Course: IST 970 Internship
Company: FedEx Custom Critical
Title: Sales Pipeline

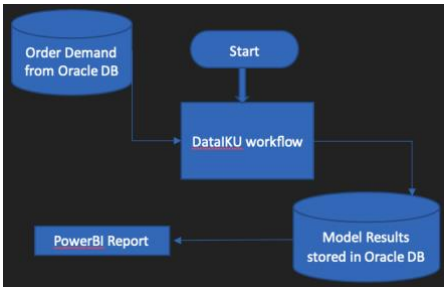
Project Description:

The Critical Shipping division of Custom Critical sources its fleet to meet its customer's transportation needs. Sourcing of vehicles on an average could take around 90 days. Hence, forecasting order demand would help in capacity planning and making sure we have enough trucks to meet the required demand in different seasons. Forecasting demand segmented by product type would aid capacity planning at a lower granularity of truck type and size.

Data Description:

To get assumed order demand data, historical order volume data along with historical missed opportunities data captured (orders about which we were approached but did not receive) is extracted from our Oracle DB. Extracted data is time-series data.
Actual demand could be much higher since all missed opportunities are not captured in our database.

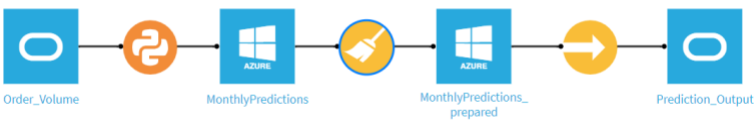
Solution Architecture:



Tools used:

1. OracleDB – Read the data and to store model results
2. Dataiku – To build project workflow
    - Data validation and cleanup
    - Build and run model to make monthly prediction
3. PowerBI – Interactive report generated to visualize model results and track model performance over time.

Steps:



- Order_Volume: Executes the SQL query to read demand from our data warehouse.

- Python recipe: Uses the best fitted exponential smoothing time-series model to forecast demand 18 months into the future segmented by product type. The data is stored into MonthlyPredictions file.
- MonthlyPredictions_prepared: Converts the dates from string format to date before writing into the database.
- Prediction_Output: Writes data into our team's database in our data warehouse.

Data Modelling and Workflow:

- 3 time-series models with different hyper-parameters were developed for each time-series algorithms used – ARIMA, TBATS and Exponential Smoothing.
- These models were developed on Dataiku development node.
- Exponential Smoothing with best fitting hyper-parameters was our final model that makes predictions.
- Dataiku workflow is triggered on the first of every month at 5am EST to make fresh monthly predictions.
- The workflow reads historical order demand from OracleDB and forecasts monthly demand segmented by product type 18 months into the future.
- The forecasted result is stored into our teams's table in OracleDB.

Visualization:

- An interactive PowerBI report was developed for the Operations Team.
- Operations team then uses these monthly forecasted sales to plan out fleet required to fulfill these forecasted order demands.
- It is refreshed daily at 5am EST.

Specific contribution to the project:

- Assisted in writing SQL query to gather order demand data.
- Assisted in model development.
- Developed PowerBI report.

Learning outcomes from the project:

- Learnt how to handle live data.
- Learnt about Dataiku tool and even got certified.
- Learnt how to develop time-series models with such huge live datasets.
- Learnt about the complete cycle of deployment of a project: development → test → production.