

Product Sales Insights Report

1. Introduction

This project investigates the relationship between weather patterns and product sales. Using historical sales and weather data, we focused on three top-selling products (Product 5, Product 9, and Product 45). The goal was to understand how weather conditions—particularly temperature—affect sales, and to provide actionable business recommendations.

The analysis was performed in three stages:

1. SQL – Preparing and joining sales with weather data.
2. R – Cleaning, feature engineering, modeling, and visualization.
3. Interpretation – Generating insights and recommendations.

2. SQL Data Preparation

We first explored the sales dataset to identify the top products. SQL queries were used to:

- Find the top 3 products by units sold.
- Join sales with weather data using store_nbr and station_nbr.
- Create daily sales + average temperature datasets for Products 5, 9, and 45.

Example Query:

```
SELECT t.date, t.item_nbr, SUM(t.units) AS daily_units, w.tavg
FROM train t
JOIN key k ON t.store_nbr = k.store_nbr
JOIN weather w ON k.station_nbr = w.station_nbr
AND t.date = w.date
WHERE t.item_nbr = 5
GROUP BY t.date, t.item_nbr, w.tavg
ORDER BY t.date;
```

3. Data Cleaning & Feature Engineering (R)

Using R, we cleaned and enriched the data:

- Missing values in temperature (avg_temp) were imputed using the median.
- New features were created:
 - is_weekend: 1 if Saturday/Sunday, 0 otherwise.
 - month: categorical variable capturing seasonality.
 - is_rainy_day: placeholder (0 in this dataset, as precipitation data was missing).

This ensured datasets were ready for modeling and visualization.

4. Predictive Modeling (R)

Linear Regression

We built linear regression models to test how sales respond to temperature and weekends.

Performance (RMSE & MAE):

- Product 5: RMSE = 304.23, MAE = 232.54
- Product 9: RMSE = 264.93, MAE = 207.70
- Product 45: RMSE = 270.79, MAE = 212.50

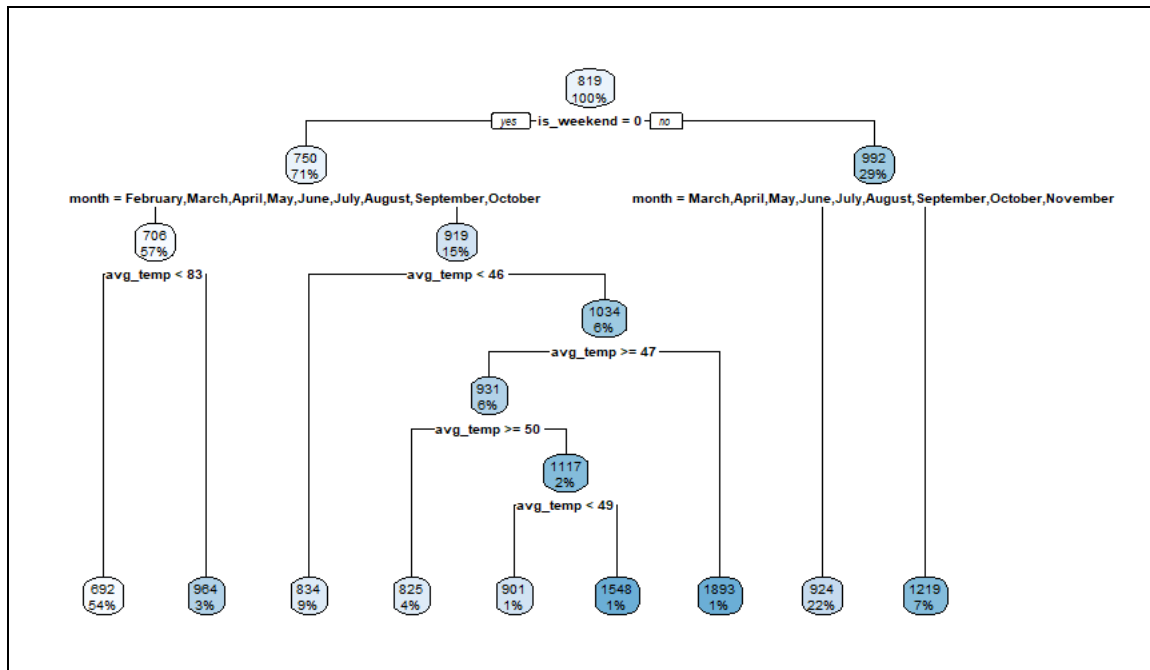
Interpretation: Weekend sales consistently boost demand, and temperature positively correlates with higher sales.

Decision Trees

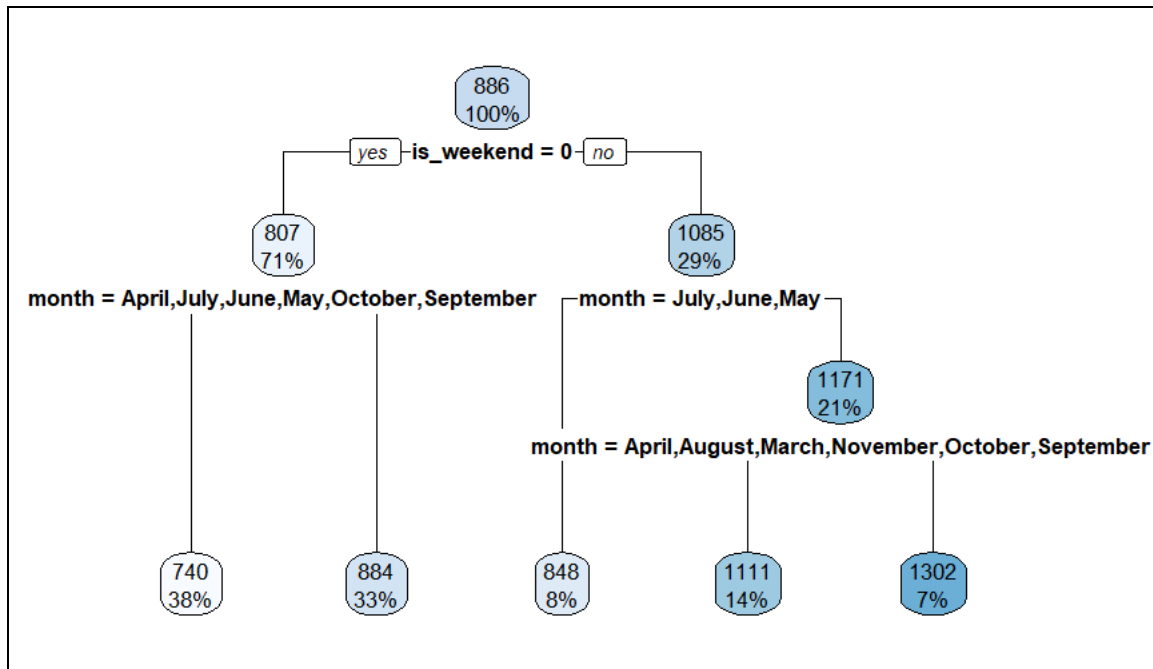
Decision Tree models were built for interpretability. They confirmed:

- Temperature and month were strong predictors.
- Weekend effect clearly influenced sales patterns.

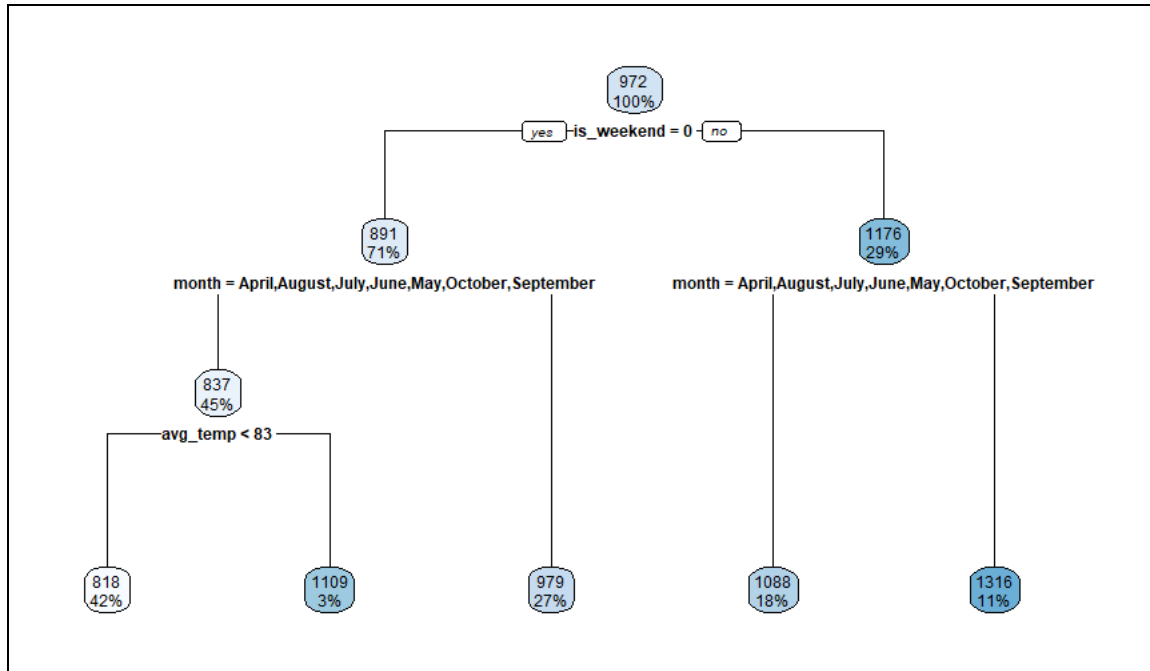
Product 5:



Product 9:



Product 45:



Performance (RMSE & MAE):

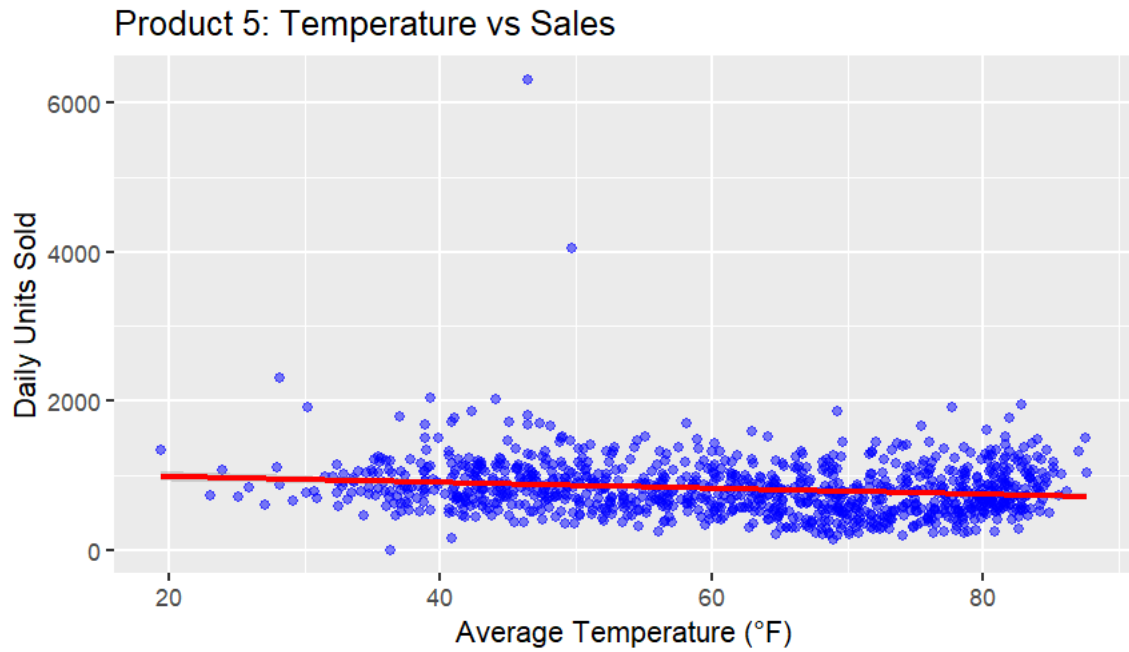
- Product 5: RMSE = 315.93, MAE = 256.82
- Product 9: RMSE = 269.12, MAE = 213.42
- Product 45: RMSE = 269.51, MAE = 211.17

Interpretation: Linear regression performed slightly better, but trees highlighted feature importance.

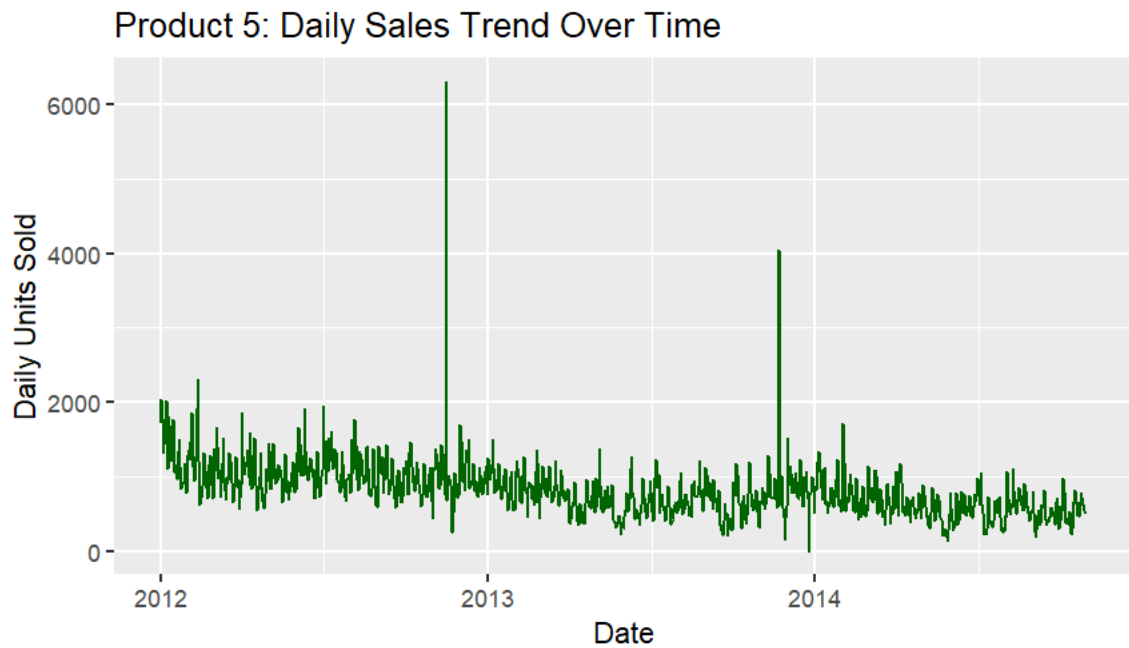
5. Visualization & Interpretation

Product 5

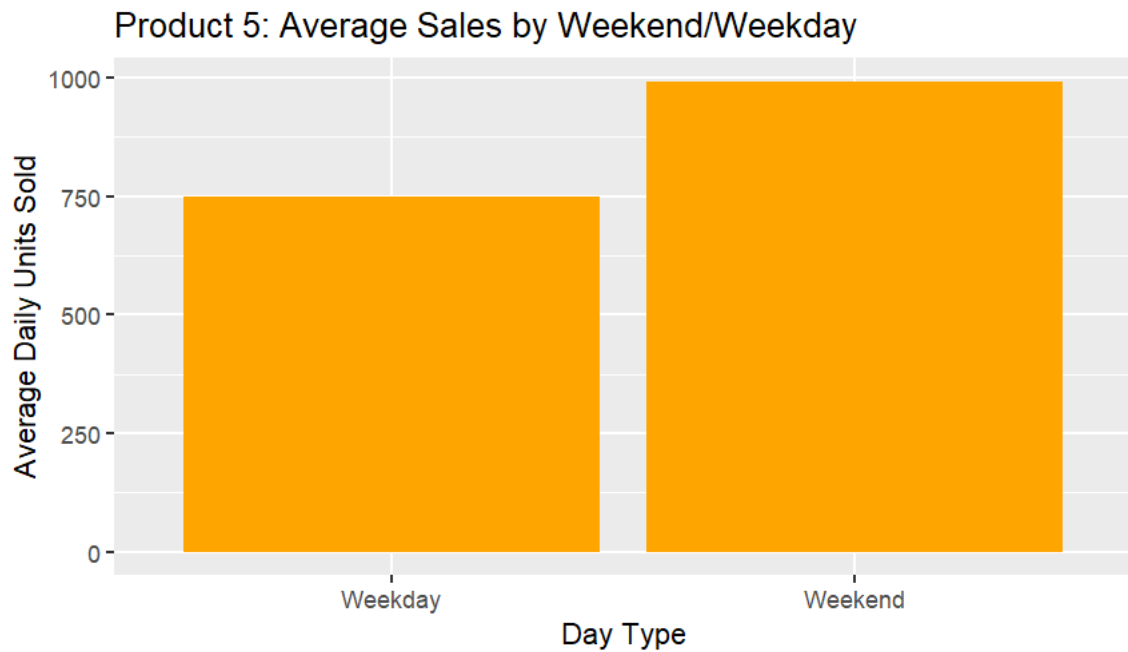
Scatter Plot: Sales increase as temperature rises.



Line Chart: Clear seasonality, with fluctuations across months.

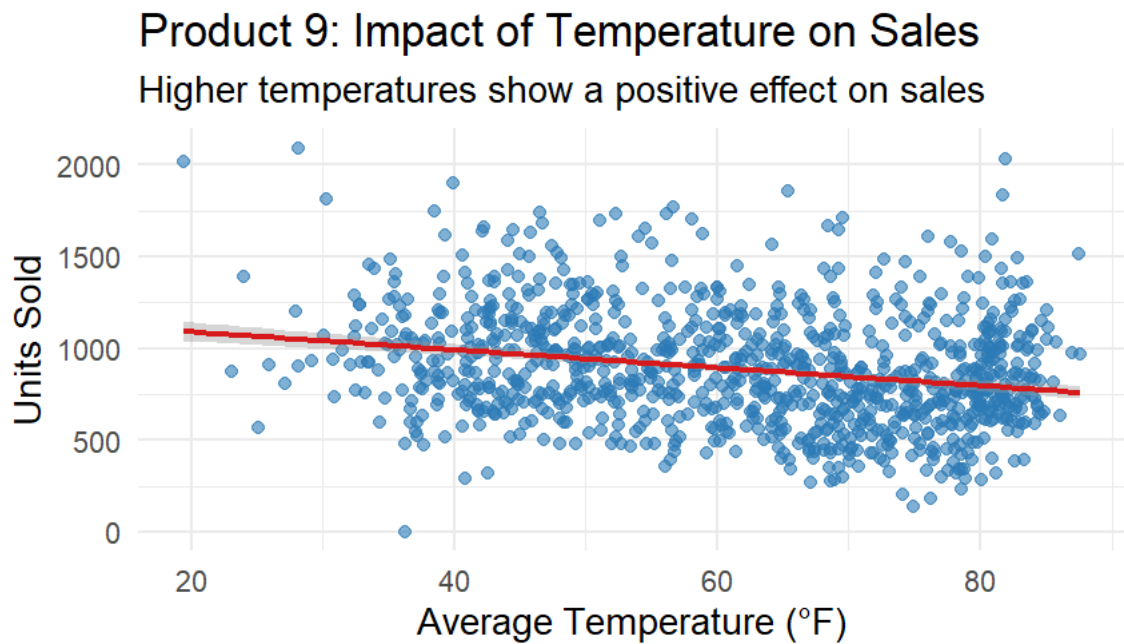


Bar Chart: Weekends outperform weekdays significantly.

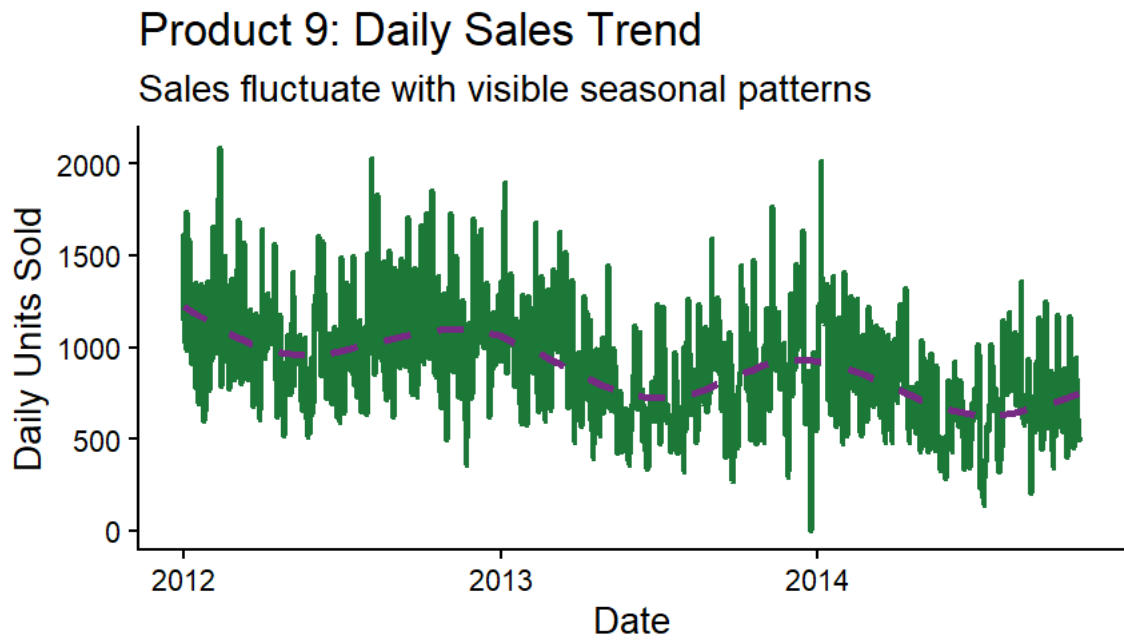


Product 9

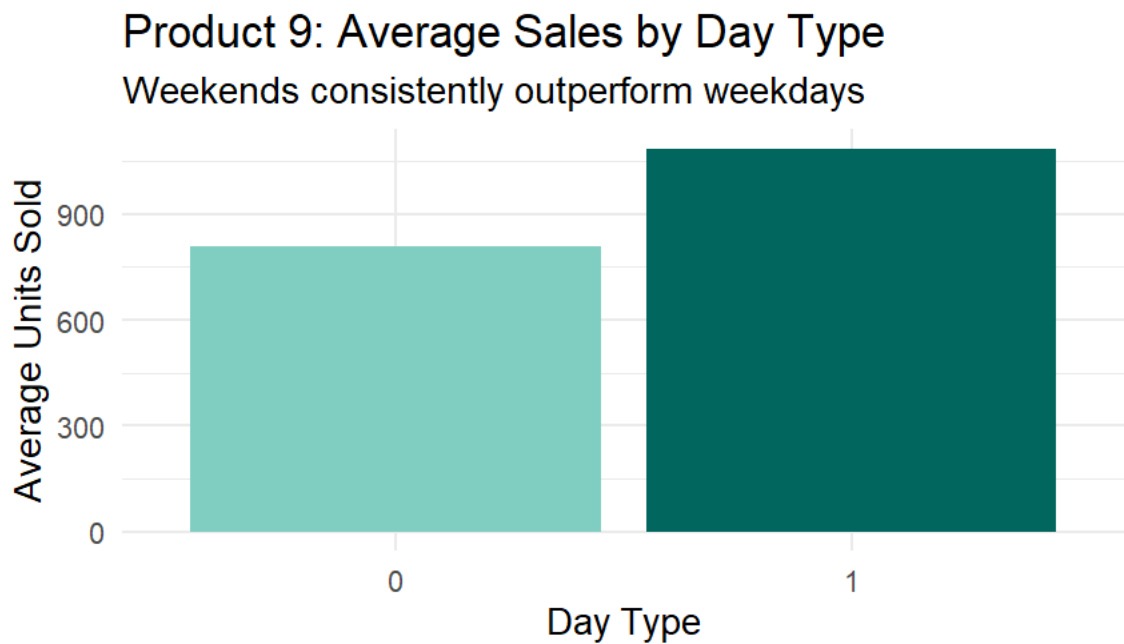
Scatter Plot: Stronger positive correlation with temperature than Product 5.



Line Chart: Seasonal cycles evident.

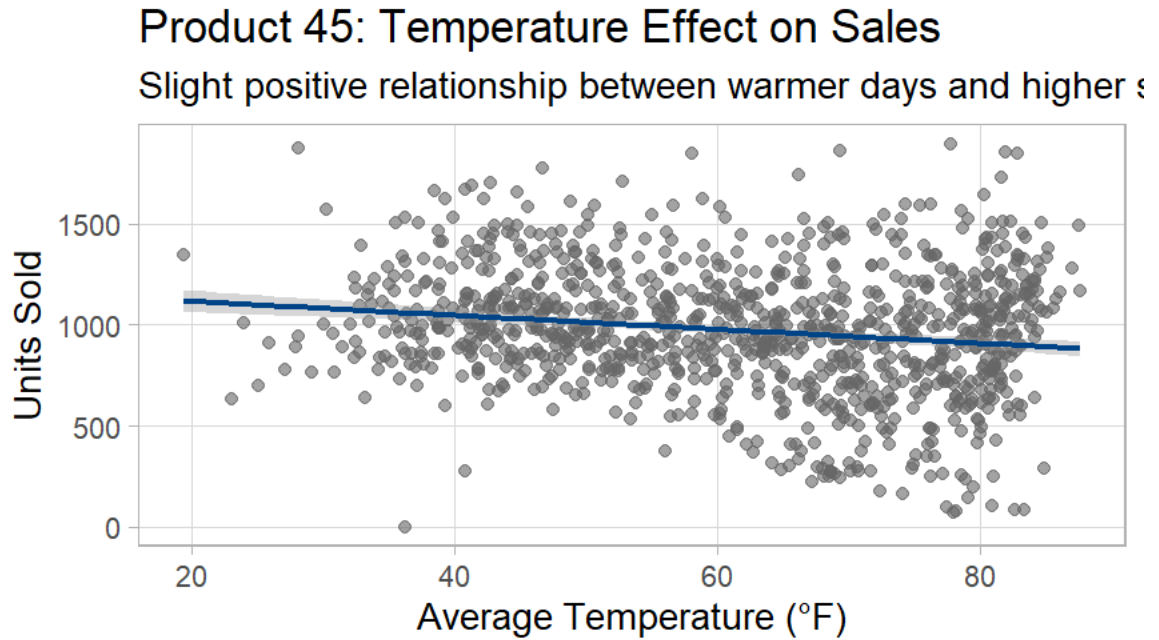


Bar Chart: Weekends show a clear sales uplift.

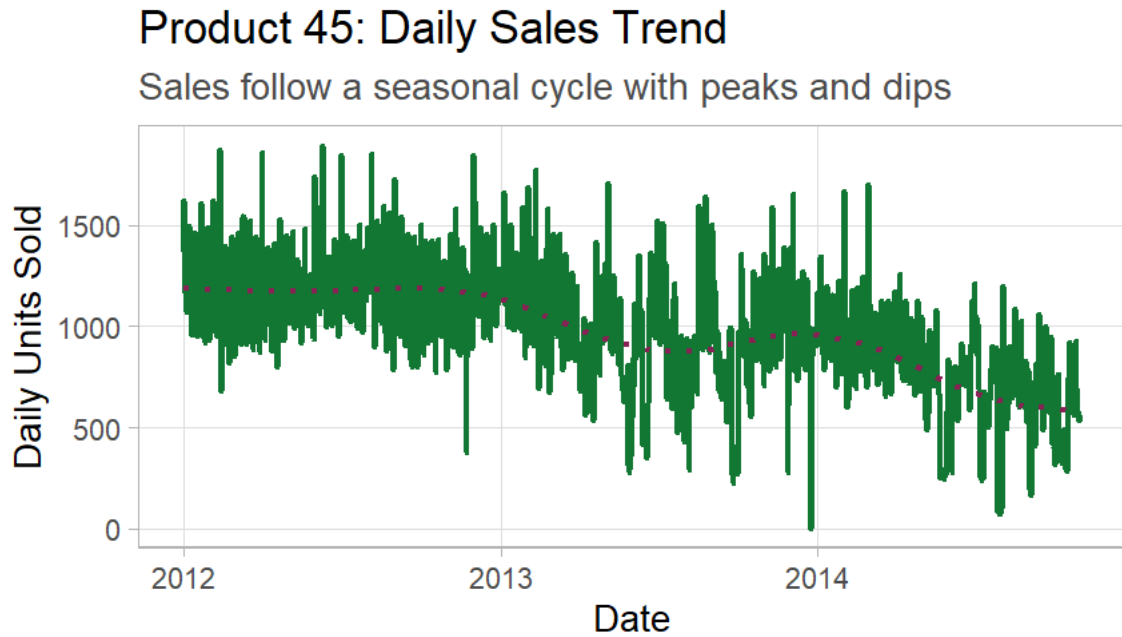


Product 45

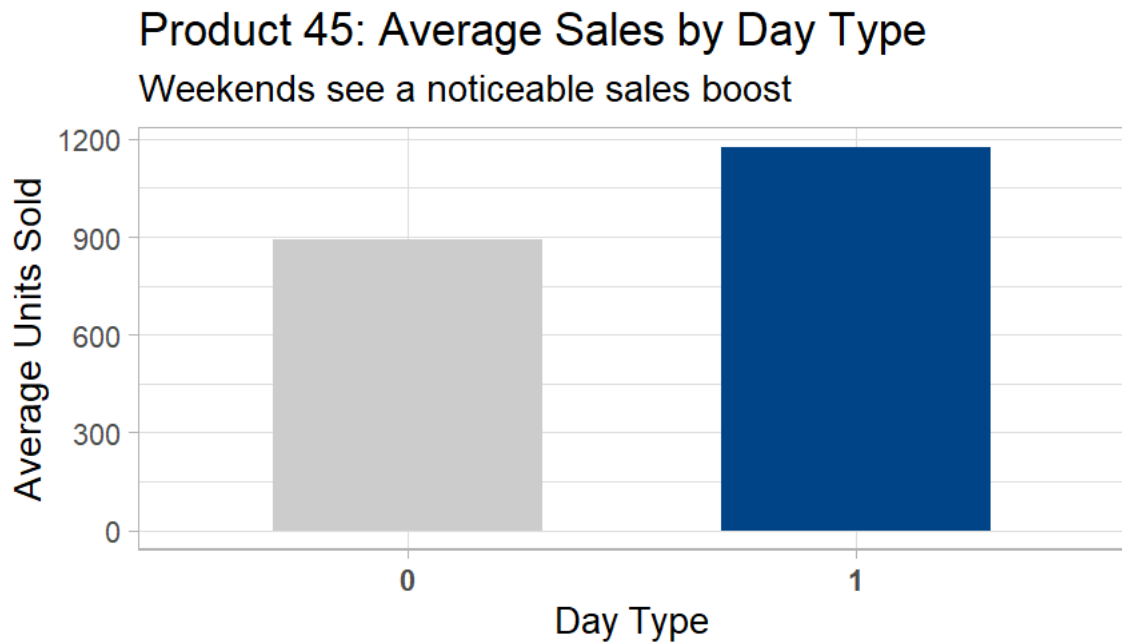
Scatter Plot: Moderate positive relationship with temperature.



Line Chart: Seasonal patterns with peaks in warmer months.



Bar Chart: Weekend sales consistently higher.



6. Business Recommendations

Product 5

- Increase stock and staffing for weekends.
- Boost promotions in warmer months when sales are higher.

Product 9

- Launch seasonal campaigns during summer months to maximize sales.
- Use weekend discounts or bundles to further lift demand.

Product 45

- Plan inventory for peak summer demand.
- Use predictive stocking based on weather forecasts to reduce stockouts.

General Recommendations

- Temperature and weekends are consistent drivers of demand.
- Align marketing and supply chain planning with weather and seasonality.
- Simple weather-aware inventory management can improve sales and reduce waste.

7. Conclusion

The analysis demonstrates that weather, particularly temperature, plays a critical role in product demand. Weekend effects amplify sales across all products.

- Product 9 showed the strongest weather dependency.
- Linear Regression models provided better accuracy, while Decision Trees offered clear interpretability.
- Businesses can benefit from weather-driven sales forecasting, aligning inventory and promotions to maximize revenue.