Forecasting Operational Performance Measures

DIPENDU CHANDA
SOUTHWEST AIRLINES

ABSTRACT

In Summer 2017, I joined the Operational Strategy & Performance department at Southwest Airlines as Operational Analyst Intern. I was part of the Simulation & Forecasting Team which is responsible for improving the operational performance through thought leadership utilizing measurement, prediction, and planning.

Until now, the team's actions towards KPI forecast, swaps, and various other performance analysis were of reactive nature. Insights were drawn just from daily, weekly and monthly reports as well as dashboards generated by the team.

In the process of helping different internal teams familiarize themselves with the latest technology and explore forecasting techniques for some of the business problems, I was assigned the task to explore the possibility of forecasting *Block Time Hit Rate*, *Turn Time Compliance*, *Departure Performance*, *Originator D0*, and finding ways to improve our *On-Time Performance* using existing historical and flight schedule data to identify the factors that influence it.

PROJECT RESULTS

- Built a forecasting model that would predict various KPI with ~80% variability explained by applying predictive analytics on combined data from Teradata, weather patterns, itineraries, Federal Aviation Administration, and location tracking system. The model was tested for 3 years of data and the average MAD score was less than 5% for all of the KPI forecasts
- Performed text mining on incoming customer tweets and survey results to detect trouble routes and stations. Even VP of Operation Strategy appreciated this project
- Analyzed flight records to detect operational anomaly at station level and potentially improved the OTP from 80.2 to 84%
- Generated Alteryx and Tableau automated interactive visualizations for monthly meetings with senior executives including VP and SWA directors

INTERNSHIP OBJECTIVES

1. Application of data analytics, statistics and predictive modeling in order to understand operational impacts of proposed business decisions
2. Assist in development of monthly on-time performance (OTP) forecast
3. Explore new data source and identify if it can be used to improve the existing predictive models
4. Create visualizations and communicate findings or recommendations to audiences with minimal understanding
5. After getting *Federal Aviation Administration* data authorization research on finding the Key parameters to improve our prediction models

METHODS & TOOLS USED

Tools used:

• Alteryx • RStudio • R • Teradata • Oracle SQL • Visio • Excel • PowerPoint • SkySYM

Forecasting 5 Key Southwest KPIs for Operational Improvement:

On-Time Performance (OTP), Block Time Hit Rate (BTHR), Turn Time Compliance, Departure Performance, Originator D0

Data gathering, cleaning & transformation:

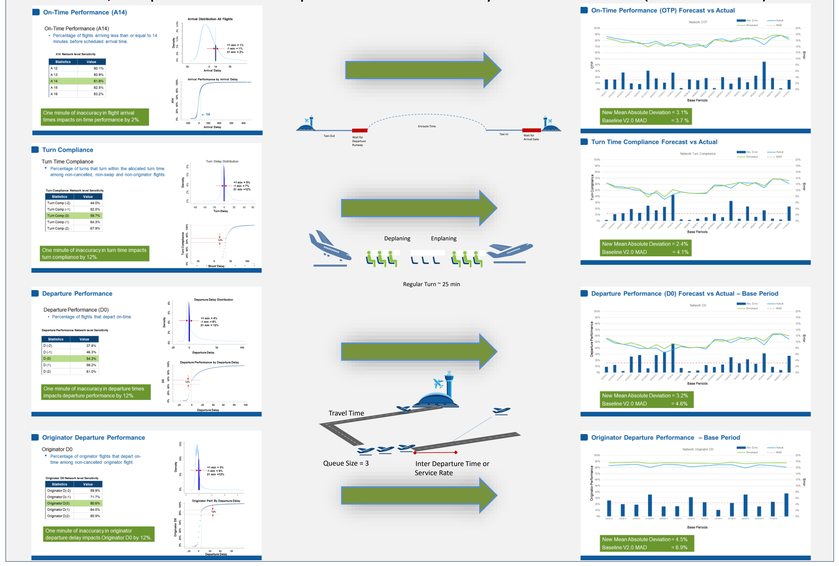
- Maintenance delay information from Oracle database and flight information data residing in the Teradata database were combined in Alteryx. For BTHR, I combined weather patterns, itineraries, Federal Aviation Administration, and location tracking info.
- Relevant business rules were applied and minor anomalies were removed or imputed using MICE (Multivariate Imputation via Chained Equations) package in R. Then I filtered on the relevant influencing factors using subject knowledge and correlation analysis.

Selecting and fine-tuning the model:

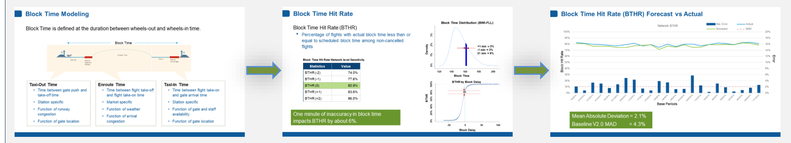
- For the KPI analysis I mostly used multivariate multiple regression. I also used counted regression (Poisson regression) model for getting station level delay events and swaps prediction since a logistic regression does not provide us the magnitude. And presented my findings to senior executives.

Results:

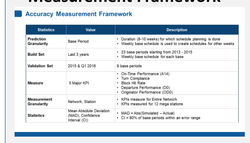
As shown, the predictions were quite accurate with very small MAD score (less deviation).



For last few base periods Block Time Hit Rate was a major concern in Operational performance. I did further analysis on it.

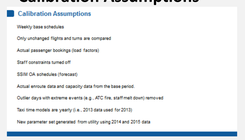


Measurement Framework



Outliers

Calibration Assumptions

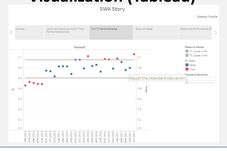


Observations:

- Outliers validation is must as several outliers dates are not factored in available
- Merger with AirTran brought a shock to the network and dynamics changed for several stations e.g. ATL
- Data limitations e.g. runway capacity available for only 68 SWA stations

Other Important Work:

Visualization (Tableau)



Data Blending (Alteryx)

Complex SQL (Teradata)

CONCLUSION

In conclusion, I really appreciate the opportunity to work at Southwest Airlines and this is a very worthy and valuable experience on my career path. The kind of freedom I was given to experiment with my work gave me a prospect to implement my newly acquired skills into real life business problems and to some extent add value to the current team with my work. At the end of this 13-week internship I feel more confident for my future career.

CONTACT

Name: Dipendu Chanda
Major: MBA & M.S., Business Analytics
Email: Dipendu.Chanda@utdallas.edu
Cell Phone: (972) 639-7718