**Methodology / Approach**

**Data Cleaning & Preparation**

Removed Duplicates: Duplicate rows were identified and removed to ensure data integrity.

Outlier Detection: Manually examined extreme values in delay fields to identify potential outliers (e.g., unusually high delay minutes).

Month Normalization:, Standardized month field entries, replacing codes with the text label "December" to improve readability.

Categorized Delay Causes: Grouped delay reasons (carrier\_ct, weather\_ct, nas\_ct, etc.) into clear, descriptive categories for analysis.

**Pivot Table Analysis**

Created PivotTables to: Analyze delays by airline and by airport

Calculate average delay minutes per arriving flight:

Formula used: = Total Delay Minutes / Total Arriving Flights

Break down delay causes as a percentage of total delays:

Example formula: = nas\_ct / SUM(all delay types) \* 100

Added a slicer (Year) to make dashboards interactive.

Used report connections to link slicers to multiple charts and pivot tables.

**Visualization & Dashboard**

Built charts to visualize: Average delay minutes per airline and per airport

Percentage of delay causes (e.g., NAS, weather, late aircraft)

Top 10 airports with the highest number of arriving flights and their delay profiles

-Limited airport comparison to those with high traffic volume (e.g., top 10 by arrivals) to ensure fairness and avoid misleading insights from small sample sizes.For example is not fair to compare an airport with 5000 flights arrived with another with 50 flights arrived.

**Issues & Data Limitations**

ExpressJet Airlines LLC: Missing data for 2020.

Missing Critical Fields: No departure or arrival times were provided, which limits deeper time-series or scheduling analyses.

Limited Date Range: Dataset includes only December 2019 and December 2020, reducing trend analysis and seasonal comparisons.

Missing Values: Some fields had nulls or blanks which were handled during cleaning but still impact completeness.

**Focused Questions Answered**

Which airlines experienced the most delays and cancellations in December 2019 and 2020?

The average minutes delayed across different airlines? Which one had the highest average minutes delay?

Which airline demonstrated superior performance, as measured by the lowest average delay time and number of cancellations during that period time ?

What were the main causes of the delays across different airlines and airports?

How did the average minutes of flight delays change across different airlines from December 2019 to December 2020 ?