# Time Series Forecasting Project – ARIMA vs SARIMA

## 1. Dataset Overview

We worked with the 'Store Item Demand Forecasting' dataset. This dataset contains daily sales data from 2013 to 2017 for 50 different items across 10 stores. Each row tells how many units of a specific item were sold in a specific store on a particular date.

## 2. Problem Setup

We focused on a single store-item pair: Store 1 – Item 1. The goal was to forecast daily sales using two different models:  
- ARIMA (AutoRegressive Integrated Moving Average): captures trends but not seasonality  
- SARIMA (Seasonal ARIMA): captures both trends and repeating seasonal patterns

## 3. Forecasting and Visualization

We trained both models on historical data and forecasted the next 30 days. The forecasts were then plotted along with the recent observed sales data. This helped visually compare how well each model performed.  
  
- ARIMA produced a flat or smoothed forecast line.  
- SARIMA produced a wavy forecast line that matched the weekly sales pattern.

## 4. Validating the Model with a Holdout Set

To evaluate model accuracy, we used the last 30 days of training data (December 2017) as a holdout set. We trained SARIMA on earlier data and predicted those 30 days. Then we plotted:  
- Actual sales (blue line)  
- SARIMA forecast (green line)  
The SARIMA model closely followed the actual trend and picked up weekly seasonality, proving it's well-suited for this task.

## 5. Conclusion

This project clearly demonstrated that SARIMA outperforms ARIMA when the data exhibits seasonal patterns. We also validated the model using real historical data and prepared to forecast future demand across all items and stores.