

# Freshflow Documentation

## Dataset

The dataset consisted of 3 items; for each item, the sales quantity was given along with other values. The task was to predict the sales quantity for the next day, given the previous day's data. The data didn't have NaNs except for the revenue column.

After plotting the moving average along with the autocorrelation for the sales quantity of all 3 items, it can be concluded that the sales quantity can be predicted very well from the timeseries data.

## Modeling

Due to the limited time constraint, I was able to apply 2-3 models with limited feature engineering. I applied random forest and gradient boosting model because of their straightforward implementation and because they require less hyperparameter tuning.

The dataset was divided into 67-33 set to test the models. The models were evaluated by visualizing the plots and checking out the mean absolute error.