**Solution Approach**

My motive was not to win this competition but see how fast can I develop a good solution if not best.

**First Attempt:**

Random Forest to predict only Sales without any Hyperparameter Tuning. But I had added Date Features such as Month, year, dayofweek, weekend etc.

Within 1 hour – I was getting 260 270 score on PB Leaderboard. The model was overfitting.

**Second Attempt:**

FB Prophet to predict only Sales for each store separately and also considered Holidays in this.

In Next Hour I could get 250 something score which is an improvement. Model did overfit in this case also.

**Third Attempt:**

I then used Pycaret for Timeseries Regression and 10 fold cv(default) to predict Sales. I compared almost all models in Pycaret and used best model.

The score suddenly improved a lot to 215 something. This was great as in another hour I could achieve this much better results.

Next I wanted to reduce Overfitting so tried other ways.

**Fourth Attempt:**

I used LGBM to predict Order after doing Hyperparameter Tuning and then used Linear Regression to Predict Sales (used only Order as X).

I though model will not overfit but again model did overfit.

PB Leaderboard score was something 255.

**Fifth Attempt:**

I then used Pycaret for Timeseries Regression and 10 fold cv(default) to predict Order and then Sales. I compared almost all models in Pycaret and used best model.

The score improved to 212 something.

**I did not try much of Feature Engineering because of my office related work.**

**Best Attempt:**

I used weighted avg of fifth attempt, Fourth Attempt, Second Attempt and First Attempt and score was something 206.

**Note:**

Important thing to notice was the models I built was not capturing the Lower and Higher values properly.

As per the Sales distribution, there were some stores which had spikes in June and July.

I think, I could have improved result by including more features such as Sales Lag values.