GTFS Brief Summary

The exploded GTFS shows that GTFS is valid from **November 17, 2024** to **December 22, 2024**.

Toronto has **215** routes and **9308** stops. Bellow, we will provide some general information.

GTFS Feeds

The exploded GTFS has **9** feeds as follows:

1. agency:
Provides information about the transit agency, including its name, URL, and contact details.

2. calendar:
Defines the service schedule (days of operation) for each service period.

3. calendar_dates:
Specifies exceptions to the regular service schedule for specific dates.

4. routes:
Contains details about the routes, such as route name, type, and associated agency.

5. stops:
Defines each transit stop's location and associated information like name and type.

6. stop_times:
Lists the scheduled arrival and departure times for each trip at each stop.

7. shapes:
Defines the geometry of a route, specifying the coordinates of each point on the route.

8. trips:
Specifies individual trips within routes, including the route and schedule details.

9. feed_info:
Provides metadata about the GTFS feed itself, including version and publisher information.

GTFS Agencies

The agency.txt file in GTFS provides essential information about the transit agencies managing the services in the feed. It includes required fields like agency_name, agency_url, and agency_timezone to identify the agency, link to its website, and specify its operating timezone. Optional fields such as agency_lang, agency_phone, agency_email, and agency_fare_url offer additional details like language preferences, contact information, and fare policies. In multi-agency feeds, the agency_id field distinguishes between agencies, making this file key for identifying and connecting transit services to their providers.In the following, we provide some information about agencies of GTFS:

Agency ID	Agency Name	Agency URL	Timezone	Phone
1703-1	Ttc	http://www.ttc.ca	America/Toronto	416-393-4636

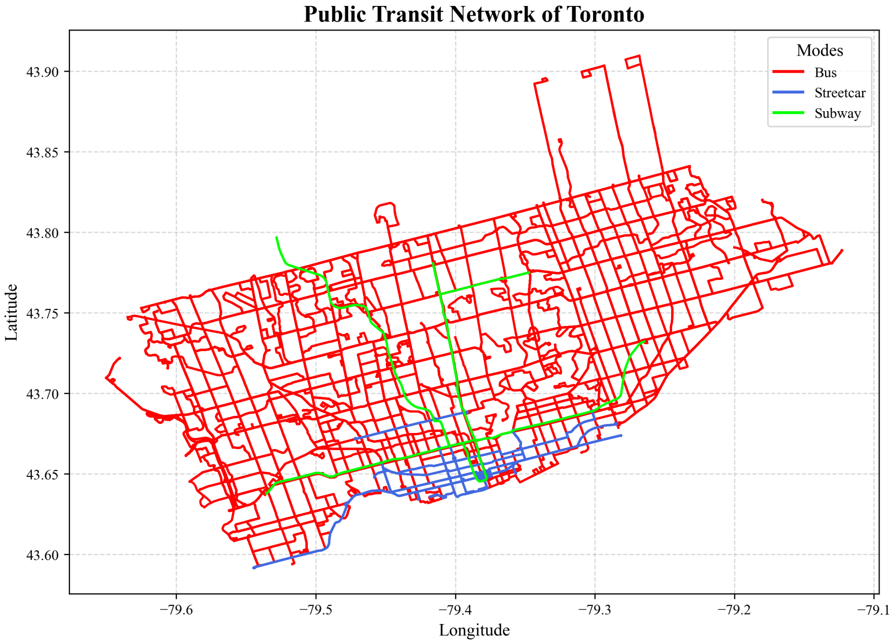
GTFS Routes Feed

The Routes feed provides information about each route, including its name, ID, and associated agency. Each route is assigned a route_type, which corresponds to the type of transportation mode. Common route_types include Bus, Streetcar, and Subway, each of which represents different forms of public transit.

No.	Mode	Number of Routes
1	Bus	194
2	Streetcar	18
3	Subway	3

Transit Network Layout

The transit network layout can be visualized using the data from 'shapes.txt', which defines the geometry of each route. The following figure presents the layout of the transit network, illustrating the routes and connections between various stops.



GTFS Day Type

In GTFS, the calendar.txt and calendar_dates.txt files define the schedules for transit services. Although these files are essential, they present challenges in identifying and grouping dates with similar service patterns due to limitations in the GTFS structure. To address this, methods have been developed to group dates with shared characteristics, improving the ability to analyze and explore transit networks and schedules effectively.

1. Day Type 1:

Dates: ['2024-11-20', '2024-11-21', '2024-11-27', '2024-11-28', '2024-12-04', '2024-12-05', '2024-12-11', '2024-12-12', '2024-12-18', '2024-12-19']

Number of trips: 39480

2. Day Type 2:

Dates: ['2024-11-19', '2024-11-22', '2024-11-26', '2024-11-29', '2024-12-03', '2024-12-06', '2024-12-10', '2024-12-13', '2024-12-17', '2024-12-20']

Number of trips: 39482

3. Day Type 3:

Dates: ['2024-11-18', '2024-11-25', '2024-12-02', '2024-12-09', '2024-12-16']

Number of trips: 39457

4. Day Type 4:

Dates: ['2024-11-23', '2024-11-30', '2024-12-07', '2024-12-14', '2024-12-21']

Number of trips: 31357

5. Day Type 5:

Dates: ['2024-11-17']

Number of trips: 27918

6. Day Type 6:

Dates: ['2024-11-24', '2024-12-01', '2024-12-08', '2024-12-15']

Number of trips: 28315

7. Day Type 7:

Dates: ['2024-12-22']

Number of trips: 397

Busiest Day

Using the calendar and calendar_dates files, we identify the busiest day of the transit network. Below, we provide an overview of this day:

* Routes in Operation: On the busiest day, **214** routes are active.

* Stops Served: Passengers can board or exit public transit at **9281** stops.

* Trips and Blocks: A total of **39482** trips are operated, managed by **2435** blocks.

* Traveled Distance: On average, blocks have traveled **188** kilometers when serving passengers.

In the following, we delve into more detailed information about the busiest day's operations.

Day Periods

Day periods play a vital role in analyzing the transit system by dividing the day into meaningful intervals based on time, allowing transit planners and analysts to observe and understand variations in transit demand and supply throughout the day. Each day period is characterized by unique transit dynamics, such as passenger demand, vehicle deployment, and service frequency. We can derive key insights from day period analysis such as:

1. Rush Hours Identification:

Rush hours typically correspond to morning and evening periods when the demand for public transit is at its peak due to work or school-related commutes. High numbers of trips during these periods indicate the system's capacity to accommodate increased ridership.

2. Off-Peak Periods:

These are periods with relatively low transit activity, often mid-morning, early afternoon, or late evening. They are critical for maintenance and adjustments to service schedules.

3. Service Optimization:

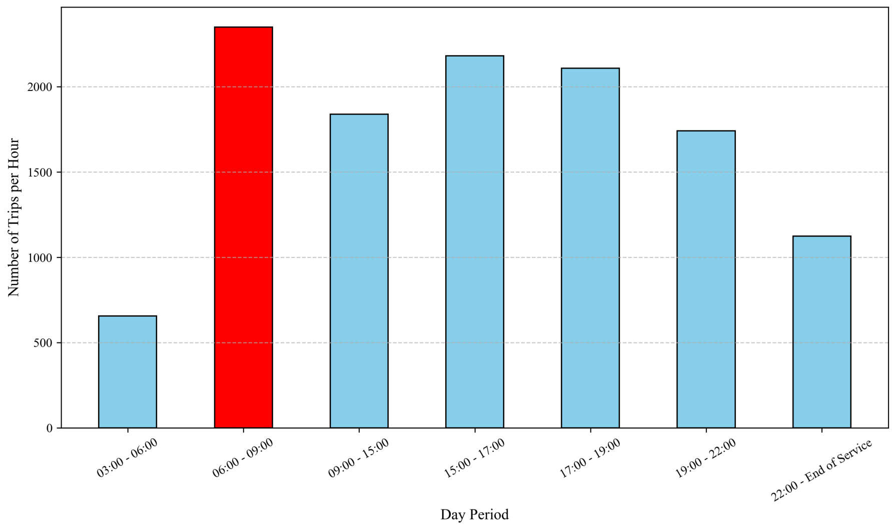
By analyzing the number of trips and their distribution across day periods, transit agencies can allocate resources (e.g., buses, drivers) more effectively. Understanding the average number of trips per hour during each period helps in identifying gaps or redundancies in service.

4. Passenger Behavior and Trends:

Day periods allow transit planners to study passenger boarding trends, which can inform service planning and fare strategies. It also aids in aligning the transit system with daily human activity patterns.

In the following table and figure, we divide the day into specific periods, showing the start and end time, duration, number of trips started in each day period, and the average number of started trips per hour. These metrics provide insights into the transit system's operational efficiency and demand patterns.

Day period id	Start time	End time	Duration (hour)	No. of trips	No. of trips per hour
1	03:00	06:00	3	1972	657
2	06:00	09:00	3	7047	2349
3	09:00	15:00	6	11033	1839
4	15:00	17:00	2	4362	2181
5	17:00	19:00	2	4217	2108
6	19:00	22:00	3	5225	1742
7	22:00	End of Service	5	5626	1125



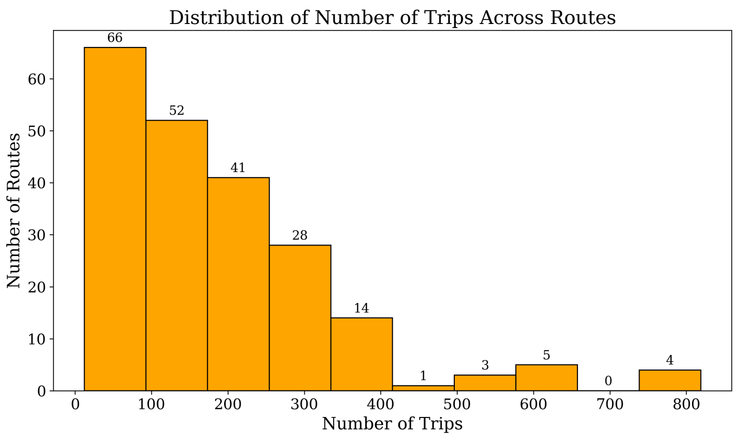
Routes

On the busiest day in the dataset, we analyzed the transit routes to determine their operational intensity by examining the number of trips and blocks assigned to each route. Each route operates with varying levels of frequency and service complexity, measured by the number of trips and blocks on that day. Notably, the route_id 1703-73152 recorded the highest number of blocks, with 78 blocks, while the route_id 1703-73080 had the maximum number of trips, totaling 819 trips. It is important to note that the routes with the most blocks and trips may not be identical, indicating differences in service structure and scheduling. In the following table, we provide detailed information on the 10 busiest routes, showcasing their trip counts, block numbers, and other relevant attributes.

route id	no. of trips	no. of blocks	short name	long name	mode
1703-73080	819	34	63	Ossington	Bus
1703-73031	800	57	36	Finch West	Bus
1703-73053	780	27	47	Lansdowne	Bus
1703-73066	740	40	510	Spadina	Streetcar
1703-73069	641	60	52	Lawrence West	Bus
1703-73153	620	45	2	Line 2 (Bloor - Danforth)	Subway
1703-73010	604	56	32	Eglinton West	Bus
1703-73152	601	78	1	Line 1 (Yonge-University)	Subway
1703-73059	588	38	504	King	Streetcar
1703-72993	553	32	25	DON Mills	Bus

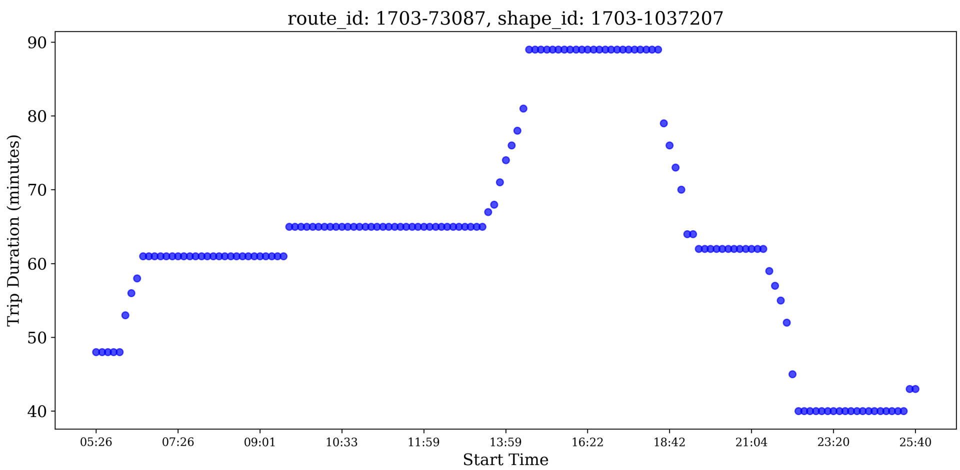
Trip frequency

A histogram is an effective way to visualize the distribution of the number of trips across all routes. It allows us to identify clusters of routes with similar trip counts, observe the spread of trip numbers, and highlight outliers, such as exceptionally busy or underutilized routes. For example, routes with a high frequency of trips likely serve high-demand corridors or peak periods, whereas routes with fewer trips might serve less-populated areas or operate during off-peak times. The interval with the highest frequency of routes is [12, 92], containing 66 routes. The last interval, [738, 819], has a frequency of 4 routes, representing the busiest routes in the network. These routes likely serve the highest-demand areas or periods, providing key insights into the most intensively used parts of the transit system.



Travel time variation

To analyze the variation in trip durations, we focused on identifying the routes and shapes with the highest standard deviation in trip durations. A high standard deviation indicates significant variation in how long trips take, which may be due to factors such as traffic congestion, irregular service patterns, or inconsistent trip durations. Understanding this variation is crucial for improving service reliability, as high variation could point to potential operational challenges that may need attention. In the case of route_id 1703-73087 and shape_id 1703-1037207, we observed the highest standard deviation in trip duration, reflecting the most variable service in the dataset. This shape had a total of 141 trips, highlighting its complexity in terms of time reliability. Below, we present a plot showing the relationship between the start time and trip duration for this specific shape, providing a visual representation of the variability in trip durations.



In the following table, we assess trips of route_id 1703-73087 and shape_id 1703-1037207 at each day period.

Day Period ID	No. of Trips	Mean Duration	Duration Std
1	4	48	0
2	24	60	3
3	49	67	7
4	11	89	0
5	12	86	6
6	18	62	3
7	23	41	3