Project Portfolio

# CAR SALES DEMAND FORECASTING

IDENTIFYING KEY BUSINESS INSIGHTS AND PREDICTING MONTHLY CAR SALES USING ARIMA & PROPHET MODELS

Presented by:
**Trika Yuliana**

Date:
**October, 15 2025**

# Outline

## Executive Summary
## Project Overview
## Methodology
## Exploratory Data Analysis (EDA)

1. Customer and Market Insight
2. Product and Brand
3. Sales Trend

## Forecasting Analysis

1. ARIMA
2. PROPHET

## Model Evaluation

MAE, RMSE, MAPE

# Executive Summary

**SUV and Hatchbacks are the most preferred body styles represented by car models Prizm, Diamante, and Silhoutte**

**Key market region: Austin and Janesville**

**Customer purchasing behavior are slightly higher on male buyers compared to female**

**Prophet model providing more accuracy for seasonal-trends product such as LHS, A6, Ram Pickup, LS400, meanwhile ARIMA excelled for more stable demand products like Jetta, Prizm, Passat, Diamante, RL, Silhouette**

**The 6-month demand forecast projects a total of 1,382 unit.**

# Project Overview

**Project Description**
This project explore key business insights regarding car sales performance across regions and forecast future demand using predictive analysis.

**Problem Addressed**
1. Perform exploratory analysis to understand sales patterns across demographic and regional segments
2. Apply time series forecasting models (ARIMA and Prophet) to predict monthly car sales
3. Evaluate and compare model accuracy to demonstrate forecasting proficiency and practical application in demand planning

**Research Study**
1. What are the key business insights to optimize car sales in upcoming months?
2. Which model provides better predictive accuracy for demand planning purposes?
3. What is the demand volume for the next 6-months to support planning allocation?

# METHODOLOGY

## Data Collection

1. The dataset was obtained through Kaggle provided by Vasu Avasthi (Licensed under Apache 2.0). The final dataset contained 23.960 rows and 17 columns.
2. Key features: Price, Annual Income, Dealer Name, Company, Model, Body Style, Dealer Region, and Gender.

## Data Wrangling

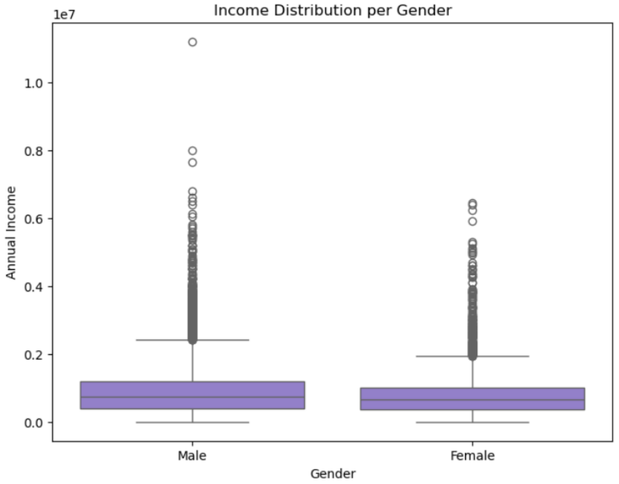
1. Data formatting (converted date column to datetime type, standardized column names and formats for consistency (numeric, categorical), checked and corrected data types for all features)
2. Missing values: verified 0 missing values
3. Duplicates values: 0 duplicate records found

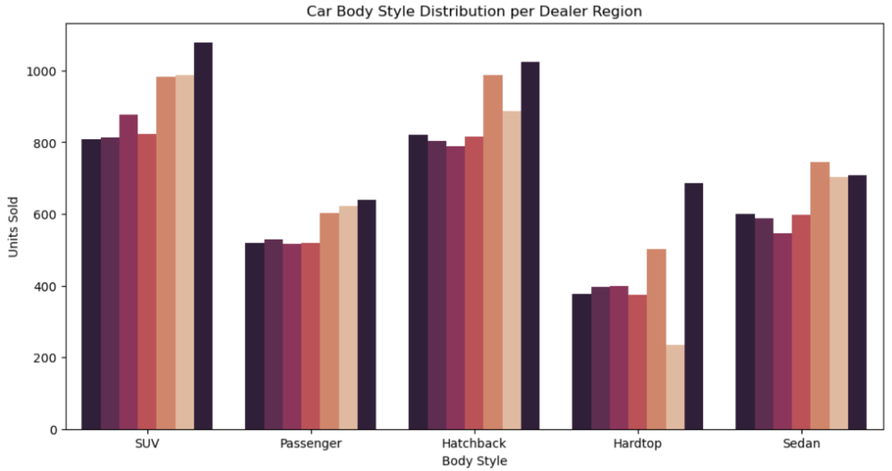
## Data Analysis

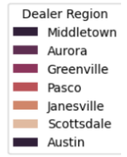
1. Utilized matplotlib and seaborn for visualization (boxplot, bar chart, line chart, stacked bar) to identify key business insights.
2. Applied ARIMA Model using statsmodels, pmdarima, with following steps:
   - Selected the input only for Top 10 car models by highest cumulative sales
   - Analyzed monthly trend series per model
   - Applied ADF (Augmented Dickey-Fuller) Test to check stationarity. Initial p-value = 0.12 (data is non-stationer), after first-order differencing, p-value= 1.26e-16 (stationarity achieved)
   - Analyzed PACF and ACF plots to determine optimal p,d,q parameters. Result: selected (p,d,q) = (4,1,2) as best performing order
   - Trained model per car model and generated 6 month ahead forecast
3. Apply Prophet Model, with following steps:
   - Selected the input only for Top 10 car models by highest cumulative sales
   - Prepared dataset in Prophet format (df, y) with monthly frequency (MS)
   - Trained model per car model and generated 6 month ahead forecast
4. Applied model evaluation using sklearn.metrics (MAE, RMSE, MAPE)
5. Conducted comparative analysis to select the best model between ARIMA and Prophet for car demand planning

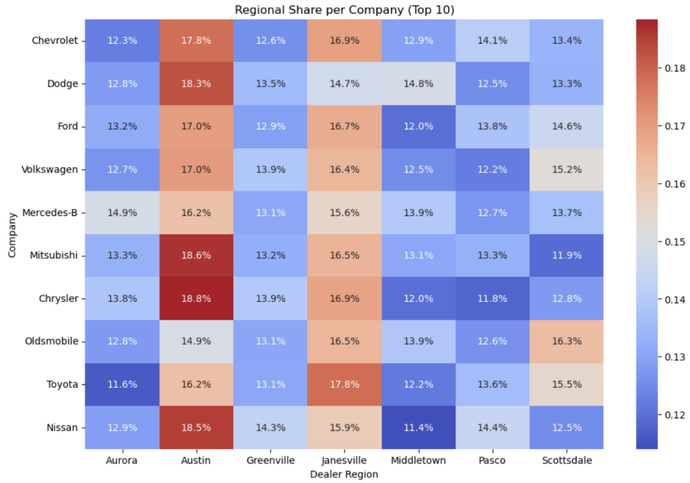
# EXPLORATORY DATA ANALYSIS

## CUSTOMER AND MARKET INSIGHTS



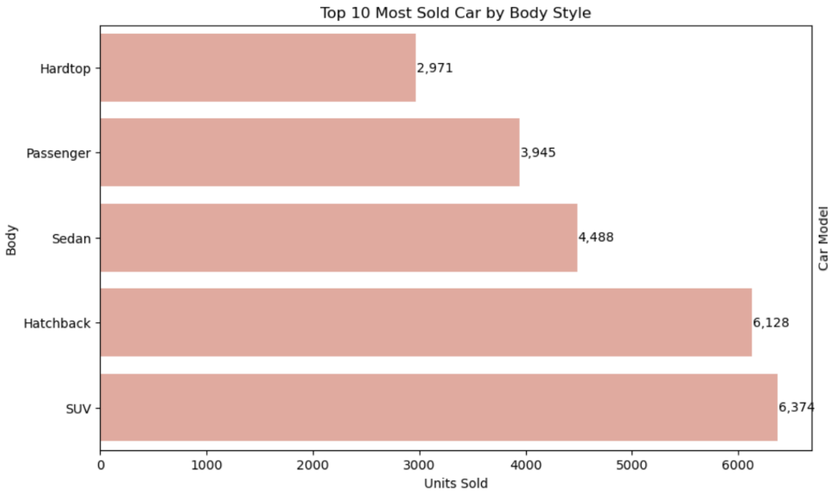



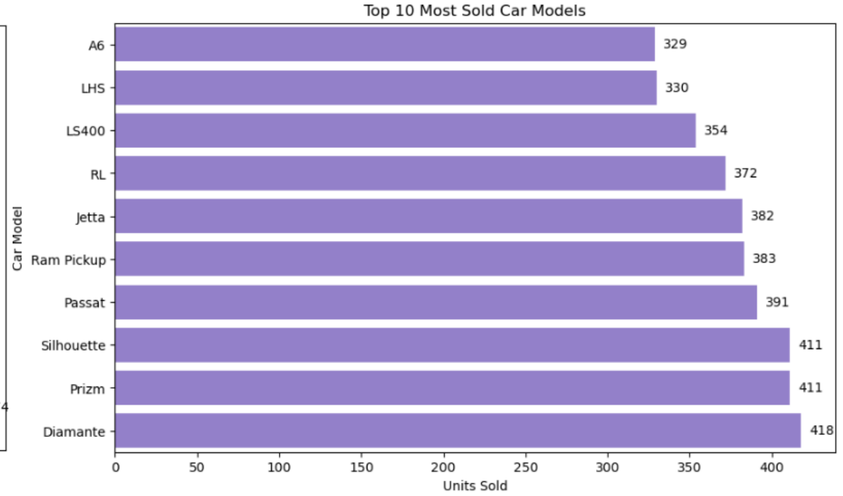

Insights:

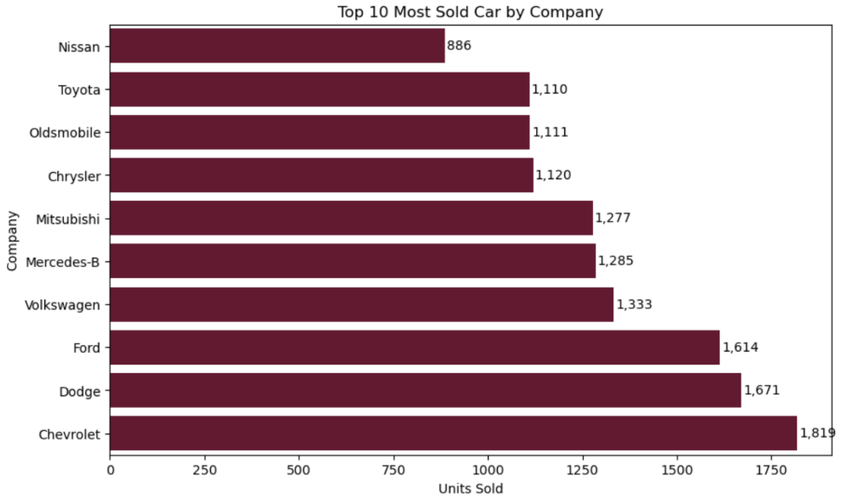
1. Male customers tend to have slightly higher purchasing power than female
2. SUV and Hatchback are the most sold by body style across regions, showing preference for versatile and family-oriented vehicles
3. Austin has the largest customer base across all income segments, indicates the strongest sales region for volume and purchasing power (suggesting promotion focus on brand awareness and loyalty). Janesvilles ranks second, good for products diversification (focus promotion: value for money and financing offers). Other regions show smaller but balanced distribution, possibly for supporting markets or dealer expansion.
4. Market share distribution spread across the region, with Austin and Janesville shares around 15-18%. This indicates a balanced and competitive landscape where no single brands hold dominant position.

# EXPLORATORY DATA ANALYSIS

## PRODUCTS AND BRANDS



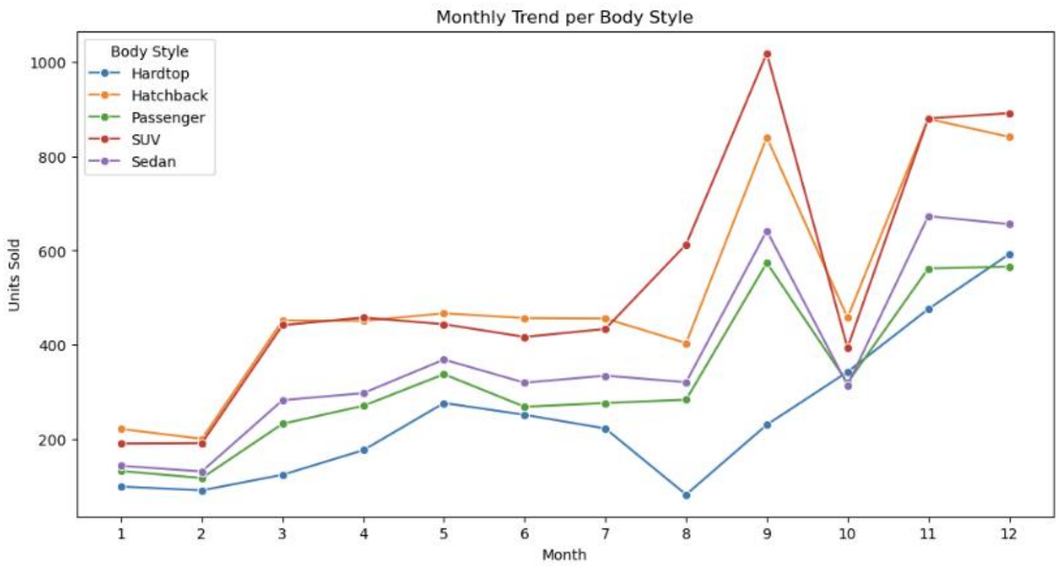



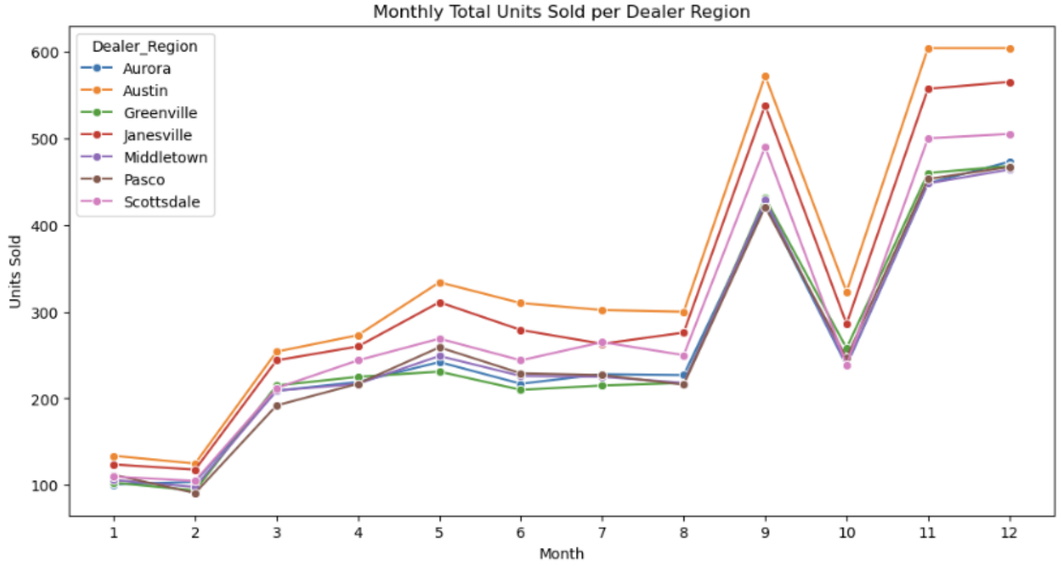

Insights:

1. SUV and Hatchback are the top-selling by body styles, showing customer preference for versatile and family-oriented vehicles
2. At the model level, Prizm and Diamante maintain consistent demand for SUV and Hatchback segments, followed by Silhoutte as a Sedan type, reflecting a balance between practicality and affordability across segments
3. In terms of brand performance, Ford, Dodge, and Chevrolet lead total units sold, indicating brand awareness and customer loyalty

# EXPLORATORY DATA ANALYSIS

## SALES TREND

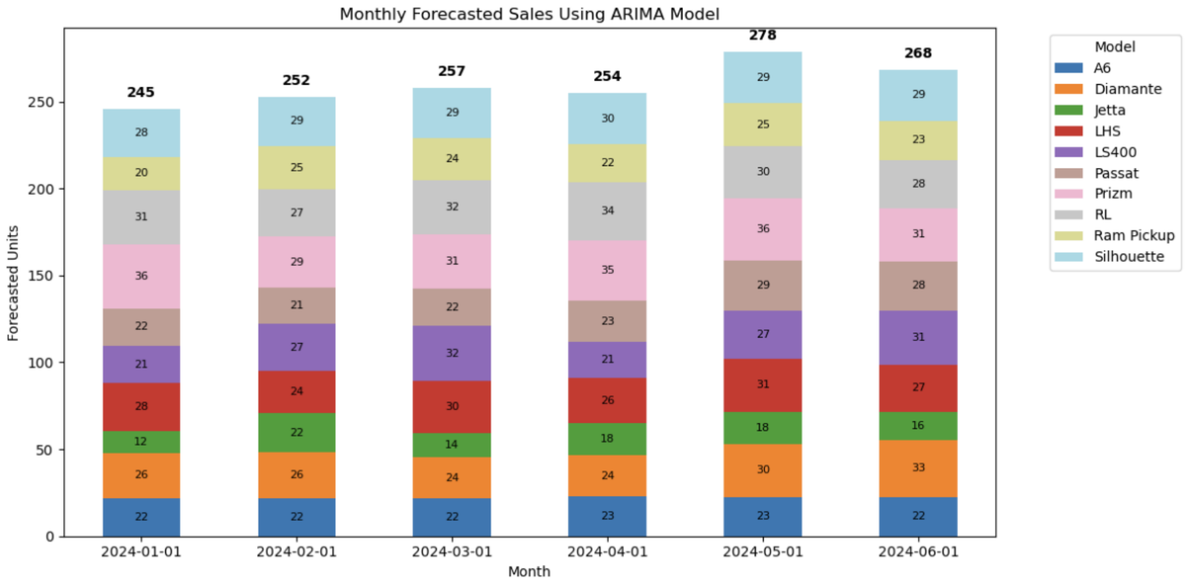


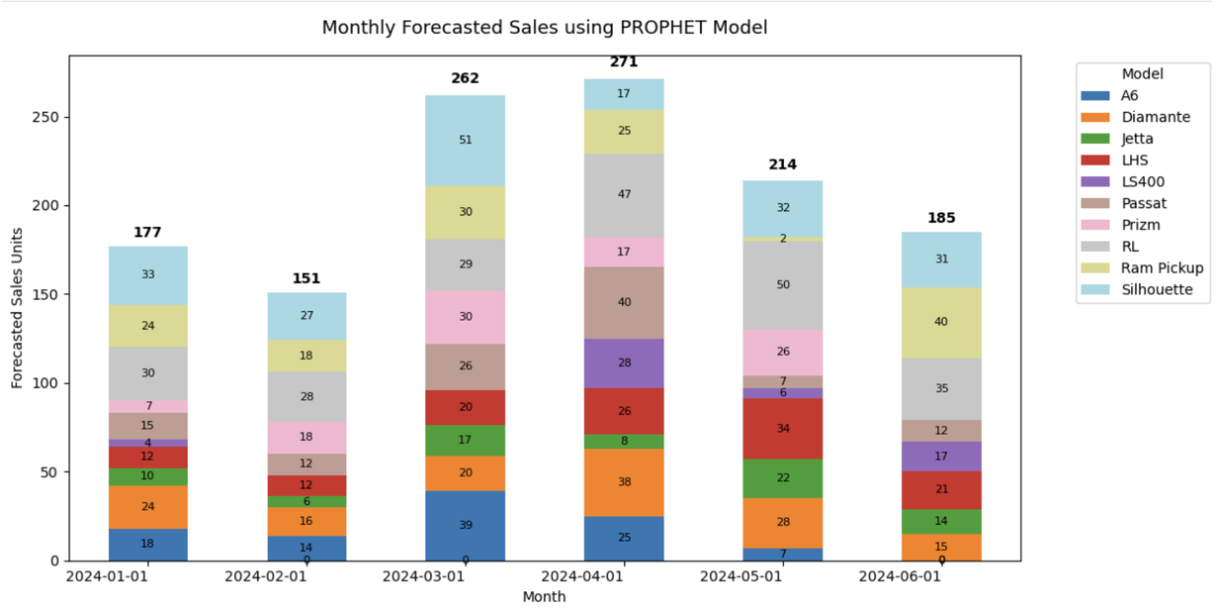

Insights:
Sales peak in Q3-Q4 particularly in September, November and December, indicating seasonal buying momentum (e.g. year-end promotions)

# MODEL COMPARISON

## ARIMA MODEL




## PROPHET MODEL



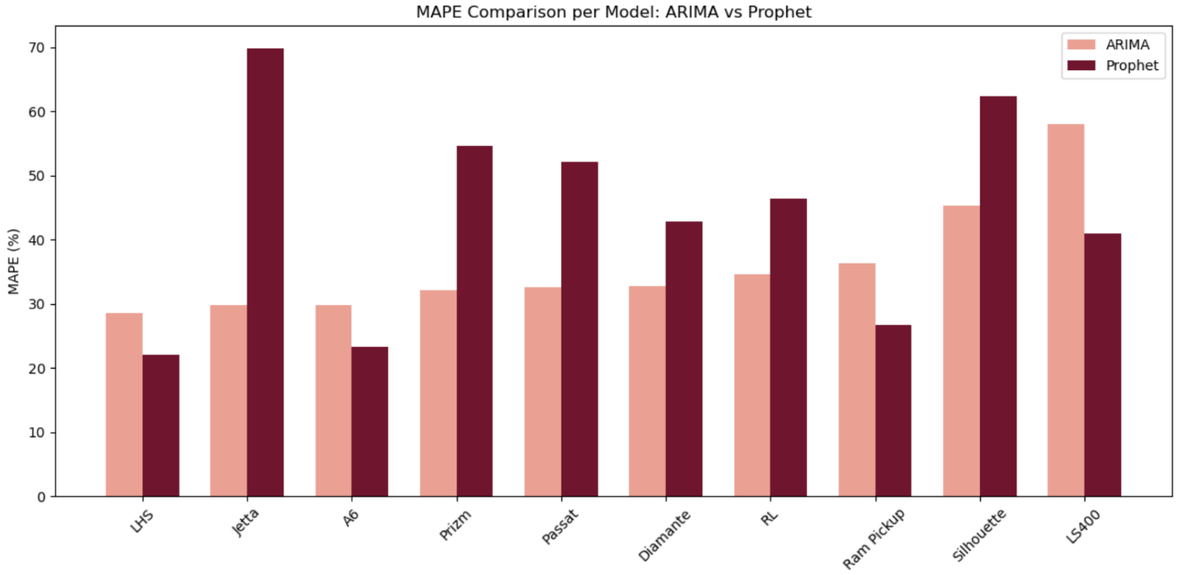

Insights:

1. ARIMA model produces more stable forecast across months, ideal for consistent demand
2. Prophet shows more fluctuations, capturing seasonal trend in historical data

# MODEL EVALUATION

## MAE, RMSE, MAPE

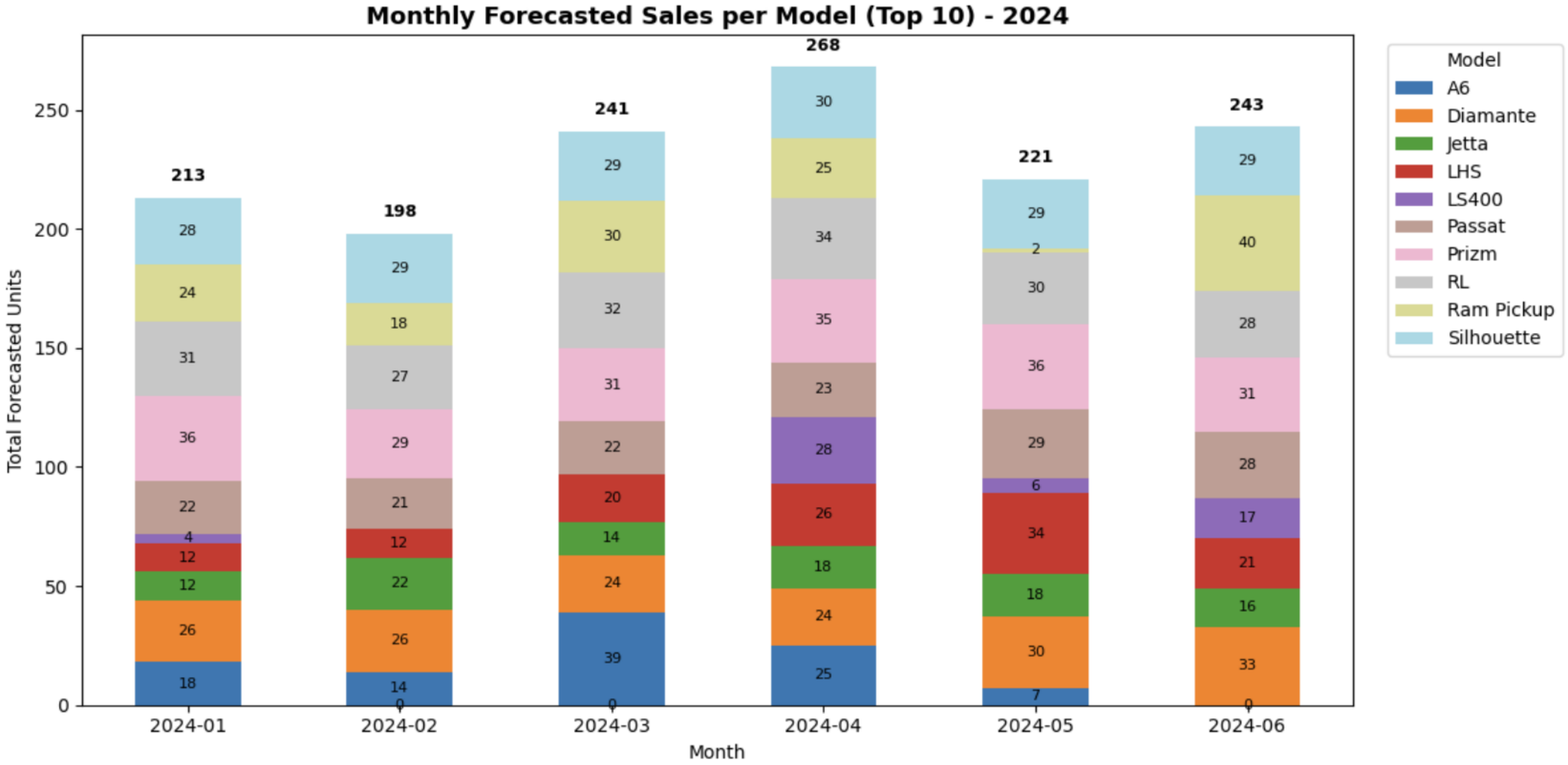
| | Model | MAE ARIMA | RMSE ARIMA | MAPE (%) ARIMA | MAE Prophet | RMSE Prophet | MAPE (%) Prophet |
|---|---|---|---|---|---|---|---|
| **0** | LHS | 6.50 | 8.14 | 28.55 | 3.96 | 4.46 | 22.03 |
| **1** | Jetta | 8.35 | 12.44 | 29.76 | 13.52 | 18.14 | 69.86 |
| **2** | A6 | 7.60 | 10.29 | 29.80 | 4.66 | 5.18 | 23.22 |
| **3** | Prizm | 10.78 | 14.69 | 32.18 | 10.87 | 13.31 | 54.61 |
| **4** | Passat | 7.73 | 8.60 | 32.57 | 9.85 | 11.01 | 52.14 |
| **5** | Diamante | 9.44 | 12.04 | 32.82 | 10.36 | 11.28 | 42.83 |
| **6** | RL | 10.66 | 14.52 | 34.55 | 10.15 | 12.43 | 46.38 |
| **7** | Ram Pickup | 10.30 | 13.04 | 36.29 | 6.39 | 7.67 | 26.70 |
| **8** | Silhouette | 12.69 | 15.00 | 45.36 | 12.48 | 16.09 | 62.37 |
| **9** | LS400 | 11.67 | 13.43 | 58.00 | 6.07 | 7.27 | 40.99 |




Based on model evaluation:

1. Prophet model provides more accurate forecast for LHS, A6, Ram Pickup, LS400, as reflected by lower MAPE values.
2. On the other hand, the ARIMA model yielded better predictive accuracy for the other car models which has stable sales trend (Jetta, Prizm, Passat, Diamante, RL, Silhouette).

# FINAL CAR SALES DEMAND FORECAST



# Conclusion

1. SUV and Hatchback remain the most preferred body styles represented by car model such as Prizm, Diamante, Silhoutte.
2. The market share distribution across region are diverse, showing that each region has competitive landscape and no brand dominates the market
3. Austin and Janesville lead regional sales, which potential as key market for brand awareness and diversification. The remaining regions show smaller but balanced distribution, indicating potential as supporting markets or dealer expansion.
4. Customer purchasing power are slightly higher on male buyer compared to female.
5. A seasonal sales patterns is observed, with sales peaking in Q3-Q4 particularly in September, November, December. Indicating year-end promotion or other incentives.
6. From the forecasting perspective, both ARIMA and Prophet Models demonstrated solid performance:
   a. Prophet providing better predictive accuracy for car model especially influenced by seasonal trends like LHS, A6, Ram Pickup, LS400
   b. ARIMA excelled for more stable demand patterns such as Jetta, Prizm, Passat, Diamante, RL, Silhouette
7. The 6-month demand forecast projects a total of 1,382 units, provides a foundation for inventory planning decisions.

**Implications:**

1. Prioritize Austin and Janesville for marketing and dealer support initiatives.
2. Focus product mix on SUVs and Hatchbacks to match market demand.
3. Leverage Prophet model for dynamic inventory planning of seasonal trends products and ARIMA model for stable demand products.
4. Prepare for Q3–Q4 promotions to maximize seasonal sales opportunities.

# Thank You