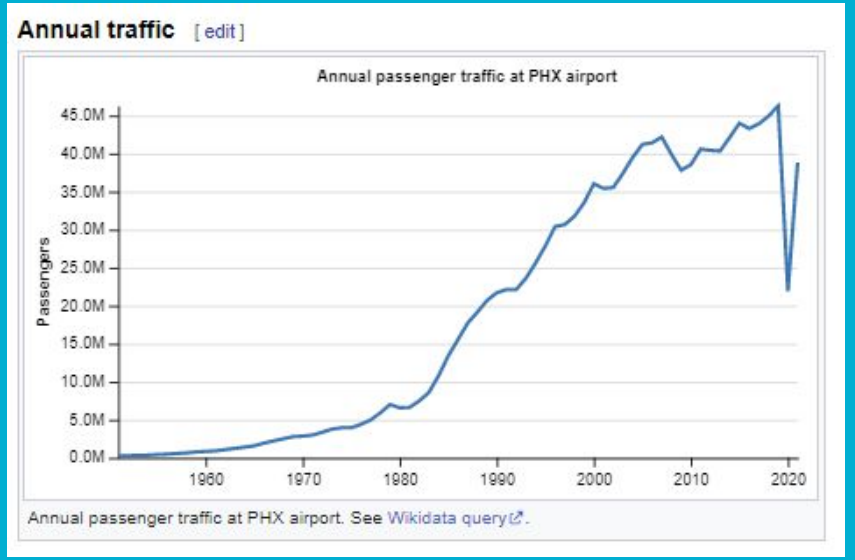
Airline Efficacy

A Primer for Phoenix Sky Harbor International Airport

Copper State Securities

Context

- In 2020, Phoenix Sky Harbor saw approximately 45 million passengers (enplaned + deplaned) through its terminals
- It serves as the hub for connecting flights, changes of airline
- It is the 11th busiest hub airport in the world, and is acutely vulnerable to frictions caused by delays and poor service



Delta & American Airlines

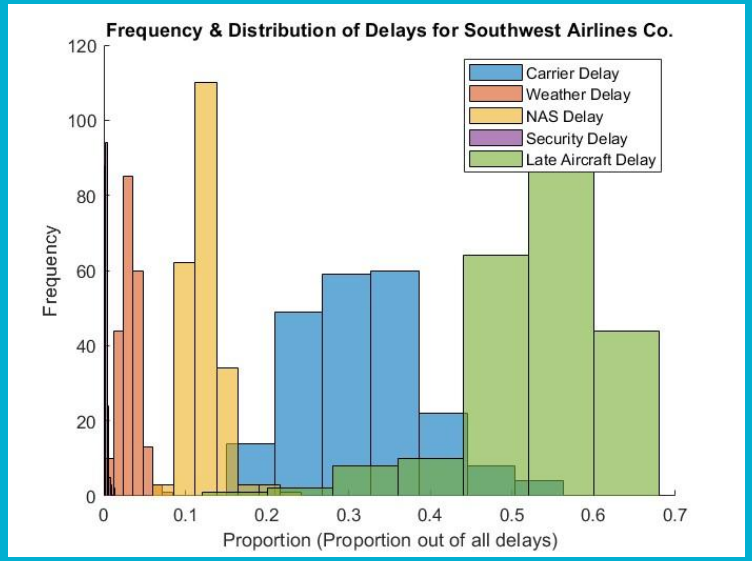
- By market share, American Airlines and Delta Air Lines are the largest carriers at Phoenix Sky Harbor
- Their success (or failure), reflects a broader success (or failure) for the airport as a whole
- Sample text

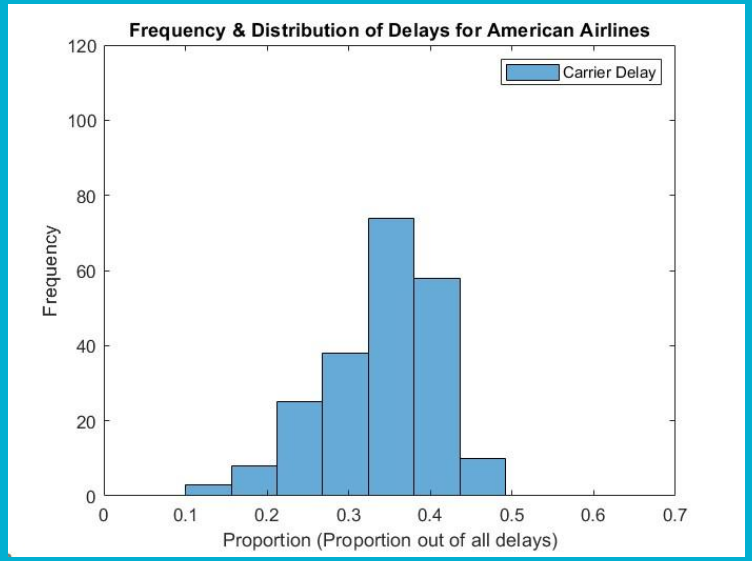
Top airlines at PHX
(August 2021 - July 2022)[87]

Rank	Airline	Passengers	Percent of market share
1	Southwest Airlines	14,574,000	35.54%
2	American Airlines	13,846,000	33.77%
3	Delta Air Lines	2,914,000	7.11%
4	United Airlines	2,321,000	5.66%
5	Skywest Airlines	1,969,000	4.80%
6	Others	5,381,000	13.12%

Comparison of Exploratory Data

A side-by-side look at the median/mean delays of each airline





Comparison of Exploratory Data II

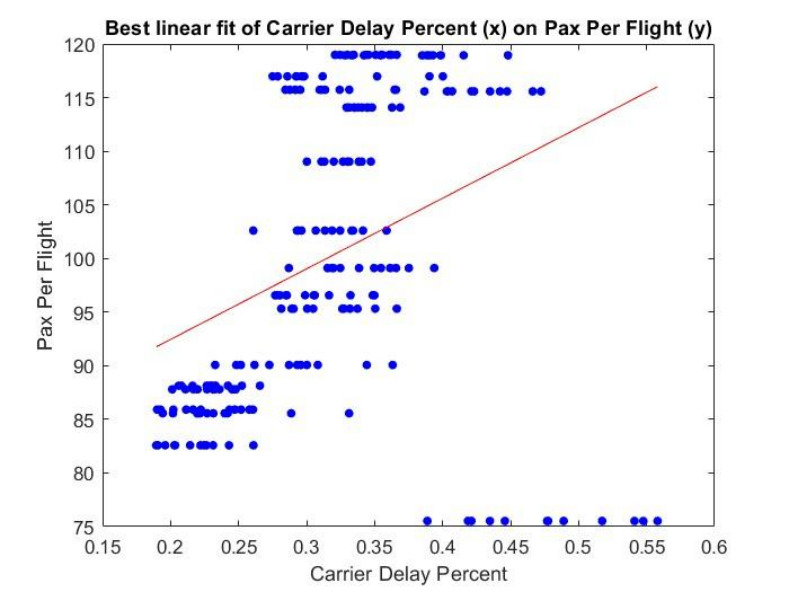
A side-by-side look at the median/mean delays of each airline

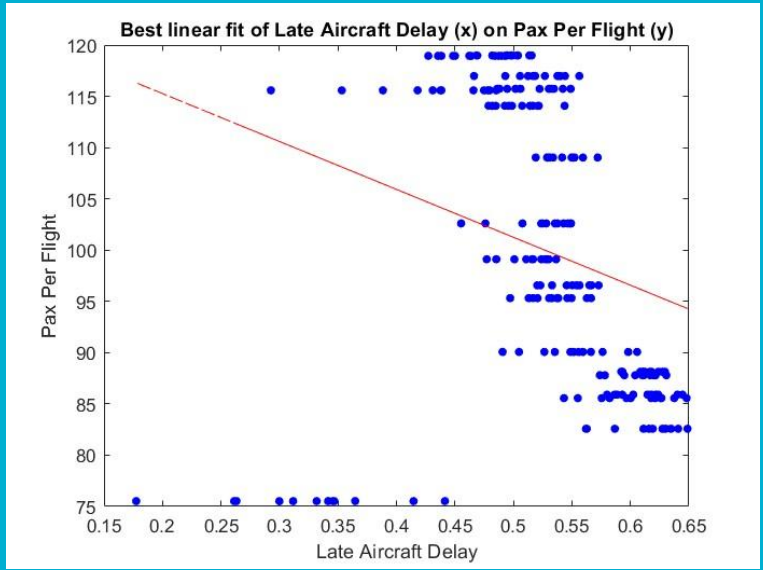
This is where we performed more analysis to really get of sense of each airport, later we move onto ML models.

Insert Graph or Figure Here

Insights from ML

A side-by-side look at the median/mean delays of each airline





Insights from ML II

A side-by-side look at the median/mean delays of each airline

Later we moved onto ML models, here's what we found

"Over 100 trials, the accuracy of our SVM model was ~95%. Therefore, we can accurately distinguish between Southwest Airline Co. and American Airlines.

This best-fit line has a correlation coefficient of -0.2587
between Late Aircraft Delay and Pax Per Flight, and a p-val of 0.0000

linfit = 216×4 table

	Late Aircraft Delay	Pax Per Flight	y-pred	y-error
1	0.5627414	82.5793915	98.3201440...	-15.74075303...
2	0.6237869	85.9055634	95.4685218...	-9.562958875...
3	0.5931904	88.1379776	96.8977787...	-8.759800747...
4	0.5781347	87.8001022	97.6010773...	-9.800975391...
5	0.6265026	85.5645828	95.3416631...	-9.777080153...
6	0.5514237	90.0695343	98.8488301...	-8.779296108...
7	0.5328086	95.3182526	99.7183974...	-4.400144435...
8	0.5329791	96.5686264	99.7104314...	-3.141805453...
9	0.5238198	99.1075745	1.00138292...	-1.030718004...

holdout =
0.200000000000000

n =
100

accuracy = 0.9186

resultTable = 86×3 table

	TrueLabel	PredictedLabel	Score
78	0	0	-0.8319
79	0	0	-1.1641
80	0	0	-0.5565
81	0	0	-1.7957
82	0	0	-1.5764
83	0	0	-1.7106
84	0	1	0.2346
85	0	1	0.2270
86	0	1	0.0124

avgAccuracy = 0.9481

Potential Losses

“Based on the assumption that passenger revenue is $x per passenger, a slope of -0.2 passengers per flight (independent variable being year, and dependent being pax_per_flight), and the # of flights in a single year being number, then…”

Loss = ($x per passenger) * (# of flights per year) * (-0.2 passengers per flight)

Additional Sources

These are the sources we used beyond the ones provided in the assignment brief