**Question:**  
What factors explain arrival delays in flights, and can carrier delays be used as a reliable predictor?

Descriptive Stats:  
I used the describe() function to get summary statistics for ARR\_DELAY and CARRIER\_DELAY, which helped understand the shape and spread of the data:

* ARR\_DELAY had a mean of 63.60 minutes and a maximum of 2692 minutes, showing a wide range of delay times.
* CARRIER\_DELAY had a mean of 18.86 minutes, but 51% of the values were zero, meaning many flights weren’t delayed by the carrier at all.

A screenshot of a computer

AI-generated content may be incorrect.

**Visualization:**A scatter plot was created to show the relationship between carrier delay and arrival delay.

* A positive trend was observed: as carrier delay increases, arrival delays also tend to increase.
* The presence of many zero values for CARRIER\_DELAY suggests that other factors, such as weather or airport congestion, may also influence delays.

A screen shot of a graph

AI-generated content may be incorrect.

Using Ordinary Least Squares (OLS) from statsmodels, we created a linear regression model with CARRIER\_DELAY as the predictor:

y = 47.95 + 0.83 \* CARRIER\_DELAY

* **R-squared: 0.322** – 32.2% of the variation in arrival delays is explained by carrier delays.
* **P-value: 0.000** – The relationship between carrier delay and arrival delay is statistically significant.
* **Intercept (const): 47.95** – Even without any carrier delay, the model predicts an average arrival delay of 47.95 minutes.
* **Coefficient: 0.83** – For every extra minute of carrier delay, the arrival delay increases by about 0.83 minutes.

A screenshot of a computer

AI-generated content may be incorrect.

**Findings:**

The analysis shows that **CARRIER\_DELAY** is an important factor in predicting **ARR\_DELAY**. For each minute of carrier delay, the arrival delay increases by about 0.83 minutes. This means that when the carrier is delayed, the arrival time is also delayed. However, only about 32% of the arrival delay can be explained by carrier delays, so other factors like weather or airport congestion likely also play a role in the delay