BUS DELAY IN TORONTO

Prepared for

Professor Enayat Rajabi

Mgsc-5127-13: Marketing Research

Cape Breton University

Sydney, Nova Scotia

Prepared By

Aniket Solanki

Darsini Unnikrishnan

Isha Sharma

Kamalpreet Kaur

BACKGROUND

Bus transportation is a crucial mode of public transportation in Toronto, Canada. However, bus delays are a common problem that commuters face daily, affecting their travel experience and productivity. Understanding the reasons for bus delays and developing effective strategies to reduce them can improve the efficiency and reliability of the public transportation system in Toronto.

Project description

As a student, we face many challenges while travelling to our destination like bus delay or bus cancellation and many more. so, we decided to prepare project on bus delay in Toronto. We collected our data from Kaggle website, and we also checked dataset on official website which we found more cleaned data. The data extracted from website is between 01 January 2022- 31 December 2022. With the help of visualization, we can gain more insight into the particular reason for delay.

Dataset:

I. The data set contains 58708 records available on the website between the dates 01 January 2022 to 31 December 2022.
II. The data contains the following fields: date, route, time , day , location, incidents, Min delay, Min gap , and vehicle.
III. The dataset can be assessed from the Kaggle website. Here is the link: https://www.kaggle.com/datasets/reihanenamdari/toronto-bus-delay-2022

Data cleaning:

- First, we removed all null values, blank records, missing data from dataset to make it more meaningful.
- We removed one column named direction as in that column most of the records were missing which is not available on portal.
- We renamed two columns from Min delay to delay in minutes and min gap to gap in minutes.
- We trimmed all extra space from the values.
- After cleaning data, around 53000 records are available for visualization.

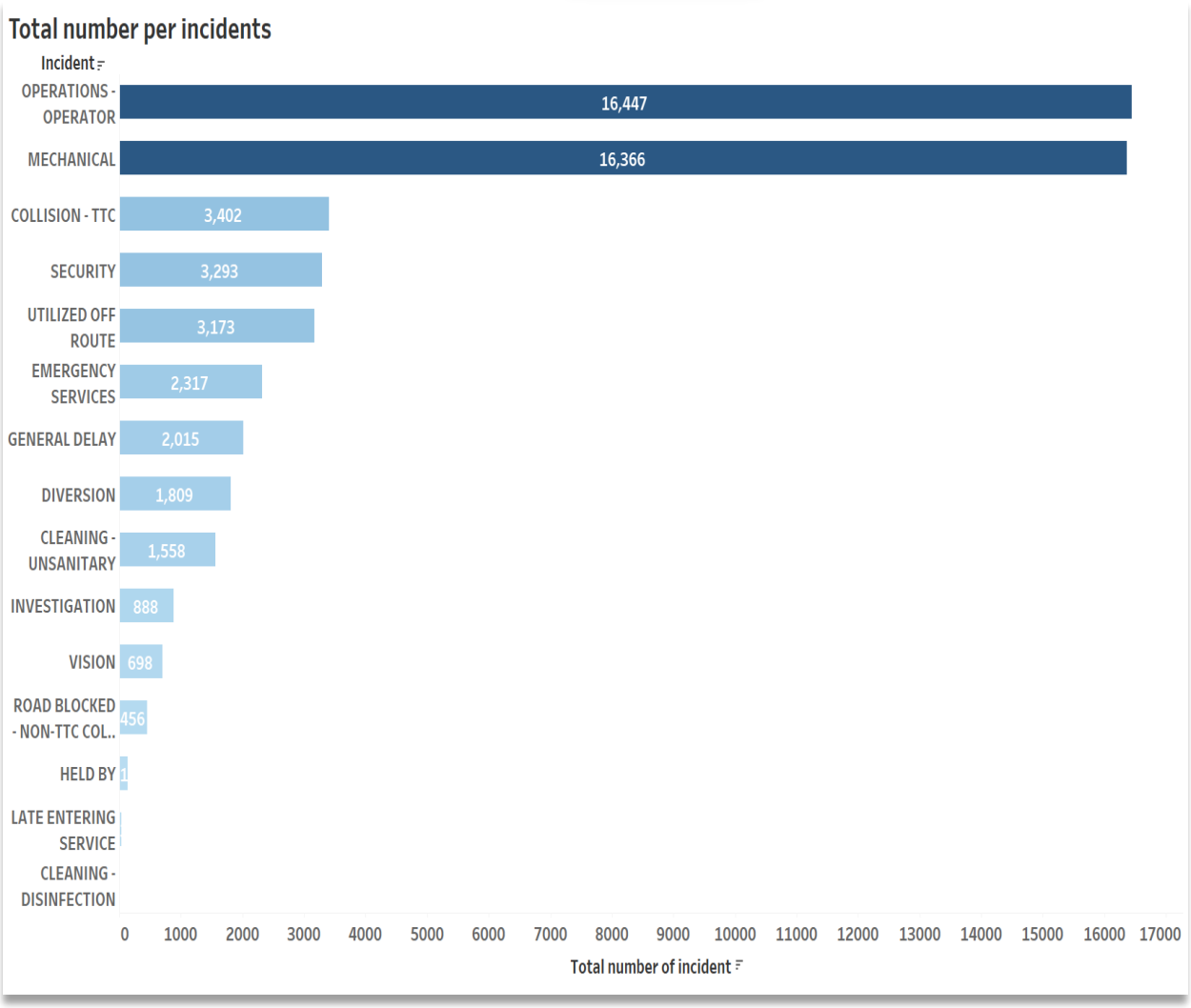
Data dictionary:

Field Name	Data Type	Definition
Date	Interval	The date (YYYY/MM/DD) when the delay-causing incident occurred
Time	Interval	The time (AM/PM) when the delay-causing incident occurred
Route	Categorical	The number of the bus route

Day of Week	Categorical	A day when the incident was happened
Location	Categorical	The location of the delay-causing incident
Incident	Categorical	The description of the delay-causing incident
Delay in minutes	Ratio	The delay, in minutes, to the schedule for the following bus
Gap in minutes	Ratio	The total scheduled time, in minutes, from the bus ahead of the following bus

Question:1

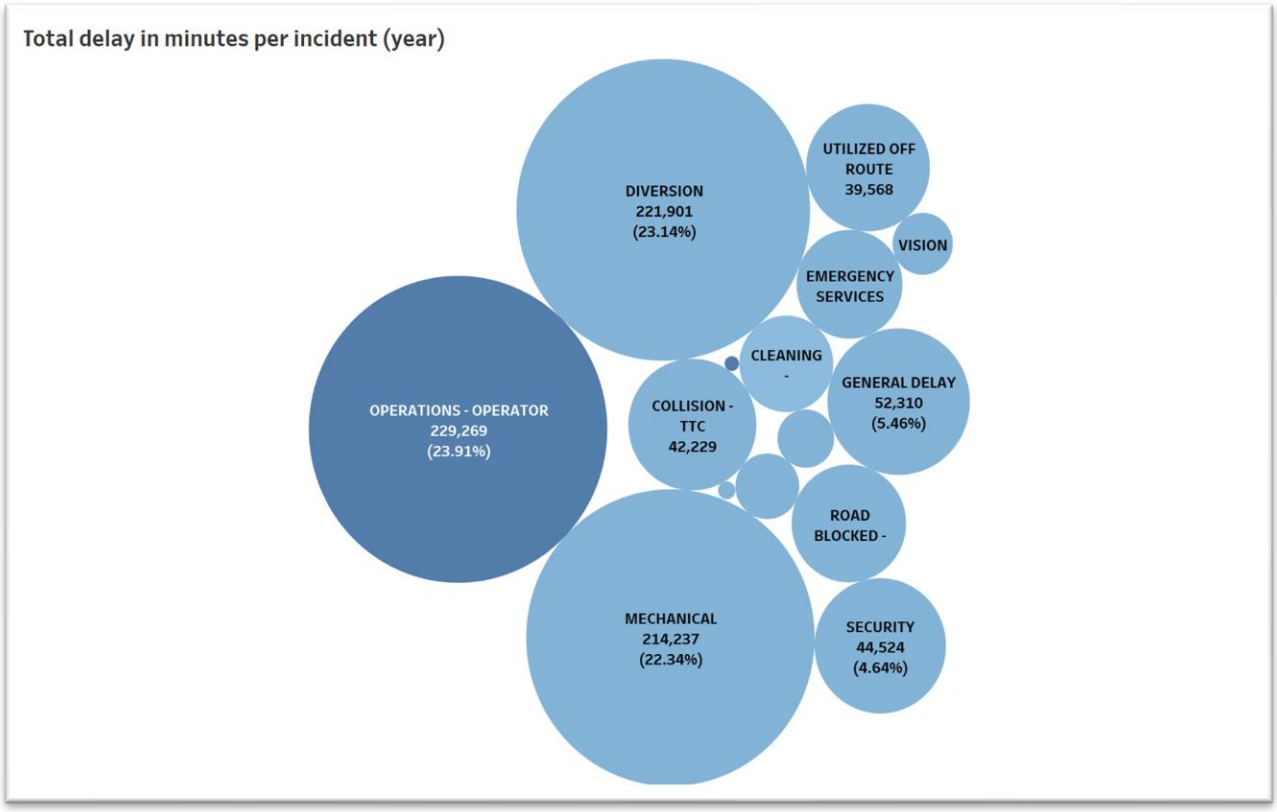
Which incident did cause the highest number of delays in 2022?



Based on the chart data, it can be inferred that the occurrence of delays in buses was attributed to the actions of the operators for a total of 16,447 instances.

Question: 2

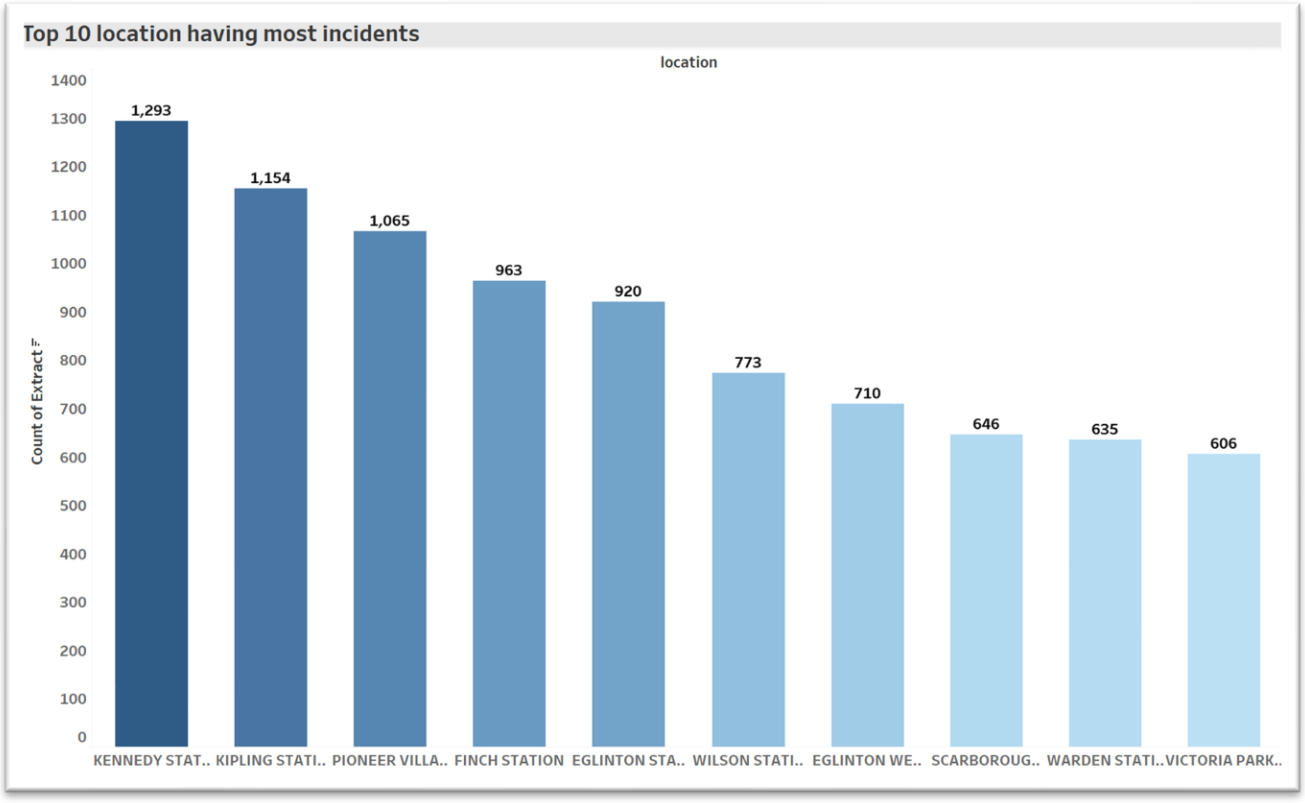
What was the total delay in minutes per incident for buses in the year 2022?



It can be concluded that the delays incurred by buses due to operational issues resulted in a cumulative delay time of approximately 229,269 minutes.

Question:3

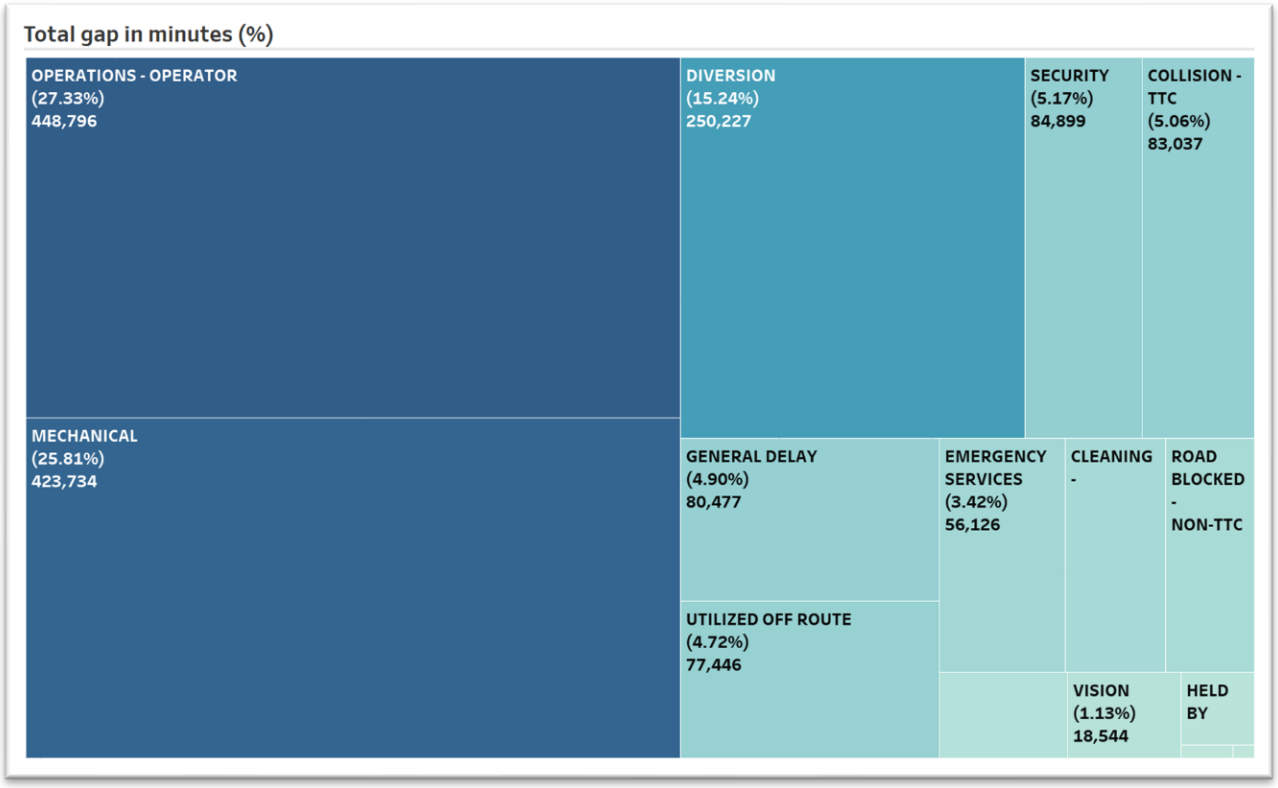
What are the top 10 locations where the most incidents occurred?



The Kennedy station was identified as the location with the highest frequency of incidents.

Question:4

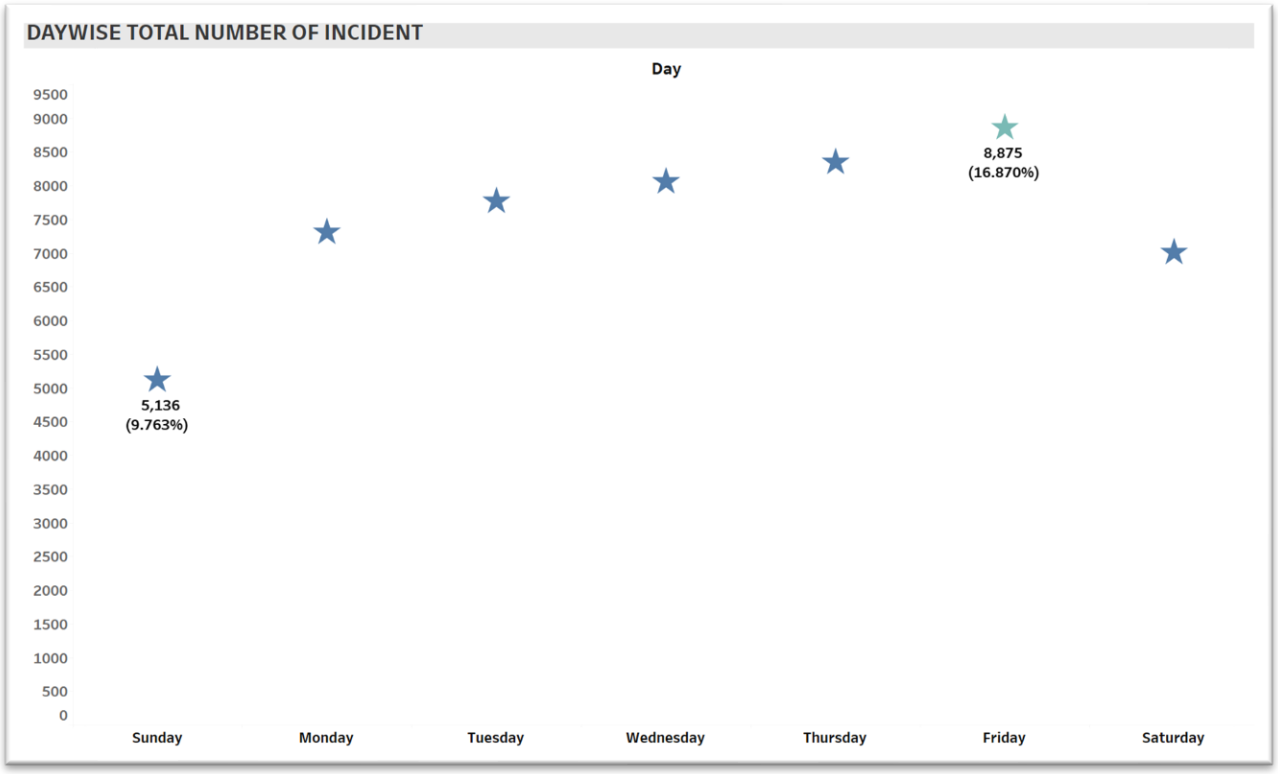
What was the total gap in minutes per incident?

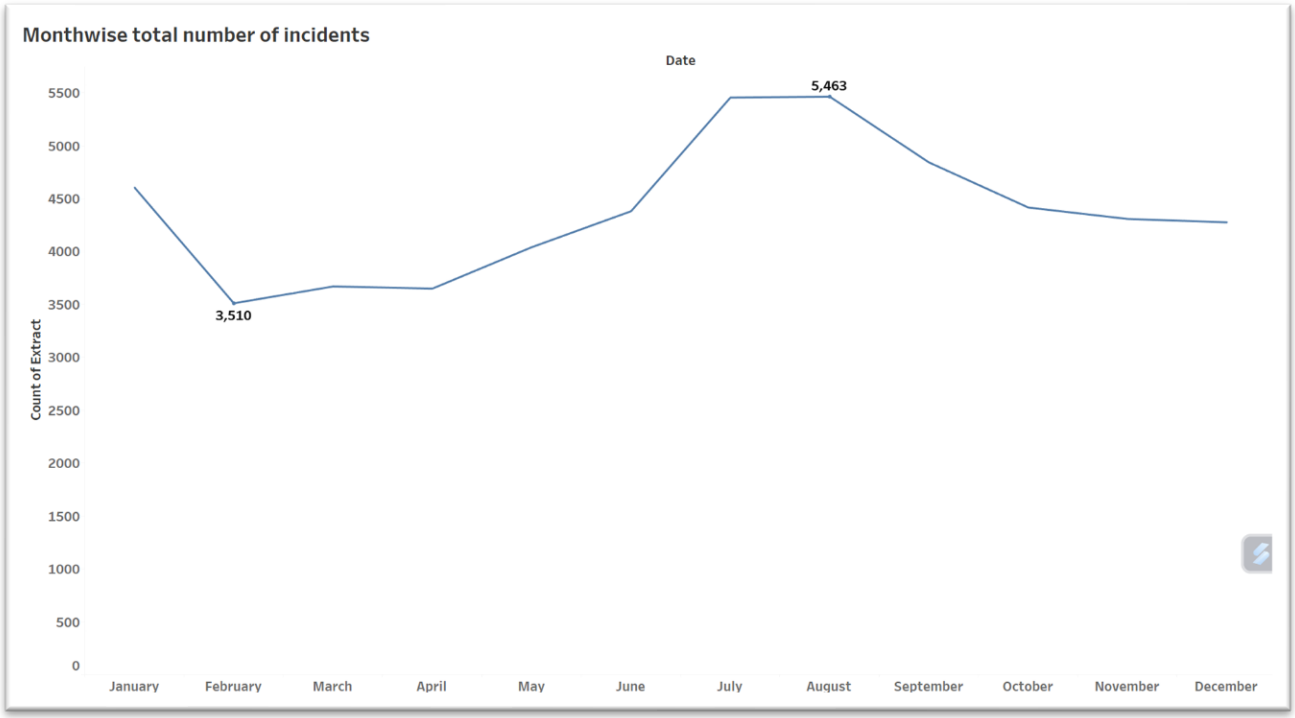


From the chart, we can say that the highest total schedule time from the bus ahead of the following bus was 448,796 minutes occurred by operations-operator.

Question: 5

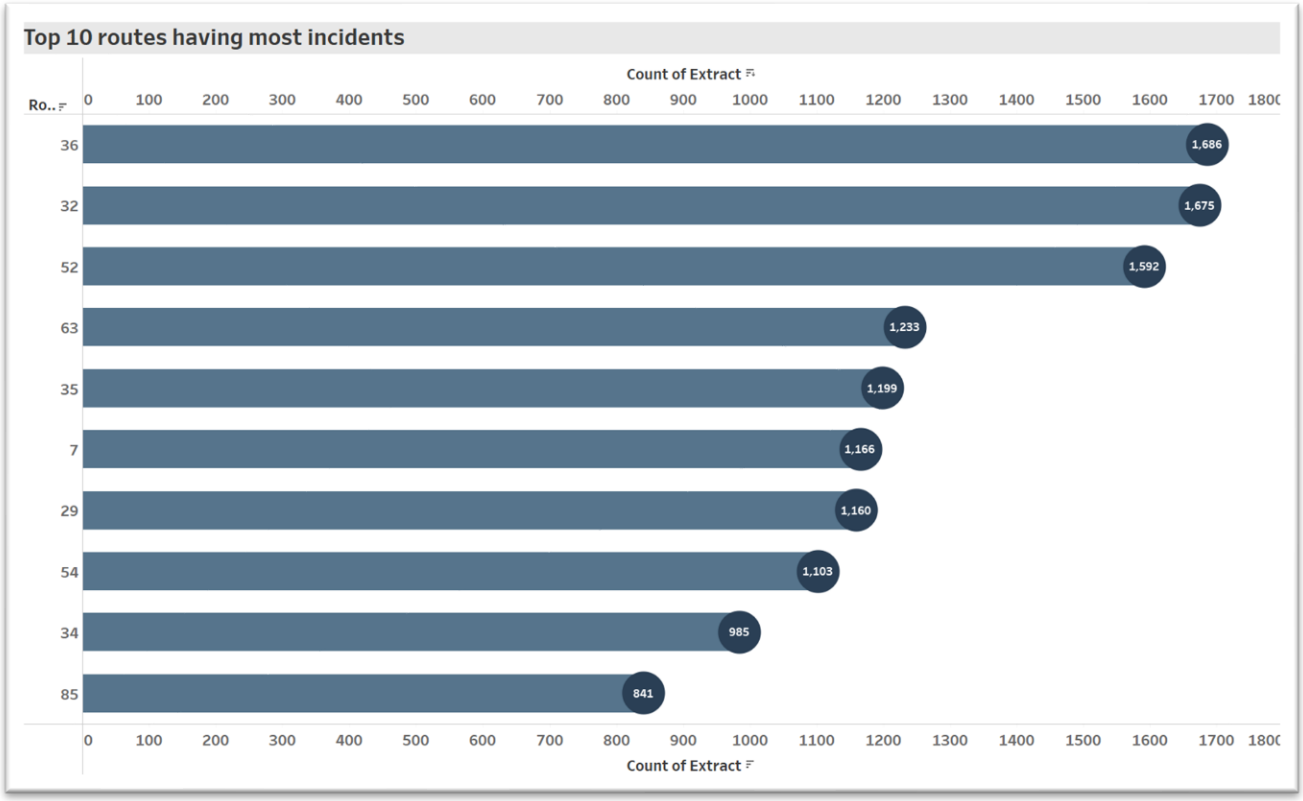
What was the total number of incidents per month? On which day, most of incidents reported?





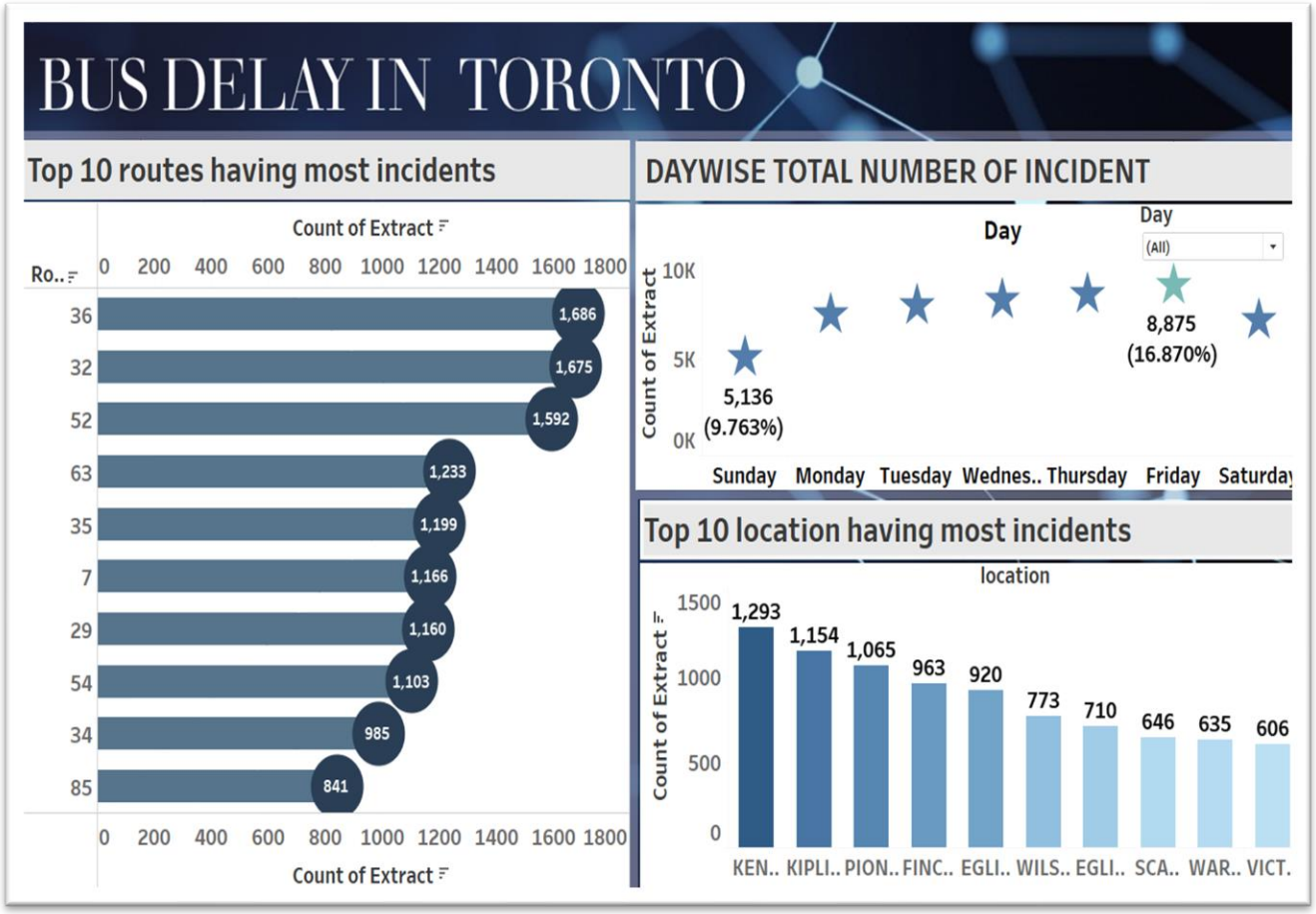
Question:6

Which were the top 10 routes where the most incident happened?



From the graph, it can be concluded that route number 36 has been the 1st route where the total incidents were reported.

Dashboard:



Thank You