



# *Comprehensive User Guide*

Version 2.0 for 2015

Except for normal in-house use, reproduction of any part of this document by any means without permission from Strava LLC is prohibited.

©2015 by Strava LLC. All rights reserved. This material is proprietary and the subject of copyright protection and other intellectual property rights owned by Strava LLC. The use of this material is subject to the terms of a license agreement. You will be held liable for any unauthorized copying or disclosure of this material.

Strava Metro files and documentation are subject to all copyrights.  
Strava Metro is a trademark of Strava LLC.

ArcGIS® is a trademark of Esri®.

Every effort has been made by Strava LLC to depict accurately the current version of this product's content and use. However, the information in this document is subject to change without notice and should not be viewed as a commitment by Strava LLC. In addition, Strava LLC is not responsible for any errors in this document.

## TABLE OF CONTENTS

---

<b>WHAT IS IN THIS DOCUMENT?</b>	<b>6</b>	
<b>WHAT IS STRAVA METRO?</b>	<b>6</b>	
Strava Metro’s Powerful Dataset	6	
Strava Metro Versus Bike/Pedestrian Counters	7	
How the Strava Process Works	8	
Step 1	8	
Step 2	8	
Step 3	8	
Advantages of Using Strava Metro	10	
Uses for Strava Metro	10	
<b>STRAVA METRO DATA PRODUCTS: DESCRIPTION AND PRICING</b>		<b>11</b>
<b>Strava Metro: Streets</b>	<b>11</b>	
Description	11	
Pricing	11	
<b>Strava Metro: Origin/Destination</b>	<b>12</b>	
Description	12	
Pricing	13	
<b>Strava Metro: Nodes</b>	<b>14</b>	
Description	14	
Pricing	14	
<b>Strava Custom Heat Maps</b>	<b>14</b>	
Description	14	
Pricing	14	
<b>WHAT IS IN EACH STRAVA METRO DELIVERY?</b>	<b>15</b>	
Core Data	15	
Roll-ups	15	

<b>Reports</b>	<b>15</b>
<b>CHANGES FOR THIS RELEASE</b>	<b>15</b>
<b>IF YOU NEED HELP</b>	<b>15</b>
<b>Contact Information</b>	<b>15</b>
<b>Strava Metro Team</b>	<b>15</b>
<b>GETTING THE MOST OUT OF STRAVA METRO DATA</b>	<b>16</b>
<b>Program Versions Required</b>	<b>16</b>
<b>Indexes</b>	<b>16</b>
<b>QUICK START GUIDE</b>	<b>ERROR! BOOKMARK NOT DEFINED.</b>
<b>The Data Tables</b>	<b>Error! Bookmark not defined.</b>
<b>Join the Street Data Tables</b>	<b>Error! Bookmark not defined.</b>
<b>Viewing Cycling Attribution</b>	<b>Error! Bookmark not defined.</b>
<b>Simple Coloring Schemes Based on Use</b>	<b>Error! Bookmark not defined.</b>
<b>Identifying Commute Data</b>	<b>Error! Bookmark not defined.</b>
<b>Dealing with Small Pieces of Geometry</b>	<b>Error! Bookmark not defined.</b>
<b>Exploring the Minute Data</b>	<b>Error! Bookmark not defined.</b>
<b>Quick views into AM/PM Commute Windows</b>	<b>Error! Bookmark not defined.</b>
<b>Pulling up a day’s worth of data for one road from the raw DBF files</b>	<b>Error! Bookmark not defined.</b>
<b>PostGIS Queries</b>	<b>16</b>
<b>PRODUCT LAYERS AND CONTENTS</b>	<b>17</b>
<b>DATA DICTIONARY</b>	<b>18</b>
<b>Metro Edges</b>	<b>19</b>

<b>Metro_&lt;datatype&gt;: Year, Season, Month</b>	<b>19</b>
<b>Metro_Minute</b>	<b>22</b>
<b>Origin Destination Data</b>	<b>23</b>
<b>Origin Destination Polygon Data</b>	<b>23</b>
<b>Intersections (Nodes) Minute Data</b>	<b>24</b>
<b>Intersections (Nodes) Rolled-Up Data</b>	<b>25</b>
<b>Demographic Files</b>	<b>27</b>
<b>GLOSSARY</b>	<b>28</b>
<b>FAQS</b>	<b>28</b>
<b>How does Strava locate commutes?</b>	<b>28</b>
<b>Can anyone be part of the Strava Metro dataset?</b>	<b>28</b>
<b>What other products is Strava Metro designing?</b>	<b>28</b>
<b>How can I help improve alternative transportation in my community?</b>	<b>29</b>
<b>I am part of an advocacy group. How can I spread the word about Strava Metro?</b>	<b>29</b>
<b>What is the difference between the Strava Heat Map and Strava Metro?</b>	<b>29</b>
<b>I am interested in local popular cycling or pedestrian routes. Is Strava Metro what I'm looking for?</b>	<b>29</b>
<b>CURRENT USE CASES</b>	<b>29</b>

## WHAT IS IN THIS DOCUMENT?

This User Guide is designed to provide key insight into the Strava Metro suite of data services, their contents and use. In this document you will find:

- Explanations of the **data**, **table structures**, **general format** and **pricing** of Strava Metro data services, including **Streets**, **Origin/Destination**, **Nodes** and **Custom Heat Maps**.
- Detailed, step-by-step information on **how to begin using the data** in an ArcGIS environment.
- A **glossary** of Strava Metro definitions.
- A **FAQ section** with answers to common questions.
- How to **contact the Strava Metro team**.

### NOTE:

Hypertext links appear in **bolded blue** throughout this user guide. Use the “Control+Click” function on the blue text to follow a link.

## WHAT IS STRAVA METRO?

Strava Metro is a suite of data services that enables cutting-edge views into cycling and pedestrian (running, walking, hiking) patterns. Our mission is to produce state-of-the-art spatial data products and services to make cycling and pedestrian activities in cities safer and more efficient.

By providing visibility into cycling and pedestrian activity at a high level of granularity, Strava Metro data enables users to analyze patterns by time of day, day of week as well as by season and local geography. It is simple to filter the data to show only commuter data, as well.

### Strava Metro’s Powerful Dataset

The Strava community is made up of all types of cyclists and pedestrians. Through Strava’s analysis of the data, it is evident that these cyclists and pedestrians all tend to use the same “best available” roads and paths while in metro areas.

- The dataset currently includes over 400 billion GPS points. In denser metropolitan areas, commutes comprise more than one-half of this data; thus Strava Metro data gives great insight into the needs of those riding for transportation purposes only.
- In metro areas, nearly everyone is a commuter: either commuting to work or commuting to the activity they’ll be doing outside the city.
- In addition, over 2.5 million—and growing—GPS-tracked activities are uploaded to Strava every week from around the globe.

These activities create billions of data points that, when aggregated, enable deep analysis and understanding of real-world cycling and pedestrian route preferences. See the examples on the following page of a user-generated “best available” route preference versus an assumed cycling/pedestrian route.

**Fig. 1: Best available route generated by Strava data. Fig. 2: Assumed cycling/pedestrian route.**

### **Strava Metro Versus Bike/Pedestrian Counters**

Strava Metro takes tracking and measurement of cycling and pedestrian activity into the digital age: it is a supplemental method for the traditional way of collecting data using human counters or electronic gates to survey key intersections within a community. The advantage is that Strava collects data down to the minute at every geographic point on a city grid—24 hours a day, 7 days per week.

## How the Strava Process Works

### **Step 1**

---

Strava users track their rides, runs, walks and hikes with the Strava app on a smartphone or with a GPS device. Whether for commuting to/from work, a desire for a healthy lifestyle, recreation or other purposes, these users record their speed, route and other useful information.

**Fig. 3: Strava App and How Data Looks When Recorded**

### **Step 2**

---

To create a Strava Metro data service for a geographic area, Strava processes the data received to remove all personal information linked to the user so that cycling and pedestrian activities cannot be associated with a specific member of Strava's community. Protecting users' privacy and not selling their raw data is very important to us. The data then is structured for compatibility with classic geographical information systems (GIS) environments and aggregated to a linear street map such that the value of the data is preserved.

### **Step 3**

---

Strava Metro data enables detailed analyses of cycling and pedestrian patterns dissected by time of day, day of week, season and local geography. See the following page for an example.

**Fig. 4: New York City Ride Image**

## **Advantages of Using Strava Metro**

By blending GIS and programming, Strava Metro makes huge amounts of data useable and ideal for deep analysis.

- Strava's custom format allows for quick data extraction and clean views of the data, bypassing problem areas such as polyline-only data that does not show direction/time or point-only data clusters that can be challenging to work with.
- Using Strava Metro provides quick data that can be used to evaluate the effects and effectiveness of changes to cycling lanes or pedestrian areas.

Strava Metro is the result of years of research and collaboration between Strava GEO and different departments of transportation (DOTs) and transportation research groups. Strava Metro:

- Works with whatever basemap you already are using.
- Is dynamic and built to provide seamless integration into a planning department.
- Data is provided in a number of different temporal formats from minute to yearly summaries.
- Is produced in classic GIS formats so it easily can be integrated into GIS environments of transportation engineering firms and DOTs.
- User-defined fields allow for a deep understanding of the data provided.

**Fig. 5: Seattle Weekday**

**Fig. 6: Seattle Weekend**

## **Uses for Strava Metro**

Too many regions around the world have no idea where anyone is riding, running, walking or hiking in their area.

- Using Strava Metro, departments of transportation and city planners, as well as advocacy groups and corporations, now can make informed and effective decisions when planning, maintaining and upgrading cycling and pedestrian corridors.
- Strava Metro data also can be used to evaluate efficacy of changes made to infrastructure within a few weeks after implementation, as opposed to waiting months or years to gather data.

Potential uses of Strava Metro include—but are not limited to—the following:

- Locating general cycling and pedestrian routes.
- Isolating commute data versus recreation data.
- Identifying pinch points.
- Auto cycling counter placement.
- Pre/Post new cycling and pedestrian infrastructure.
- Weekend versus weekday behaviors.
- Tourism trends.
- Blending with counter data for extrapolated road use.
- Mountain bike network use and preferred flow.

## STRAVA METRO DATA PRODUCTS: DESCRIPTION AND PRICING

The Strava Metro suite of data services includes **Streets**, **Origin/Destination**, **Nodes** and **Custom Heat Maps**. Organizations can license Strava Metro data and tools for use with GIS mapping software.

- Licensing costs are based on the number of Strava members in the requested geographic area and the time span of data required.
- Strava Metro is offered in two Temporal Plans: Previous 12 months (one-time delivery) and/or the Next 12 months (12 monthly deliveries).

### **Strava Metro: Streets**

Strava Metro Streets is the core data service we offer. It is defined as a minute-by-minute reflection of all cycling and pedestrian activity over your specified data range in a defined region at the street level.

#### **Description**

---

- Unique cyclists counts per direction on each street at the minute.
- Total cycling counts per direction on each street at the minute.
- Ability to extract commute/utilitarian trips from the total collection of trips.
- Time in seconds of how long it takes the cycling population to cross each piece of road in each direction.
- Custom built on the preferred polyline basemap.

#### **Pricing**

---

- License fees are \$0.80 per unique Strava user that started an activity in the AOI (Area of Interest) for the first 10,000 users (subject to a \$1,000 minimum).
- After the first 10,000 users the rate drops to \$0.70 per user and provides a point of flexibility in the pricing.

Fig. 11: Utah Streets and Trails

### **Strava Metro: Origin/Destination**

Strava Metro: Origin/Destination builds off of the Streets data service by reporting the starting and ending polygons of activities. The Origin/Destination service was the logical first new data feed for the Strava Metro suite. The data in this format shows great functionality of zonal transportation by bike and on foot.

When paired with the minute routing data, the data then shows the key paths used between each zone. This data is reported by the minute and covers beyond your street network to get polygons of users starting outside of the AOI.

#### **Description**

---

- Complete path of all polygons intersected included in the data.
- Trips are reported at the minute and have a commute flag allowing for the separation of trip types.
- Custom built on the preferred polygon network (Blocks, Wards).

## Pricing

---

- The cost to license this data is \$0.10 per unique user, subject to a \$500 minimum. It has a lower cost because it does not require the entire activity GPS points to be pulled and takes a fraction of the time to process.
- This data can be licensed separately from the Streets service.

**Fig. 12: NYC Origin/Destination**

## **Strava Metro: Nodes**

Intersections are a key piece of any street network. How long will a cyclist or pedestrian have to wait? What are the busiest intersections in the city? Do cyclists or pedestrians stay away from them? These all are the questions that the Nodes product will answer. The Node data service is designed to focus on these intersections and primarily report use and wait duration.

### **Description**

---

- Collections of waiting times at every intersection with the basemap network.
- Perfect for cycling transportation modeling (impedance).

### **Pricing**

---

- The cost to license this data is \$0.10 per unique user, subject to a \$500 minimum.
- Please note that Nodes requires a Streets license for the same AOI and time frame.

## **Strava Custom Heat Maps**

Customers wanting greater detail and higher quality of visualization of a large collection of GPS points recorded by Strava members should consider purchasing the Strava Metro Custom Heat Map.

### **Description**

---

The Strava Heat Map is a way to visualize the world of Strava. For example, a popular street for cyclists looks like a bright line in the Strava Heat Map.

### **Pricing**

---

Custom pricing is available, depending upon a number of factors.

**NOTE:**

Advocacy organizations and the general public now can access high-resolution heat map visualizations of the data free of charge at [Strava Labs](#).

## WHAT IS IN EACH STRAVA METRO DELIVERY?

There are two temporal formats in which Strava Metro is delivered: Previous 12 months (one-time delivery) and/or the Next 12 months (12 monthly deliveries). Whether the data is in a moving-forward contract or historical, the data delivered is in the same format.

### Core Data

- Street level database file (sql or csv) with minute-to-minute cycling/pedestrian information on every street for the time period of the delivery.
- Intersection database file (sql or csv) with minute to minute cycling/pedestrian information on node (street intersection) for the time period of the delivery.
- Origin/Destination file of pairs for all trips during the time period of the delivery.

### Roll-ups

- Street/Nodes: Each core dataset will have a set of summarized views that show monthly use, weekend versus weekday, seasonality, hours groupings and total counts. This typically results in over 50 rolled up views for the cycling data.
- OD: Trip tables that show collective trip matches.

### Reports

- Demographics: Report of Strava users included in each data delivery (age, gender).
- Tourism: Breakdown of the home location (county level) of all Strava users over the time frame.

## CHANGES FOR THIS RELEASE

No changes from the last release.

## IF YOU NEED HELP

We at Strava are always happy to help you understand our data. Please feel free to contact us at any of the following e-mail addresses:

### Contact Information

Strava

35 S. Main Street

Hanover, NH

Please send feedback and requests to [geo@strava.com](mailto:geo@strava.com).

### Strava Metro Team

Brian Riordan, Manager, Strava Metro: [brian@strava.com](mailto:brian@strava.com)

Davis Kitchel, GEO Team Lead: [dkitchel@strava.com](mailto:dkitchel@strava.com)

Michael Horvath, President: [michael@strava.com](mailto:michael@strava.com)

## GETTING THE MOST OUT OF STRAVA METRO DATA

Strava Metro is designed for use by geographic information system (GIS) professionals familiar with GIS software and engaged in city planning. To get the most out of Strava Metro data, you need a spatially enabled GIS program: Esri's ArcMap or QGIS. To use the sql raw files you need to have postgres/PostGIS installed.

### Program Versions Required

The files are provided as shapefiles and .dbf data files so they can be used by QGIS or ArcGIS. The project files (.mxd) are built in ArcGIS 10.2+ and oracle spatial.

- Postgres Version: 9.1+.
- PostGIS Version: 2.0+.
- ArcGIS Version: 10.1.+.
- QGIS.
- Google Earth (to view KML).

### Indexes

Shapefile indexes often become corrupt and need to be rebuilt. By indexing the ID fields, the joins between all the files will be much faster. In addition, make sure there is a spatial index on the street network.

### PostGIS Queries

The power of this data is at the database level. Being able to generate custom views of the data is key to being able to fully use the data. We provide a set of rolled-up views but ultimately these are not going to be the views that will work for all planning groups. Below are a few queries that can be used to build new rolled-up views and tables in Postgres/PostGIS.

**Build Cycling Data for Times and Days Example:** This SQL builds a new table that can be exported as a new .dbf file. The key values that can be modified here are the hour windows. By changing these values you can change to time range you are rolling up. The other value is at the end of the query *where day > 0 and day < 32* by moving the day range you can look at specific weeks or different groups of months.

**NOTE:**

Postgres does not have a native Median calculation included. You might need to add this function:

[https://wiki.postgresql.org/wiki/Aggregate\\_Median](https://wiki.postgresql.org/wiki/Aggregate_Median)

**Hour by Day Summaries**

```

select edge_id,
sum(rider_count) as TRiderCnt,
sum(rev_rider_count)as TRRiderCnt,
sum(ride_count) as TRideCnt,
sum(rev_ride_count)as TRRideCnt,
sum(bike_count) as BikeCnt,
median(CASE when (ride_count > 0) then bike_time end)as BikeTime,
median(CASE when (rev_ride_count > 0) then rev_bike_time end)as RBikeTime,
sum(commute_count) as CommuteCnt,
sum(CASE when (hour between 6 and 10) then rider_count end) as AMRider2013,
sum(CASE when (hour between 6 and 10) then rev_rider_count end) as AMRRider2013,
sum(CASE when (hour between 6 and 10) then ride_count end) as AMRide2013,
sum(CASE when (hour between 6 and 10) then rev_ride_count end) as AMRRide2013,
sum(CASE when (hour between 6 and 10) then bike_count end) as AMBikeCount2013,
median(CASE when (hour between 6 and 10 and ride_count > 0) then bike_time end)as AMBikeTime2013,
median(CASE when (hour between 6 and 10 and rev_ride_count > 0) then rev_bike_time end)as AMBikeTime2013,
sum(CASE when (hour between 6 and 10) then commute_count end) as AMCommute2013,
sum(CASE when (hour between 16 and 19) then rider_count end) as PMRider2013,
sum(CASE when (hour between 16 and 19) then rev_rider_count end) as PMRRider2013,
sum(CASE when (hour between 16 and 19) then ride_count end) as PMRide2013,
sum(CASE when (hour between 16 and 19) then rev_ride_count end) as PMRRide2013,
sum(CASE when (hour between 16 and 19) then bike_count end) as PMBikeCount2013,
median(CASE when (hour between 16 and 19 and ride_count > 0) then bike_time end)as PMBikeTime2013,
median(CASE when (hour between 16 and 19 and rev_ride_count > 0) then rev_bike_time end)as PMBikeTime2013,
sum(CASE when (hour between 16 and 19) then commute_count end) as PMCommute2013
from edges_data_balt
where day > 0 and day < 32
group by edge_id

```

**PRODUCT LAYERS AND CONTENTS**

This section outlines what is contained in a delivery of Strava Metro. The Strava Metro product is constantly evolving as we locate and build in key features. The delivery contains the following data files:

- .csv/.sql raw hourly data file.
- .dbf rolled-up views in the cycling data (listed in the table below).
- Origin/Destination data table raw.
- Origin/Destination Polygon.
- Nodes data table raw.
- Nodes point file.
- Demographics document.
- Product description document.
- ESRI .MXD project file.
- ESRI .MPK map package file.

## DATA DICTIONARY

Below is the list of the core data files contained in Strava Metro:

TABLE NAME	TYPE
<identifier>_edges	Line
<identifier>_nodes	Point
<identifier>_od_polygons	Polygon
<identifier>_<ride/run>_rollup_month_<year>_<month>_weekend	Table
<identifier>_<ride/run>_rollup_month_<year>_<month>_weekday	Table
<identifier>_<ride/run>_rollup_season_<month>_weekend	Table
<identifier>_<ride/run>_rollup_season_<month>_weekday	Table
<identifier>_<ride/run>_rollup_total_weekend	Table
<identifier>_<ride/run>_rollup_total_weekday	Table
<identifier>_<ride/run>_rollup_total	Table
<identifier>_edges_<ride/run>_data	.csv/.sql/Table
<identifier>_edges_nodes_<ride/run>_data	.csv/.sql/Table
<identifier>_<ride/run>_od_polygons_data	.csv/.sql/Table

**Table 1 definition:** There is a naming structure to the files so that a name would look similar to joondalup\_city\_edges\_nodes\_ride\_rollup\_month\_2013\_6\_weekend.

**<identifier>**=Unique name of the region or group that is purchasing the data; for example, joondalup\_city. The identifier is referenced in the e-mail that is sent with the product.

**<year>**=The numeric year of the rollup in conjunction with the following month.

**<month>**=A numeric value of 1–12 that represents the months (Jan.–Dec.).

### Metro Edges

**Name:** <identifier>\_edges.shp  
**File Type:** Shapefile, PostGIS, OSM  
**Datum:** WGS84 [SRID = 4326]

If the map used for the data came from Open Street Map, the OSM street network is converted from the .pbf (xml) format into a PostGIS table during the process. This table is then exported to a shapefile with corresponding related data tables. The id field joins the street network to the rolled up data. Not all OSM attributes are carried through because the Strava Metro process does not require them.

#	FIELD	TYPE	DESCRIPTION
1	Id	Double	Unique and permanent OSM Edge ID number for delivery.
2	osm_id	Double	Preserved OSM street id; this is not stable and is intended only to be used for historical reference.
3	osm_name	Text	Street name from the OSM data.
4	osm_meta	Text	Metadata from OSM about the street.
5	Clazz	Double	Street class according to OSM.
6	Flags	Double	Any key flags from the OSM data.
7	source	Double	Starting nodes for the line.
8	target	Double	Ending nodes for the line.
9	Km	Double	Distance of the line in km.
10	length	Double	Distance of the line in meters.
11	the_geom**		Binary internal representation of geometry.

\*\*the\_geom field only in PostGIS formats

### Metro <datatype>: Year, Season, Month

**Name:** <identifier>\_<ride/run>\_rollup\_month\_<year>\_<month>\_weekend.dbf  
**File Type:** DBF  
**Datum:** N/A

To facilitate the use of the Strava Metro product, we provide a set of rolled up data files that show pre-determined temporal ranges. These ranges are extracted into monthly, peak riding season and yearly files. These temporal ranges vary with latitude and season. Below are the date and hourly ranges for the AM/PM data used for a typical Strava Metro product.

- On-Season: April 1–Oct 1
- Very Early AM hours: 12am–3:59am
  - Labeled as \_0
- Early AM hours: 4am–5:59am
  - Labeled as \_1
- AM Peak Hours: 6am–8:59am
  - Labeled as \_2
- Mid-Day Hours: 9am–2:59pm
  - Labeled as \_3
- PM Peak Hours: 3pm–5:59pm

- Labeled as \_4
- Early Evening Hours: 6pm–7:59pm
  - Labeled as \_5
- Late Evening Hours: 8pm–11:59pm
  - Labeled as \_6

For each data file there is a sum for that time frame, and a group for all requested time frame roll-ups. This creates numerous fields per data file and allows for fast pinpointed analysis. If there is a different time frame you are interested, in please let us know. These time frames are built-up from what was defined in the contract.

**NOTE:**

When a FIELD name starts with an “R” (as in “RACTCNT”) or with “REV” (as in REV\_ATHLET), this indicates the number **going against** the direction the street was digitized.

#	FIELD	TYPE	DESCRIPTION
1	EDGE_ID	Double	Unique and permanent Street ID number for delivery.
2	ATHCNT	Double	Count of unique cyclists on the piece of street for the rolled-up date frame. This number represents the number of cyclists going the direction the street was digitized (see Glossary).
3	RACTCNT	Double	Count of unique cyclists on the piece of street for the rolled-up date frame. This number represents the number of cyclists going against direction the street was digitized (see Glossary).
4	ACTCNT	Double	Count of bike trips (regardless of unique riders) on the piece of street for the rolled-up date frame. This number represents the number of cyclists going the direction the street was digitized (see Glossary).
5	RACTCNT	Double	Count of bike trips (regardless of unique riders) on the piece of street for the rolled-up date frame. This number represents the number of cyclists going the direction the street was digitized (see Glossary).
6	TATHCNT	Double	Total number of unique cyclists on the piece of street regardless of direction of travel for the rolled-up date frame.
7	TACTCNT	Double	Total number of bike trips on the piece of street regardless of direction of travel for the rolled-up date frame.
8	ACTTIME	Double	Median time in seconds that bike trips on the piece of street for the rolled-up date frame. This number represents the time of cyclists going the direction the street was digitized (see Glossary).
9	RACTTIME	Double	Median time in seconds that bike trips on the piece of street for the rolled-up date frame. This number represents the time of cyclists going against direction the street was digitized (see Glossary).
10	CMTCNT	Double	Total number of commute bike trips on the piece of street regardless of direction of travel for the rolled-up date frame.
11	ATHCNT_X	Double	Count of unique cyclists on the piece of street for the rolled-up date frame between the predefined time frames as noted above (where X=the numbered time frame). This number represents the number of cyclists going the direction the street was digitized (see Glossary).

12	RATHCNT_X	Double	Count of unique cyclists on the piece of street for the rolled-up date time frame between the predefined time frames as noted above (where X=the numbered time frame). This number represents the number of cyclists going against direction the street was digitized (see Glossary).
13	ACTCNT_X	Double	Count of bike trips (regardless of unique riders) on the piece of street for the rolled-up time frame between the predefined time frames as noted above (where X=the numbered time frame). This number represents the number of cyclists going the direction the street was digitized (see Glossary).
14	RACTCNT_X	Double	Count of bike trips (regardless of unique riders) on the piece of street for the rolled-up time frame between the predefined time frames as noted above (where X=the numbered time frame). This number represents the number of cyclists going the direction the street as digitized (see Glossary).
15	TATHCNT_X	Double	Total number of unique athletes on the piece of street regardless of direction of travel for rolled-up time frame between the predefined time frames as noted above (where X=the numbered time frame).
16	TATHCNT_X	Double	Total number of bike trips on the piece of street regardless of direction of travel for rolled-up time frame between the predefined time frames as noted above (where X=the numbered time frame).
17	ACTTIME_X	Double	Median time in seconds for bike trips on the piece of street during the rolled-up time frame between the predefined time frames as noted above (where X=the numbered time frame). This number represents the time of cyclists going the direction the street was digitized (see Glossary).
18	RACTTIME_X	Double	Median time in seconds for bike trips on the piece of street during the rolled-up time frame between the predefined time frames as noted above (where X=the numbered time frame). This number represents the time of cyclists going against direction the street was digitized (see Glossary).
19	CMTCNT_X	Double	Total number of commute bike trips on the piece of street regardless of direction of travel for the rolled-up time frame between the predefined time frames as noted above (where X=the numbered time frame).
XXX	Pattern Cont.		The naming pattern continues for all predefined time frames; for example, if the day had 4 time frames then it would be _0, _1, _2, _3.

## Metro Minute

**Name:** <identifier>\_edges\_<ride/run>\_data.csv, <identifier>\_edges\_<ride/run>\_data.dbf

**File Type:** Comma-delimited, DBF (only if the file is less than 2 gigs)

**Datum:** N/A

The minute data is the granular format of the Strava Metro product. This file contains a row for every street edge and minute that a cyclist crossed it over the defined time. These tend to be very large datasets in the tens and hundreds of millions of records for an annual review.

**NOTE:**

When a FIELD name starts with an “R” (as in “RACTCNT”) or with “REV” (as in REV\_ATHLET), this indicates the number **going against** the direction the street was digitized.

#	FIELD	TYPE	DESCRIPTION
1	edge_id	Integer	Unique and permanent Open Street Map Edge ID number for delivery.
2	year	Integer	Numerical year format (yyyy).
3	day	Integer	Numerical day format (1–365).
4	hour	Integer	Numerical hour format (0–24).
5	minute	Integer	Numerical minute format (0–59).
6	athlete_count   ATHLETE_CO	Integer	Count of unique cyclists on the piece of street for the day, hour and minute. This number represents the number of cyclists going the direction the street was digitized (see Glossary).
7	rev_athlete_count   REV_ATHLET	Integer	Count of unique cyclists on the piece of street for the day, hour and minute. This number represents the number of cyclists going against direction the street was digitized (see Glossary).
8	activity_count   ACTIVITY_C	Integer	Count of bike trips (regardless of unique riders) on the piece of street for the day, hour and minute. This number represents the number of cyclists going the direction the street was digitized (see Glossary).
9	rev_activity_count   REV_ACTIVI	Integer	Count of bike trips (regardless of unique riders) on the piece of street for the day, hour and minute. This number represents the number of cyclists going the direction the street was digitized (see Glossary).
10	total_activity_count   TOTAL_ACTI	Integer	Total number of bike trips on the piece of street regardless of direction of travel for the day, hour and minute.
11	activity_time   ACTIVITY_T	Integer	Median time in seconds that bike trips on the piece of street for the day, hour and minute. This number represents the time of cyclists going the direction the street was digitized (see Glossary).
12	rev_activity_time   REV_ACT_1	Integer	Median time in seconds that bike trips on the piece of street for the day, hour and minute. This number represents the time of cyclists going against direction the street was digitized (see Glossary).

13	commute_count   COMMUTE_CO	Integer	Total number of commute bike trips on the piece of street regardless of direction of travel for the day, hour and minute.
----	-------------------------------	---------	---

**Origin Destination Data**

**Name:** <identifier>\_<ride/run>\_od\_polygons\_data.csv; <identifier>\_<ride/run>\_od\_polygons\_data.dbf

**File Type:** .csv

**Datum:** N/A

This data has the starting and ending polygon for all the rides that made up the data delivery. It also has an array of all of the polygons that a trip touched along the way. This is a new way of looking at recreation and commute trips while further protecting a user’s privacy.

We do not provide roll-ups of the OD because there is not a roll-up we can provide that cannot be built quickly in a GIS program.

**NOTE:**

To use the array you cannot load it into a .dbf file, as the .dbf has a 255-character limit. It works best in a database that can take advantage of an array.

#	FIELD	TYPE	DESCRIPTION
1	polygon_id	Integer	Unique Starting Polygon ID number.
2	year	Integer	Numerical year format (yyyy).
3	day	Integer	Numerical day format (1–365).
4	hour	Integer	Numerical hour format (0–24).
5	minute	Integer	Numerical minute format (0–59).
6	commute	Integer	0 -1 Flag for locating commutes (1=commute).
7	dest_polygon_id	Integer	Unique Ending Polygon ID number.
8	intersected_polygons	Integer	This is an array of polygons that were intersected during this activity. Only present in the .csv format.

**Origin Destination Polygon Data**

**Name:** <identifier>\_od\_polygons.shp

**File Type:** Polygon Shapefile, postgres sql file

**Datum:** N/A

This data has the starting and ending polygon for all the rides that made up the data delivery. It also has an array of all of the polygons that a trip touched along the way. This is a new way of looking at recreation and commute trips while further protecting a user’s privacy.

**Intersections (Nodes) Minute Data****Name:** <identifier>\_edges\_nodes\_<ride/run>\_data**File Type:** Point Shapefile**Datum:** N/A

The minute data is the granular format of the Strava Metro product. This file contains a row for every node and minute that a cyclist crossed that node over the defined time. These tend to be very large datasets in the tens and hundreds of millions of records for an annual review.

#	FIELD	TYPE	DESCRIPTION
1	node_id	Integer	Unique and permanent Node ID number for delivery.
2	year	Integer	Numerical year format (yyyy).
3	day	Integer	Numerical day format (1–365).
4	hour	Integer	Numerical hour format (0–24).
5	minute	Integer	Numerical minute format (0–59).
6	athletes	Integer	Count of unique cyclists at the intersection for the day, hour and minute. This number represents the number of cyclists that meet at the intersection.
7	activities	Integer	Count of activities at the intersection for the day, hour and minute. This number represents the number of activities that meet at the intersection.
8	median_wait	Integer	Median wait time at the intersection for that minute.
9	max_wait	Integer	Maximum wait time at the intersection for that minute. This can be very high at times, as sometimes cyclists can stop and cause longer than normal wait times.
10	min_wait	Integer	Minimum wait time at the intersection for that minute.
11	commute	Integer	Count of commute trips at the intersection for the day, hour and minute. This number represents the number of trips that meet at the intersection that are defined as a commute.

### Intersections (Nodes) Rolled-Up Data

**Name:** <identifier>\_nodes\_<ride/run>\_rollup\_month\_<year>\_<month>\_weekend.dbf

**File Type:** DBF

**Datum:** N/A

To facilitate the use of the Strava Metro Nodes product, we provide a set of rolled-up data files that show pre-determined temporal ranges. These ranges are extracted into weekend, weekday and yearly files. Below are the date and hourly ranges for the AM/PM data typically used when generating this product.

- Very Early AM hours: 12am–3:59am
  - Labeled as `_0`
- Early AM hours: 4am–5:59am
  - Labeled as `_1`
- AM Peak Hours: 6am–8:59am
  - Labeled as `_2`
- Mid-Day Hours: 9am–2:59pm
  - Labeled as `_3`
- PM Peak Hours: 3pm–5:59pm
  - Labeled as `_4`
- Early Evening Hours: 6pm–7:59pm
  - Labeled as `_5`
- Late Evening Hours: 8pm–11:59pm
  - Labeled as `_6`

For each data file there is a sum for that time frame, and a group for all requested time frame roll-ups. This creates many fields per data file and allows for fast pinpointed analysis. If there is a different time frame you are interested in, please let us know. These time frames are built-up from what was defined in the contract.

#	FIELD	TYPE	DESCRIPTION
1	NODE_ID	Integer	Unique and permanent Node ID for delivery.
2	ATHCNT	Integer	Count of unique cyclists at the intersection summarized for the time frame.
3	ACTCNT	Integer	Count of activities at the intersection summarized for the time frame.
4	MEDMEDT	Double	This is a bit more complicated as it is the median value of a set of Median values. In this case it the Median of the Median wait time of the activities at the intersection summarized for the time frame.
5	MEDMAXT	Double	This is a bit more complicated as it is the median value of a set of median values. In this case it the Median of the Maximum wait time of the activities at the intersection summarized for the time frame.
6	MEDMINT	Double	This is a bit more complicated as it is the median value of a set of Median values. In this case it the Median of the Minimum wait time of the activities at the intersection summarized for the time frame.
7	CMTCNT	Double	Sum of the commute activities at the intersection summarized for the time frame.

11	ATHSCNT_X	Double	Count of unique cyclists at the intersection summarized for the predefined time frames as noted above (where X=the numbered time frame).
12	ACTSCNT_X	Double	Count of activities at the intersection summarized for the rolled-up date time frame between the predefined time frames as noted above (where X=the numbered time frame).
13	MEDMEDT_X	Double	This is a bit more complicated as it is the median value of a set of Median values. In this case it the Median of the Median wait time of the activities at the intersection summarized between the predefined time frames as noted above (where X=the numbered time frame).
14	EDMAXT_X	Double	This is a bit more complicated as it is the median value of a set of median values. In this case it the Median of the Maximum wait time of the activities at the intersection summarized for the rolled-up time frame between the predefined time frames as noted above (where X=the numbered time frame).
15	MEDMINT_X	Double	This is a bit more complicated as it is the median value of a set of Median values. In this case it the Median of the Minimum wait time of the activities at the intersection summarized for rolled-up time frame between the predefined time frames as noted above (where X=the numbered time frame).
16	CMTCNT_X	Double	Sum of the commute activities at the intersection summarized for the time frame for rolled-up time frame between the predefined time frames as noted above (where X=the numbered time frame).
XXX	Pattern Cont.		The naming pattern continues for all predefined time frames; for example, if the day had 4 time frames then it would be _0, _1, _2, _3.

## Demographic Files

**Name:** <identifier>\_Demographics.txt

**File Type:** text files

We understand how important demographic data is to analysis. Therefore, we built a file that is designed to help you have a deeper understanding of the users who make up the data Strava is supplying. The Demographics document shows demographic detail generated by the data sample at the region requested. Below you can find the fields and their definitions.

FIELD	DESCRIPTION
County Name	The name of the polygon region to which the data is being summarized.
Athlete ID Count	Number of unique athlete IDs that had a ride start in the region.
Activity Count	Number of unique cycling activity IDs that had a ride start in the region.
Male Count	Number of male athletes that occupy that region as assigned by Strava.
Male Count Under 25	Number of male athletes that occupy that region as assigned by Strava: Under 25.
Male Count 25 - 34	Number of male athletes that occupy that region as assigned by Strava: 25–34.
Male Count 35 - 44	Number of male athletes that occupy that region as assigned by Strava: 35–44.
Male Count 45 - 54	Number of male athletes that occupy that region as assigned by Strava: 45–54.
Male Count 55 - 64	Number of male athletes that occupy that region as assigned by Strava: 55–64.
Male Count 65 - 74	Number of male athletes that occupy that region as assigned by Strava: 65–74.
Male Count 75 - 84	Number of male athletes that occupy that region as assigned by Strava: 75–84.
Male Count 85 - 94	Number of male athletes that occupy that region as assigned by Strava: 85–94.
Male Count No Bday	Number of male athletes that occupy that region as assigned by Strava: Unspecified Birth Day.
Female Count	Number of female athletes that occupy that region as assigned by Strava.
Female Count Under 25	Number of female athletes that occupy that region as assigned by Strava: Under 25.
Female Count 25 - 34	Number of female athletes that occupy that region as assigned by Strava: 25–34.
Female Count 35 - 44	Number of female athletes that occupy that region as assigned by Strava: 35–44.
Female Count 45 - 54	Number of female athletes that occupy that region as assigned by Strava: 45–54.
Female Count 55 - 64	Number of female athletes that occupy that region as assigned by Strava: 55–64.
Female Count 65 - 74	Number of female athletes that occupy that region as assigned by Strava: 65–74.
Female Count 75 - 84	Number of female athletes that occupy that region as assigned by Strava: 75–84.
Female Count 85 - 94	Number of female athletes that occupy that region as assigned by Strava: 85–94.
Female Count No Bday	Number of female athletes that occupy that region as assigned by Strava: Unspecified Birth Day.
Blank Gender Count	Number athletes that occupy that region as assigned by Strava with no specified gender.

## GLOSSARY

TERM	DESCRIPTION
Athletic Location	Athletes' locations are generated by a process of computing a locus based on the start points of current activities.
Digitization Direction	The direction the street network was drawn is important to understand when working with Strava Metro. That is how we assign ride/pedestrian counts to the different travel directions. An extra 'R' in the header name indicates the opposite of digitization direction. In other words, the direction it is referring to is from the end node to the start node.
Edge	An edge is a piece of road geometry between nodes (intersections).
Rolled Up	Rolled up indicates a view showing totals, sums and averages of the minute data for a given temporal range.
Time	Time is the median value of all recorded times for that edge. We do not use average time because there are instances where cyclists or pedestrians will stop on a piece of road to go into a store or talk with a friend while their GPS is still running.
Zero Value	A zero value happens when the edge is very small. For example, if the edge is less than 1 meter long and a cyclist is moving at 20mph, then the GPS pulse will show no change for that piece of road. This means that from available time data, the cyclist is on both ends of the road at the same time.

## FAQS

We at Strava are happy to answer your questions, hear your opinions about our data or offer suggestions for using our data. Always feel free to contact us (see If You Need Help earlier in this User guide).

Below are some typical frequently asked questions we have received from people just like you.

### How does Strava locate commutes?

Commuter data is derived by three methods:

- The first is the commute flag that is native to the Strava experience.
- The second is an automated process that locates point-to-point cycling and pedestrian trips that are within duration and distance constraints.
- The third is fuzzy name matching from the activity titles.

### Can anyone be part of the Strava Metro dataset?

Anyone using Strava to publicly track their rides, walks, runs and other fitness activities is anonymously contributing to the Strava Metro dataset with every upload. It is a way for you to vote with your ride, run, walk or hike for better cycling and pedestrian infrastructure within your community.

### What other products is Strava Metro designing?

We see many potential products in the future for Strava Metro. In the near-term we are working on enhanced visualizations and tools for use with Strava Metro data by users without access to and experience in GIS software. We also are looking into providing tools to facilitate the analysis of the data so groups can use it as effectively as possible.

### **How can I help improve alternative transportation in my community?**

If you are a cyclist, runner, walker or hiker, use Strava to track your activities. In addition to the fun and motivation Strava provides, your data will be in the Strava Metro system and can help inform alternative transportation systems in your community. And tell your friends who ride and run—Strava grows by word of mouth.

### **I am part of an advocacy group. How can I spread the word about Strava Metro?**

Tell your DOT and city planning colleagues about Strava Metro. We have successfully partnered with several advocacy organizations to jointly approach the planning authorities in their area in presenting what Strava Metro can do.

### **What is the difference between the Strava Heat Map and Strava Metro?**

The Strava Heat Map is a way to visualize the world of Strava. Strava Metro is a product that contains the data behind the visualization for the purposes of analysis and infrastructure planning. For example, a popular street for cyclists looks like a bright line in the Strava Heat Map; Strava Metro provides data about how many cyclists rode in which direction on that street minute-by-minute.

### **I am interested in local popular cycling or pedestrian routes. Is Strava Metro what I'm looking for?**

No. You are looking for Strava Routes. We've counted the "votes" of millions of cyclists and pedestrians globally and made them available via a route creation and discovery tool.

## **CURRENT USE CASES**

To see how others currently are using Strava Metro data, just click on Join the User Community @: <https://strava.zendesk.com/forums/21670164-Strava-Metro-Discussion>. This link will take you to customer-generated case studies, maps and tools based on their Strava Metro experience.