DEBRE BERHAN UNIVERSITY

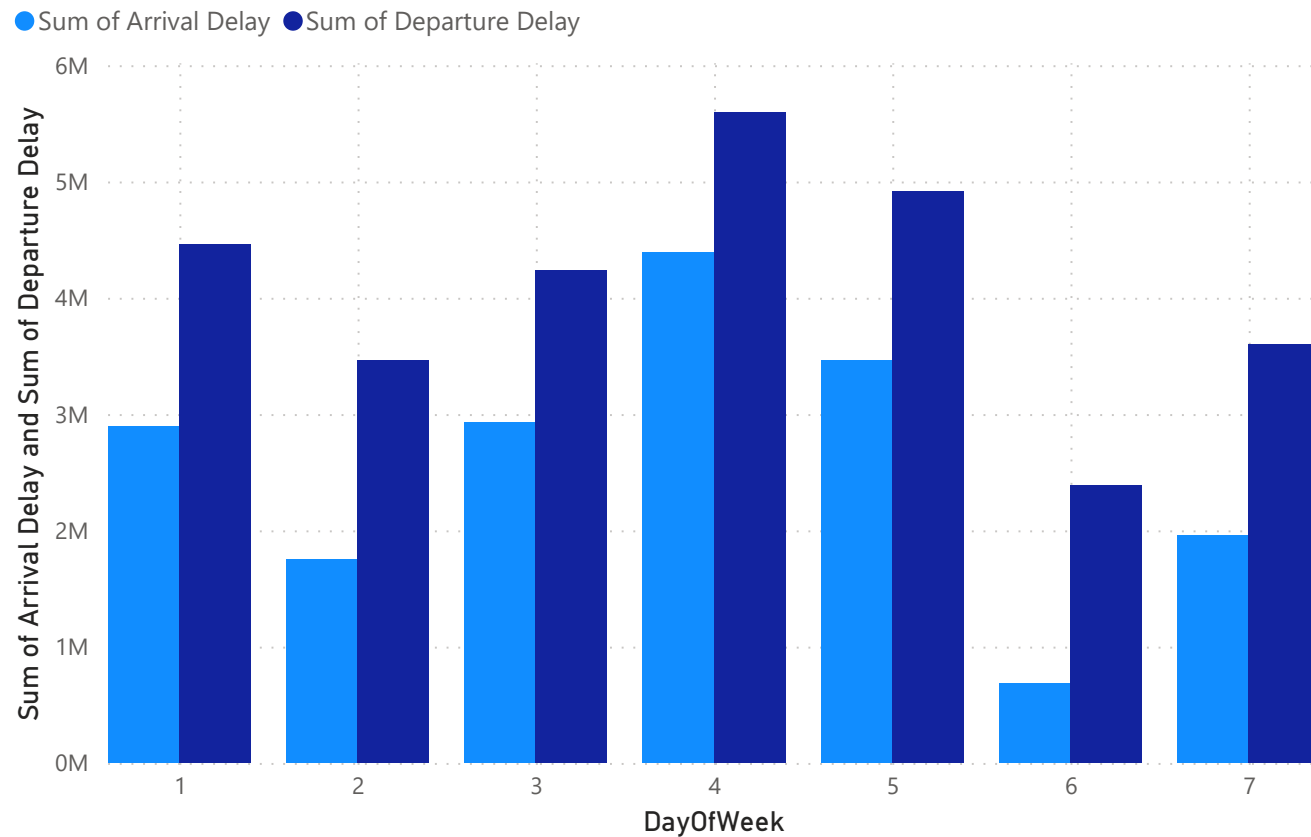## College of Computing

Department of Software Engineering

Fundamentals of Big Data Analysis and BI Group Assignment

Name	ID
1. Bereket Zewude	1594/12
2. Biniyam Sileshi	1286/12
3. Biruk Yosef	1386/12
4. Daniel Kebede	0438/12
5. Dawit Gebeyehu	1399/12
6. Hiro Worku	1400/12

Sum of Arrival Delay and Sum of Departure Delay by DayOfWeek

● Sum of Arrival Delay ● Sum of Departure Delay

- The **graph** represents data related to arrival delays and the sum of departure delays over the days of the week.
- We're interested in understanding how these delays vary across different days.

- X-Axis (Day Of Week): The x-axis is labeled with numbers from 1 to 7, representing days of the week (although it doesn't specify which number corresponds to which day).

- Y-Axis (Sum of Arrival Delay and Sum of Departure Delay): The vertical axis is labeled from 0M to 6M, indicating the scale for measuring the sum of delays.
- Bars: There are two sets of bars:
- Lighter Blue Bars (Arrival Delays): Each bar represents the total arrival delay for a specific day of the week.
- Darker Blue Bars (Departure Delays): Each bar represents the total departure delay for the same day.
- By comparing the heights of the two bars for each day, we can analyze patterns or trends in flight delays throughout a typical week.

Sum of Arrival Delay by Destination AirportID

Destination AirportID
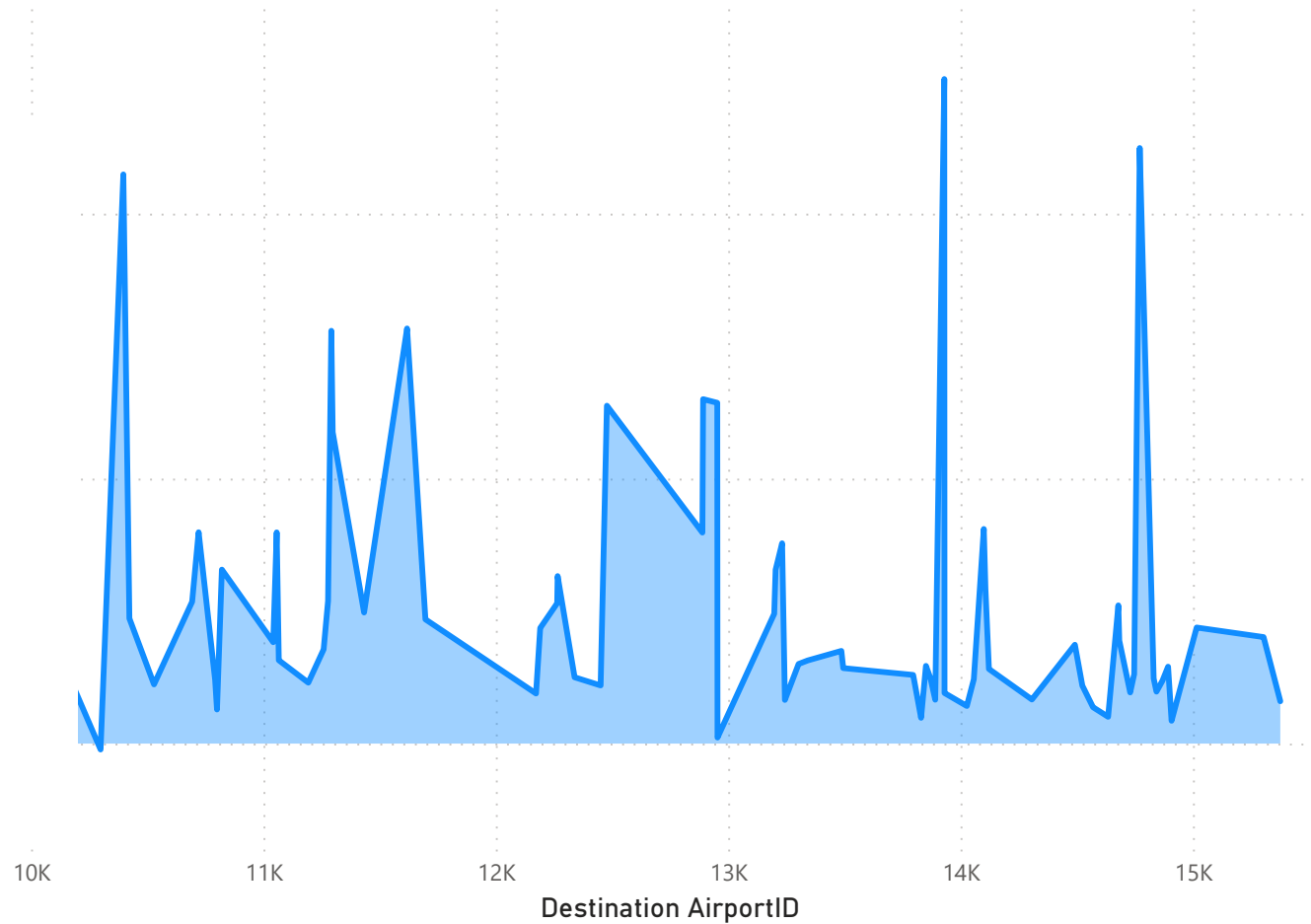
- The **graph** represents the sum of arrival delays at various destination airports.
- We're interested in understanding which airports experience higher or lower arrival delays.

- X-Axis (Destination Airport ID): The x-axis lists destination airport IDs, ranging from 10,000 to 15,000. Each ID corresponds to a specific airport.

- Y-Axis (Sum of Arrival Delay): The vertical axis quantifies the total arrival delays, ranging from 0.0M to 1.0M.
- Blue Bars: Each blue bar corresponds to a specific Destination Airport ID and represents the total arrival delay at that airport.
- Notable **peaks** are observed around IDs 10K, 14K, and close to 15K, indicating significantly higher arrival delays at those airports.
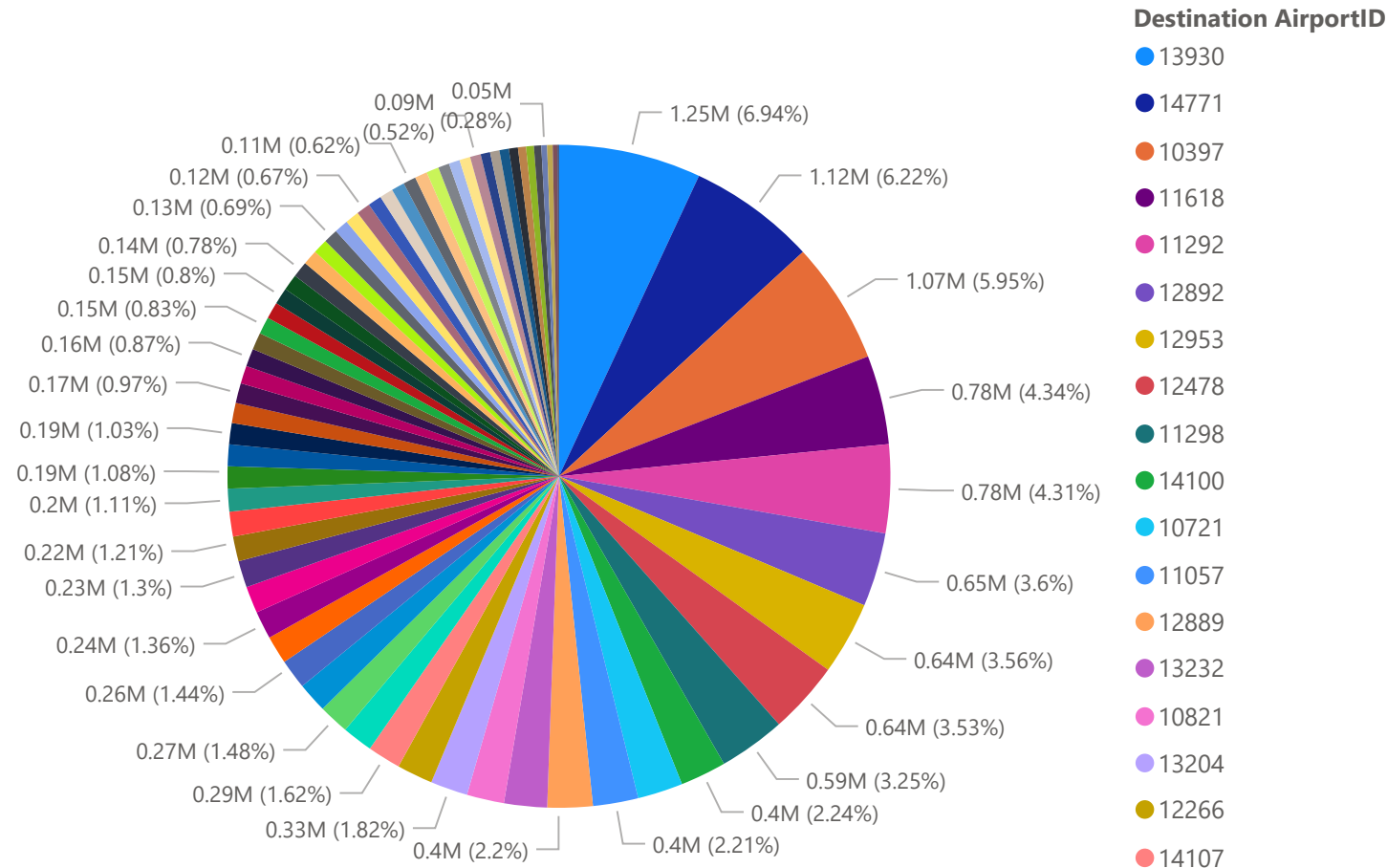
ArrivalDelay and Sum of Departure Delay by DayofMonth
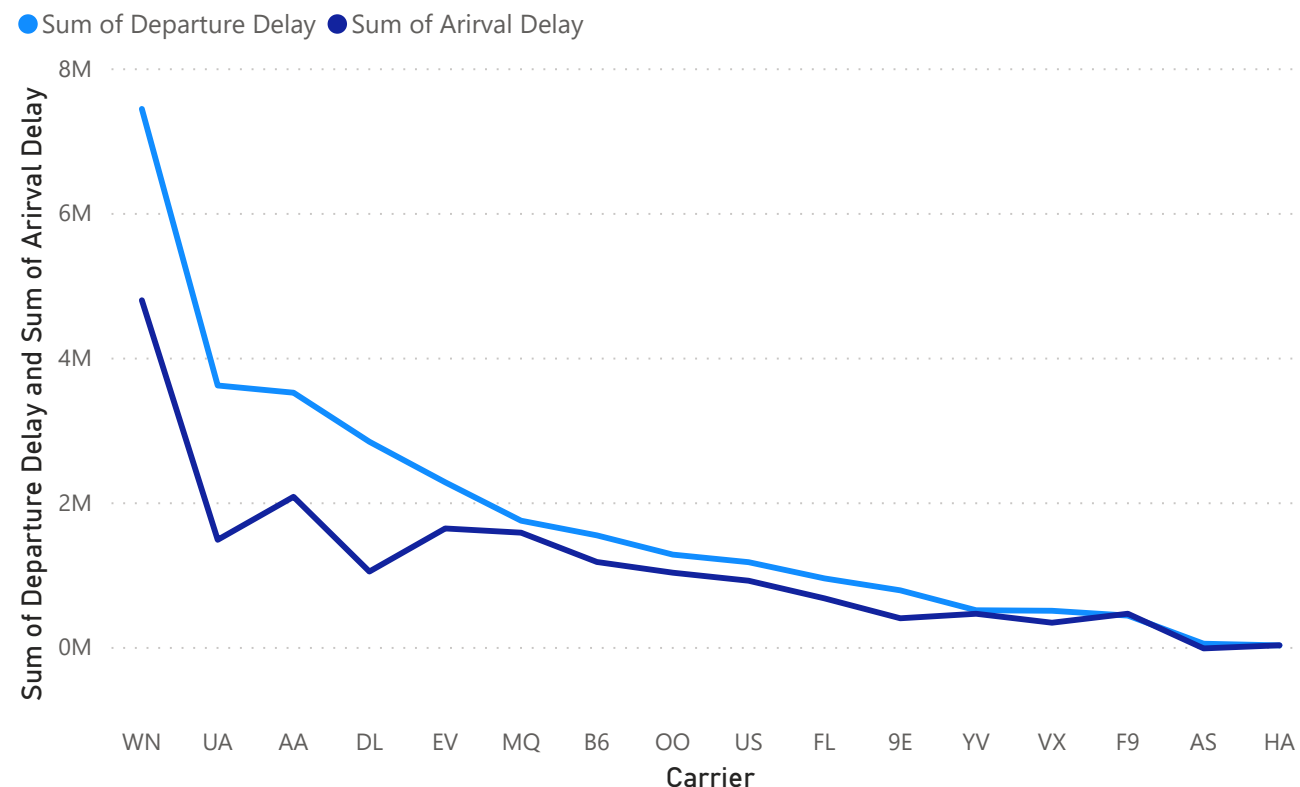
● ArrivalDelay ● Sum of Departure Delay
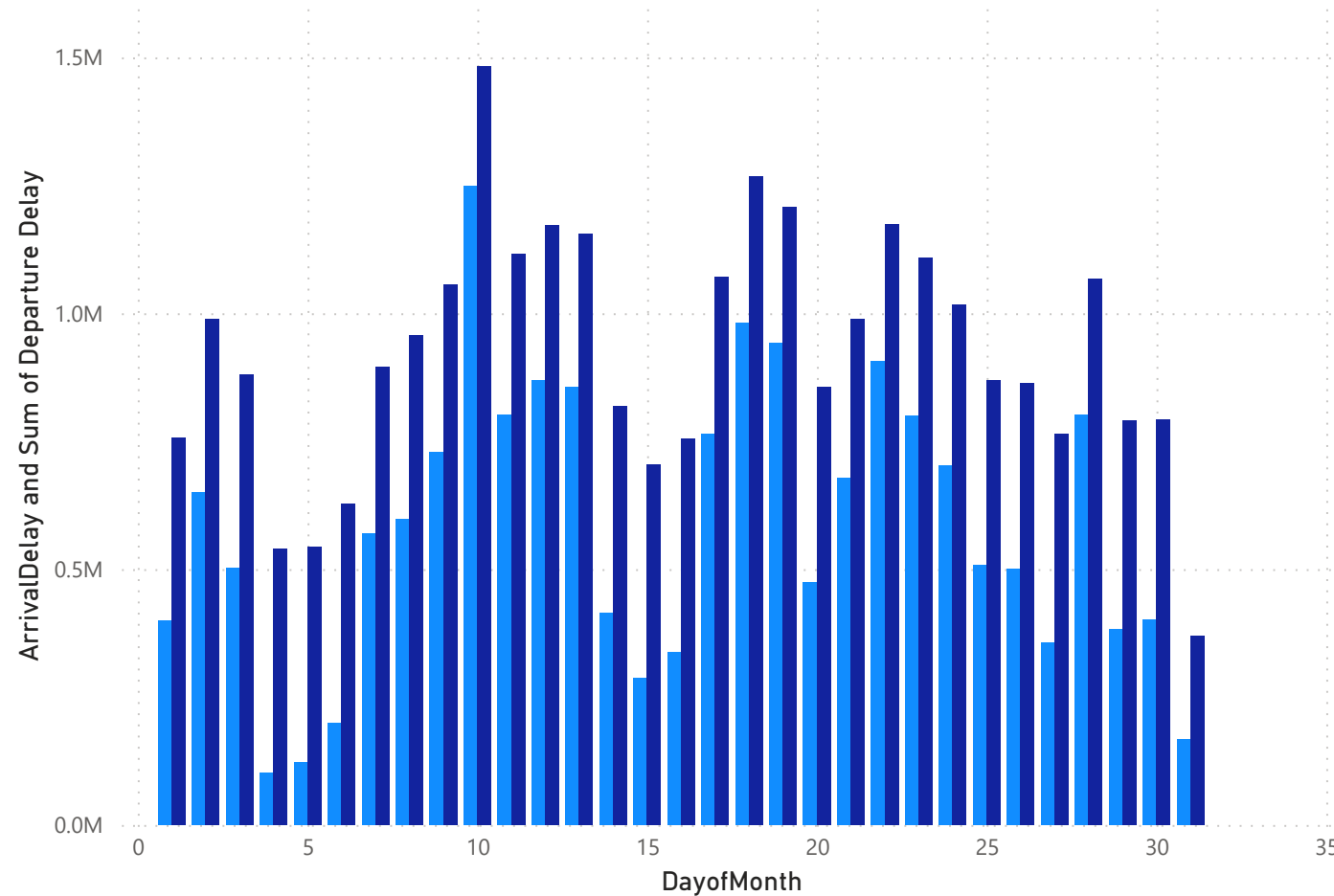
- The graph represents data related to arrival delays and the sum of departure delays over the days of a month.

- We're interested in understanding how these delays vary across different days.
- X-Axis (Day of Month): The x-axis is labeled "Day of Month," likely representing numerical values from 0 to slightly over 30, corresponding to days in a month.
- Y-Axis (Arrival Delay and Sum of Departure Delay): The vertical axis quantifies the delays, ranging from 0M to 1.5M.

- Blue Bars: Each blue bar corresponds to a specific day of the month. The height of each bar indicates the magnitude of arrival delay and sum of departure delay for that day.
- There is significant variability in delays throughout the month.
- A noticeable peak in delay occurs around the tenth day, reaching close to 1.5 million units.

Sum of Arrival Delay by Destination AirportID

Destination AirportID
- 13930
- 14771
- 10397
- 11618
- 11292
- 12892
- 12953
- 12478
- 11298
- 14100
- 10721
- 11057
- 12889
- 13232
- 10821
- 13204
- 12266
- 14107

- The **chart** represents data related to arrival delays at various airports, identified by their AirportID.
- We're interested in understanding how much delay occurred at each airport as a part of the total delay captured in this dataset.
-
- **Pie Chart**: The chart is divided into numerous slices, each representing a different airport.
- Each slice is colored differently for distinction and has labels indicating the sum of arrival delays and their percentage of the total.
- For example, one slice labeled "0.12M (6.22%)" indicates that there were 0.12 million minutes of delays, accounting for 6.22% of the total delays represented in this chart.

- Data Labels :The data labels range from "0.08M (4%)" to "1.25M (6.94%)," showing a variety in the extent of delays at different airports.
- Airport IDs :On the right side, there's a list of Destination Airport IDs with numerical identifiers like "13930," "14771," etc., presumably corresponding to specific airports.

Sum of Departure Delay and Sum of Arirval Delay by Carrier

● Sum of Departure Delay ● Sum of Arirval Delay

- The **chart** represents data related to departure delays and arrival delays for various airline carriers.
- We're interested in understanding how these delays accumulate for different carriers.

- X-Axis (Carrier): The x-axis lists airline carrier codes.
- Y-Axis (Sum of departure delay and Sum of arrival delay): The vertical axis shows the total delay in millions (ranging from 0M to 8M).

- Each point on the line represents the combined sum of departure and arrival delays for a specific carrier.
- Carrier WN (Southwest Airlines) has the **highest** combined delay, reaching close to 8 million units.
- The delay sums decrease after WN, with minor fluctuations from carriers AA to HA.
- Carrier HA (Hawaiian Airlines) has the lowest delay sum depicted on this graph.

Count of Departure Delay by OriginAirportID
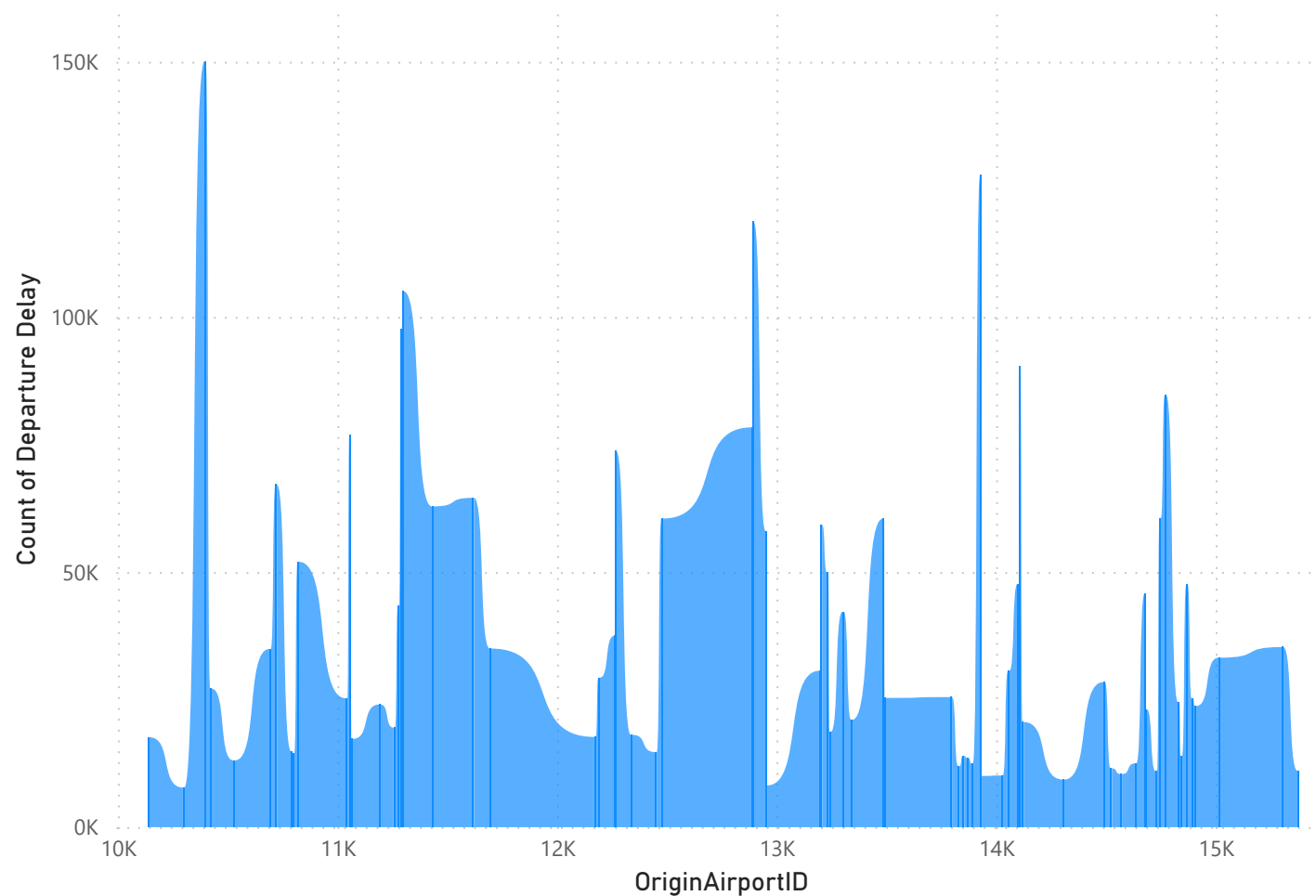
- Count of **Departure** Delay by Origin Airport ID indicating that it visualizes the number of departure delays based on different airport IDs.

- X-Axis (Origin Airport ID): This axis represents the airport IDs, ranging from 10,000 to 15,000. Each ID corresponds to a specific airport.

- Y-Axis (Count of Departure Delay): The vertical axis shows the count of departure delays. The range extends from 0 to approximately 150,000.

- Blue Area: The shaded blue area represents the frequency of departure delays at each airport ID.
- Notable peaks in the blue area indicate airports with a **higher** number of delays.