-2022)

Data 606 Capstone in Data Science

Under the guidance of Professor Dr. Chaojie (Jay) Wang

Alisha Minj – WG56858

Introduction

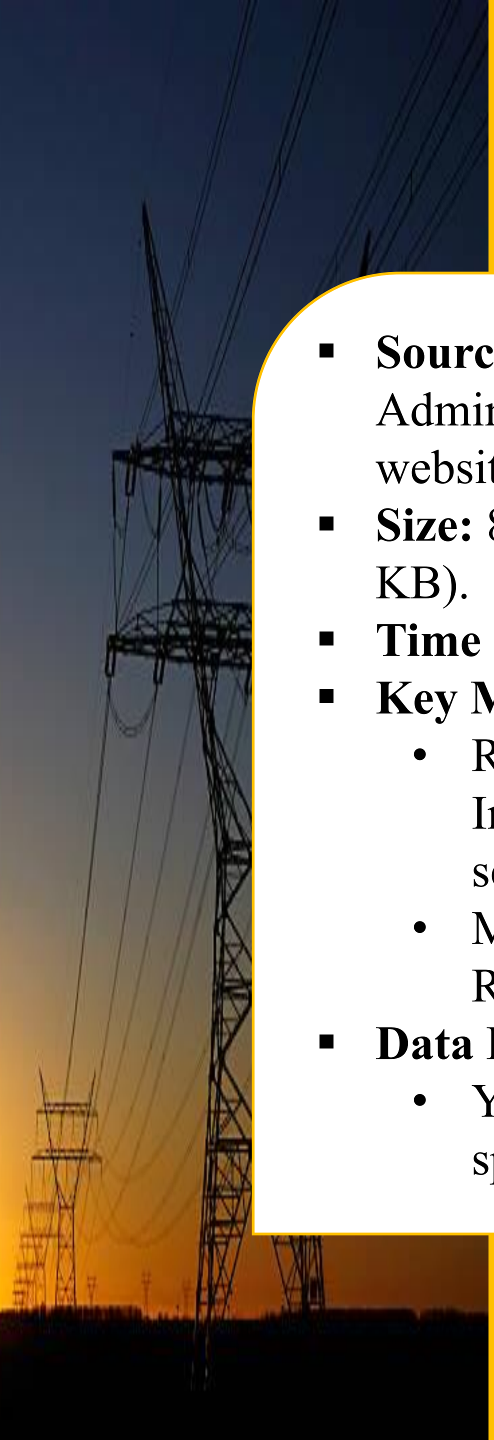
Purpose: Delve into the intricacies of Maryland's electricity consumption patterns from 2010 to 2022

Economic Insights: Understanding how electricity consumption aligns with economic activities.

Policy Decisions: Guiding policymakers in framing energy policies, infrastructure development, and sustainability initiatives.

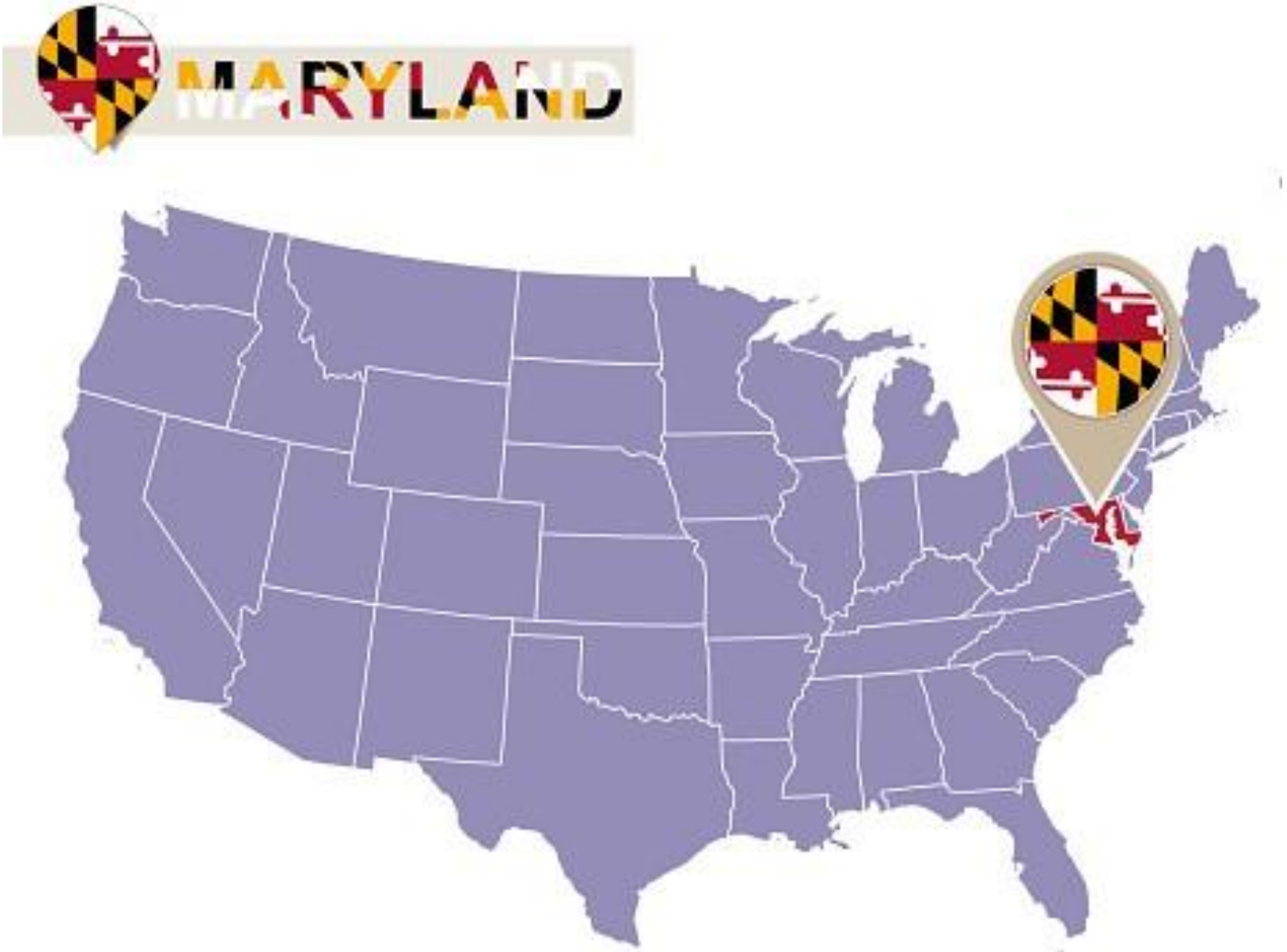
Future Planning: Assisting utility companies and state planners in forecasting future demand for a sustainable energy future.

Project



Data Overview

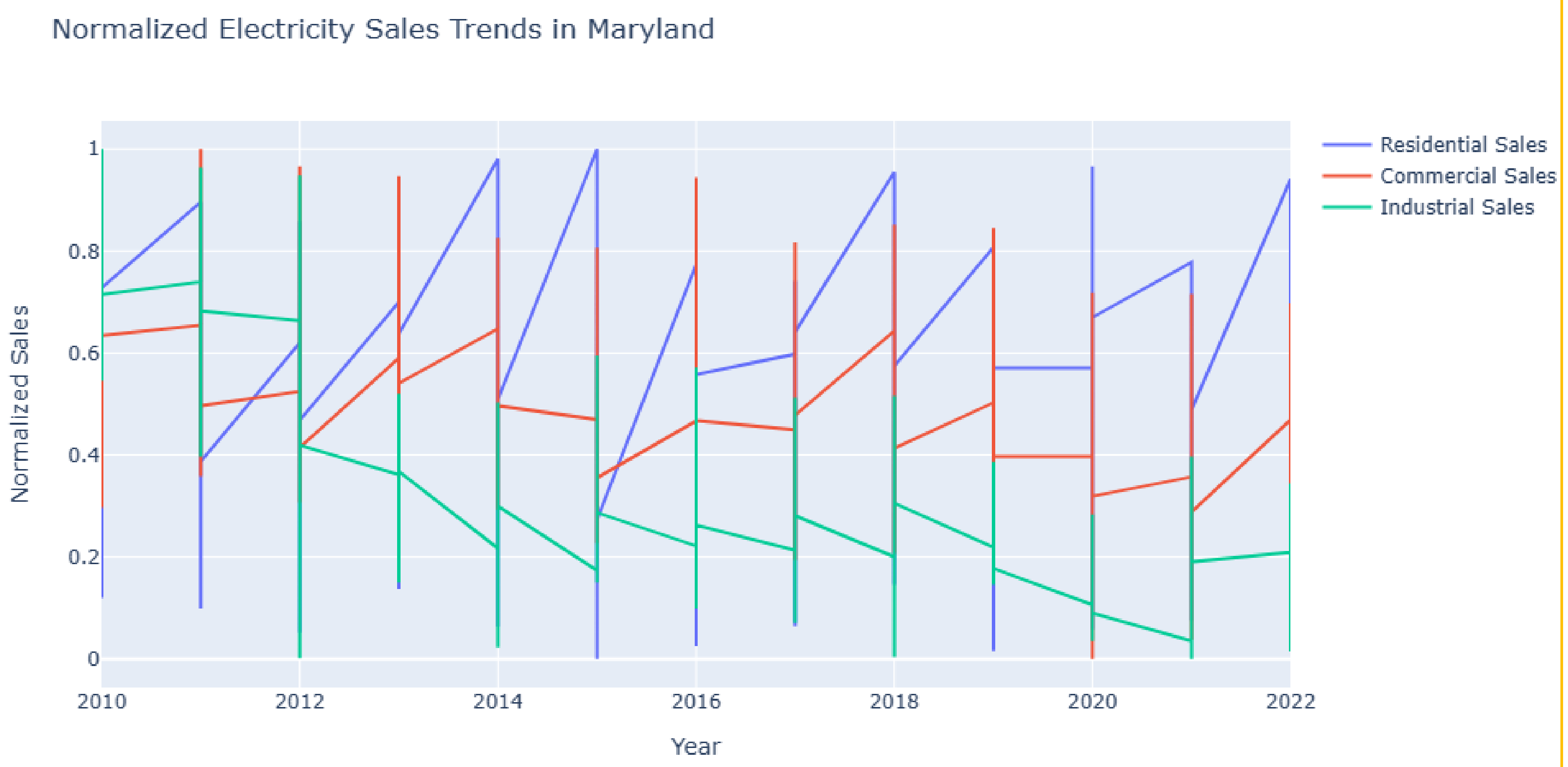
- **Source:** U.S. Energy Information Administration (EIA) official website.
- **Size:** 8,313 rows, 24 columns (1655 KB).
- **Time Period:** 2010 to 2022.
- **Key Metrics:**
  - Residential, Commercial, Industrial, Transportation sectors.
  - Metrics include Sales, Revenue, Customers, Price.
- **Data Dictionary:**
  - Year, Month, State, and sector-specific details.



Project

Project

Overall Electricity Sales Trend



Key Points

General Increase in Residential Consumption

- **Population Growth**: The rise in residential consumption aligns with the growing population.
- **Housing Expansion:** Increased housing developments contribute to higher electricity demand in residential areas.
- **Widespread Device Usage**: The pervasive use of electronic devices in households is a significant driver.

Fluctuations in Commercial and Industrial Sectors:

- **Economic Factors**: Commercial electricity consumption shows fluctuations influenced by economic conditions and business activities.
- **Seasonal Changes**: Both commercial and industrial sectors exhibit variations related to seasonal demands and operational changes.
- **Technological Advancements**: The industrial sector experiences fluctuations driven by technological advancements and operational adjustments.

Project

Seasonal Patterns in Electricity Consumption

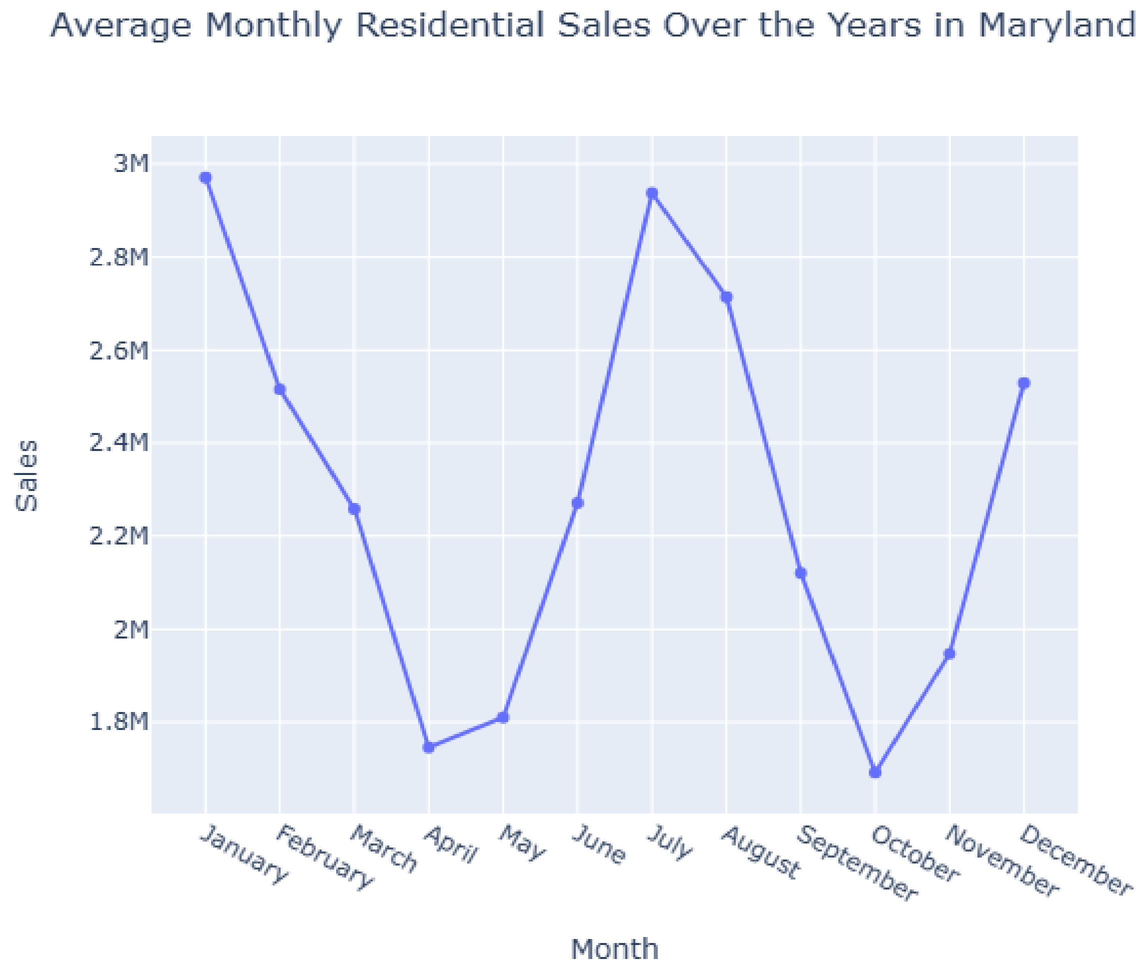
Key Points

Summer Peaks and Winter Lulls:

- Increased Air Conditioning Use in Summer
- Winter Months Reflecting Lower Consumption due to Heating Systems

Impact on Resource Planning:

- Effective Resource Allocation
- Stable Supply during Peak Seasons
- Optimization of Energy Generation to Match Demand Variations



Project

Bivariate Analysis

Sales vs. Revenue

Sales vs. Price Dynamics

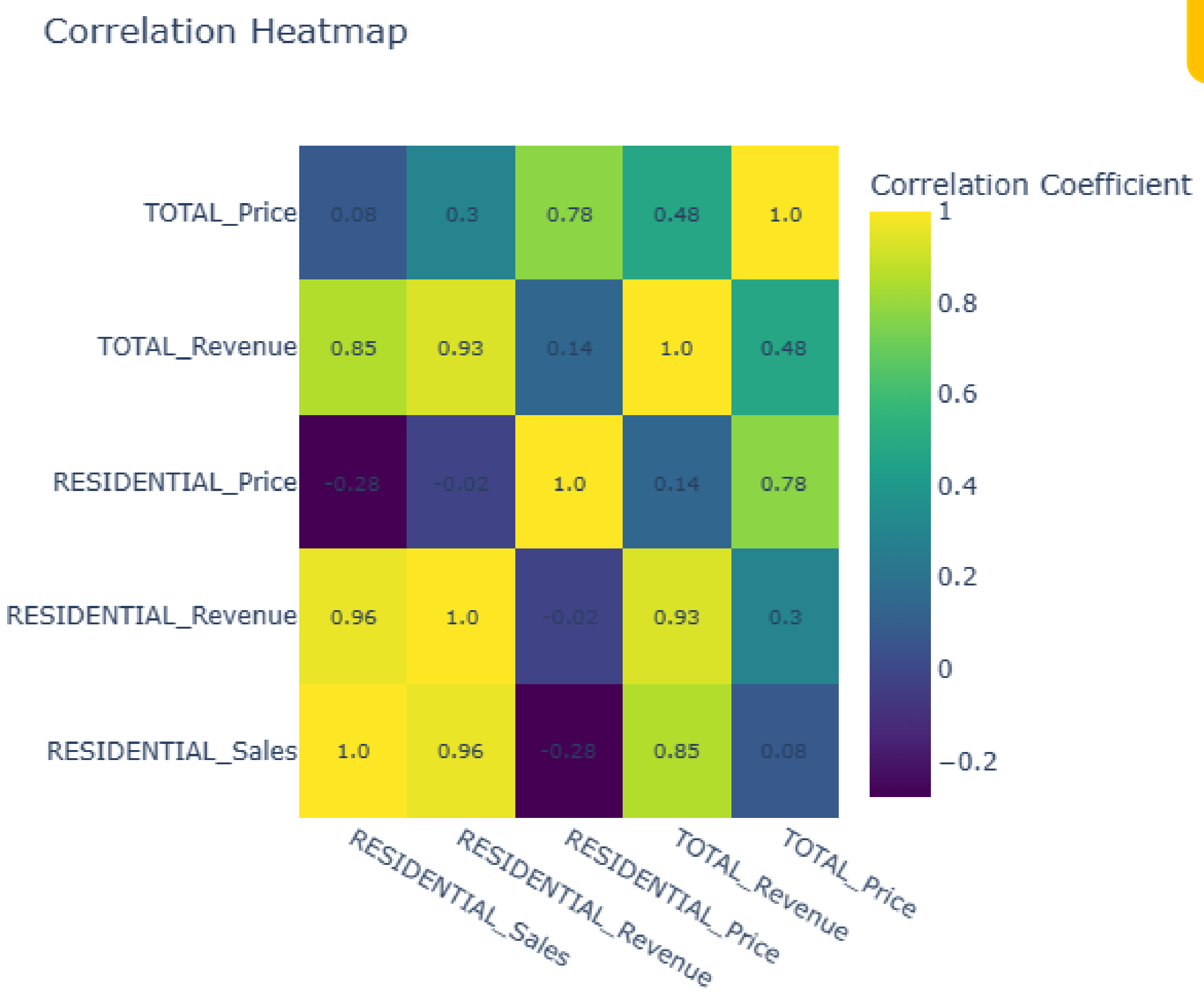
Key Points

- Revenue varies from 225,123 to 472,807.
- Sales show fluctuations ranging from 1,536,738 to 3,316,855.
- General increasing trend observed in both revenue and sales over time.

- Prices vary from 12.22 to 15.89.
- Sales values exhibit an inverse relationship with prices, indicating higher prices correspond to lower sales.
- Some fluctuation observed in prices over time.

Project

Correlation Heatmap



Key Points

Positive Correlations:

- Strong positive correlations observed between several predictors.

TOTAL_Revenue and TOTAL_Price:

- Notably high correlation (0.85).
- Implies a consistent pricing strategy across sectors.

RESIDENTIAL_Sales:

- Significant positive relationships with RESIDENTIAL_Revenue and RESIDENTIAL_Price.
- Suggests residential sales influenced by both price and revenue dynamics.

Feature Engineering Insights

Temporal Features Extraction:

Month and Year Analysis:

- Revealed distinct seasonality trends.
- Identified peak electricity consumption months, aiding future demand predictions.

Lag Features and Rolling Statistics for RESIDENTIAL_Sales:

Lag Features Impact:

- Found correlations between past residential sales and the current month's performance.

Rolling Statistics Significance:

- Smoothed out fluctuations, highlighting long-term consumption patterns.

Outcome:

- Improved predictive accuracy, enabling proactive energy resource planning.

Project

Data Splitting

- **Utilized Standard Split**: We partitioned the dataset into training (80%) and testing (20%) sets, a common practice for robust model development.
- **Randomized Selection**: Implemented randomized selection to eliminate bias, ensuring a fair representation of patterns in both sets.
- **Enhancing Generalization**: This approach enhances the model's adaptability to new, unseen data by preventing overfitting to specific training set patterns.
- **Balanced Approach**: The 80/20 split strikes a balance between providing ample data for training, allowing the model to learn historical patterns, and reserving a portion for testing to evaluate performance on unseen data.

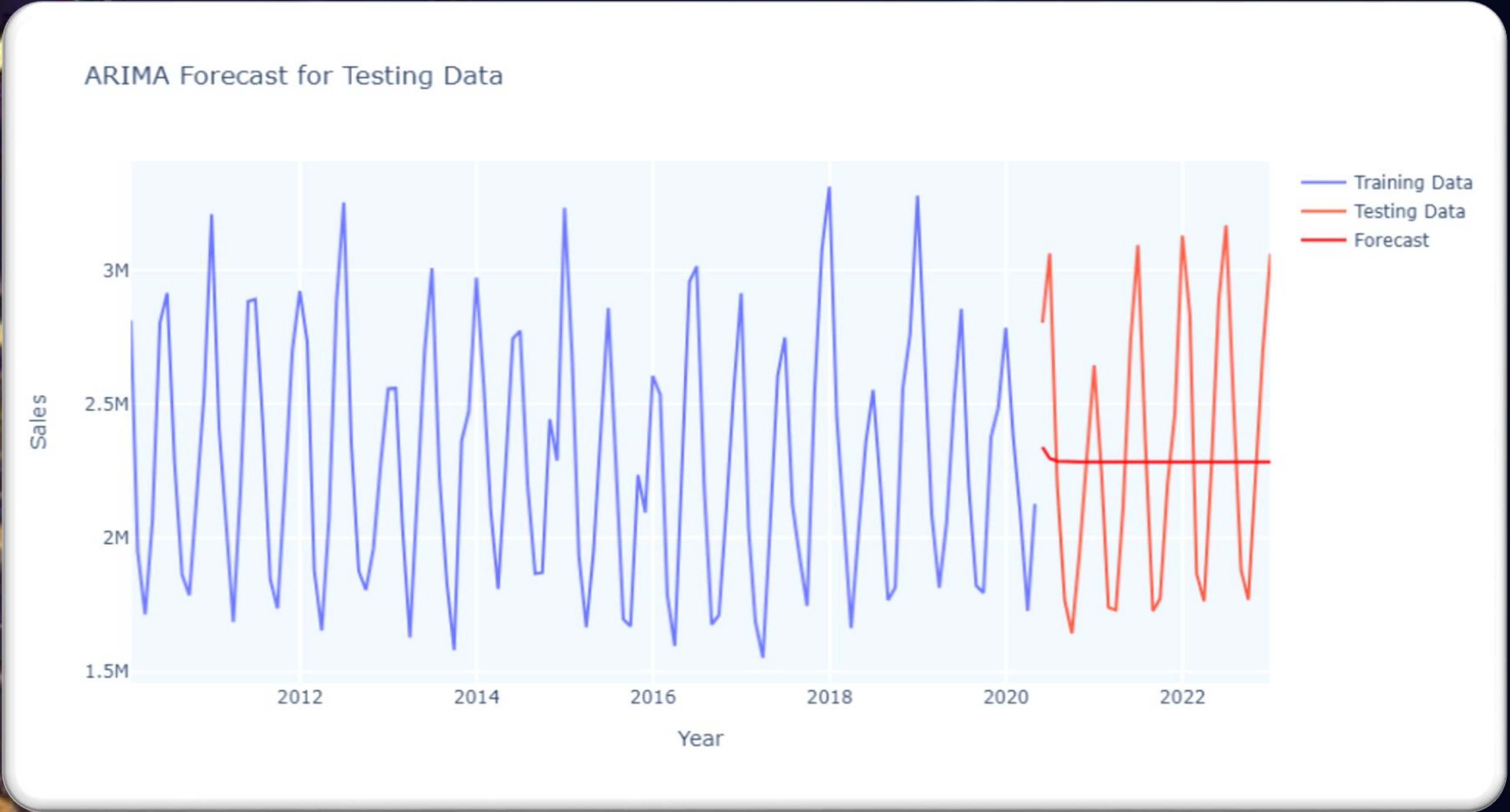
Project

Unveiling Tomorrow: Time Series Forecasting for Electricity Sales

Project

ARIMA

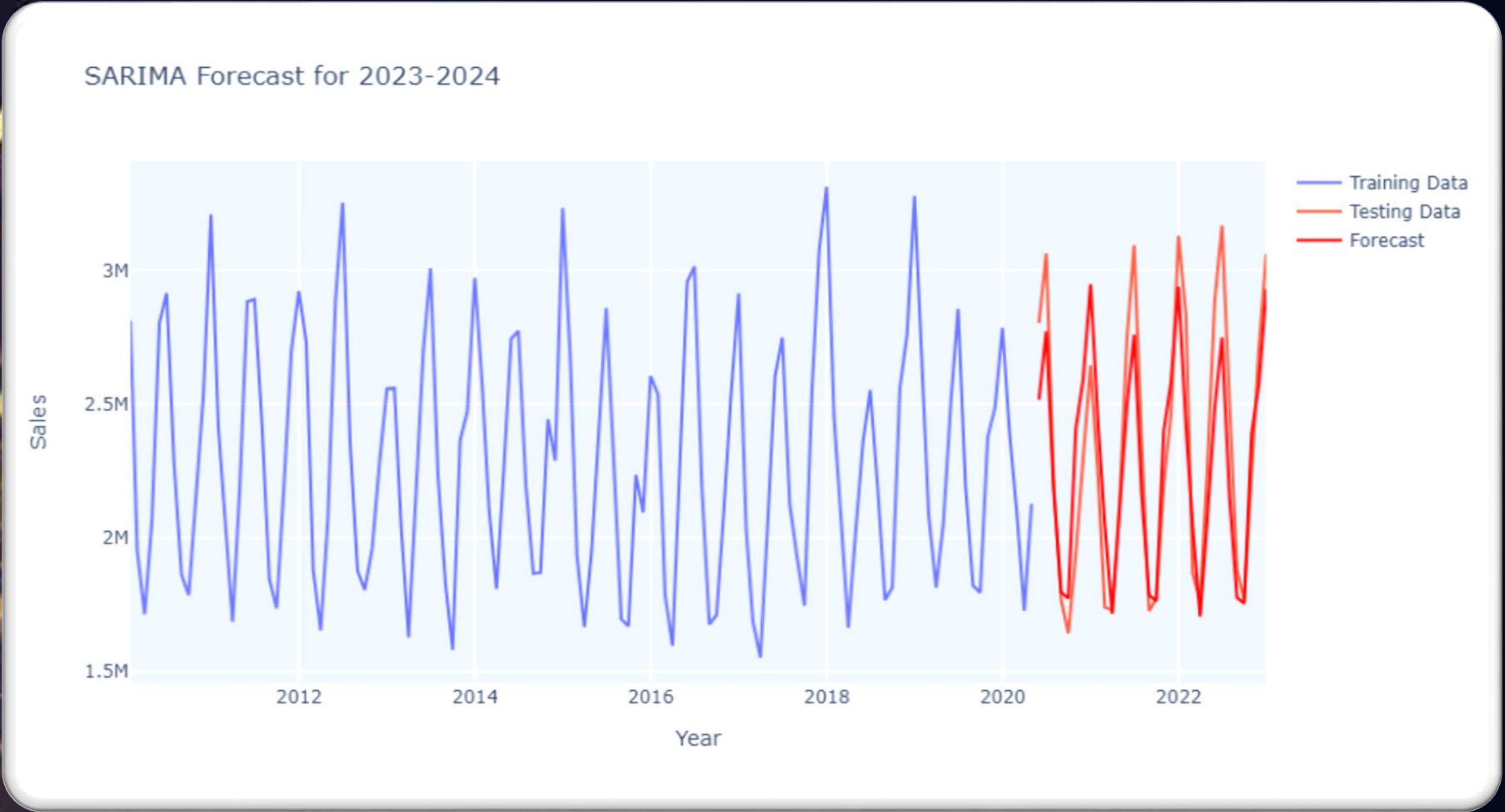
- **Mean Absolute Error (MAE):** 418,794.74
- **Mean Squared Error (MSE):** 241,358,590,866.97
- **Root Mean Squared Error (RMSE):** 491,282.60
- **R-squared (R2):** 0.00



Project

SARIMA

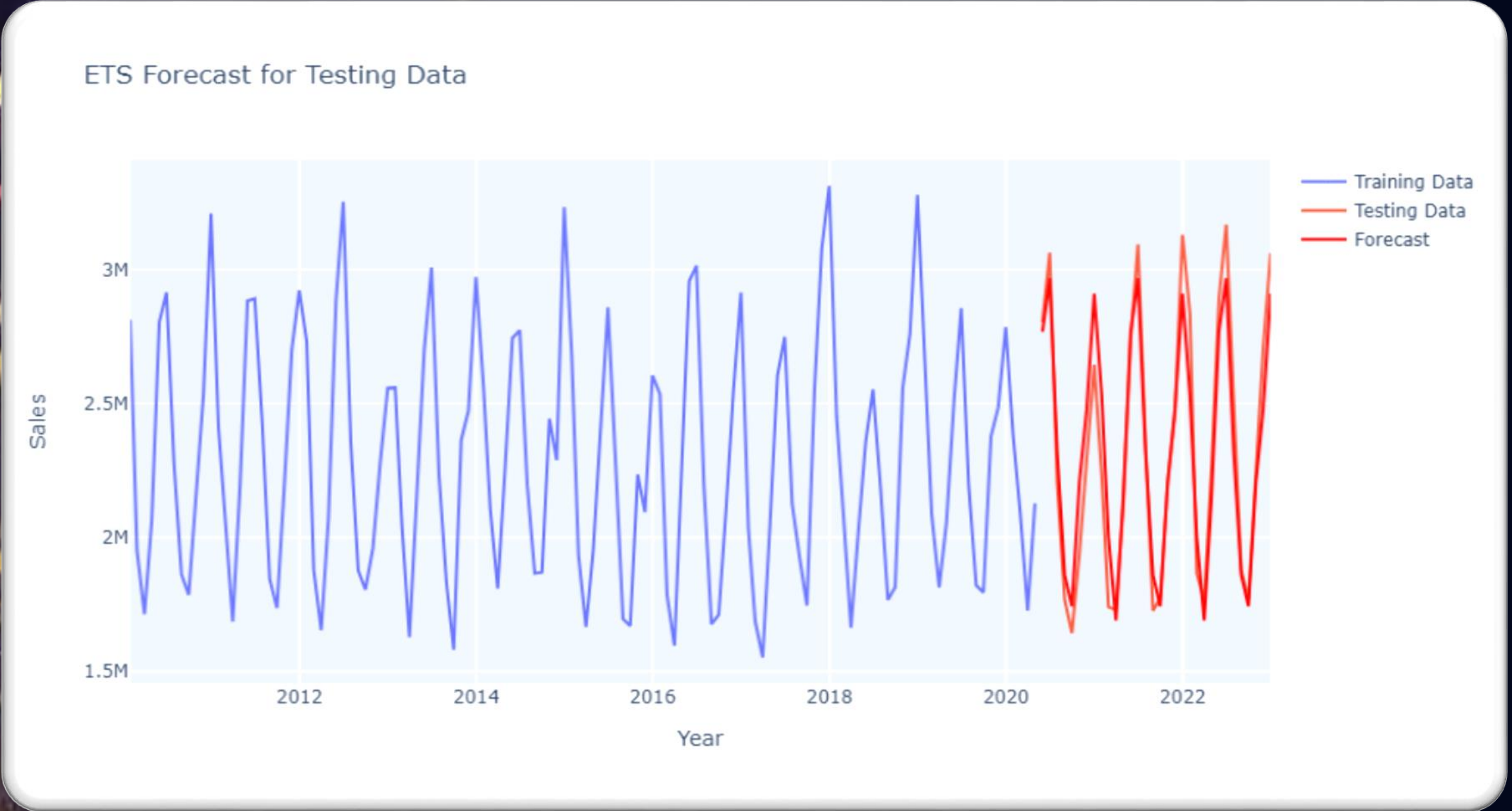
- **Mean Absolute Error (MAE):** 195,646.46
- **Mean Squared Error (MSE):** 55,809,529,344.26
- **Root Mean Squared Error (RMSE):** 236,240.41
- **R-squared (R2):** 0.77



Project

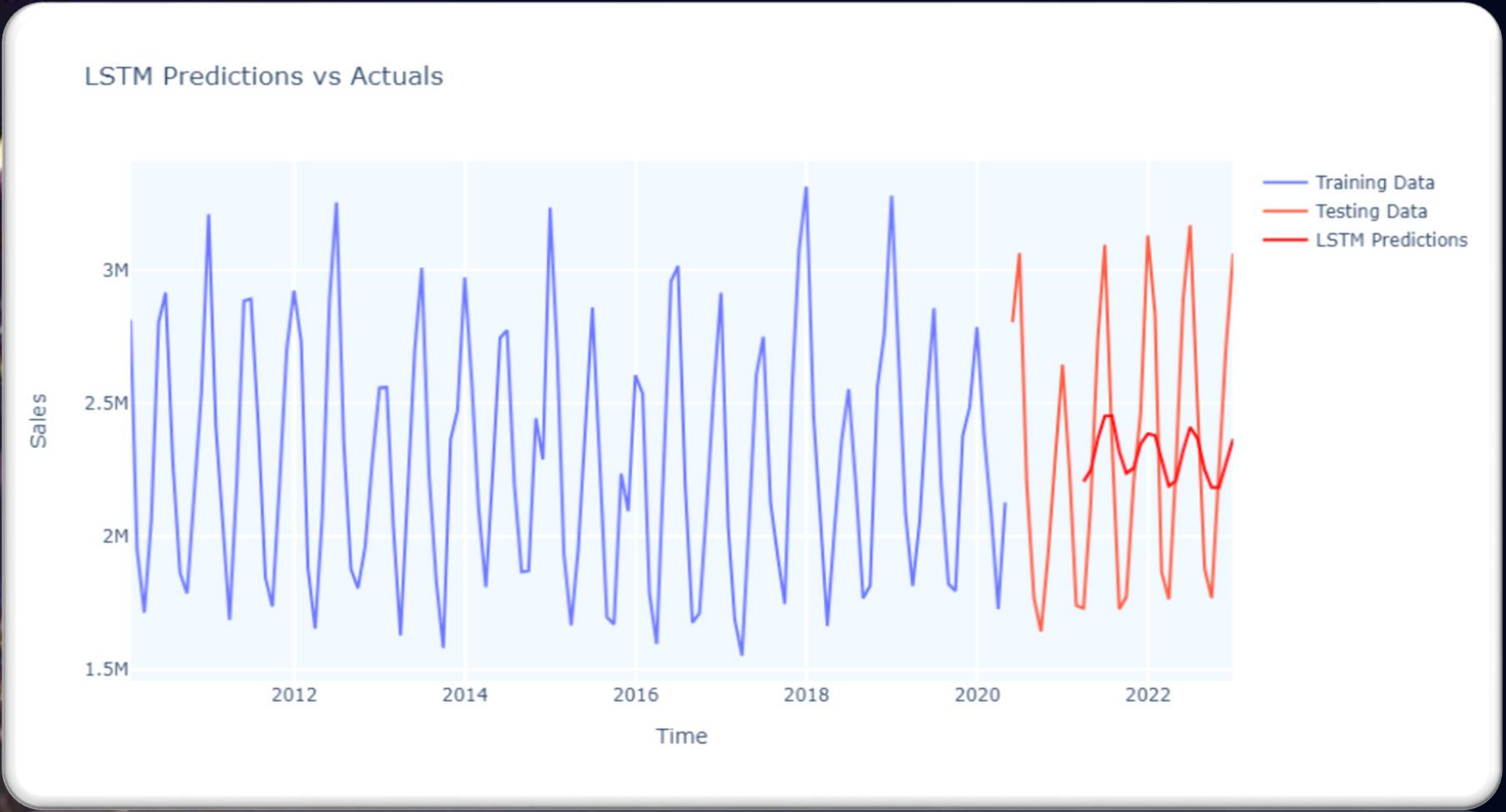
ETS Model

- **Mean Absolute Error (MAE):** 124,074.09
- **Mean Squared Error (MSE):** 23,632,254,389.68
- **Root Mean Squared Error (RMSE):** 153,727.86
- **R-squared (R2):** 0.90



Project

LSTM Model



- **Mean Absolute Error (MAE):** 385,076.84
- **Mean Squared Error (MSE):** 200,148,281,997.56
- **Root Mean Squared Error (RMSE):** 447,379.35
- **R-squared (R2):** 0.20

Project

Conclusion

Insights Unveiled

Deep insights gained into Maryland's electricity consumption trends.

Residential Surge

Increased residential consumption driven by population growth and tech reliance.

ETS Model Accuracy

ETS model proved most accurate, a robust predictor of consumption trends.

Reliable Models

Our approach ensures model reliability for real-world adaptability.

Practical Utility

Balanced methodology enhances models' practical utility in diverse scenarios.

Future Exploration

Findings pave the way for ongoing exploration and optimization.

THANK YOU