<u>GTFS Brief Summary</u>

The exploded GTFS shows that GTFS is valid from **December 15, 2024** to **July 04, 2025**.

Washington has **192** routes and **9230** 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
100347-1	Wmata	https://www.wmata.com	America/New_York	(202) 637-7000

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	192

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.

Public Transit Network of Washington
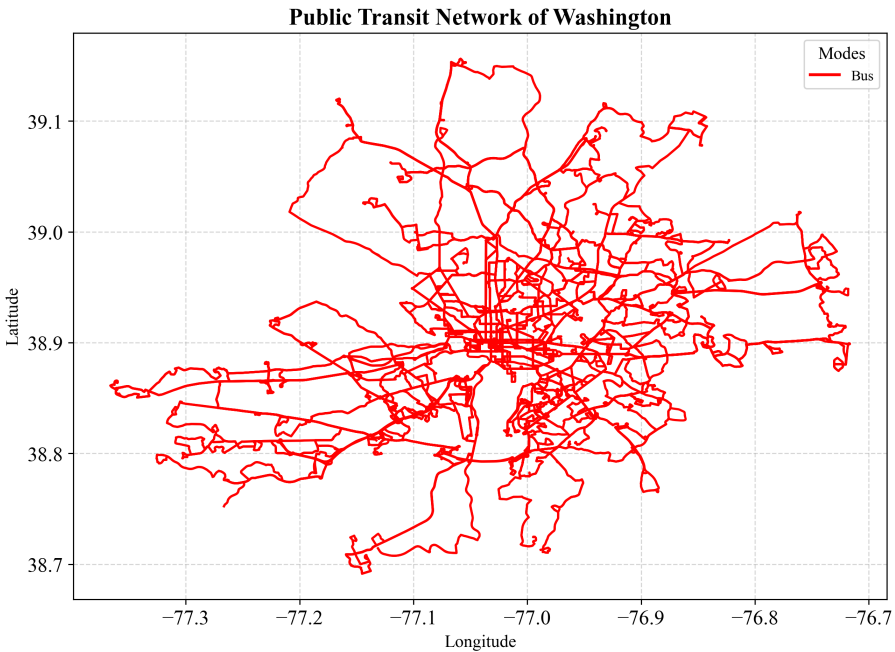
GTFS Day Type

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: ['2025-06-29']

Number of trips: 0

2. Day Type 2:
Dates: ['2025-03-31']

Number of trips: 13748

3. Day Type 3:
Dates: ['2024-12-23', '2024-12-24', '2024-12-26']

Number of trips: 14871

4. Day Type 4:
Dates: ['2024-12-31', '2025-01-27', '2025-03-13', '2025-04-07', '2025-04-14', '2025-04-15',

'2025-04-16', '2025-04-17', '2025-06-17', '2025-06-18', '2025-06-23', '2025-06-24', '2025-06-25',

'2025-06-26', '2025-06-30', '2025-07-01', '2025-07-02', '2025-07-03']

Number of trips: 13801

5. Day Type 5:
Dates: ['2025-01-20', '2025-02-17', '2025-06-19']

Number of trips: 10002

6. Day Type 6:
Dates: ['2025-01-04', '2025-01-11', '2025-01-18', '2025-01-25', '2025-02-01', '2025-02-08',

'2025-02-15', '2025-02-22', '2025-03-01', '2025-03-08', '2025-03-15', '2025-03-22', '2025-03-29',

'2025-04-05', '2025-04-12', '2025-04-19', '2025-04-26', '2025-05-03', '2025-05-10', '2025-05-17',

'2025-05-24', '2025-05-31', '2025-06-07', '2025-06-14', '2025-06-21', '2025-06-28']

Number of trips: 9687

7. Day Type 7:
Dates: ['2024-12-21']

Number of trips: 10569

8. Day Type 8:

Dates: ['2024-12-28']

Number of trips: 10569

9. Day Type 9:

Dates: ['2024-12-30']

Number of trips: 14871

10. Day Type 10:

Dates: ['2024-12-15', '2025-01-01', '2025-01-05', '2025-01-12', '2025-01-19', '2025-01-26',

'2025-02-02', '2025-02-09', '2025-02-16', '2025-02-23', '2025-03-02', '2025-03-09', '2025-03-16',

'2025-03-23', '2025-03-30', '2025-04-06', '2025-04-13', '2025-04-20', '2025-04-27', '2025-05-04',

'2025-05-11', '2025-05-18', '2025-05-25', '2025-05-26', '2025-06-01', '2025-06-08', '2025-06-15',

'2025-06-22']

Number of trips: 8964

11. Day Type 11:

Dates: ['2024-12-22', '2024-12-25']

Number of trips: 9798

12. Day Type 12:

Dates: ['2024-12-29']

Number of trips: 9798

13. Day Type 13:

Dates: ['2025-07-04']

Number of trips: 4178

14. Day Type 14:

Dates: ['2025-01-17', '2025-02-14', '2025-03-14', '2025-04-18', '2025-06-20', '2025-06-27']

Number of trips: 13806

15. Day Type 15:

Dates: ['2024-12-27']

Number of trips: 14924

16. Day Type 16:

Dates: ['2024-12-16', '2024-12-17', '2024-12-18', '2024-12-19', '2025-01-02', '2025-01-06',

'2025-01-07', '2025-01-08', '2025-01-09', '2025-01-13', '2025-01-14', '2025-01-15', '2025-01-16',

'2025-01-21', '2025-01-22', '2025-01-23', '2025-01-28', '2025-01-29', '2025-01-30', '2025-02-03',

'2025-02-04', '2025-02-05', '2025-02-06', '2025-02-10', '2025-02-11', '2025-02-12', '2025-02-13',

'2025-02-18', '2025-02-19', '2025-02-20', '2025-02-24', '2025-02-25', '2025-02-26', '2025-02-27',

'2025-03-03', '2025-03-04', '2025-03-05', '2025-03-06', '2025-03-10', '2025-03-11', '2025-03-12',

'2025-03-17', '2025-03-18', '2025-03-19', '2025-03-20', '2025-03-24', '2025-03-25', '2025-03-26',

'2025-03-27', '2025-04-01', '2025-04-02', '2025-04-03', '2025-04-08', '2025-04-09', '2025-04-10',

'2025-04-21', '2025-04-22', '2025-04-23', '2025-04-24', '2025-04-28', '2025-04-29', '2025-04-30',

'2025-05-01', '2025-05-05', '2025-05-06', '2025-05-07', '2025-05-08', '2025-05-12', '2025-05-13',

'2025-05-14', '2025-05-15', '2025-05-19', '2025-05-20', '2025-05-21', '2025-05-22', '2025-05-27',

'2025-05-28', '2025-05-29', '2025-06-02', '2025-06-03', '2025-06-04', '2025-06-05', '2025-06-09',

'2025-06-10', '2025-06-11', '2025-06-12', '2025-06-16']

Number of trips: 13876

17. Day Type 17:

Dates: ['2025-01-03', '2025-01-10', '2025-01-24', '2025-01-31', '2025-02-07', '2025-02-21',

'2025-02-28', '2025-03-07', '2025-03-21', '2025-03-28', '2025-04-04', '2025-04-11', '2025-04-25',

'2025-05-02', '2025-05-09', '2025-05-16', '2025-05-23', '2025-05-30', '2025-06-06', '2025-06-13']

Number of trips: 13881

18. Day Type 18:

Dates: ['2024-12-20']

Number of trips: 14999

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, **189** routes are active.

* Stops Served: Passengers can board or exit public transit at **9220** stops.

* Trips and Blocks: A total of **14999** trips are operated, managed by **2292** blocks.

* Traveled Distance: On average, blocks have traveled **85** 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	00:00	06:00	5	1026	194
2	06:00	09:00	3	2945	982
3	09:00	15:00	6	4200	700
4	15:00	17:00	2	1983	992
5	17:00	19:00	2	1882	941
6	19:00	22:00	3	1797	599
7	22:00	End of Service	5	1166	233
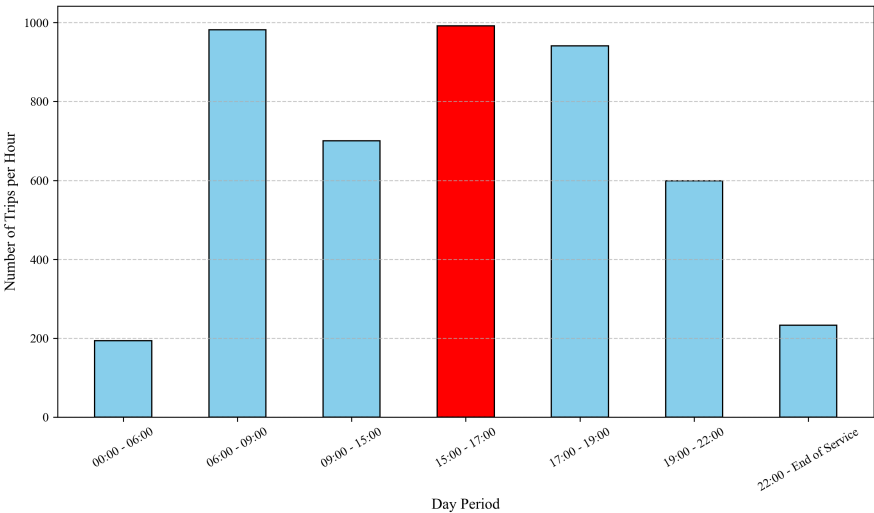
Routes

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 100347-32 recorded the highest number of blocks, with 84 blocks, while the route_id 100347-EXP1 had the maximum number of trips, totaling 324 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
100347-LOC1	324	16	LOC1	Foggy Bottom-Federal Triangle	Bus
100347-EXP2	324	12	EXP2	Foggy Bottom-Gallery Place	Bus
100347-EXP1	324	12	EXP1	Foggy Bottom-Federal Triangle	Bus
100347-33	267	51	33	Wisconsin Avenue	Bus
100347-X2	235	36	X2	Benning Road-H Street	Bus
100347-B2	229	41	B2	Bladensburg Road-Anacostia	Bus
100347-70	220	33	70	Georgia Avenue-7TH Street	Bus
100347-S9	212	33	S9	16th Street Metroextra	Bus
100347-W4	210	47	W4	Deanwood-Alabama Ave	Bus
100347-80	206	35	80	North Capitol Street	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 [33, 65], containing 49 routes. The last interval, [291, 324], has a frequency of 3 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.

Distribution of Number of Trips Across Routes
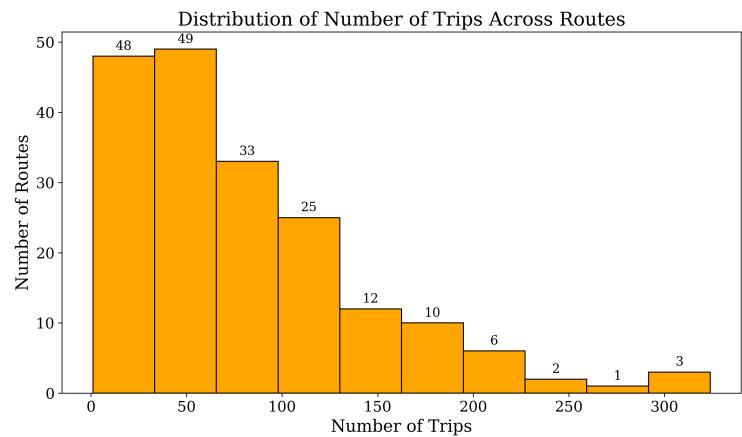
Travel time variation

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 100347-D6 and shape_id 100347-D6:01, we observed the highest standard deviation in trip duration, reflecting the most variable service in the dataset. This shape had a total of 51 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.
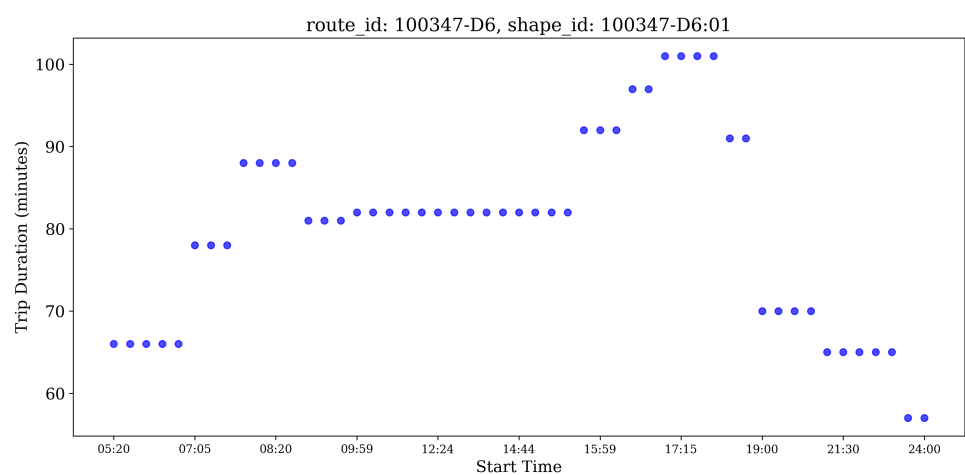
In the following table, we assess trips of route_id 100347-D6 and shape_id 100347-D6:01 at each day period.

Day Period ID	No. of Trips	Mean Duration	Duration Std
1	2	66	0
2	11	79	9
3	13	82	0
4	8	90	7
5	6	98	5
6	6	68	3
7	5	62	4