

Predicting Flight Delays

LHL Data Mid_term_Project

*By: Furqan Ejaz, Shiv Soloman, & Jesse
Randolph*



Flight Delays Cost Everyone \$33B in 2019*

Airlines



\$8.3B

According to the FAA, Flight delays costed airlines a combined \$8.3 Billion in 2019.

Passengers



\$18.1B

This results from time lost due to schedule buffers, delayed flights, flight cancellations etc. All coming out of customers pockets.

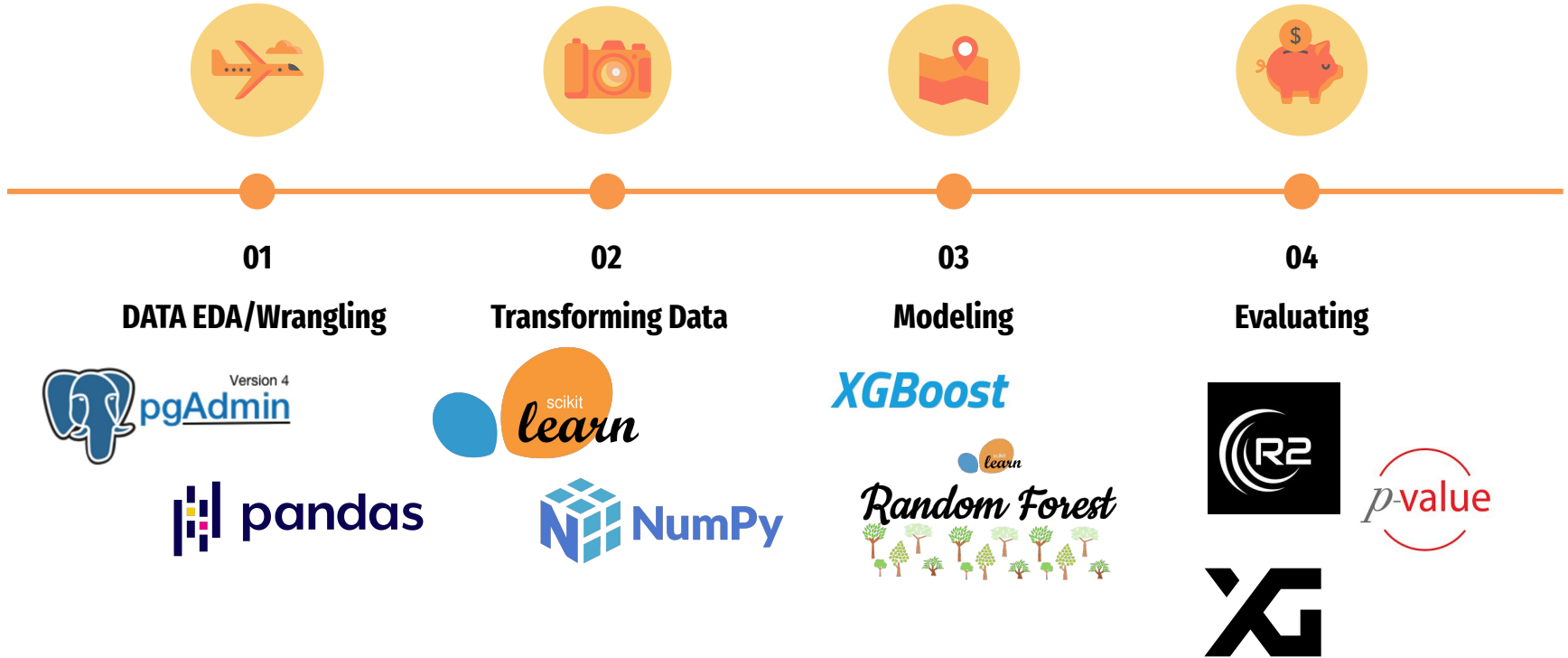
Indirect



\$4.2B

The increased cost of doing business for other, non-airline, sectors, making the associated business less productive.

Machine Learning Workflow (Tool Showcase)



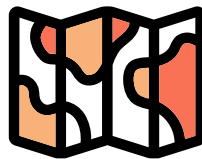
Overview Of Our Data

FLIGHT DATA

Flight data was provided by Lighthouse Labs.



This was a very large Data set.



WEATHER DATA

We found weather data and imported it into our data From Kaggle.



2016-2020

The data had information from 2016-20

Key Insights for Best Model Accuracy



Data

- Clean Data
- Using sample data gave us faster process times



ML Algorithm

- Classification approach was more accurate (**%70 accuracy score**) using Logistic Regression



Features

- Choosing different features gave us different accuracy scores
- Using Weather Data gave us better insights



Our Top 4 Features for SKLearn Logistic Regression

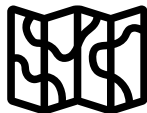
Carrier Type

The type of plane made a big difference as bigger passenger planes were more likely to be delayed!



Taxi/Runway Strip

This has to do with organizing the runway, and the order the planes take off.



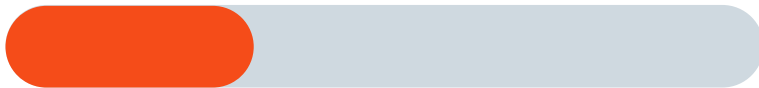
Destination

Yes, this is the obvious one. Certain flights have more history of being delayed...*cough *cough JFK -> LAX



Distance

Certain airports and cities are more likely to have delays than others!



How Airlines Would Save \$ With This Info

Customer Loyalty

I'd rather Fly on select airlines with less delays

Crew

....good thing they're paid by the hour hey?

Fuel/Maintenance

Longer runtimes mean more fuel, and breakdowns



Flight Cancellations

Longer delays result in more flight cancellations*

Missed Connections

Having to re route planes is costly and time consuming

Indirect

Associated businesses will have to pay more because of delays



NOVEMBER

Data Science Is an Art

We had a lot of fun learning more about the different tools under our belts.

Peers/Mentors

Thank you everyone for collaborating, teaching, and reaching out to our group! It really made us feel a sense of community!

Thanks for Listening Everyone.

LHL Data
Mid_term_Project

By: Furqane Ejaz, Shiv Soloman, & Jesse Randolph

****FAA.gov**
****Kaggle.com**
****Slides by SlidesGo**