GTFS Brief Summary

The exploded GTFS shows that GTFS is valid from **September 01, 2024** to **January 25, 2025**.

San Diego has **105** routes and **4360** stops. Bellow, we will provide some general information.

GTFS Feeds

The exploded GTFS has **12** 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. transfers:

Defines transfer rules between stops or routes for seamless connections.

8. shapes:

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

9. fare_rules:

Describes fare rules, linking fare attributes to specific routes or zones.

10. fare_attributes:

Provides information about fares, including price, currency, and payment method.

11. trips:

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

12. 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
561-MTS	Mts	http://www.sdmts.com	America/Los_Angeles	619-233-3004
561-SAN	San Diego International Airport	http://www.san.org	America/Los_Angeles	619-694-7427

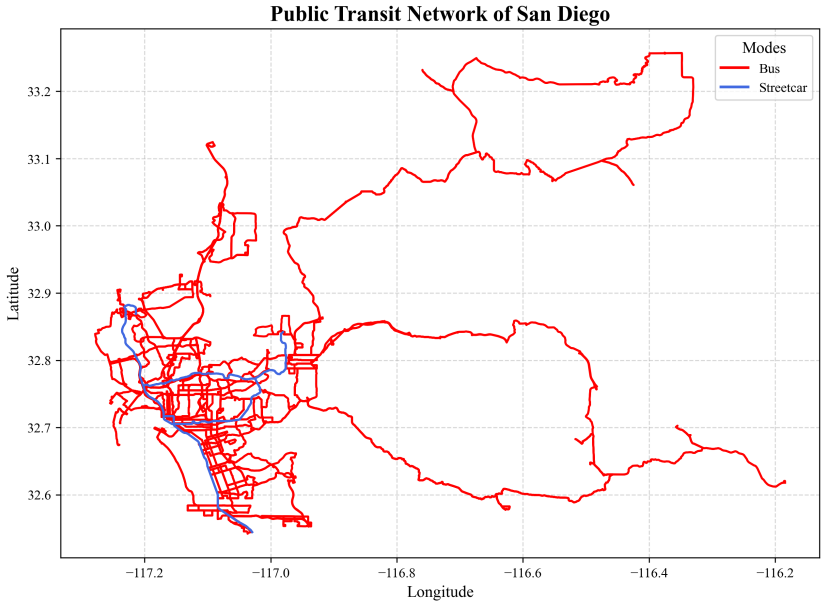
GTFS Routes Feed

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	99
2	Streetcar	5
3	Ferry	1

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: ['2025-01-01']

Number of trips: 3568

2. Day Type 2:

Dates: ['2024-12-25']

Number of trips: 3320

3. Day Type 3:

Dates: ['2024-10-04', '2024-10-11', '2024-10-18', '2024-10-25', '2024-11-01', '2024-11-08',

'2024-11-15', '2024-11-22', '2024-11-29', '2024-12-06']

Number of trips: 7872

4. Day Type 4:

Dates: ['2024-10-01', '2024-10-02', '2024-10-08', '2024-10-09', '2024-10-15', '2024-10-16',

'2024-10-22', '2024-10-23', '2024-10-29', '2024-10-30', '2024-11-05', '2024-11-06', '2024-11-12',

'2024-11-13', '2024-11-19', '2024-11-20', '2024-11-26', '2024-11-27', '2024-12-03', '2024-12-04']

Number of trips: 7862

5. Day Type 5:

Dates: ['2024-09-30', '2024-10-07', '2024-10-14', '2024-10-21', '2024-10-28', '2024-11-04',

'2024-11-11', '2024-11-18', '2024-11-25', '2024-12-02']

Number of trips: 7864

6. Day Type 6:

Dates: ['2024-10-03', '2024-10-10', '2024-10-17', '2024-10-24', '2024-10-31', '2024-11-07',

'2024-11-14', '2024-11-21', '2024-11-28', '2024-12-05']

Number of trips: 7864

7. Day Type 7:

Dates: ['2024-12-13']

Number of trips: 7782

8. Day Type 8:
Dates: ['2024-12-10', '2024-12-11']

Number of trips: 7772

9. Day Type 9:
Dates: ['2024-12-09']

Number of trips: 7774

10. Day Type 10:
Dates: ['2024-12-12']

Number of trips: 7774

11. Day Type 11:
Dates: ['2024-12-16']

Number of trips: 7699

12. Day Type 12:
Dates: ['2025-01-10', '2025-01-17', '2025-01-24']

Number of trips: 7779

13. Day Type 13:
Dates: ['2025-01-07']

Number of trips: 7768

14. Day Type 14:
Dates: ['2025-01-08', '2025-01-14', '2025-01-15', '2025-01-21', '2025-01-22']

Number of trips: 7769

15. Day Type 15:
Dates: ['2025-01-13']

Number of trips: 7771

16. Day Type 16:
Dates: ['2025-01-09', '2025-01-16', '2025-01-23']

Number of trips: 7771

17. Day Type 17:
Dates: ['2024-12-20']

Number of trips: 7688

18. Day Type 18:
Dates: ['2024-12-17', '2024-12-18']

Number of trips: 7679

19. Day Type 19:

Dates: ['2024-12-19']

Number of trips: 7681

20. Day Type 20:
Dates: ['2025-01-06']

Number of trips: 7769

21. Day Type 21:
Dates: ['2024-12-27', '2025-01-03']

Number of trips: 7686

22. Day Type 22:
Dates: ['2024-12-24', '2024-12-31']

Number of trips: 7676

23. Day Type 23:
Dates: ['2024-12-23', '2024-12-30']

Number of trips: 7678

24. Day Type 24:
Dates: ['2024-12-26', '2025-01-02']

Number of trips: 7678

25. Day Type 25:
Dates: ['2025-01-20']

Number of trips: 7666

26. Day Type 26:
Dates: ['2024-10-05', '2024-10-12', '2024-10-19', '2024-10-26', '2024-11-02', '2024-11-09',

'2024-11-16', '2024-11-23', '2024-11-30', '2024-12-07', '2024-12-21', '2024-12-28', '2025-01-04',

'2025-01-11', '2025-01-18', '2025-01-25']

Number of trips: 4596

27. Day Type 27:
Dates: ['2024-12-14']

Number of trips: 4605

28. Day Type 28:
Dates: ['2024-09-29', '2024-10-06', '2024-10-13', '2024-10-20', '2024-10-27', '2024-11-03',

'2024-11-10', '2024-11-17', '2024-11-24', '2024-12-01', '2024-12-08', '2024-12-15', '2024-12-22',

'2024-12-29', '2025-01-05', '2025-01-12', '2025-01-19']

Number of trips: 3559

29. Day Type 29:

Dates: ['2024-09-27']

Number of trips: 7154

30. Day Type 30:
Dates: ['2024-09-26']

Number of trips: 7146

31. Day Type 31:
Dates: ['2024-09-06', '2024-09-13', '2024-09-20']

Number of trips: 7064

32. Day Type 32:
Dates: ['2024-09-03', '2024-09-04', '2024-09-10', '2024-09-11', '2024-09-17', '2024-09-18',

'2024-09-24', '2024-09-25']

Number of trips: 7054

33. Day Type 33:
Dates: ['2024-09-09', '2024-09-16', '2024-09-23']

Number of trips: 7056

34. Day Type 34:
Dates: ['2024-09-05', '2024-09-12', '2024-09-19']

Number of trips: 7056

35. Day Type 35:
Dates: ['2024-09-02']

Number of trips: 138

36. Day Type 36:
Dates: ['2024-09-01', '2024-09-08', '2024-09-15', '2024-09-22']

Number of trips: 3010

37. Day Type 37:
Dates: ['2024-09-07', '2024-09-14', '2024-09-21', '2024-09-28']

Number of trips: 4047

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, **103** routes are active.

* Stops Served: Passengers can board or exit public transit at **4232** stops.

* Trips and Blocks: A total of **7872** trips are operated, managed by **631** blocks.

* Traveled Distance: On average, blocks have traveled **207** 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:

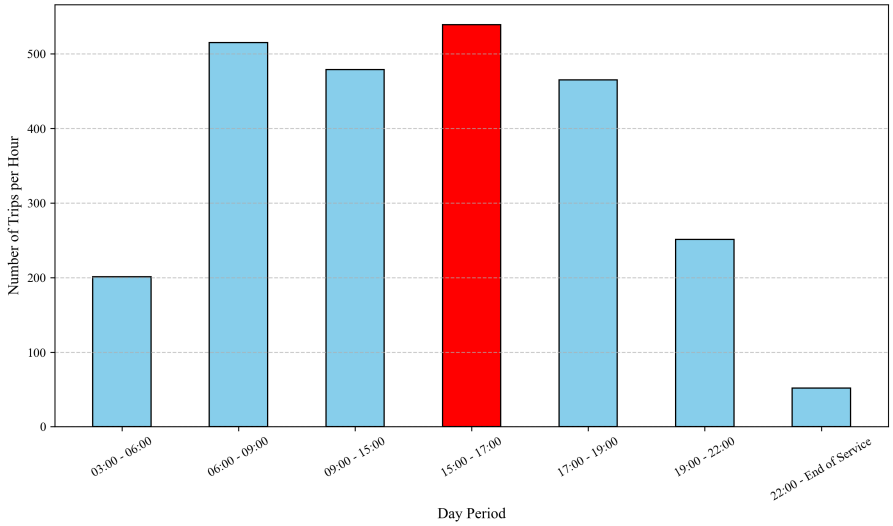
By analyzing the number of trips and their distribution across day periods, transit agencies can allocate resources (e.g., buses, drivers) more effectively. Understanding the average number of trips per hour during each period helps in identifying gaps or redundancies in service.

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	2	428	201
2	06:00	09:00	3	1546	515
3	09:00	15:00	6	2876	479
4	15:00	17:00	2	1078	539
5	17:00	19:00	2	930	465
6	19:00	22:00	3	752	251
7	22:00	End of Service	5	262	52



Routes

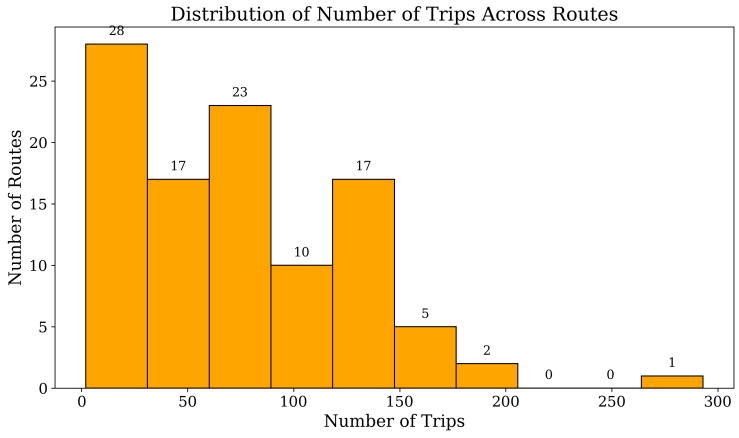
On the busiest day in the dataset, we analyzed the transit routes to determine their operational intensity by examining the number of trips and blocks assigned to each route. Each route operates with varying levels of frequency and service complexity, measured by the number of trips and blocks on that day. Notably, the route_id 561-30 recorded the highest number of blocks, with 28 blocks, while the route_id 561-510 had the maximum number of trips, totaling 293 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
561-510	293	21	Blue	San Ysidro - Utc	Streetcar
561-227	194	24	227	Imperial Beach - Otay Mesa Tc	Bus
561-7	183	20	7	Downtown San Diego - University/Colleg	Bus
561-215	174	12	215	Mid-City Rapid	Bus
561-13	158	17	13	Kaiser Hospital - 24th St Transit Cente	Bus
561-3	151	12	3	Ucsd Hospital - Euclid Transit Center	Bus
561-12	151	20	12	City College - Skyline Hills	Bus
561-41	148	26	41	Fashion Valley - Ucsd	Bus
561-530	144	9	Green	Santee - 12th & Imperial	Streetcar
561-44	143	21	44	Old Town - Kearny Mesa	Bus

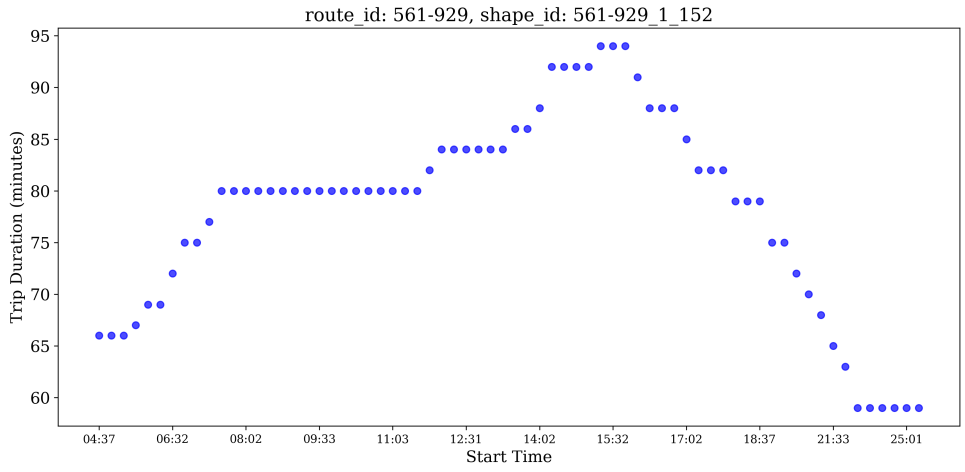
Trip frequency

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 [2, 31], containing 28 routes. The last interval, [263, 293], has a frequency of 1 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 561-929 and shape_id 561-929_1_152, we observed the highest standard deviation in trip duration, reflecting the most variable service in the dataset. This shape had a total of 68 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 561-929 and shape_id 561-929_1_152 at each day period.

Day Period ID	No. of Trips	Mean Duration	Duration Std
1	4	66	0
2	12	76	4
3	24	83	4
4	8	91	3
5	7	81	2
6	6	71	4
7	7	60	2