

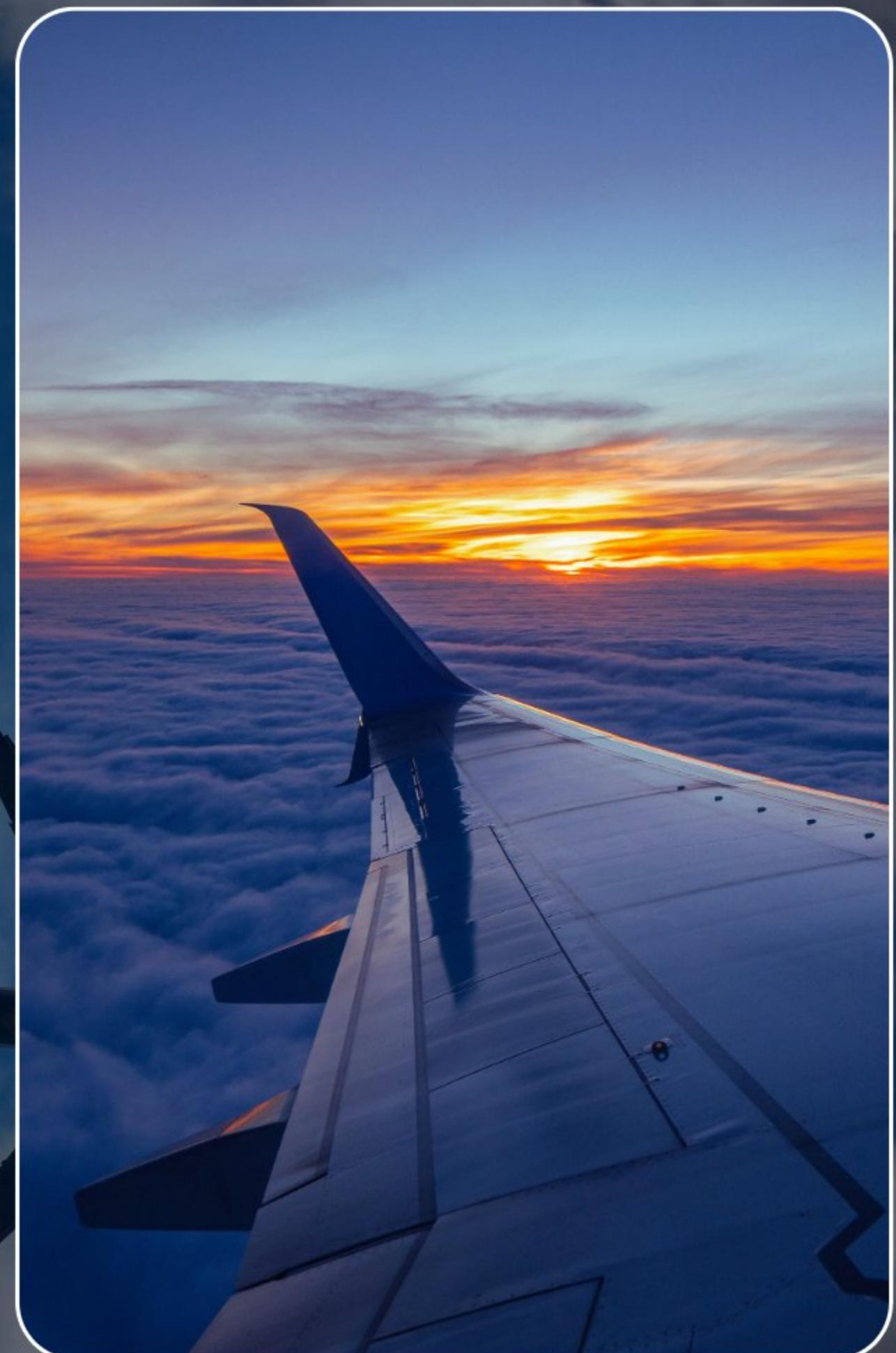


Final Project

U.S. AIRLINE PERFORMANCE & DELAY ANALYSIS

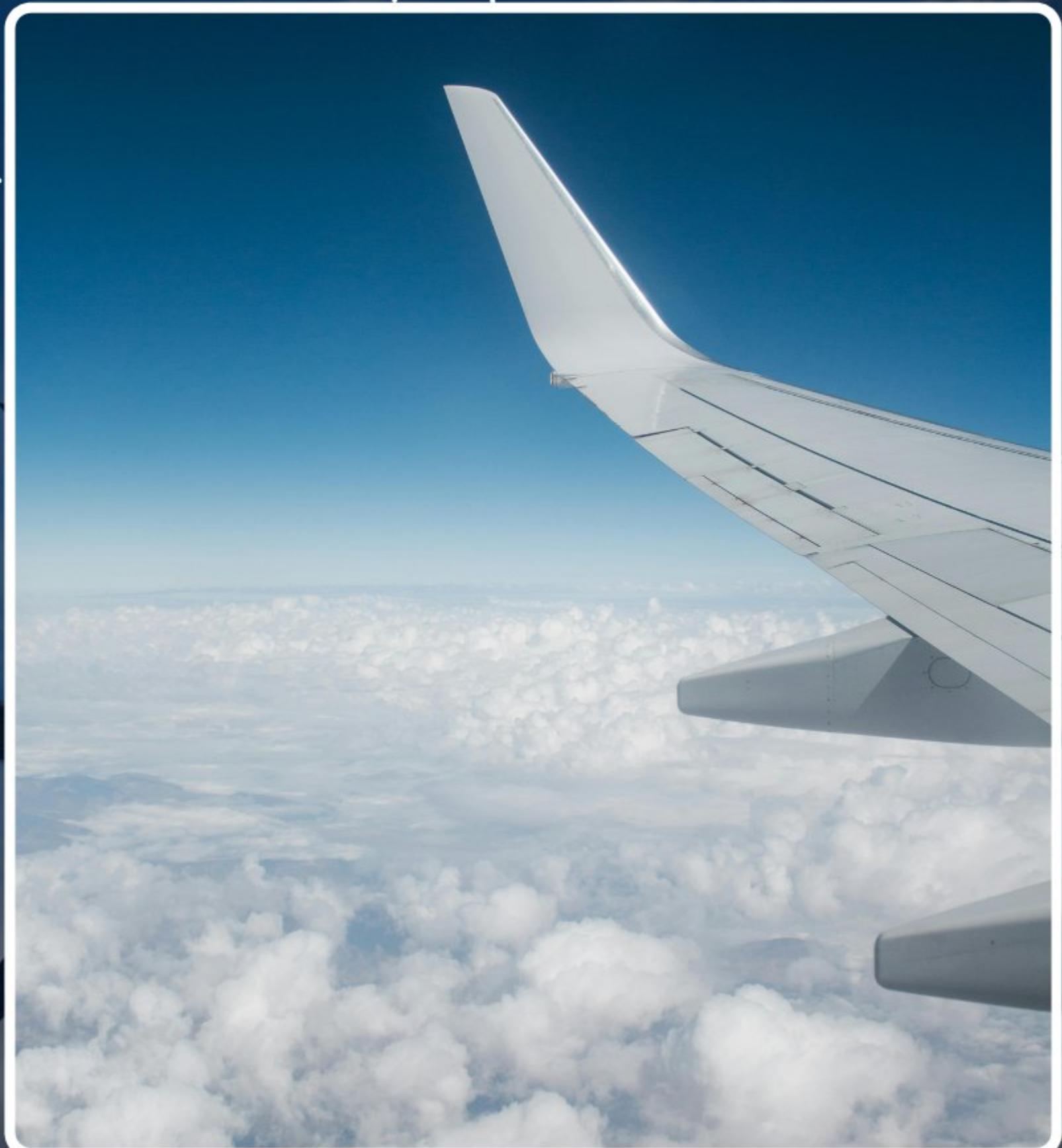
By,
Karthikeyan.S

verti0077@gmail.com
9597150808



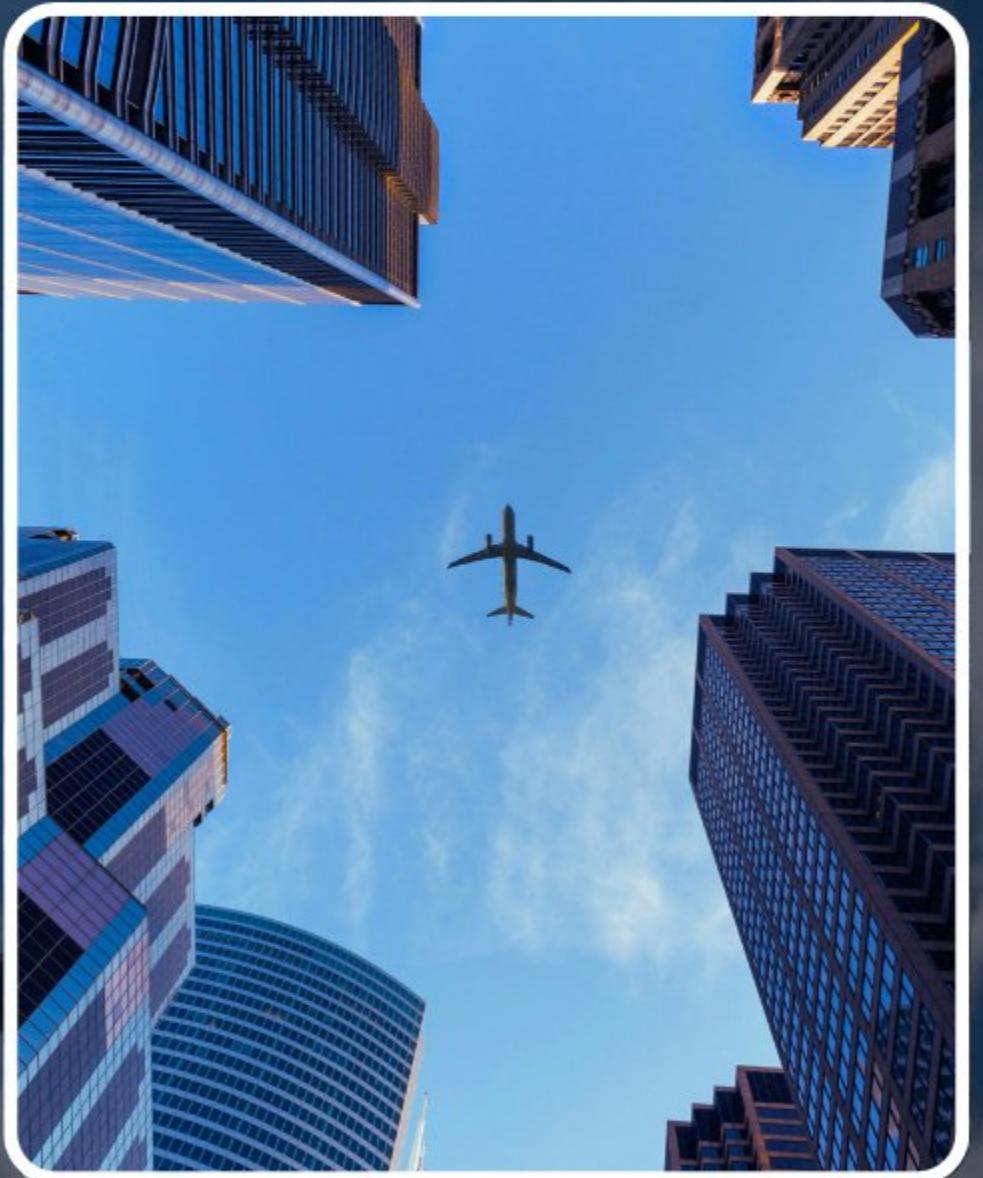
PROBLEM STATEMENT

- Flight delays and cancellations are significant issues in the U.S. aviation industry, impacting passengers, airlines, and the economy.
- Task is to perform an in-depth analysis of historical flight data to identify key drivers of these disruptions, assess performance, and propose actionable insights.



OBJECTIVE

- 1) Analyze the primary causes and patterns of flight delays and cancellations.
- 2) Benchmark the on-time performance, delay severity, and cancellation rates of different airlines.
- 3) Evaluate the operational performance of various U.S. airports.
- 4) Investigate how factors like time of day, day of week, month, and route affect flight operations.
- 5) Translate my findings into meaningful recommendations for stakeholders.



First Page of my dashboard

GUVI FINAL PROJECT - AVIATION ANALYSIS

AIRLINES PERFORMANCE

Airlines Performance

Evaluates how different airlines perform across On-Time Arrival Rate, Cancellation Rate, Delay Distribution, Avg Arrival Rate, Avg Air Time ,Trip Distance.

AIRPORTS PERFORMANCE

Airports Performance

Section focuses on Avg Arrival & Departure Delay Trends, Taxi Time Efficiency, Cancelled & Diverted Flights,Top 10 Busy Airports Performance

TIME ANALYSIS

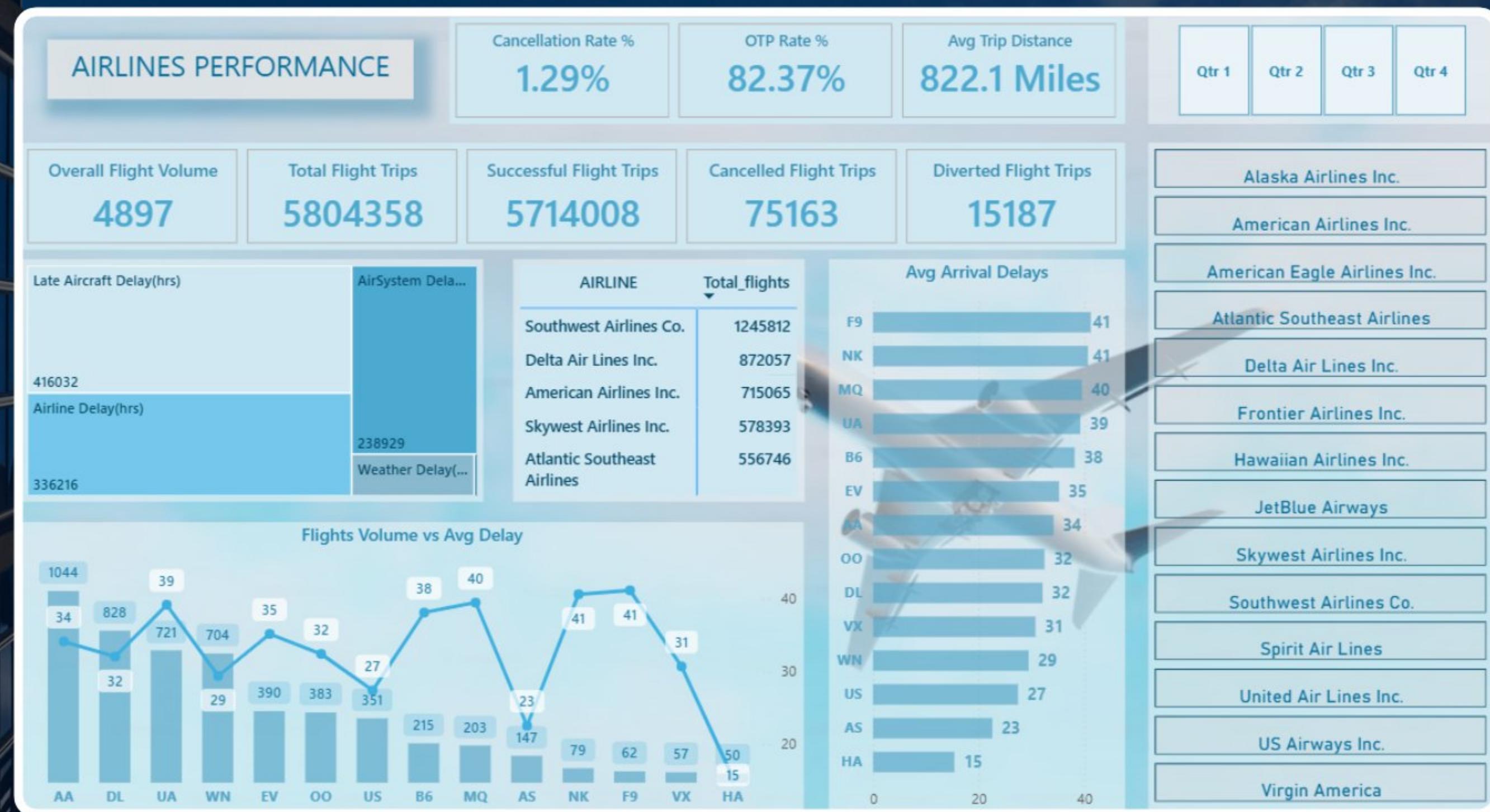
Time Analysis

Matrix patterns reveals when delays and cancellations are most likely happen, Hourly Delay Trends, Day-of-Week Analysis, Quarterly Performance

Mail Id - verti0077@gmail.com

Mobile No - 9597150808

Second Page of my dashboard



Third Page of my dashboard

AIRPORTS PERFORMANCE

Avg_Departure_Delays	Avg_Arrival_Delays	Taxi Time Efficiency	OTP Rate %	Cancellation Rate %
32.6 Mins	33.1 Mins	23.2 Mins	82.37%	1.29%

Top 10 Airports	TFT	AAD (mins)	ADD (mins)	TTE (mins)	CFT	DFT
Hartsfield-Jackson Atlanta International Airport	346638	31.05	29.30	23.11	2359	773
Chicago O'Hare International Airport	284634	39.57	37.08	25.35	7298	782
Dallas/Fort Worth International Airport	239042	34.17	33.35	23.14	5745	650
Denver International Airport	195486	32.91	30.43	22.54	1554	530
Los Angeles International Airport	194292	30.51	30.40	23.81	1783	506
San Francisco International Airport	147504	33.87	34.15	25.19	1644	369
Phoenix Sky Harbor International Airport	146550	26.86	26.74	21.94	637	361
George Bush Intercontinental Airport	146017	32.98	32.91	24.65	1525	473
McCarran International Airport	133022	31.76	30.94	21.75	777	308
Minneapolis-Saint Paul International Airport	112021	30.05	30.60	24.51	712	254

Qtr 1	Qtr 2	Qtr 3	Qtr 4
STATE	CITY	DFT	CFT
All	All		

Distributions of Cancellation Reasons

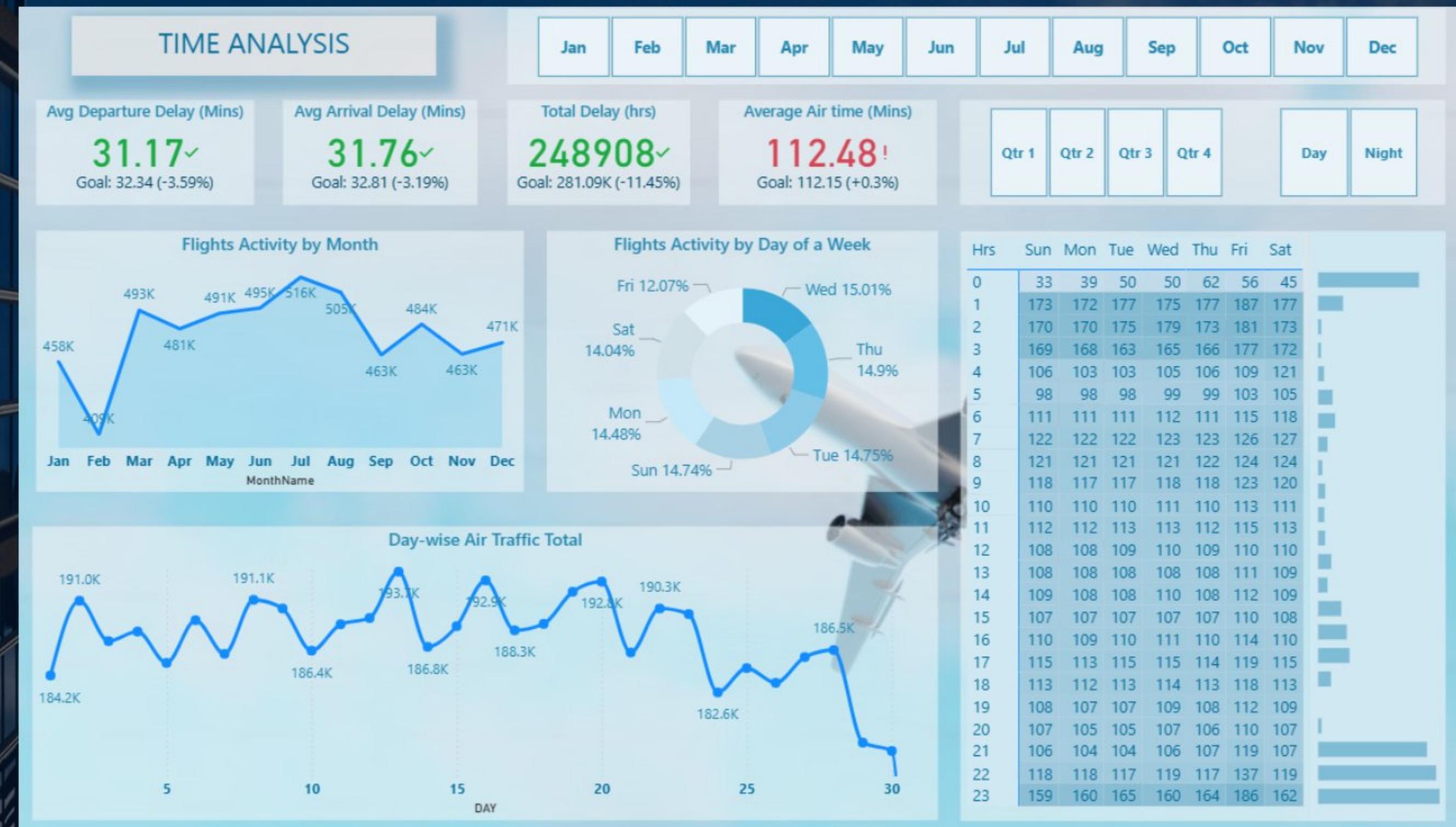
CC_AIRPORT_NAME.AIRPORT	DFT
Chicago O'Hare International Airport	782
Hartsfield-Jackson Atlanta International Airport	773
Dallas/Fort Worth International Airport	650
Denver International Airport	530
Los Angeles International Airport	506

Avg Arrival Delays

Avg Arrival Delays	
Chicago O'Hare International Airport	40
Dallas/Fort Worth International Air...	34
San Francisco International Airport	34
George Bush Intercontinental Airport	33
Denver International Airport	33
McCarran International Airport	32
Hartsfield-Jackson Atlanta Internati...	31
Los Angeles International Airport	31
Minneapolis-Saint Paul Internationa...	30
Phoenix Sky Harbor International A...	27

ADD - Avg Arrival Delay ; ADD - Avg Departure Delay ; TTE - Taxi Time Efficiency ; CFT - Cancelled Flight Trips ; DFT - Diverted Flight Trips

Fourth Page of my dashboard





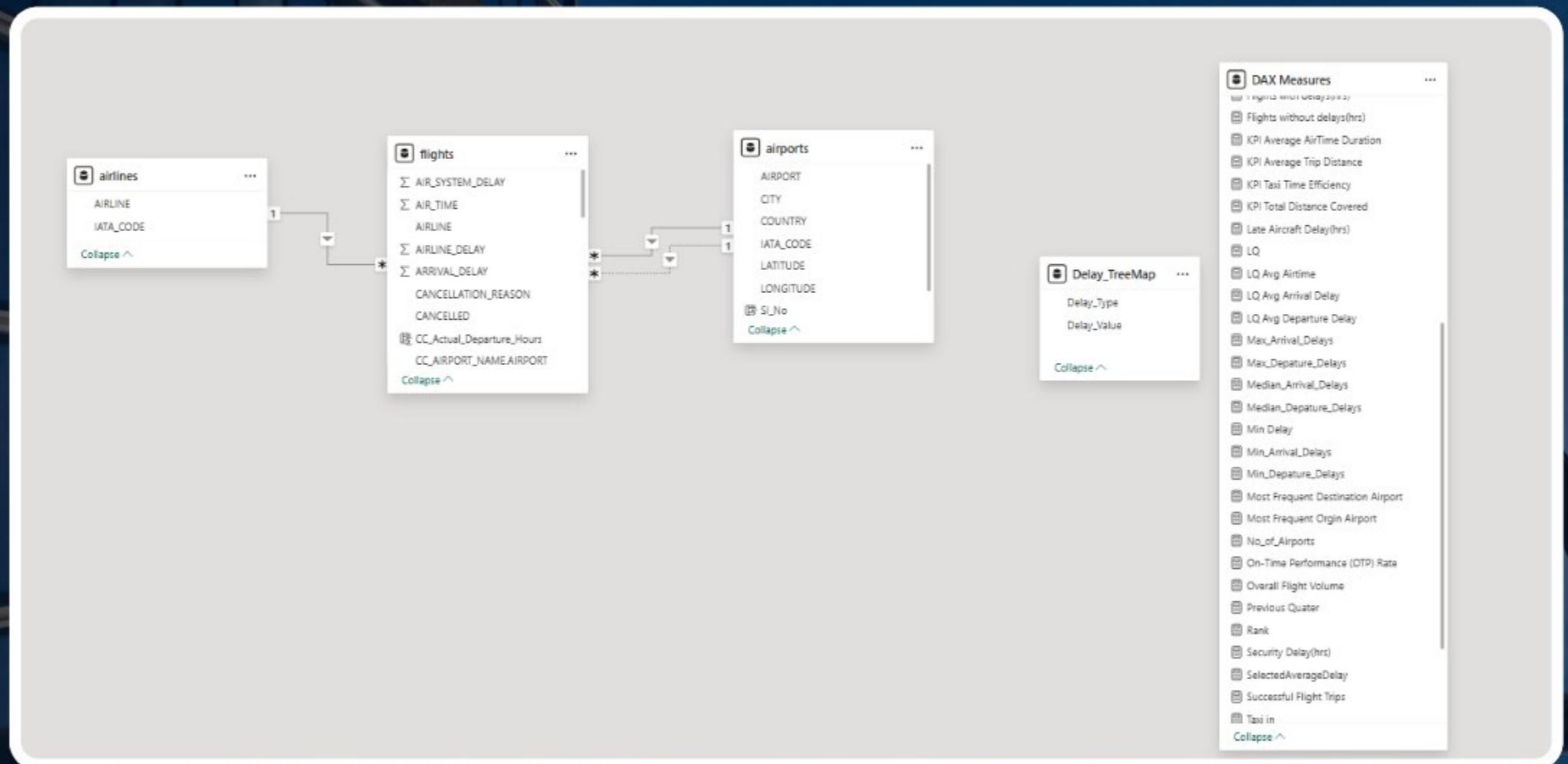
Technical Tools used :

- MS - EXCEL

- SQL

- POWER BI

Dataset Overview



3 tables – Airlines, Airports & Flights

- Airlines – 2 columns, 14 rows.
- Airports – 7 columns, 322 rows.
- Flights – 33 columns, 58,19,079 rows.

Data Ingestion

- Initial Verification in excel.
- Finding Primary Keys.
- Writing down Appropriate Data Types.
- In Excel - 10,48,576.
- In SQL - 6,59,664.
- In Power Bi - Entire Dataset.



Data Cleaning

- 1)Appropriate Data Types - Whole Number,Text,Time,Date etc.
- 2)Handling Null Values - Removed & Replaced .
 - 2.1)Removed - Tail Number rows - blank..
 - 2.2)Replaced - Null, Black values with 0 & 'Not known'.
 - 2.2.1)Zero for - Departure Delay, Taxi out, Scheduled Time etc.
 - 2.2.2)Not known for Cancelation Reason.
- 3)Created Calculated Columns - Departure Date, Departure Delay Desc, Airport Name, Actual Departure Airport Name, Month name, Quater, Day/Night..
- 4)Formatted Columns - 0005 → 00:05 hh:mm - Scheduled departure, Actual departure, Wheels OFF, Wheels ON, Scheduled Arrival, Actual Arrival.

Key Findings from Airlines Performance



Key Findings - Airlines Performance

- **Total Flight Trips:** 5.8M indicates substantial air traffic volume.
- **Successful Flight Trips:** 5.7M with only 75,163 cancellations confirms robust flight execution.
- **Diverted Trips:** 15,187 – suggests minor rerouting but within control limits.

Overall Flight Volume	Total Flight Trips	Successful Flight Trips	Cancelled Flight Trips	Diverted Flight Trips
4897	5804358	5714008	75163	15187

Key Findings - Airlines Performance

- **Low Cancellation Rate:** At just 1.29%, this indicates overall a strong operational reliability.
- **High On-Time Performance (OTP):** 82.37% reflects that most flights are adhering to schedules.
- **Average Trip Distance** - 822 Miles

Cancellation Rate %

1.29%

OTP Rate %

82.37%

Avg Trip Distance

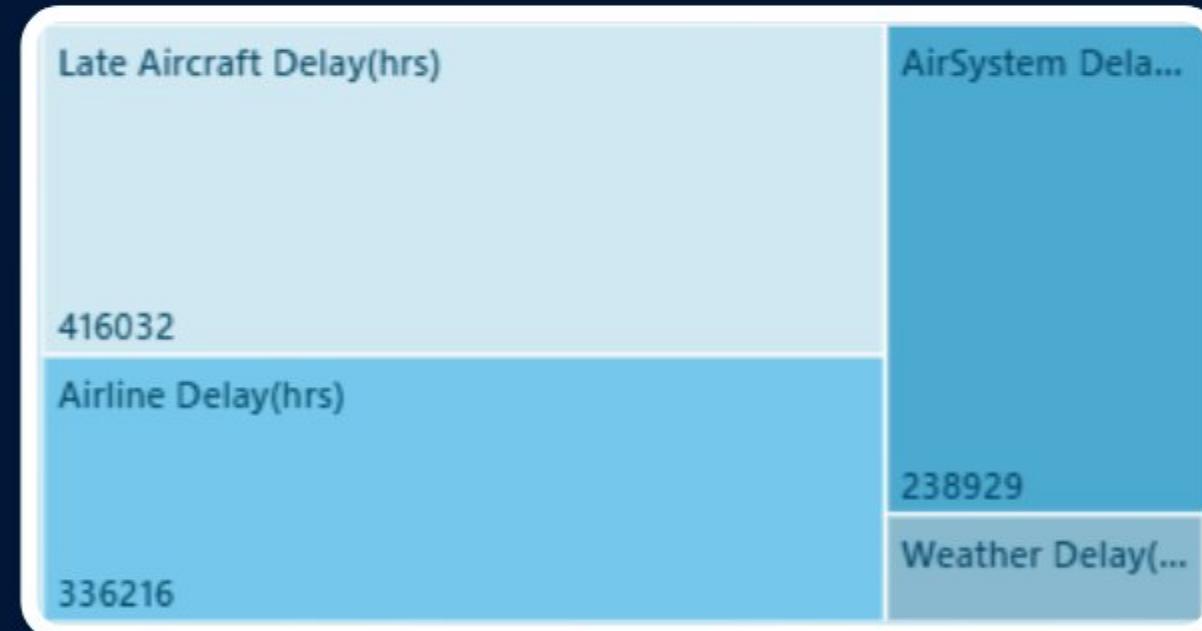
822.1 Miles

Key Findings - Airlines Performance

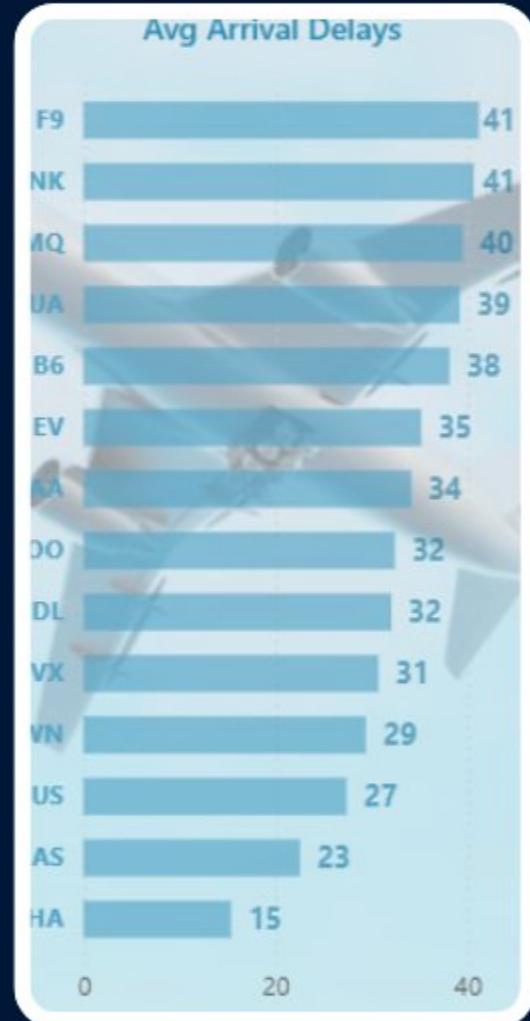


- **Delay Attribution Insights:**

- Late Aircraft, Airline-related issues, Air system delay dominate delay reasons.
- Weather Delay and Security delays are minimal, suggesting external infrastructure isn't a major bottleneck.
- Airlines with Longer Flight Distances tend to show slightly higher arrival delays, indicating possible routing or turnaround inefficiencies.



Key Findings - Airlines Performance



Avg Arrival Delays

- Frontier Airlines and Spirit Airlines lead in average delays – potential areas for strategic improvement.
- Hawaiian Airlines shows best delay control, a best practice benchmark.

Key Findings - Airlines Performance



Area to be Focused

American Eagle Airlines has highest rate and lowest avg trip distance, Avg Delay - 40 mins.

Cancellation Rate %

5.08%

OTP Rate %

80.10%

Avg Trip Distance

422.3 Miles

Overall Flight Volume

203

Total Flight Trips

294584

Successful Flight Trips

278791

Cancelled Flight Trips

14977

Diverted Flight Trips

816

Key Findings from Airports Performance



Key Findings - Airports Performance



Performance Metrics :

- **Average Departure Delay:** 32.6 minutes → Indicates moderate congestion affecting both ends of flight operations.
- **Average Arrival Delay:** 33.1 minutes. Causing inconvenience to passengers.
- **Taxi Time Efficiency (TTE):** 23.2 minutes → Suggests relatively smooth gate-to-runway processes, though some airports show higher values.

Avg_Departure_Delays	Avg_Arrival_Delays	KPI Taxi Time Efficiency
32.6 Mins	33.1 Mins	23.2 Mins

Key Findings - Airports Performance



Analyzed the behaviour of TOP 10 busiest airports in diff metrices.

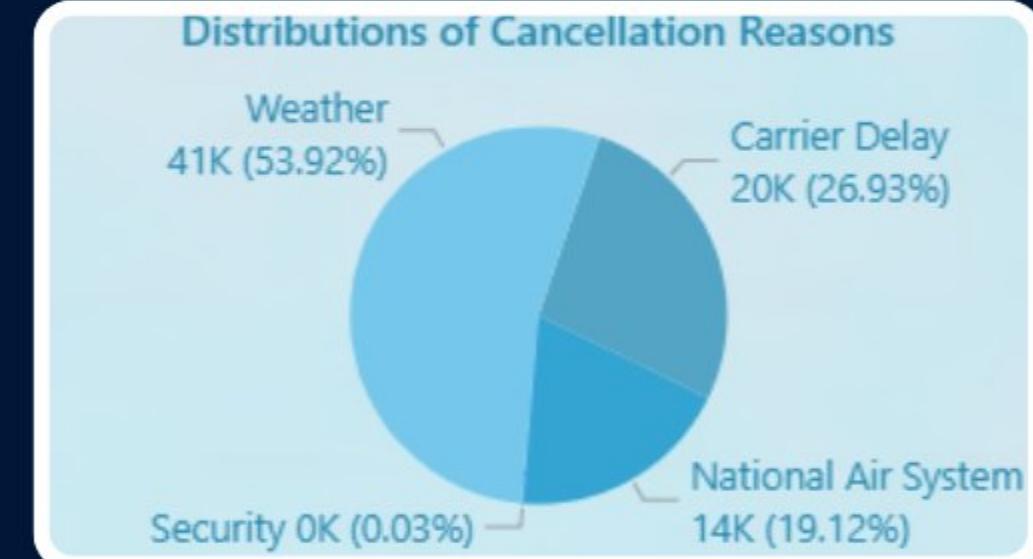
Top 10 Airports	TFT	AAD (mins)	ADD (mins)	TTE (mins)	CFT	DFT
Hartsfield-Jackson Atlanta International Airport	346638	31.05	29.30	23.11	2359	773
Chicago O'Hare International Airport	284634	39.57	37.08	25.35	7298	782
Dallas/Fort Worth International Airport	239042	34.17	33.35	23.14	5745	650
Denver International Airport	195486	32.91	30.43	22.54	1554	530
Los Angeles International Airport	194292	30.51	30.40	23.81	1783	506
San Francisco International Airport	147504	33.87	34.15	25.19	1644	369
Phoenix Sky Harbor International Airport	146550	26.86	26.74	21.94	637	361
George Bush Intercontinental Airport	146017	32.98	32.91	24.65	1525	473
McCarran International Airport	133022	31.76	30.94	21.75	777	308
Minneapolis-Saint Paul International Airport	112021	30.05	30.60	24.51	712	254

- **Congestion vs. Delay:** Airports like 2 and 3 show a clear link between high traffic and elevated delays and cancellations. These hubs might be nearing infrastructure or scheduling thresholds.
- **Taxi Time Impact on Delays:** Airports with longer taxi times (2 & 6) also have higher departure delays. Suggests gate allocation and runway sequencing may be contributing factors.
- **Flight Volume ≠ Delay Severity:** While 1 has the most flights, its delay and cancellation metrics are more moderate than 2 or 3.

Key Findings – Airports Performance

Distributions of Cancellation Reasons

- Weather dominates cancellations, reinforcing the need for resilient schedules and dynamic re-routing.
- Carrier and NAS issues remain significant, pointing to internal and systemic challenges.



Key Findings – Airports Performance

- Consistent OTP above 80% across major airports — indicates successful prioritization of schedule integrity.
- Arrival and departure delays trend upward in larger hubs — emphasizing a correlation between traffic density and operational lag.

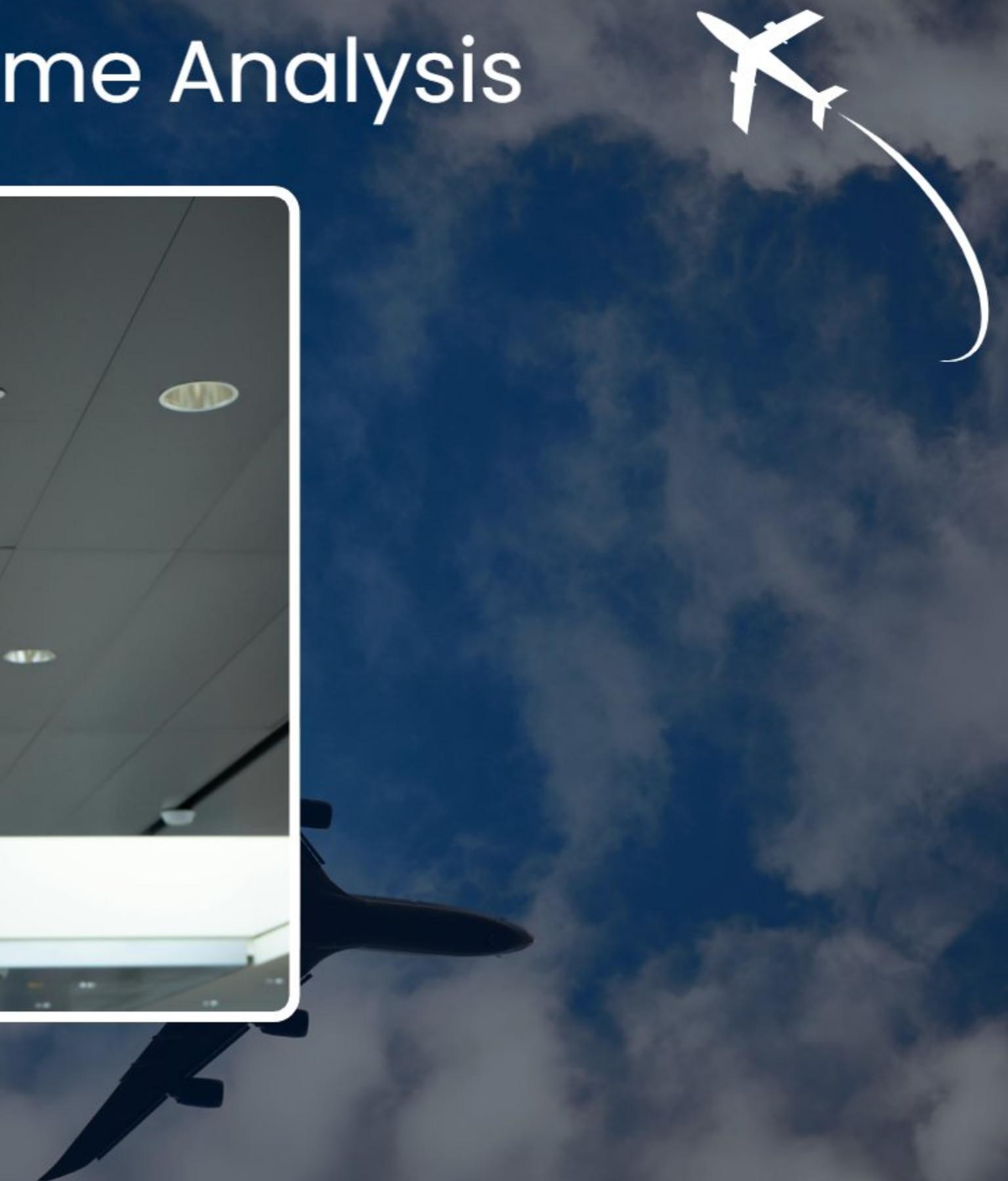
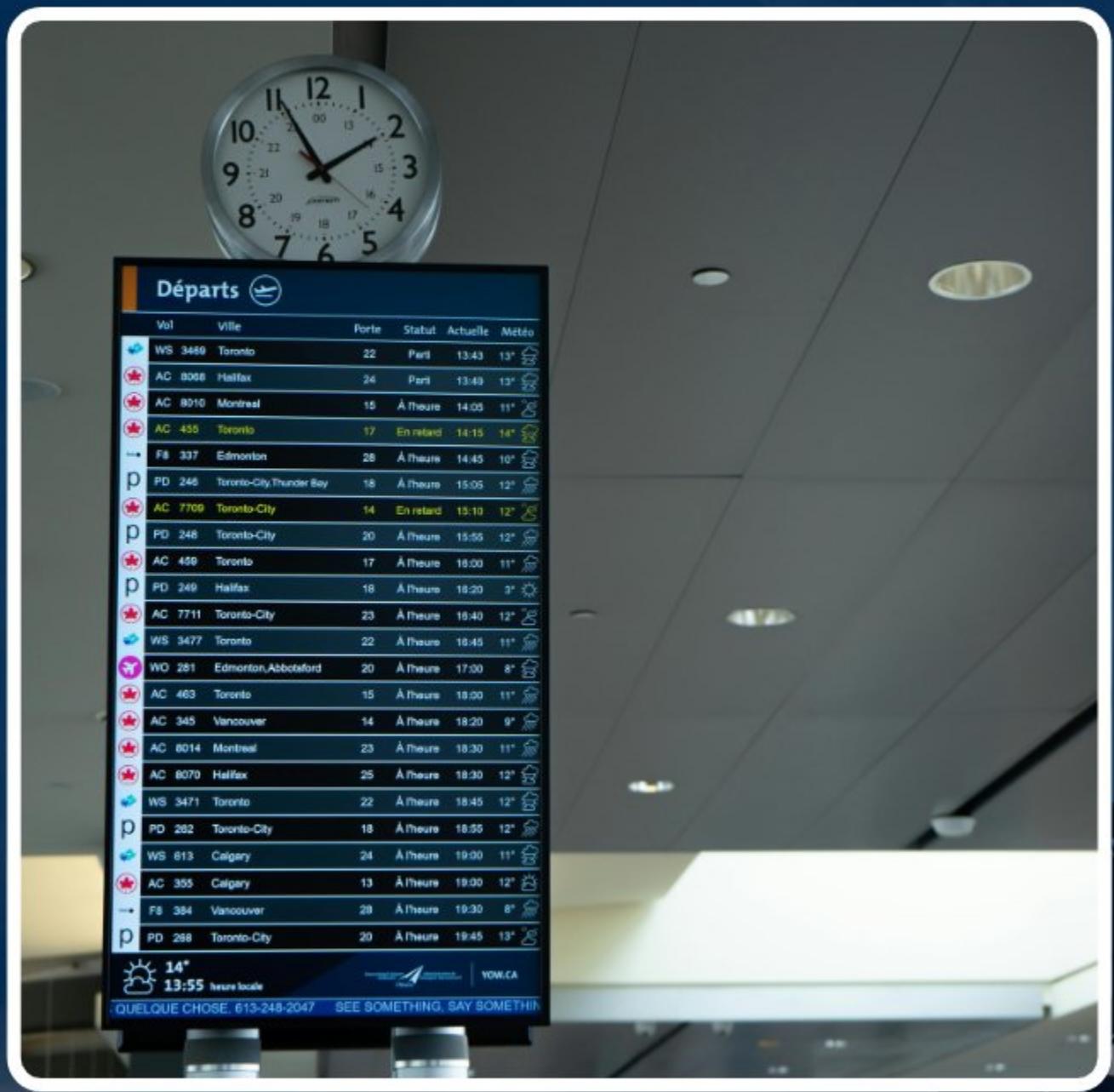


Key Findings – Airports Performance

- Taxi time is inversely related to OTP – airports with lower taxi time often exhibit better punctuality.
- Weather-related cancellations are seasonal risks – high during winter and monsoon months; warrants preventive contingency planning.



Key Findings from Time Analysis



Key Findings - Time Analysis

Flights Activity by Month :

- **Flights Peak by Month:** March, June, and July show seasonal spikes, each exceeding 500K flights.
- **Lowest Activity Months:** January (458K) and February (409K) suggest post-holiday lulls.

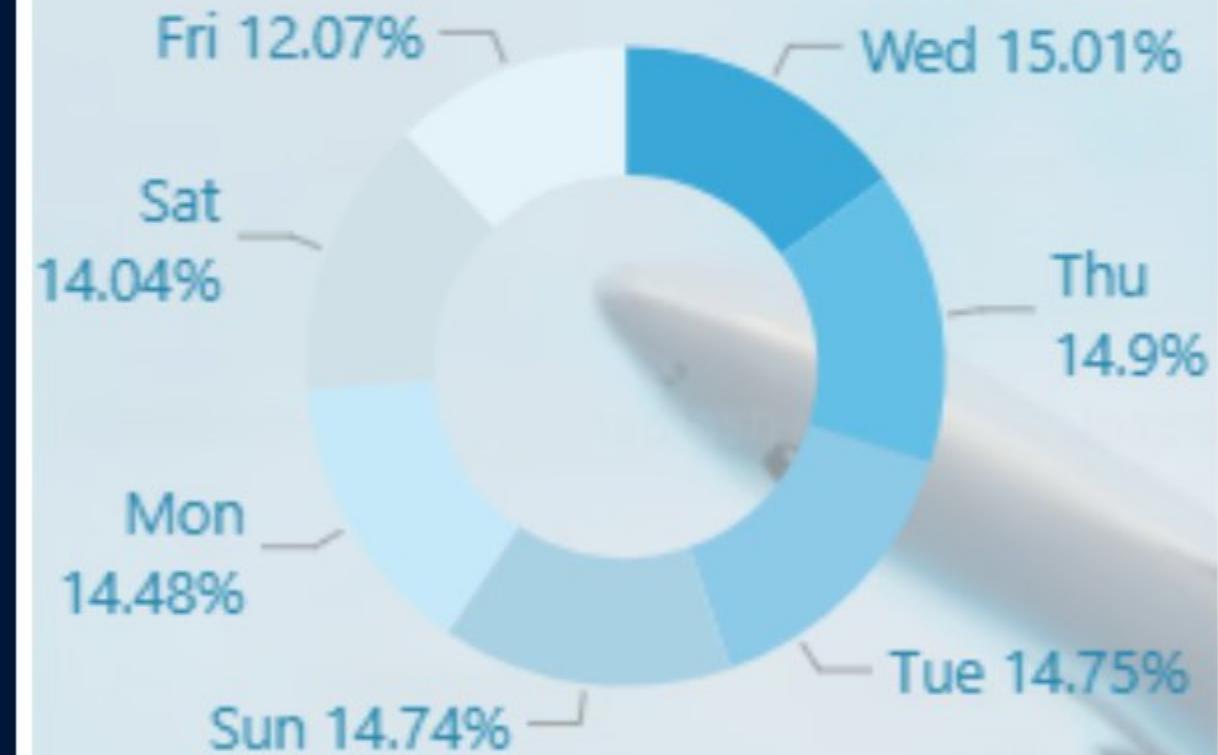


Key Findings - Time Analysis

Flights by Day of Week:

- Wednesday (15.01%) is the busiest, followed closely by Tuesday and Sunday.
- Friday (12.07%) sees the lowest flight activity—possibly a strategic lull before weekend surges.

Flights Activity by Day of a Week



Key Findings - Time Analysis

Hourly Behavior :

- Avg Air Time (mins) is high in early morning (1–3 AM):
 - Friday at 1 AM has the highest Avg Air Time (187 mins).
- Air time Activity begins tapering in morning (6-9 AM), ramping again in late night (22-23 PM) – suggests three-peak scheduling.

Hrs	Sun	Mon	Tue	Wed	Thu	Fri	Sat
0	33	39	50	50	62	56	45
1	173	172	177	175	177	187	177
2	170	170	175	179	173	181	173
3	169	168	163	165	166	177	172
4	106	103	103	105	106	109	121
5	98	98	98	99	99	103	105
6	111	111	111	112	111	115	118
7	122	122	122	123	123	126	127
8	121	121	121	121	122	124	124
9	118	117	117	118	118	123	120
10	110	110	110	111	110	113	111
11	112	112	113	113	112	115	113
12	108	108	109	110	109	110	110
13	108	108	108	108	108	111	109
14	109	108	108	110	108	112	109
15	107	107	107	107	107	110	108
16	110	109	110	111	110	114	110
17	115	113	115	115	114	119	115
18	113	112	113	114	113	118	113
19	108	107	107	109	108	112	109
20	107	105	105	107	106	110	107
21	106	104	104	106	107	119	107
22	118	118	117	119	117	137	119
23	159	160	165	160	164	186	162

Key Findings - Time Analysis



Peak Traffic Days:

- Day 15: 193.1K – highest traffic recorded.
- Day 20: 192.3K – closely follows as another major surge.
- Day 5: 191.1K – early-period spike worth monitoring.

Mid-Month Surge: Days 10–20 consistently outperform early and late month totals, indicating seasonal or operational escalation.



Key data-driven recommendations to potential stakeholders



Key Recommendation :

Target Delay Reduction in Top Categories:

- Focus on Late Aircraft (416K hrs) and Airline-driven delays (336K hrs).
- Introduce tighter gate scheduling, predictive maintenance, and faster turnaround protocols.

Late Aircraft Delay(hrs)

416032

Airline Delay(hrs)

336216

AirSystem Dela...

238929

Weather Delay(...



Key Recommendation :

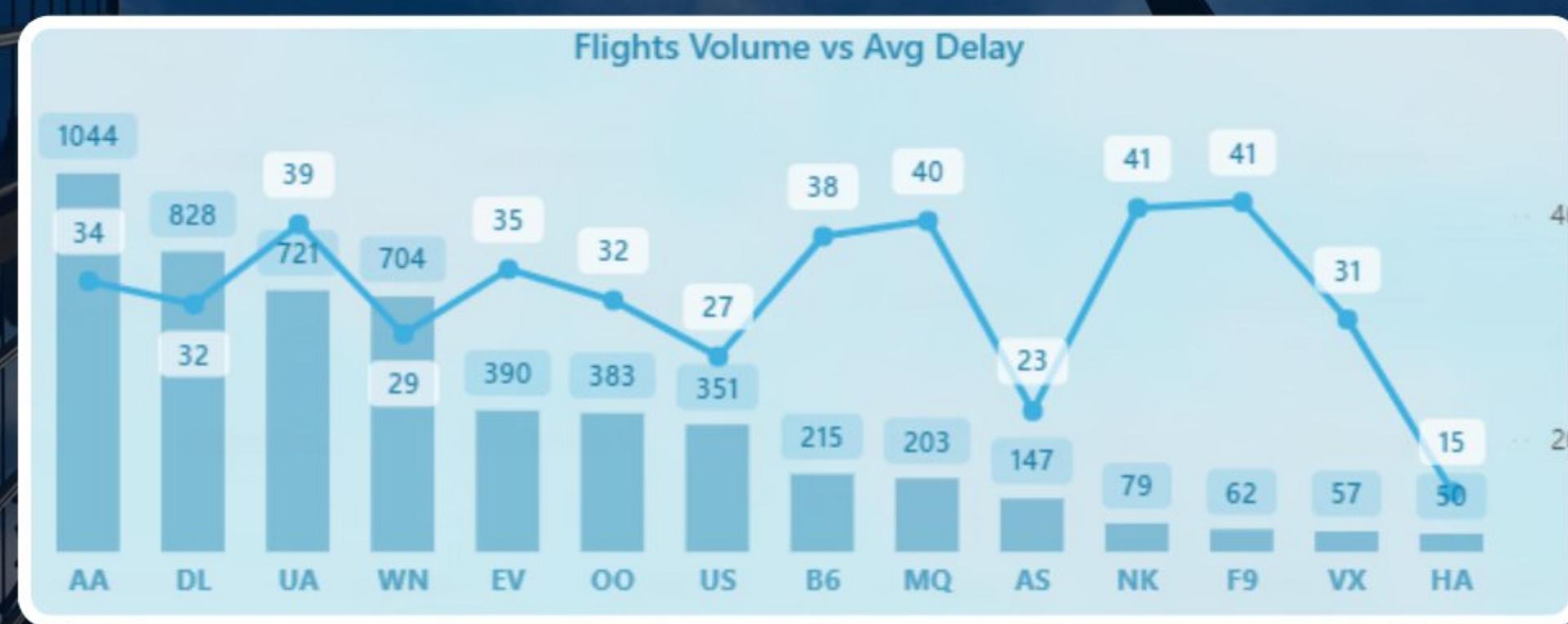


Flights volume vs Avg Delay:

- **Inverse Delay-to-Volume Trend:** As flight count decreases, delay tends to rise—possibly due to shorter turnaround buffers or fewer scheduling options.

For High-Volume Airlines (AA, UA, DL):

- UA should audit delay hotspots—given 39 mins average.
- DL should reinforce delay-prevention protocols to match WN's low numbers.



Key Recommendation:

- 2 & 3 should audit taxi routing, gate allocations, and runway to reduce TTE and associated departure delays.
- PHX's playbook on minimizing delays and taxi times should be studied and replicated in similar-sized hubs.

Top 10 Airports	TFT	AAD (mins)	ADD (mins)	TTE (mins)	CFT	DFT
Hartsfield-Jackson Atlanta International Airport	346638	31.05	29.30	23.11	2359	773
Chicago O'Hare International Airport	284634	39.57	37.08	25.35	7298	782
Dallas/Fort Worth International Airport	239042	34.17	33.35	23.14	5745	650
Denver International Airport	195486	32.91	30.43	22.54	1554	530
Los Angeles International Airport	194292	30.51	30.40	23.81	1783	506
San Francisco International Airport	147504	33.87	34.15	25.19	1644	369
Phoenix Sky Harbor International Airport	146550	26.86	26.74	21.94	637	361
George Bush Intercontinental Airport	146017	32.98	32.91	24.65	1525	473
McCarran International Airport	133022	31.76	30.94	21.75	777	308
Minneapolis-Saint Paul International Airport	112021	30.05	30.60	24.51	712	254

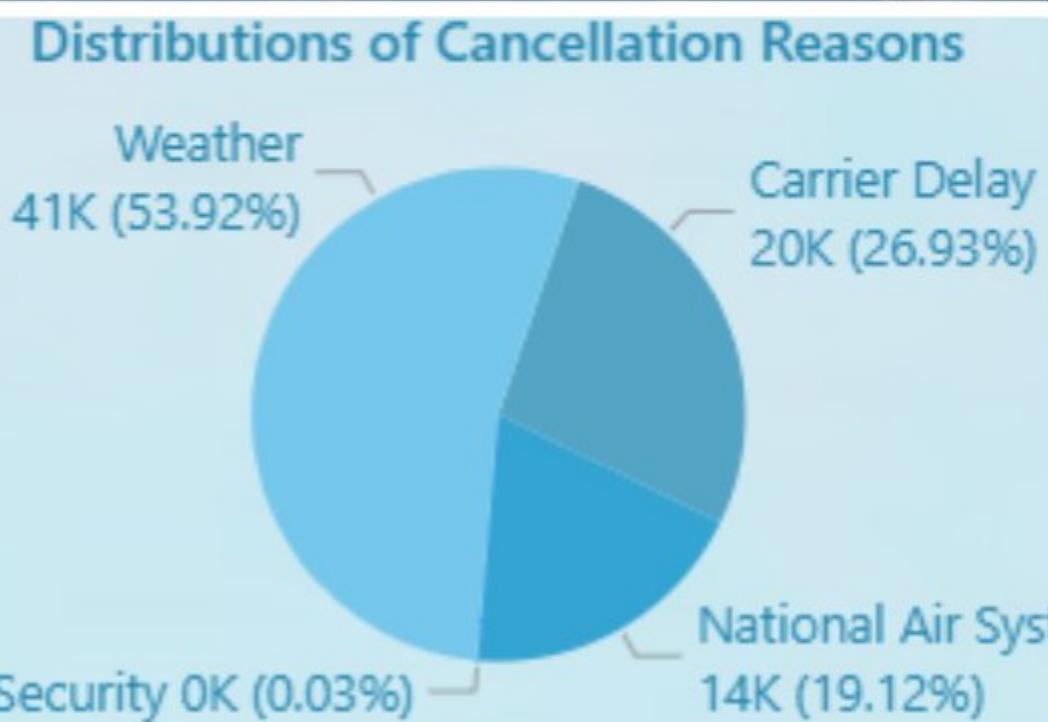
Key Recommendation:

Weather-Responsive Planning:

- With 53.92% of cancellations caused by weather, airlines should invest in:
 - Dynamic re-routing tools.
 - Real-time weather prediction models.
 - Contingency slot allocations at alternate airports.

Carrier Delay Mitigation:

- Responsible for 26.93% of cancellations
- Action points:
 - Improve crew scheduling buffers
 - Standardize aircraft readiness checklists
 - Enhance turnaround protocols at high-volume hubs



Key Recommendation:

Boost Capacity in Peak Months:

- **Prioritize Staffing During Early Morning Surge:** 1–3 AM slots maintain 170+ flight counts—ensure gate availability, baggage crews, and fuel services are scaled accordingly.

Late-Night Resilience:

- 22:00–23:00 hours show another surge (160–186 flights).
- Action: Implement late-night operational continuity plans and shift rotations.

Hrs	Sun	Mon	Tue	Wed	Thu	Fri	Sat
0	33	39	50	50	62	56	45
1	173	172	177	175	177	187	177
2	170	170	175	179	173	181	173
3	169	168	163	165	166	177	172
4	106	103	103	105	106	109	121
5	98	98	98	99	99	103	105
6	111	111	111	112	111	115	118
7	122	122	122	123	123	126	127
8	121	121	121	121	122	124	124
9	118	117	117	118	118	123	120
10	110	110	110	111	110	113	111
11	112	112	113	113	112	115	113
12	108	108	109	110	109	110	110
13	108	108	108	108	108	111	109
14	109	108	108	110	108	112	109
15	107	107	107	107	107	110	108
16	110	109	110	111	110	114	110
17	115	113	115	115	114	119	115
18	113	112	113	114	113	118	113
19	108	107	107	109	108	112	109
20	107	105	105	107	106	110	107
21	106	104	104	106	107	119	107
22	118	118	117	119	117	137	119
23	159	160	165	160	164	186	162

Key Recomendation :

Boost Capacity in Peak Months:

- July (516K) and August (505K) show peak flight volumes—prioritize crew allocation, aircraft availability, and slot optimization in these months.
- These reflect seasonal travel demand or business cycle spikes.



Key Recomendation :

Conclusion

- By analyzing over 33 key data points—including on-time performance, turnaround efficiency, and air system disruptions, I uncovered both operational inefficiencies and actionable improvement areas.
- My analysis elevates data beyond reporting—towards predictive intelligence and operational strategy.
- Future enhancements could include integrating real-time data feeds, forecasting passenger loads, and benchmarking across airports globally.