**Executive Summary**

This project presents a comprehensive data analysis pipeline built on the **2008 U.S. Domestic Flight dataset**, structured according to the **Medallion Architecture** with Bronze, Silver, and Gold layers. The primary objective was to assess airline performance, uncover patterns in delays, and derive actionable insights for operational improvements.

Using over **2.3 million flight records**, we engineered features such as flight status, route identifiers, and enriched the data with airline and aircraft metadata. Key performance metrics such as **on-time performance**, **average arrival delays by carrier**, and **flight frequency by route and day of the week** were derived in the Gold layer.

**Key Findings:**

* **On-Time Performance Rate:** 73.53%
* **Cancellation Rate:** 2.93%
* **Most Frequent Route:** SFO → LAX with 4,686 flights
* **Best Delay Performance:** AQ and HA (Avg. delays near or below 0 mins)
* **Worst Delay Performance:** AA and UA (Avg. delays over 15 mins)
* **Flight Volume Peak Days:** Tuesday to Thursday

This analysis provides a robust foundation for optimizing scheduling, enhancing airline reliability, and supporting data-driven decisions in the aviation industry. The final outputs include clean, transformed data and curated performance metrics, ready for further visualization, modeling, or dashboarding.