STRAVA | METRO

Comprehensive User Guide

Version 8.0 Effective February 03, 2019

Contents

Data Delivery

Your data will be delivered after the initial onboarding meeting with the Customer Success Team, in which we will help you customize your data and help answer any questions you might have before the data is built. If you are using your own basemap (instead of Open Street Map), please be sure to follow the guidelines in the Basemap Requirements document. If you need a copy of that document, please reach out to your Customer Success team member.

When your data is available, you will receive two emails. The first will be from your Customer Success team member, letting you know that your data is available for download. The second email will be from ExaVault, our FTP client, giving you your login credentials.

If you have purchased access to an ongoing subscription of Strava Metro data, your data will be delivered quarterly. Only the most recent delivery of data will be stored on our FTP.

Enterprise

Product - What's Included

- **Edges (Streets/Trails)**
  - Shapefile - Two options, as determined during the contract process
    - Open Street Map
    - Customer-provided
  - Rollups
    - Total
    - Monthly & Seasonal
    - Hourly Groupings (as subsets of Total, Monthly, & Seasonal)

- **Origin / Destination Polygons (OD)**
  - Shapefile - Two options, as determined during the contract process
    - 350 meter hexgrid
    - Customer-provided
  - Rollups
    - Total
    - Monthly & Seasonal
    - Hourly Groupings (as subsets of Total, Monthly, Seasonal, Weekday, & Weekend)
  - Trip Table Matrix (Hourly)

Note that all counts are rounded up to the nearest multiple of 5. For instance, counts of 4 show as 5. Counts of 8 show as 10, etc. All counts must meet a minimum of 3 users before showing counts on the street or between the OD pair.

File Structure

File Type	Timeframe	File Name
Roll Ups	Monthly	Region_Date_Activity_Rollup_Month_Year_{MonthNumber}_total Region_Date_Activity_Rollup_Month_Year_{MonthNumber}_weekend Region_Date_Activity_Rollup_Month_Year_{MonthNumber}_weekday
	Seasonal (Monthly and Hourly)	Region_Date_Activity_Rollup_Season_total Region_Date_Activity_Rollup_Season_weekend Region_Date_Activity_Rollup_Season_weekday
	Total (Yearly, Monthly, and Hourly)	Region_Date_Activity_Rollup_total Region_Date_Activity_Rollup_weekend Region_Date_Activity_Rollup_weekday
Hourly	YYYY/MM/DD/HH (Hourly Only)	Region_Date_Activity_Edge_Hourly.csv Region_Date_Activity_OD_Hourly.csv

Field Definitions

Edges - Rollups

Field	Definition
Edge_ID	Field used to join to the ID field in the shapefile. (Note: the ID field name in your shapefile may vary. Check your delivery email for the name of the ID field for your shapefile)
ATHCNT	Count of people on the street/trail segment, traveling in the direction of line digitization.
RATHCNT	Count of people on the street/trail segment, traveling in the opposite direction of line digitization.
ACTCNT	Count of trips on the street/trail segment, traveling in the direction of line digitization.
RACTCNT	Count of trips on the street/trail segment, traveling in the opposite direction of line digitization.
TATHCNT	Total count of people on the street/trail segment, regardless of the direction of travel.
TACTCNT	Total count of trips on the street/trail segment, regardless of the direction of travel.
ACTTIME	Median time in seconds of trips on the street/trail segment, traveling in the direction of line digitization.
RACTTIME	Median time in seconds of trips on the street/trail segment, traveling in the opposite direction of line digitization.
CMTCNT	Count of commute trips (point-to-point activities where the starting and ending point are more than 1km apart) traveling in the direction of line digitization.
RCMTCNT	Count of commute trips (point-to-point activities where the starting and ending point are more than 1km apart) traveling in the opposite direction of line digitization.
TCMTCNT	Total count of commute trips (point-to-point activities where the starting and ending point are more than 1km apart) regardless of the direction of travel.
Columns with _X	These column headers, when included with an underscore and a number (ie: _0, _1, etc), represent the hourly grouping rollups. For instance, _0 is typically the early morning hours, _1 is typically the morning commute hours, etc. These timeframes are determined by the customer during the contract process. Please see your initial data delivery email for details.

Edges - Hourly

Field	Definition
Edge_ID	Field used to join to the ID field in the shapefile. (Note: the ID field name in your shapefile may vary. Check your delivery email for the name of the ID field for your shapefile).
Year	Numerical year format (yyyy).
Day	Numerical day format (1-365).
Hour	Numerical hour format (0-24).
Minute	Numerical minute format. For hourly data, the value will be 0.
Athlete_Count	Count of people on the street/trail segment for that year, day, and hour, traveling in the direction of line digitization.
Rev_Athlete_Count	Count of people on the street/trail segment for that year, day, and hour, traveling in the opposite direction of line digitization.
Activity_Count	Count of trips on the street/trail segment for that year, day, and hour, traveling in the direction of line digitization.
Rev_Activity_Count	Count of trips on the street/trail segment for that year, day, and hour, traveling in the opposite direction of line digitization.
Total_Activity_Count	Total count of trips on the street/trail segment for that year, day, and hour, regardless of the direction of travel.
Activity_Time	Median time in seconds of trips on the street/trail segment for that year, day, and hour, traveling in the direction of line digitization.
Rev_Activity_Time	Median time in seconds of trips on the street/trail segment for that year, day, and hour, traveling in the opposite direction of line digitization.
Commute_Count	Total count of commute trips (point-to-point activities where the starting and ending point are more than 1km apart) for that year, day, and hour, regardless of the direction of travel.

OD - Rollups

Field	Definition
Origin	Field used to join to the ID field in the shapefile, indicating the polygon starting polygon (Note: the ID field name in your shapefile may vary. Check your delivery email for the name of the ID field for your shapefile).
Destination	Field used to join to the ID field in the shapefile, indicating the ending polygon (Note: the ID field name in your shapefile may vary. Check your delivery email for the name of the ID field for your shapefile).
ATHCNT	Count of people that traveled between that Origin/Destination pair.
ACTCNT	Count of trips made between that Origin/Destination pair.
CMTCNT	Count of trips classified as commutes made between that Origin/Destination pair.
Columns with _X	These column headers, when included with an underscore and a number (ie: _0, _1, etc), represent the hourly grouping rollups. For instance, _0 is typically the early morning hours, _1 is typically the morning commute hours, etc. These timeframes are determined by the customer during the contract process. Please see your initial data delivery email for details.

OD - Hourly

Field	Definition
Polygon_ID	Field used to join to the ID field in the shapefile, indicating the polygon starting polygon (Note: the ID field name in your shapefile may vary. Check your delivery email for the name of the ID field for your shapefile).
Dest_Polygon_ID	Field used to join to the ID field in the shapefile, indicating the ending polygon (Note: the ID field name in your shapefile may vary. Check your delivery email for the name of the ID field for your shapefile).
Year	Numerical year format (yyyy).
Day	Numerical day format (1-365).
Hour	Numerical hour format (0-24).
Minute	Numerical minute format. For hourly data, the value will be 0
Commute_Count	Total count of commute trips (point-to-point activities where the starting and ending point are more than 1km apart) for that year, day, and hour between that Origin/Destination pair.
Athlete_Count	Count of people for that year, day, and hour that traveled between that Origin/Destination pair.
Activity_Count	Count of trips made for that year, day, and hour, between that Origin/Destination pair.
Distance	Median distance traveled between that Origin/Destination pair, indicated in meters.
Duration	Median distance traveled between that Origin/Destination pair, indicated in seconds.

Reports

Demographics

Field	Definition
Activities	Count of activities included in the total timeframe of the product.
Athletes	Count of people included in the total timeframe of the product.
Age/Gender	Average and median distance of all activities in the total timeframe of the product, provided in meters.
Average/Median Distance	Count of trips made between that Origin/Destination pair.
Average/Median Time	Average and median time of all activities in the total timeframe of the product, provided in seconds.

File Geodatabase

Data is also offered as a file geodatabase. For users working in the ESRI suite of products, accessing the data via a file geodatabase has several performance and usability advantages over individual shapefiles and tables. If you selected to receive a file geodatabase delivery during the contract process, it will be included in your quarterly data delivery.

If you are interested in receiving a file geodatabase, please let your Customer Success team member know.

Working With The Enterprise Data

Shapefiles and Rollups

The shapefiles for Edges and OD pairs can all be joined with their respective rollups to do detailed geospatial analysis on your data. By joining the id field from the rollup (“EDGE_ID””, “ORIGIN”, or “DESTINATION”) to the “ID” field in the shapefile, you can associate all of the rolled up data with the respective features in the shapefile.

(Note: the ID field name in your shapefile may vary. Check your delivery email for the name of the ID field for your shapefile).

Hourly Files

Because the hourly data has a many-to-one relationship, the best way to work with this data is in a relational database. By loading these files into a relational database, you can create a wide range of custom queries related to temporal analysis (activities per hour, activities per day of the week, etc). The examples below use a Postgres database with the PostGIS extension enabled, as well as the pgAdmin user interface. The hourly data is designed to work in any relational database system.

Edges Hourly

Loading data into a relational database:

1) Create table

```
CREATE TABLE denver_sample_edges_ride_data
(edge_id INTEGER,
year INTEGER,
day INTEGER,
hour INTEGER,
minute INTEGER,
Athlete_count INTEGER,
rev_athlete_count INTEGER,
activity_count INTEGER,
rev_activity_count INTEGER,
total_athlete_count INTEGER,
total_activity_count INTEGER,
activity_time NUMERIC,
rev_activity_time NUMERIC,
commute_count INTEGER);
```

2) Import data into table

3) Create datetime column

```
ALTER TABLE denver_sample_edges_ride_data
ADD COLUMN datetime TIMESTAMP WITHOUT TIME ZONE;
```

4) Populate datetime column

```
UPDATE denver_sample_edges_ride_data
SET datetime = CAST(TIMESTAMP '2017-12-31' +
interval '1 day' * day  + interval '1 hour' * hour as
TIMESTAMP)
WHERE year = 2018;
```

5) Create indexes

```
CREATE INDEX denver_sample_datetime_idx
ON denver_sample_edges_ride_data
USING btree (datetime);

CREATE INDEX denver_sample_edge_id_idx
ON denver_sample_edges_ride_data
USING btree (edge_id);

CREATE INDEX denver_sample_hour_idx
ON denver_sample_edges_ride_data
USING btree (hour);
```

6) Query the data

a) Counts by day - check that your data is loaded for the full timeframe

```
SELECT day, count(*)
FROM denver_sample_edges_ride_data
GROUP BY day
ORDER BY day;
```

b) Sum activity counts by day on specific edge

```
SELECT day, SUM(total_activity_count)
FROM denver_sample_edges_ride_data
WHERE edge_id = 33
GROUP BY day
ORDER BY day;
```

b) Sum activity counts on specific edge by day of week

```
SELECT EXTRACT(dow from datetime)dow,
SUM(total_activity_count)
FROM denver_sample_edges_ride_data_hourly
WHERE edge_id = 33
GROUP BY EXTRACT(dow from datetime)
ORDER BY dow;
```

Origin/Destination hourly

Loading data into a relational database:

1) Create table

```
CREATE TABLE dc_sample_od_polygons_hex
(polygon_id INTEGER,
year INTEGER,
day INTEGER,
hour INTEGER,
minute INTEGER,
commute_count INTEGER,
athlete_count INTEGER,
activity_count INTEGER,
dest_polygon_id INTEGER,
distance NUMERIC,
duration INTEGER);
```

2) Import data

3) Query the data

a) Counts by origin polygon

```
CREATE TABLE dc_sample_polygon_starts as
SELECT polygon_id, sum(activity_count)
FROM denver_sample_od_hourly
GROUP BY polygon_id
```

b) Join query to shapefile

```
CREATE TABLE dc_sample_polygon_starts_gis as
SELECT * from dc_sample_od_polygons a, dc_sample_polygon_starts b
WHERE a.gid = b.polygon_id
```

If You Need Help

At Strava Metro, we are always happy to help our customers understand our data. Please contact us at any of the methods below:

cs@strava.com

Appendix

OSM Clazz Definitions

1) Priority
2) Clazz identifier (1-127)
3) Default speed in KMH
4) Allowed transportation type

Highway.motorway:	1, 11, 120, car
Highway.motorway_link:	1, 12, 30, car
Highway.trunk:	1, 13, 90, car
Highway.trunk_link:	1, 14, 30, car
Highway.primary:	1, 15, 70, car
Highway.primary_link:	1, 16, 30, car
Highway.secondary:	1, 21, 60, car
Highway.secondary_link:	1, 22, 30, car
Highway.tertiary:	1, 31, 40, car|bike
Highway.residential:	1, 32, 50, car|bike
Highway.road:	1, 41, 30, car|bike
Highway.unclassified:	1, 42, 30, car|bike
Highway.service:	1, 51, 5, car|bike
Highway.living_street:	1, 63, 7, car|bike|foot
Highway.pedestrian:	1, 62, 5, bike|foot
Highway.track:	1, 71, 10, bike|foot
Highway.path:	1, 72, 10, bike|foot
Highway.bridleway:	1, 73, 10, bike|foot
Highway.cycleway:	1, 81, 15, bike
Highway.footway:	1, 91, 5, foot
Highway.steps:	1, 92, 5, foot
Route.ferry:	2, 1, 10, ferry
Route.shuttle_train:	2, 2, 50, rail|car
Railway.rail:	3, 3, 50, rail
Railway.abandoned:	1, 74, 10, bike|foot