Teaching & Learning
Step-by-Step Guide:

Mapping Walking Patterns with Smartphones and GIS

Amila Jayasinghe

Harini Sawandi

Samith Madusanka

Teaching & Learning Step by -Step Guide:

Mapping Walking Pattern with Smartphones and GIS

Authors

Amila Jayasinghe

Harini Sawandi

Samith Madusanka

Publisher

University of Moratuwa

Author contribution

1. Amila Jayasinghe (Supervision, Conceptualisation, Methodology, Validation), Department of Town & Country Planning, University of Moratuwa, Sri Lanka.
2. Harini Sawandi (Formal Analysis, Writing—original draft preparation), Department of Town & Country Planning, University of Moratuwa, Sri Lanka.
3. Samith Madusanka (Project Administration, Review and Editing), Department of Town & Country Planning, University of Moratuwa, Sri Lanka.

All authors have read and agreed to the published version of the book.

Contact authors amilabj@uom.lk

This book was produced with the valuable support of the Erasmus+ Capacity Building in Higher Education (CBHE) project 'Curricula Enrichment for Sri Lankan Universities delivered through the application of Location-Based Services to Intelligent Transport Systems' (LBS2ITS https://lbs2its.net/)

Project Number: 618657-EPP-1-2020-1-AT-EPPKA2-CBHE-JP
Programme: Erasmus+
Key Action: Cooperation for innovation and the exchange of good practices
Action Type: Capacity Building in Higher Education
Co-funding: Erasmus+ Programme of the European Union

This book was reviewed as an Open Education Resource for University students by Prof Retscher Günther (Vienna University of Technology, Austria) under the LBS2ITS project.

LBS2ITS
Curricula Enrichment delivered through
the Application of Location-based Services
to Intelligent Transport Systems

lbs2its.net

Co-funded by the Erasmus+ Programme of the European Union

Edition

First Edition - May 2025

Copyright

ISBN 978-955-9027-93-5 (ebook)

Citation

Jayasinghe, A., Sawandi, H., & Madusanka, S. (2025). *Teaching & learning step-by-step guide—Mapping walking pattern with smartphones and GIS* (1st ed.). University of Moratuwa.

Disclaimer

Publisher

University of Moratuwa

PREFACE

This book serves as open educational material for both undergraduate and postgraduate degree programs, offering a detailed, step-by-step guide to mapping walking patterns using smartphones and GIS technology. Designed to bridge the gap between theoretical knowledge and practical application, this guide is meticulously crafted to meet the needs of students, educators, and practitioners alike.

Within the book, readers will find comprehensive instructions on using smartphones to collect data, employing GIS software to analyze spatial information, and interpreting this data to understand walking patterns in various contexts. The book not only enhances learning in academic settings by providing real-world applications and case studies but also equips industry professionals with the skills necessary to conduct advanced spatial analysis and contribute meaningful insights in their fields.

Whether you are a student aiming to master mobile mapping and geographic information systems, a teacher looking for robust educational tools, or a practitioner in need of refining your technical expertise, this book offers invaluable guidance and support. It ensures that users at all levels gain proficiency in leveraging modern technologies to explore and solve geographic challenges effectively.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

1. MAPPING PROCESS

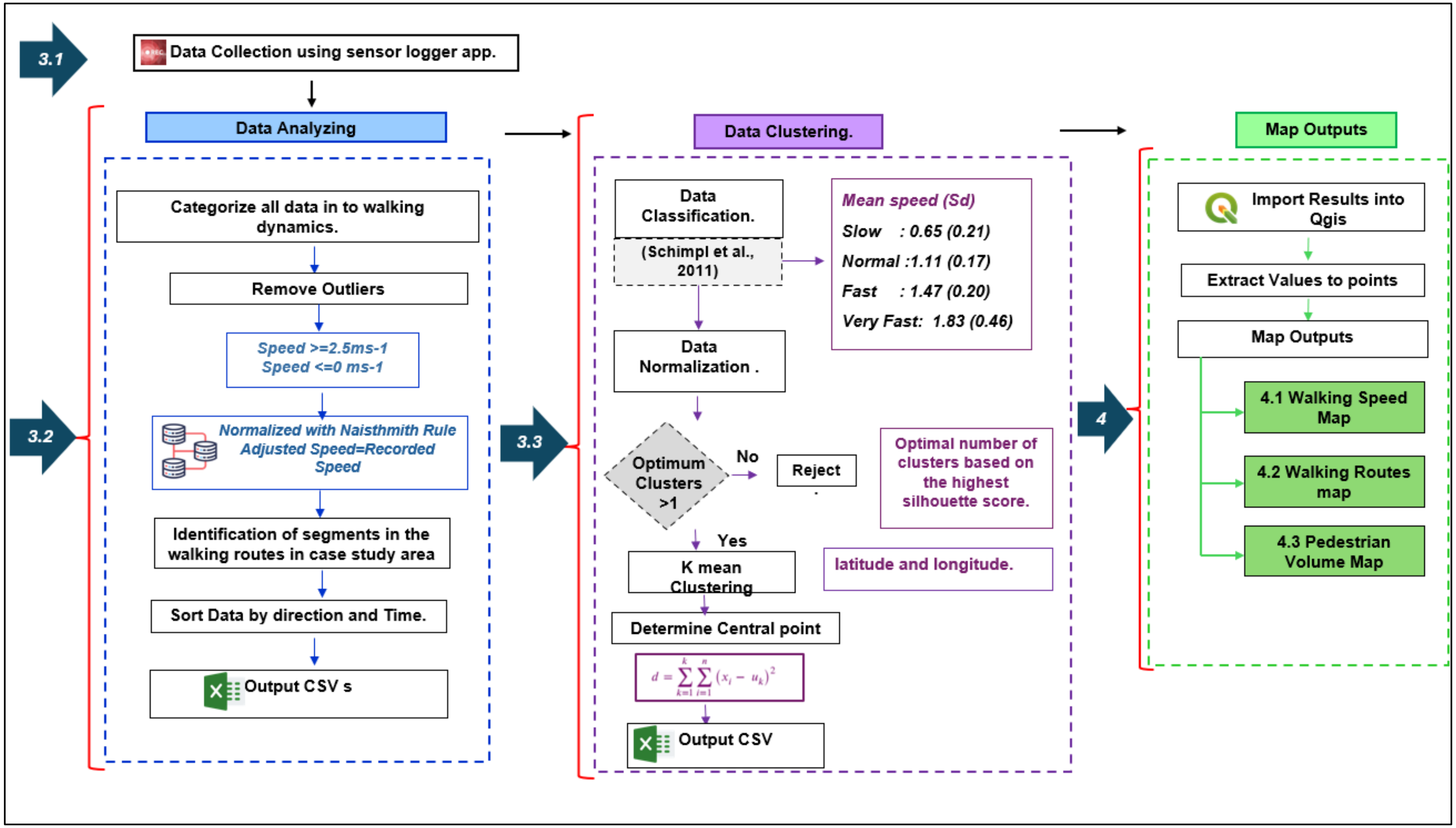


Figure 1 -Mapping Process

2. REQUIRED TOOLS / SOFTWARE AND EQUIPMENT

EQUIPMENT'S	SOFTWARE'S	INPUTS
Multi-constellation GNSS phones • Select a multi-constellation GNSS smartphone with the capability to log location and/or raw data. • Data collection activities can drain the battery quickly, so choose a phone with a long-lasting battery. • Storage Capacity and Processing Power	**Q GIS** • Version: QGIS 3.0 or later • Installation: Available for Windows, macOS, Linux	 **Base Map of the Study Area** • Students need to have a proper base map of the study area in any format preferred. This could include Satellite Imagery, Open Street Maps, Custom Maps Online Maps, etc.
 **Smartphone Holder** • Adjustable holder to secure smartphone to the waistband	**Google Colab** • Cloud-based Jupyter Notebook environment for Python • Web Browser: Google Chrome, Firefox, Safari, etc. • Internet Connection: Required for real-time collaboration	

Table 1 – Required Items for study.

3. STEPS

3. 1 Data Collection

3.1.1 Experiment Details

As pointed out in the literature which investigated several mobile phone sensor placements including hip (belt), wrist, upper arm, ankle, and thigh, using numerous accelerometers aids in activity identification. Thus, as illustrated in Figure 2 in this study participants were advised to secure their mobile phones to their waistbands with designated holders. This strategy guaranteed the phones remained in a consistent and steady position throughout the data collection process. They were advised to keep the mobile phone in a vertical orientation within the holder to preserve consistency. This system enabled the accurate gathering of data on walking speed, acceleration, and GPS position during the study period. To effectively capture speed data when walking, hold the phone upright in portrait orientation by ensuring that the X-axis (horizontal) is parallel to the direction of movement whereas the Y-axis (vertical) is perpendicular to the ground.

Why waistband?

- Reflects natural body movement during walking, enhancing speed detection.
- Minimizes interference from movements or shifting positions.
- Minimizes signal obstruction and improves GPS position accuracy.

Why this orientation?

This orientation enables the phone's sensors to better capture forward movement (along the X-axis) and up-and-down motion (along the Y-axis), both of which are important for estimating speed.

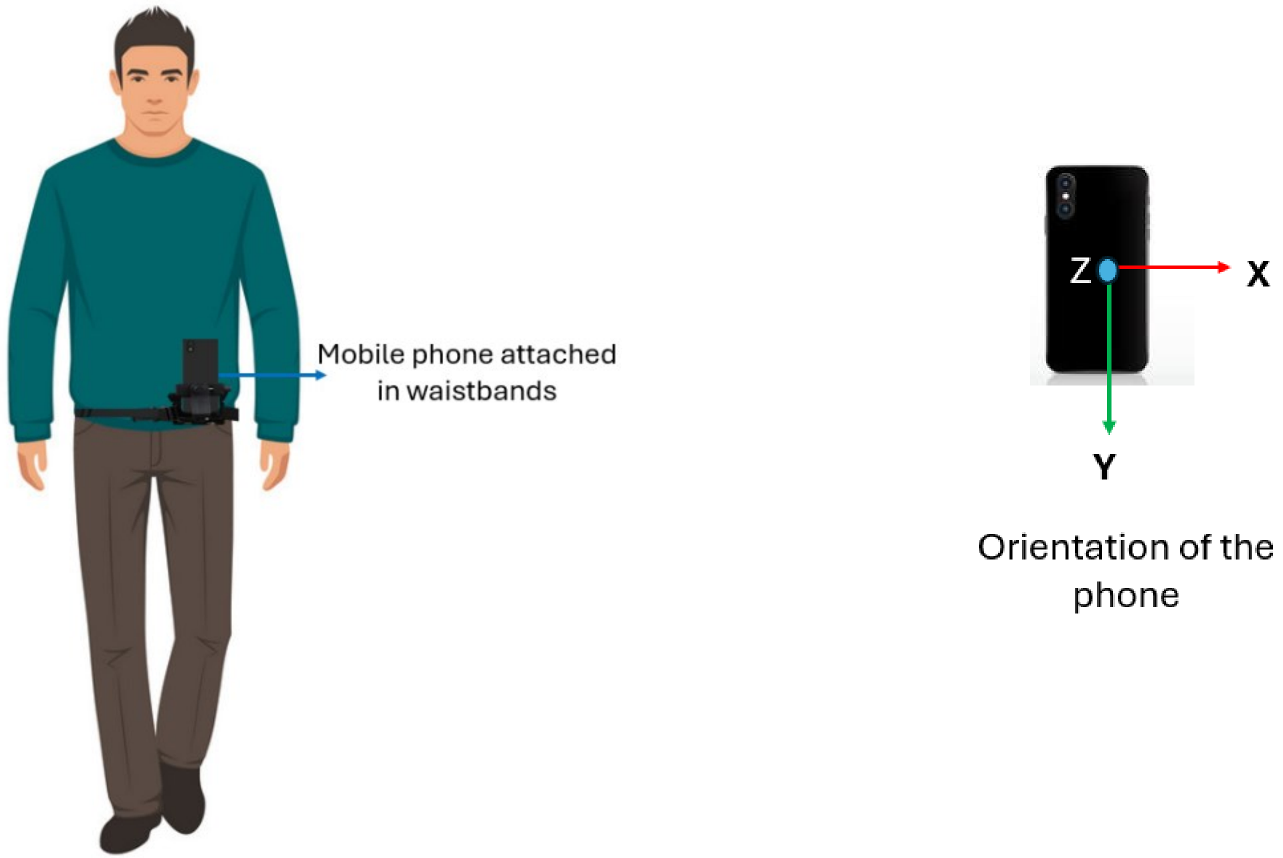


Figure 2 -Experiment Design

3.1.2 Other body placement areas to hold their mobile devices.

Body placement	Suitable data types for capturing [2]
Thigh	Ambulation, walking, standing, sitting, and running
Hip	Ambulation, posture
Chest, thigh, wrist, forearm	Ambulation, posture, typing, talking, bicycling
Wrist	Arm movements, heart rate (if the wearable device includes a heart rate monitor), step count
Head (wearable)	Head orientation (for VR/AR devices), Eye movement
Footwear	Gait Analysis (using pressure sensors or imus)

Table 2 - Body Placement areas for attached sensors.

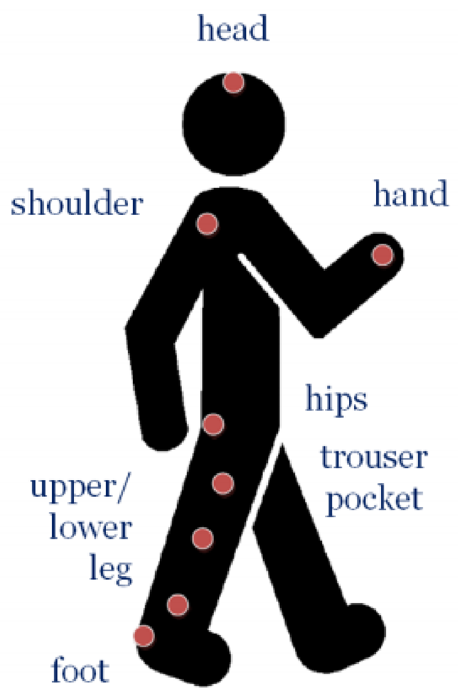


Figure 3 -Mounting of IMUs on a human body(1)

1. *Retscher, G. (2022). Indoor navigation—User requirements, state-of-the-Art and developments for smartphone localization. Geomatics, 3(1), 1-46. doi:10.3390/geomatics3010001*
2. *Twomey, N., Diethe, T., Fafoutis, X., Elsts, A., McConville, R., Flach, P., & Craddock, I. (2018). A comprehensive study of activity recognition using accelerometers.* Informatics, 5(2), *27. doi:10.3390/informatics5020027*

was developed by Kelvin Choi" (Sensor logger. Kelvin Choi. (n.d.). https://www.tszheichoi.com/sensorlogger)

“Sensor Logger” app that can be used to collect objective participant data. It mostly records the participants' geographical locations and time stamps as they walk the path, Sensor Logger is an open-source mobile application that helps organizations author, collect, and manage mobile data. Sensor Logger is a cross-platform data logger that logs readings from common motion-related sensors on smartphones. Once completed, recordings can be exported as a zipped CSV file, JSON, SQLite, or KML, or be viewed within the app via interactive plots. The iOS version comes with a free companion watch app that allows you to log heart rate & wrist motion, and control recording sessions remotely from your wrist.

*Note * If you are performing cross-platform analysis using Sensor Logger, there are minor unit differences and important coordinate system differences between iOS and Android.*

Figure 4 -Interface of the Mobile application

Source - https://www.tszheichoi.com/sensorlogger.

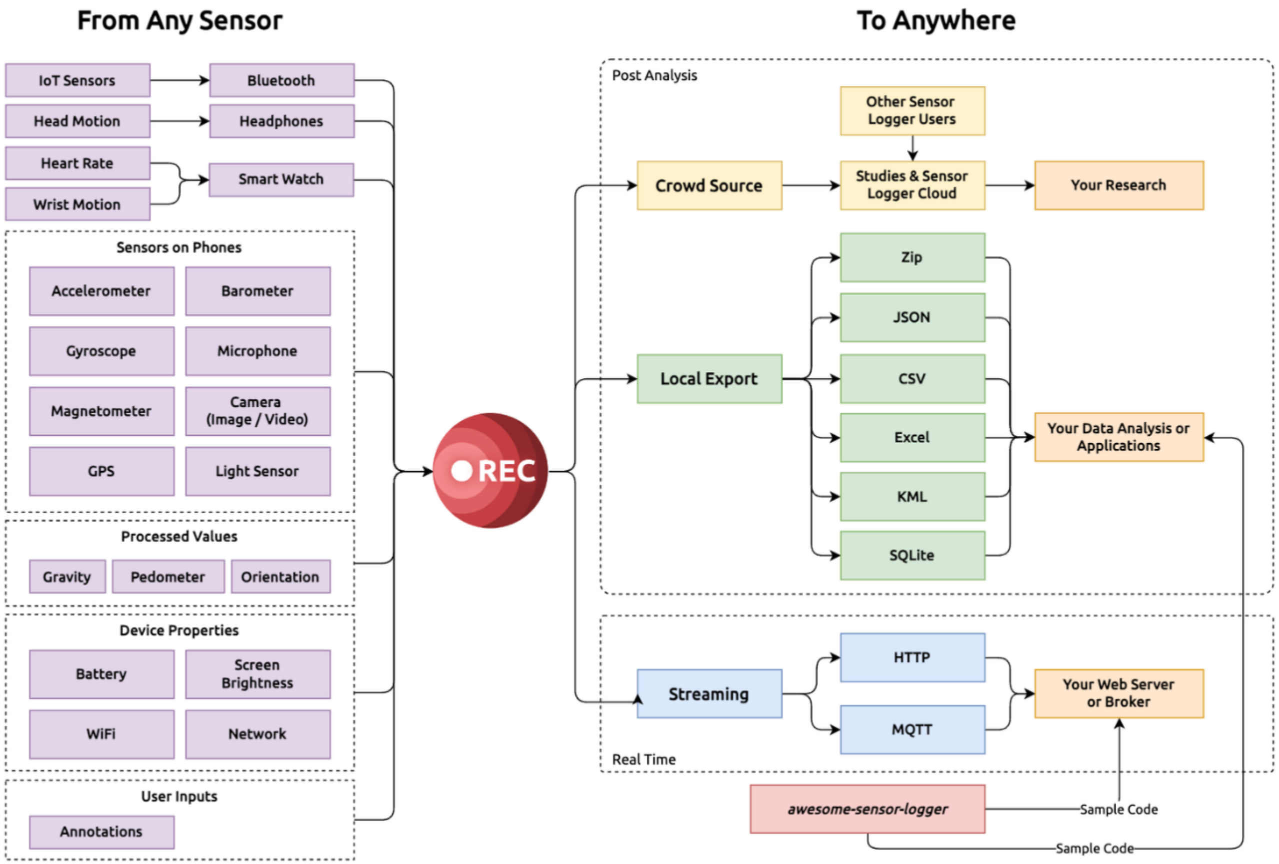


Figure 5 -The Complete system Mobile application

Source - https://www.tszheichoi.com/sensorlogger

Sensor Logger connects every sensor to anywhere you want your data to be — both in real-post-analysis.

Why Sensor Logger?

- Multifunctional Sensor Integration
- User-Friendly Interface
- Real-Time Visualization
- Open-Source and Cross-Platform Compatibility
- Recorded data can be exported in various formats.
- Both android and iOS
- Terrain Variation Adaptation

Minimum Specifications for Installation?

Required OS – Android 5.0 and up.

Required iOS\ iPad\ iPod touch – IOS 13.0 or later.

Apple vision - Vison OS 1.0 or Later

Apple watch OS 8.0 or later

Data Privacy and Ethical Considerations?

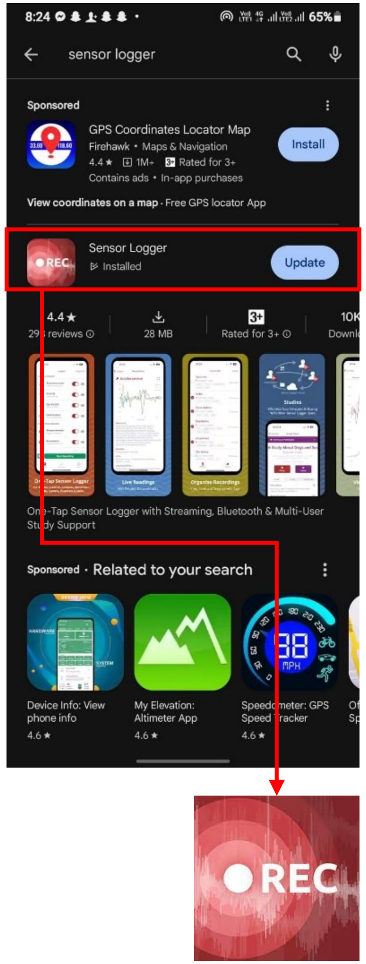
The privacy of participants will be thoroughly respected during the data collection procedure. The data collected will be only utilized for research objectives and will not be disclosed to any outside parties. In addition, participants own the right to voluntarily stop their involvement in the study.

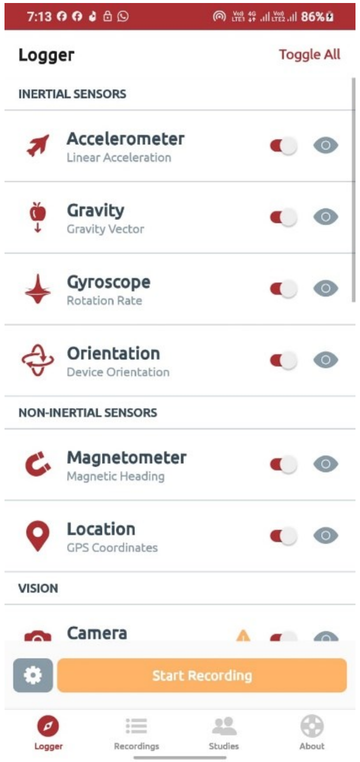
Feature Highlights

- Comprehensive sensor support
- One-tap logging
- Background recording
- Stream data via HTTP / MQTT
- Zipped CSV, JSON, Excel, KML & SQLite exports
- Resample and aggregate measurements.
- Enable & disable specific sensors.
- Scan and log nearby Bluetooth devices.
- Add timestamp-synchronized annotations during recording.
- Adjust sampling frequency for sensor groups.
- Custom recording workflow
- Raw and calibrated measurements
- Live plots and readings for sensors
- Organize, sort, and filter recordings.
- Control & track recording sessions from Watch.
- Bulk export & delete recordings.
- Resources to help you analyses your data.

Step 01 –

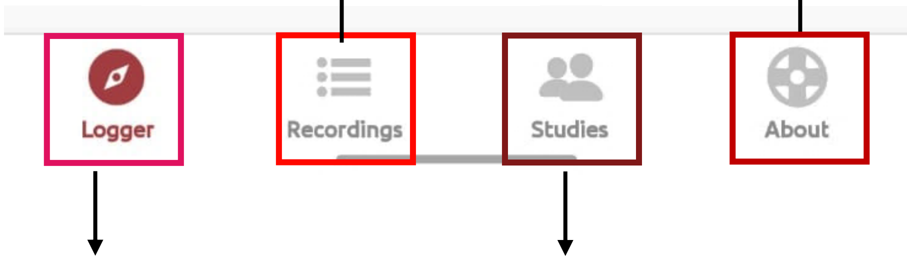
Installation: Download ‘Sensor Logger’ from the respective app store for your smartphone (iOS or Android).





On the Recordings Screen, you can view all your past recordings. By default, they sorted from newest to oldest, but you can sort them differently by tapping the sort of icon on the upper left. You can view, rename, export, or delete recordings by selecting them.

The About Screen provides information about the app, including its version number and the developer’s contact information. Additionally, you can find links to additional resources to help you analyses your data, and the privacy policy of the app.



The Logger Screen allows you to start logging sensor data with a single tap. You can toggle which sensors you want to include in your recording.

You can create or join Studies. Studies allow you to collect data from or contribute your data to other Sensor On this screen, you can see all the Studies you have joined as a participant, or that you own as an investigator.

Figure 6 -Icon of the app

Figure 7 -Interface of the app

Figure 8 -Main tabs of the app

Step 02

Enable GPS (Location) on Mobile Phone

Note - Some sensors require additional permissions. If they are not granted yet, there will be a yellow warning icon next to it. When prompted for location permission, the app will ask you to choose from options such as ***'Allow Once,' 'Allow While Using the App,' or 'Don't Allow.'*** *For this study, please select* ***'Allow Once'*** *to enable location access only for the duration of the data collection session. However, depending on your phone's security settings, you may see variations like* ***'Allow All the Time****.' In such cases, consider the security of your device and select the appropriate method accordingly.*

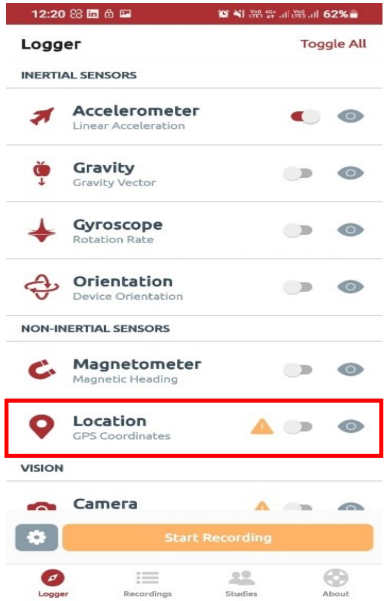


Figure 10 – Location Icon

This mobile application is equipped with several types of sensors, including inertial sensors, non-inertial sensors, and vision sensors, as depicted in Figure 9. In addition to the location feature, you have the flexibility to enable specific sensors within the app based on the requirements of your study. This allows for customization and adaptation of sensor usage according to the objectives of future research.

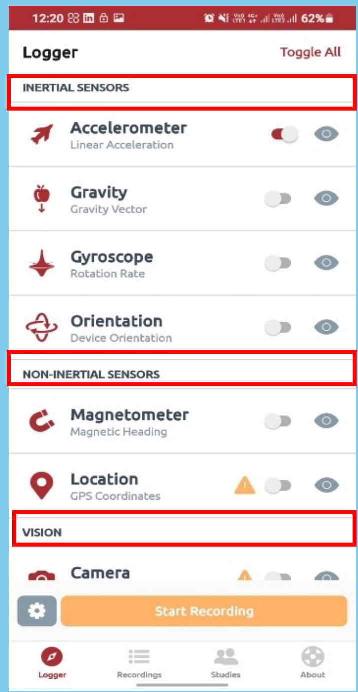


Figure 9 - Different Sensors

What options does this study use for mapping pedestrian walking behavior?

This study Enable the "Location" option.

Note – In the experiment you can ask participants to start the recording when they start there walking.

Step 03 –

Click "Start Recording".

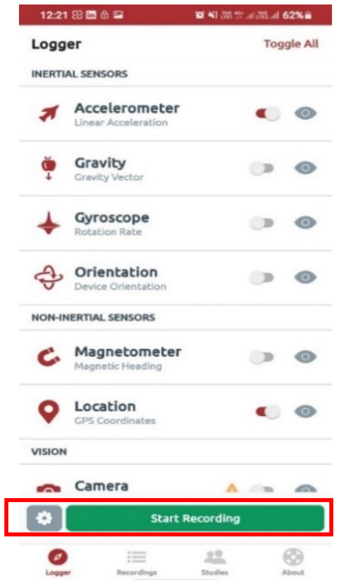


Figure 11: Start recording icon.

Instructions To Participants

Before initiating the recording process, participants briefed on the proper handling of their mobile phones, as illustrated in Figure 2. They were instructed to securely hold the device during the walking session.

- Participants were trained to hold their mobile phones securely in a designated manner.
- It was recommended for participants to lock their phones after initiating the recording. This helps maintain stability and consistency in data collection.
- Participants were instructed to walk along their designated routes as they would normally do during their daily activities.
- To minimize potential data errors, participants were highly encouraged not to answer phone calls or engage in other activities that might disrupt the walking session.

Step 04

Click "End Recording"

Note – Participants should end the recording once they have completed the walking session.

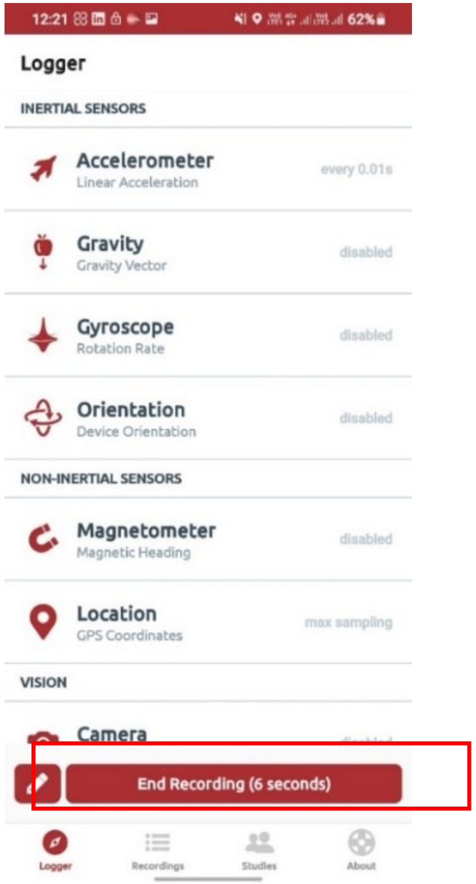


Figure 12 -End Recording Option

Step 05

Click "Recording" Tab

The recorded data is automatically saved within the "Recording" tab of the Sensor Logger app.

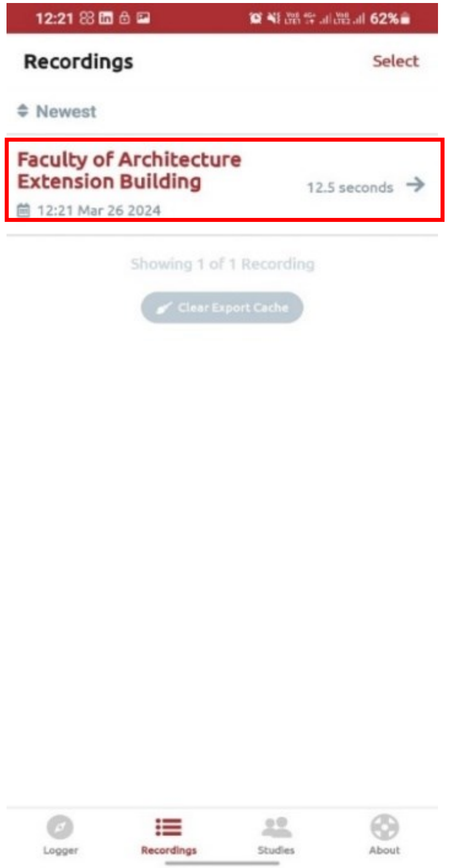


Figure 13 -Saved File

Step 06

Click on the saved file under the Recording Tab : When you are click on the saved file there will be 4 option available for you namely “Rename, Tag, Export, Delete”.

To export the saved file, click on “**Export” icon.**

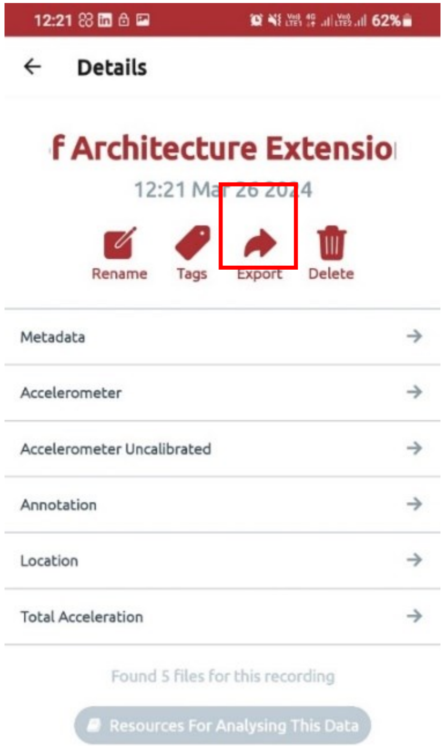


Figure 14 -Export icon

Step 07

Chose the Export format: For this study, the CSV (Comma-Separated Values) format has been selected for data export.

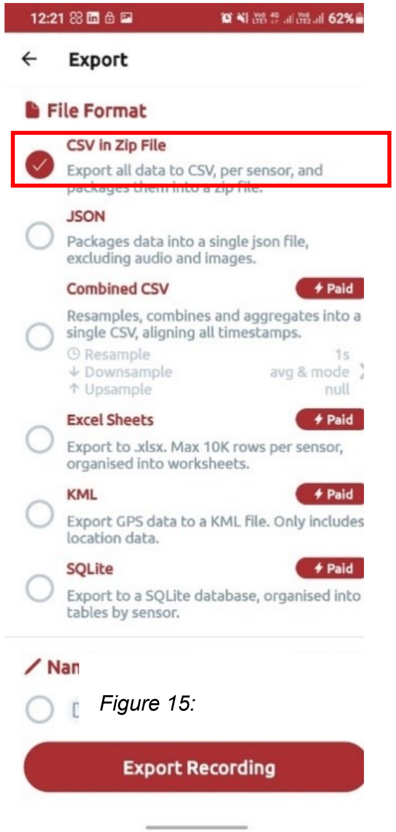


Figure 15:

Figure 15 -Export Formats

Step 08

Click on “Export Recording”

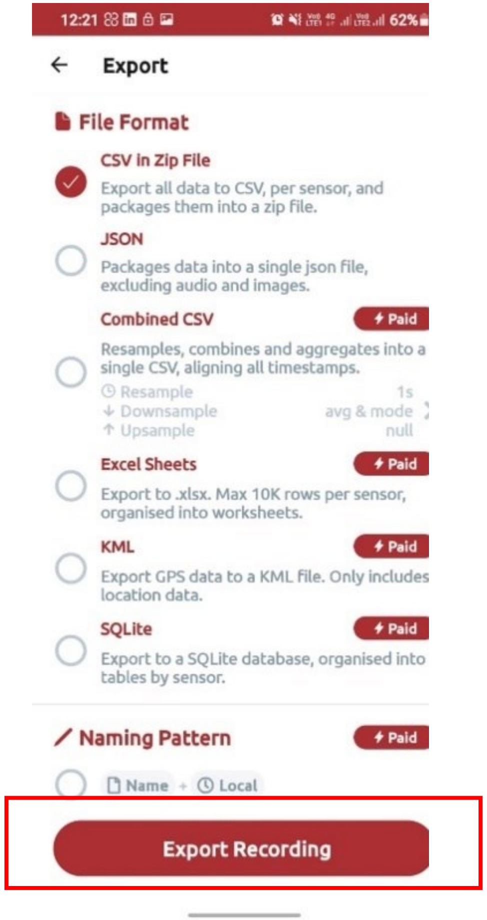


Figure 16 -Export Options

Step 09

Once you click on "Export," you can seamlessly share the data through various platforms, as illustrated in the figure. The exported file automatically saves to your device's internal storage, making it readily accessible for further analysis or sharing. This streamlined process allows you to quickly transfer the CSV file to external locations such as email, messaging apps, cloud storage services, or any other platform supported by your device.

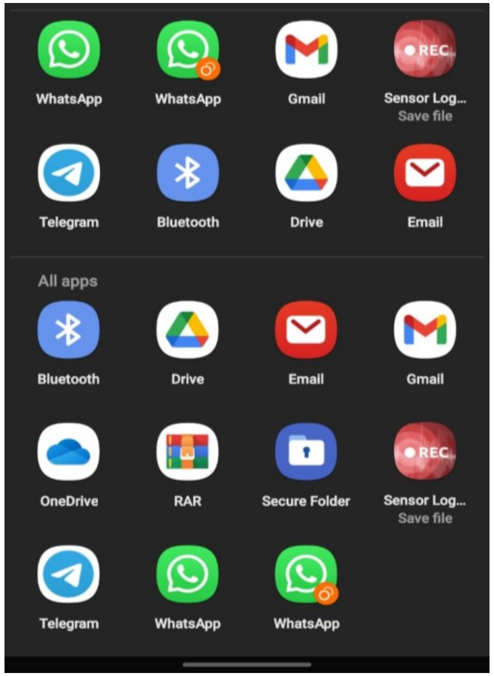


Figure 17 -Export Platforms

3. 2 Data Analyzing

Step 01

After exporting the CSV files, compile all participant data into a single CSV file, as depicted in Figure 18. Each compiled CSV file comprises the following columns.

Time	Seconds_elapsed	BearingAccuracy	SpeedAccuracy	VerticalAccuracy	HorizontalAccuracy	Speed	Bearing	Altitude	Longitude	Latitude
1.70E+18	6.176000244	0	0	0	14.59799957	0	0	-85.9	79.90104	6.800591
1.70E+18	11.102	0	0	0	13.9630003	0.147502	74.54628	-85.6	79.90105	6.800593
1.70E+18	16.171	0	0	0	14.61999989	0.689701	204.9736	-85.9	79.90103	6.800563
1.70E+18	21.177	0	0	0	26.39999962	2.040156	180.4714	-90.1	79.90103	6.800467
1.70E+18	26.213	0	0	0	14.39999962	0.825003	57.36399	-83.5	79.90109	6.800576

Figure 18 -Outputs of Data

Based on the purpose of the study, you may add the necessary columns once you receive the outputs from each participant. Since this study includes walking speed, direction, and time, add separate columns for the participant's name, gender, time, and direction manually as shown in Figure 18.

This study considered the direction and the time.

Time	Seconds_elapsed	BearingAccuracy	SpeedAccuracy	VerticalAccuracy	HorizontalAccuracy	Speed	Bearing	Altitude	Longitude	Latitude	Name	Gender	Time	Direction
1.70E+18	6.176000244	0	0	0	14.59799957	0	0	-85.9	79.90104	6.800591	Chamath	Male	Evening	From B hostel to TCP
1.70E+18	11.102	0	0	0	13.9630003	0.147502	74.54628	-85.6	79.90105	6.800593	Chamath	Male	Evening	From B hostel to TCP
1.70E+18	16.171	0	0	0	14.61999989	0.689701	204.9736	-85.9	79.90103	6.800563	Chamath	Male	Evening	From B hostel to TCP
1.70E+18	21.177	0	0	0	26.39999962	2.040156	180.4714	-90.1	79.90103	6.800467	Chamath	Male	Evening	From B hostel to TCP
1.70E+18	26.213	0	0	0	14.39999962	0.825003	57.36399	-83.5	79.90109	6.800576	Chamath	Male	Evening	From B hostel to TCP

Figure 19 -Manually added columns.

Step 02

Remove the outliers.

What is outlier?

Outliers are samples that are inconsistent with the overall trend of the GPS signal. They could be peaks, discontinuities, saturation, and so on. To accurately assess a signal, it must remove without affecting the rest of the data.

In the context of analyzing walking behavior patterns, outliers were defined as speed values in the CSV that were either lower or equal to 0 ms-^{1}. These examples typically depict negligible movement, such as waiting still. And higher, or equivalent to 2.5 ms-^{1}. These speeds can be incorrect owing to GPS signal inaccuracies.

Remove,

Speed >= 2.5 ms^{-1} & Speed <= 0ms^{-1}

In Excel to remove the outliers the below process has been followed

To remove Speed lower than or equal to 0ms-^{1} and greater than or equal to 2.5 ms-^{1},

Click on the Filter tool under the data tab. → Click on Speed Column Header in Excel → Click on the drop down in the speed → Number Filters > Greater Than To > Enter 0 And Less Than or Equal To > Enter 2.5

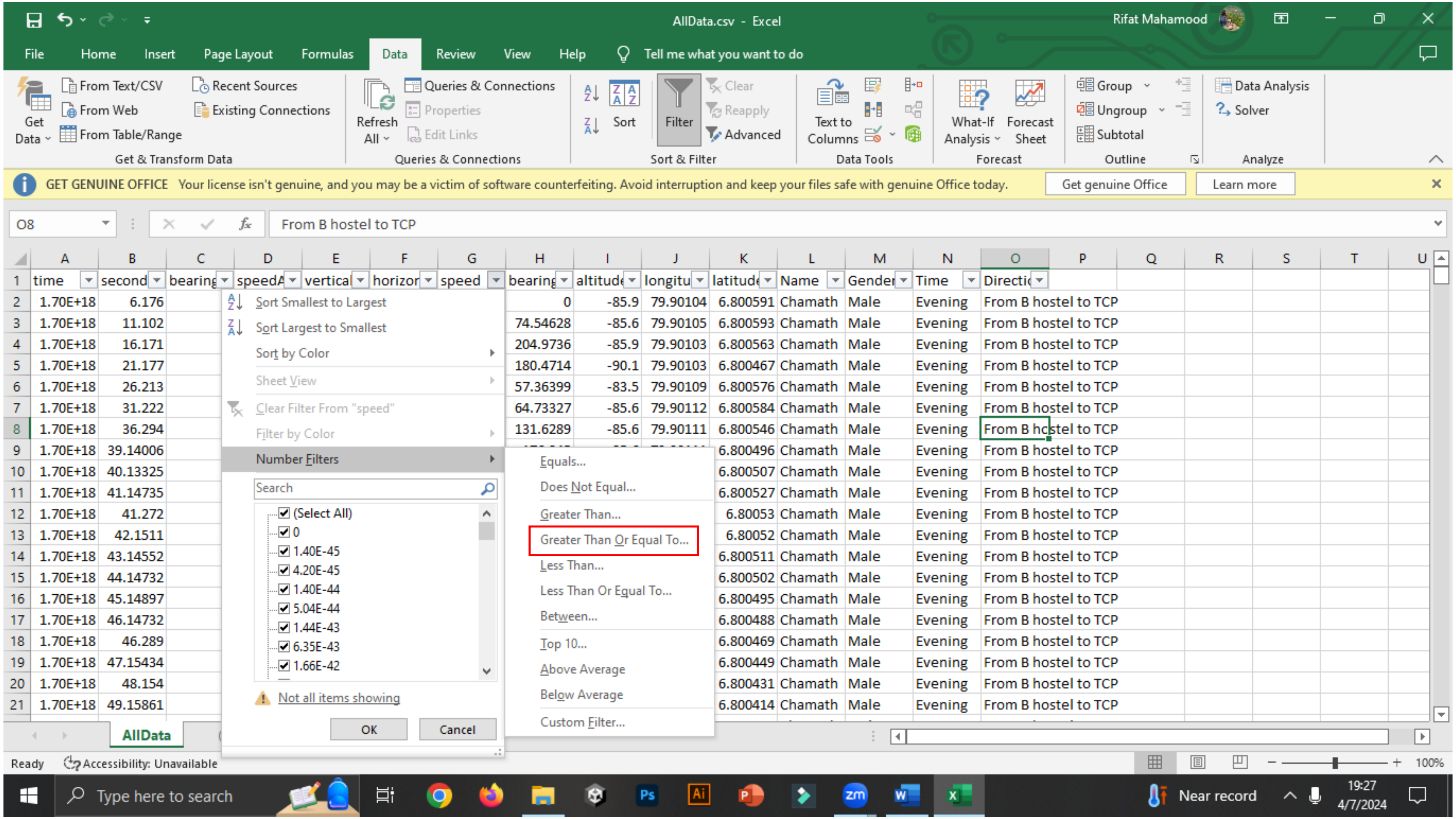


Figure 20 -Process of outlier Remove step 01

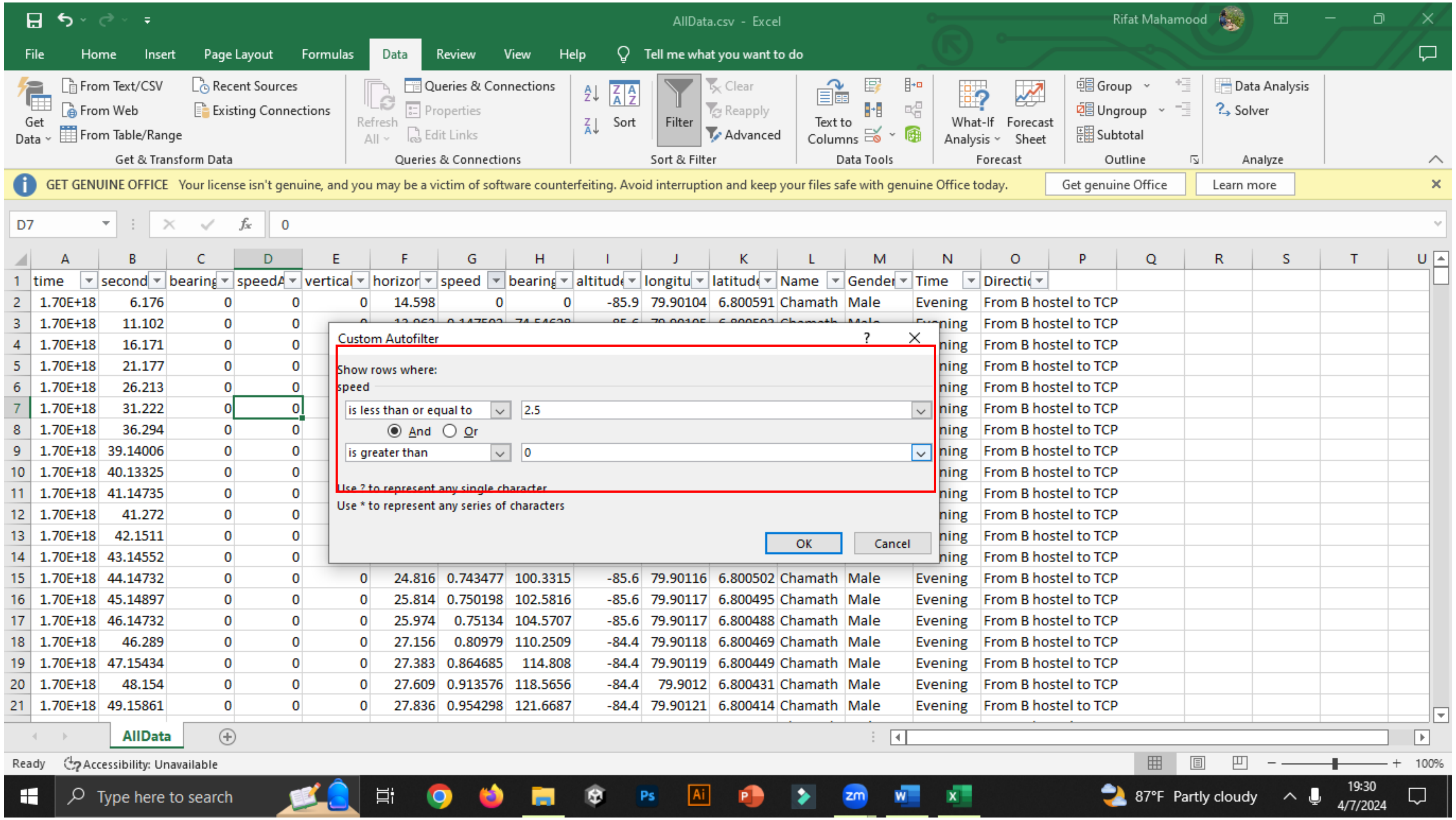


Figure 21 -Process of outlier Remove step 02.

Step 03 –

Normalize with Naismith Rule

The entire dataset was initially normalized using the Naismith method to ensure consistency in the measurements of walking speed across various terrains. The Naismith rule is a common technique for calculating the actual working speed when facing slopes or uneven surfaces. Naismith's Rule, named after the Scottish mountaineer William W. Naismith, is a widely used formula in hiking and mountaineering.

Naismith Rule,

Equivalent distance = Horizontal distance + (Vertical distance * α),

Horizontal distance refers to the distance at which an object moves on a level surface, whereas 'Vertical Distance' indicates the rate of ascent or descent. The character 'α' represents Naismith's number, a constant coefficient that quantifies the additional exertion needed because of variations in altitude, which is 7.92.

Why Naismith Rule,

The rationale behind using Naismith's Rule in this study is to provide participants with a general guideline for estimating the duration of their walking routes. By factoring in terrain variations such as elevation changes, roughness of paths, and obstacles encountered along the way, participants can make informed decisions about their walking routes and plan their activities accordingly. Naismith's Rule is particularly useful in urban environments where pedestrians may encounter diverse terrain conditions, such as steep slopes, staircases, or uneven surfaces.

OPTION 01	OPTION 02
When working with **sequential data points**, such as speed measurements taken in succession without regard for segments or overall values, it may not be necessary to apply the Naismith Rule. Within this framework, the emphasis is frequently placed on the comparative alterations in speed as time progresses, with each data point symbolizing a progression in the travel. **Given that the data points are contiguous and accurately represent the immediate fluctuations in speed caused by terrain or other factors, it may not be essential to apply normalization using Naismith's Rule**. The basic sequential nature of the data allows for the capture of speed changes, making it more relevant to analyze current trends without the need to adjust for overall lengths or segments.	When **analyzing segmented data** that considers the **whole distance and length of travel, the application of Naismith's Rule** can be advantageous. This rule modifies the walking speeds according to the increase in elevation, which is especially important when calculating the total hiking or walking duration for specific parts of a route. Naismith's Rule is a useful tool for assessing the impact of terrain elevation on travel time by considering the overall values. When examining divided paths or tracks, modifying walking speeds using Naismith's Rule offers a more precise depiction of the influence of different terrains on the total time of the journey.

Table 3 - Options in Naismith Rule

Step 04

Categorize the study area into segments.

Categorizing the study area into segments and time is crucial, especially when considering walking directions and the analysis of time-based data. The segmentation of the dataset into different formats can aid in performing a comprehensive analysis. However, the choice of how to categorize the data into segments depends on the specific purposes and goals of the study.

QGIS,

QGIS is a geographic information system software that is free and open source. For the mapping Purposes this study Used Q GIS. It supports viewing, editing, printing, and analysis of geospatial data in a range of data formats.

https://www.youtube.com/watch?v=CLuSZB95ly0

How to install QGIS on a Windows 10 compute

https://www.youtube.com/watch?v=kCnNWyl9qSE

This video is a basic look at QGIS for Absolute Beginners.

Step 4 .1

Open QGIS on your desktop. After opening, you will be received with the interface as shown in Figure 22 - Then go to the Data Source Manager in QGIS which is highlighted in the interface, this action will open the Data Source Manager dialog, allowing you to import the CSV file into QGIS. It is important to note that the CSV file you are importing should be normalized using the Naismith Rule and should have the outliers removed beforehand, ensuring the accuracy and reliability of the data for further analysis.

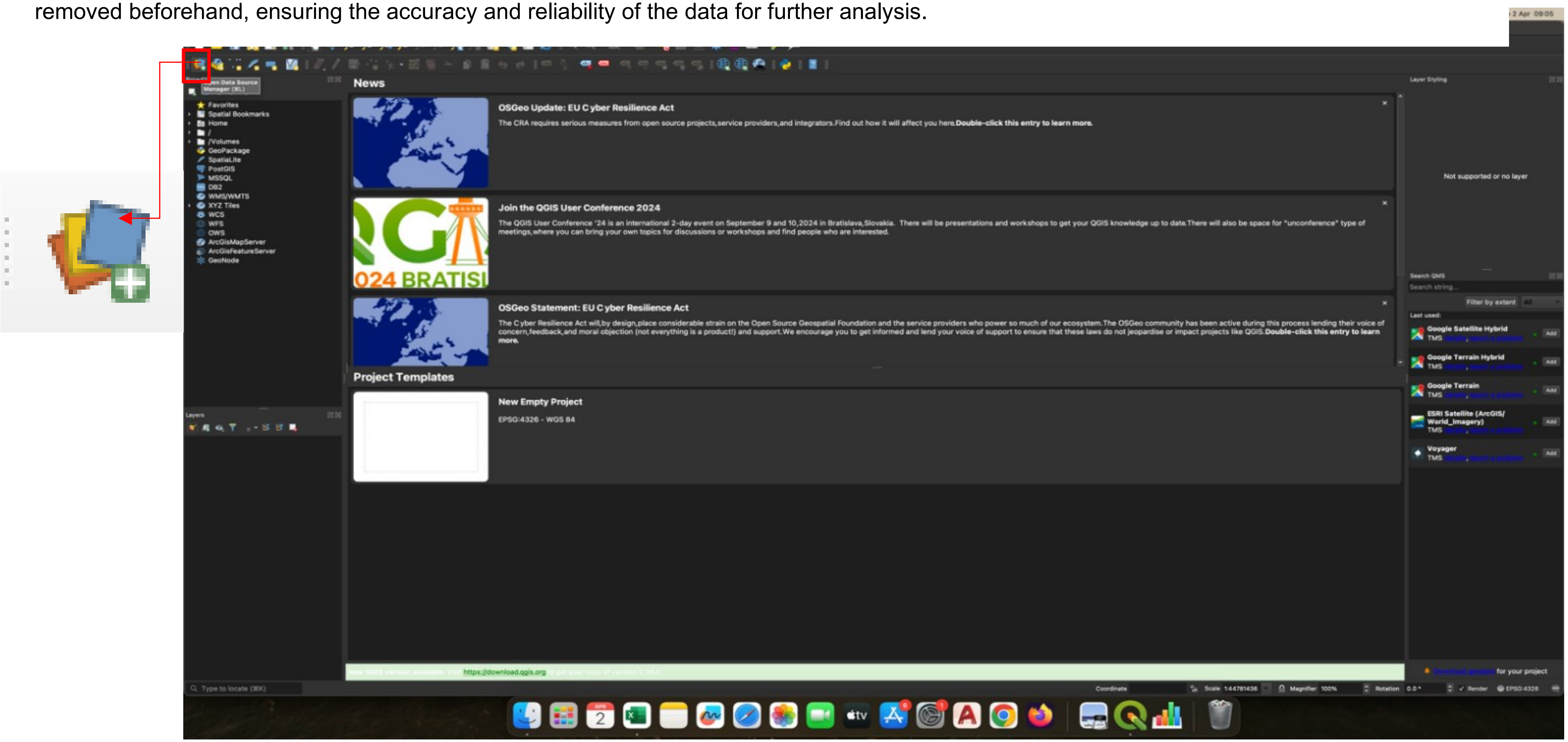


Figure 22 -Q GIS interface and the Data Management Tool

Step 4 .2

After clicking on the "Data Source Manager" icon, you will see a dialog box appear, like the one shown in Figure 23. Here is what you need to do:

1. Click on "Delimited Text" in the dialog box.
2. Next, you will need to browse for the CSV file that you have already cleaned up and prepared by removing outliers.
3. If the X and Y fields (which represent longitude and latitude) are not automatically detected, you will need to manually select them. Look for options to choose the longitude and latitude values from your CSV file and assign them to the X and Y fields, respectively. And you need to enable the option called "points coordinate."

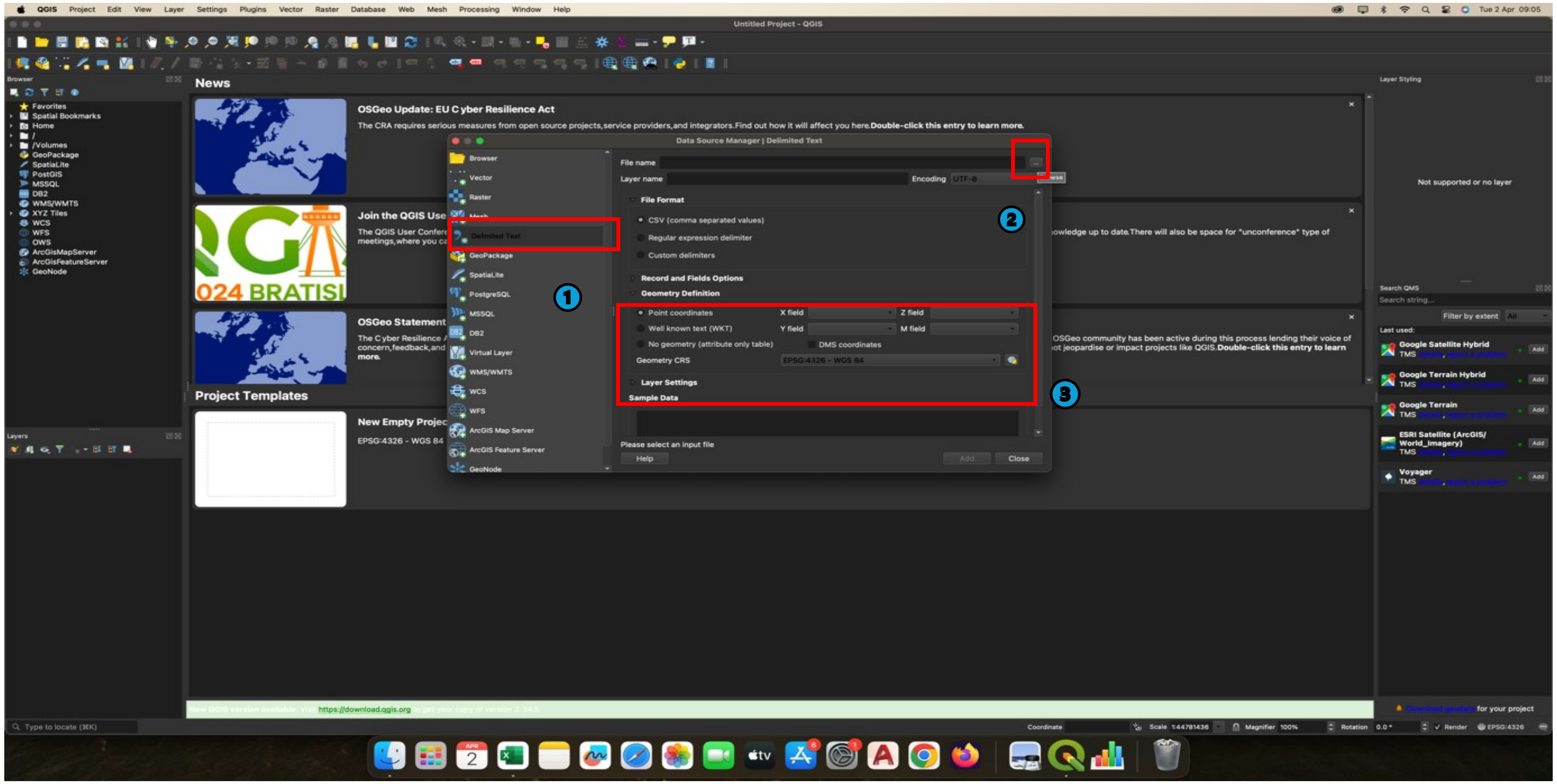


Figure 23 -Process of import CSV

After following the steps in figure 22 &23, you will see the output as depicted in Figure 23. This image will show your imported CSV data displayed on the map in QGIS.

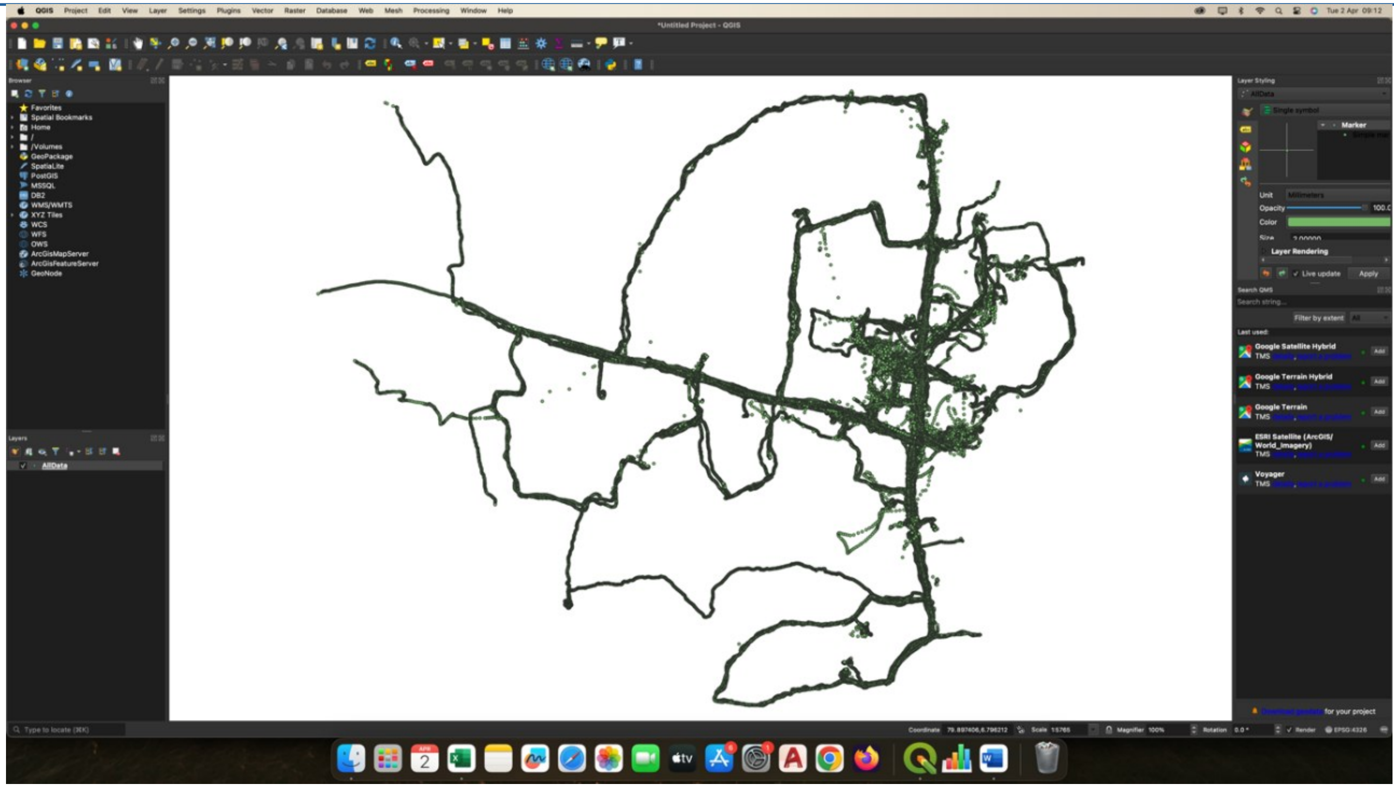


Figure 24 -Location Points Map

Input

ALL Data CSV File

Converting the raw GPS data stored in the CSV file into point features, each representing a specific geographic location, to visualize and analyze spatial patterns within QGIS.

Output

All users GPS Points

Step 4.3

Adding the Base Map:

1. Click on the "Plugin" tab in the top menu bar of QGIS.
2. From the dropdown menu, select "Manage and Install Plugins."

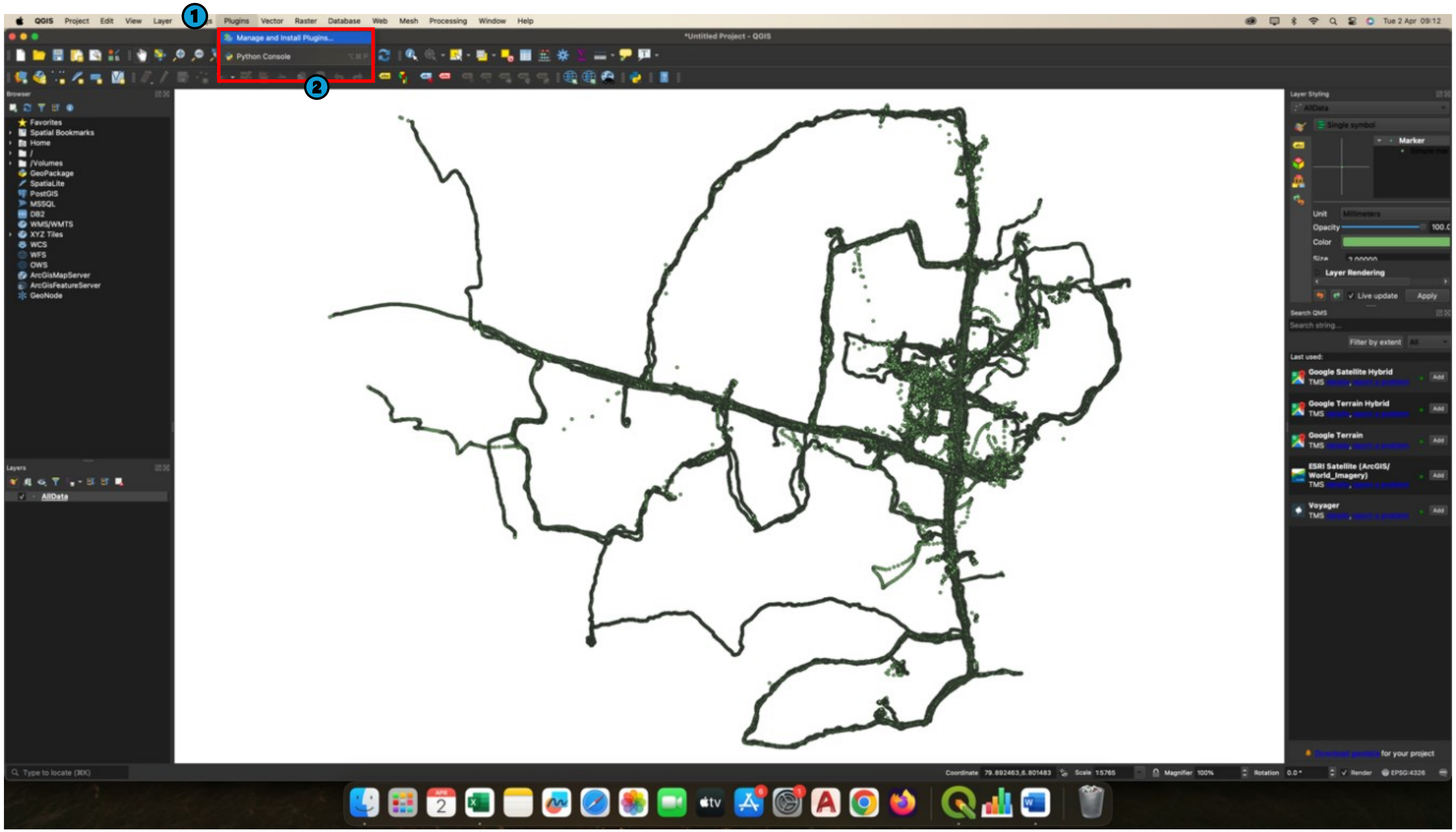


Figure 25 – Adding Base Map

Step 4.4

Clicking on the "Manage and Install Plugins" option as shown in Figure 26, a dialog box will appear.

1. In the search bar of the dialog box, type "QuickMapServices" and press Enter. The "QuickMapServices" plugin should appear in the search results.
2. Click on the "Install" or "Update" button next to the "QuickMapServices" plugin to install or update i

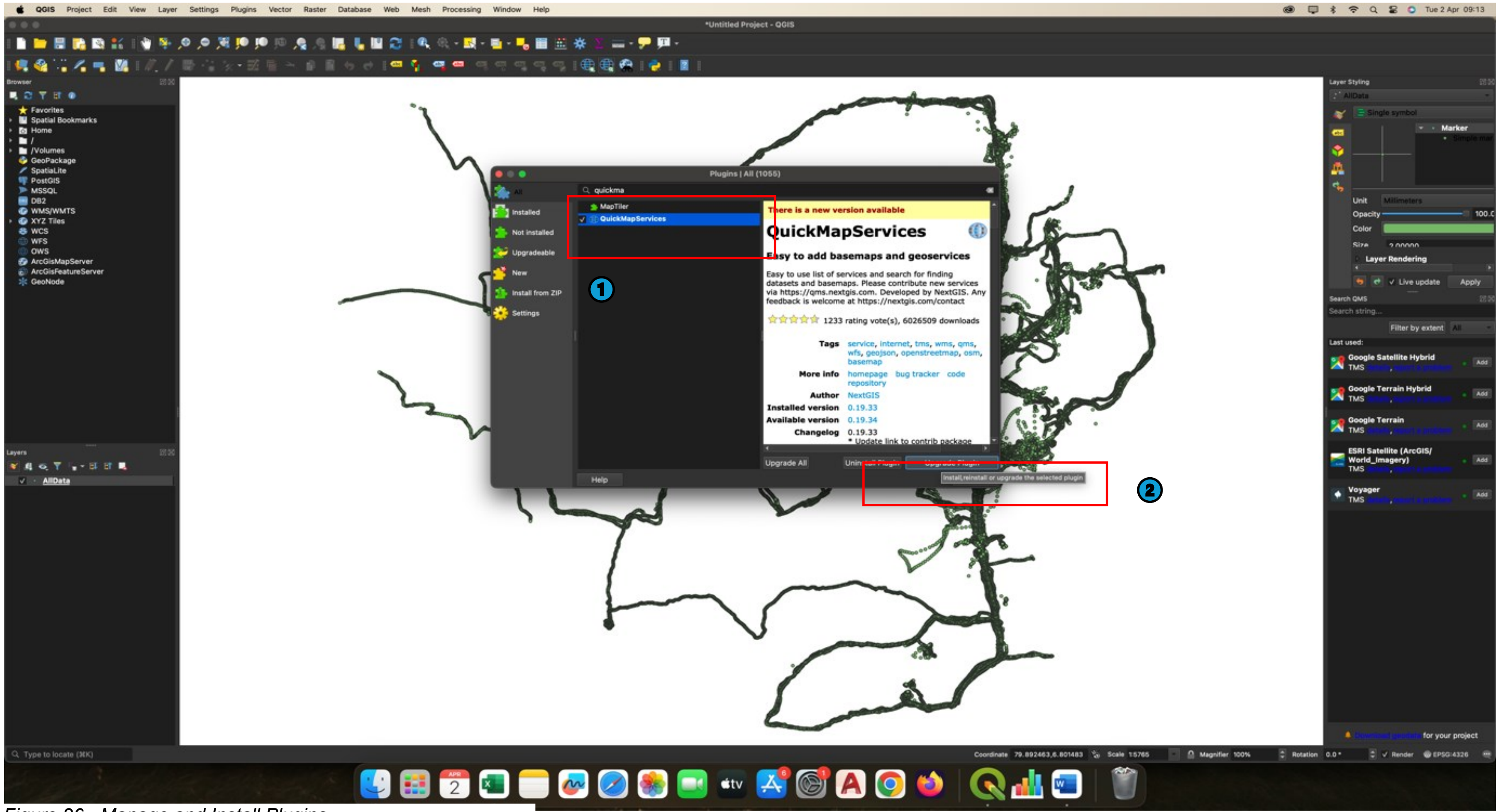


Figure 26 - Manage and Install Plugins

Step 4.5

Adding Base Map using QuickMapServices (QMS):

1. After installing the "QuickMapServices" plugin (Step 4.4), you should see a new window appear on the right side of the screen. If you don't see the QMS window, you can open it by going to the "Web" option in QGIS.
2. In the QMS window, search for a suitable base map. For this tutorial, we will select the "Voyager (no labels) (retina)" base map

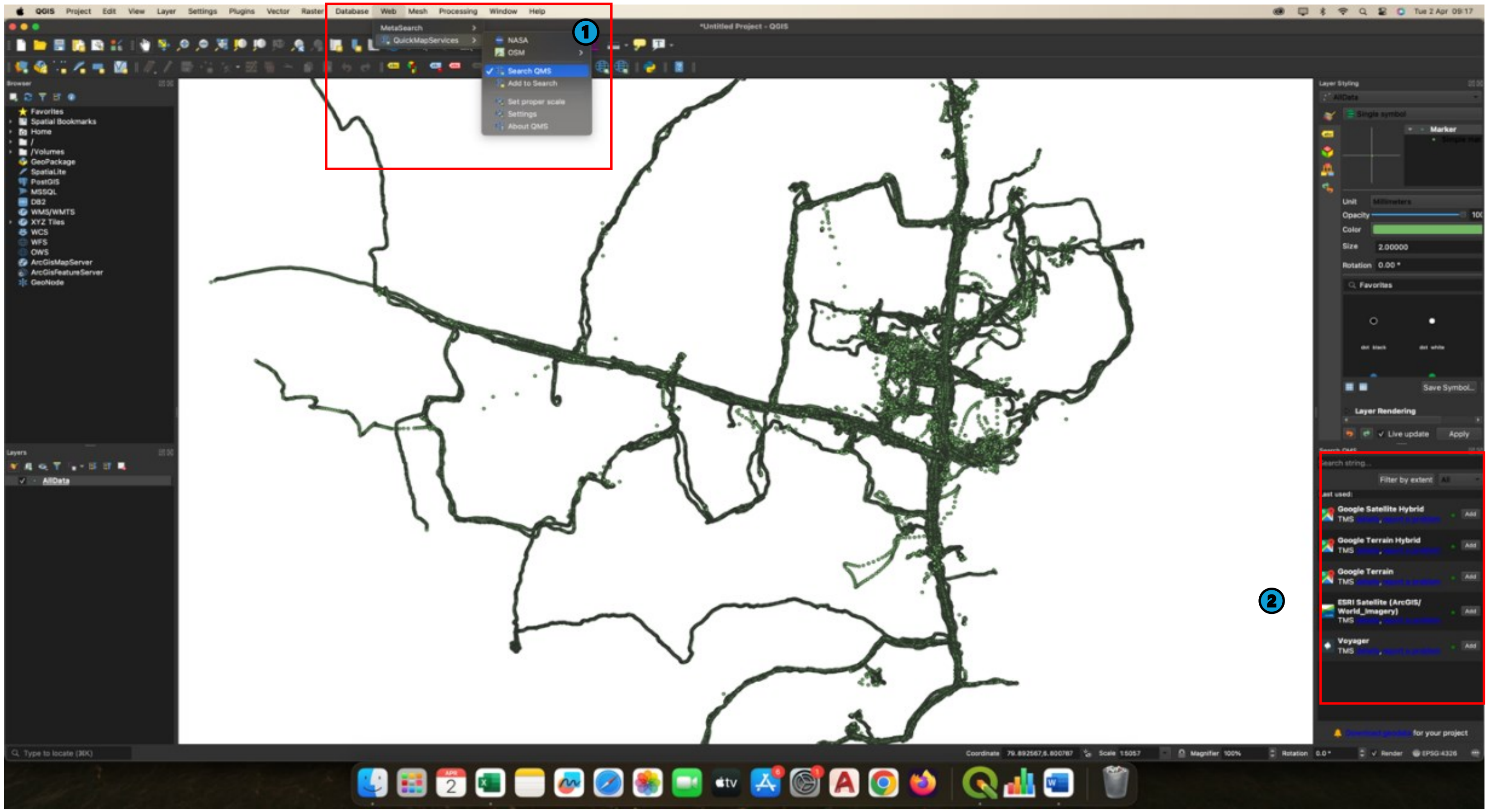


Figure 27 - QuickMapServices

After following the steps in 4.3 – 4.5, you will see the output as depicted in Figure 27. There you can edit the base map by click on the properties.

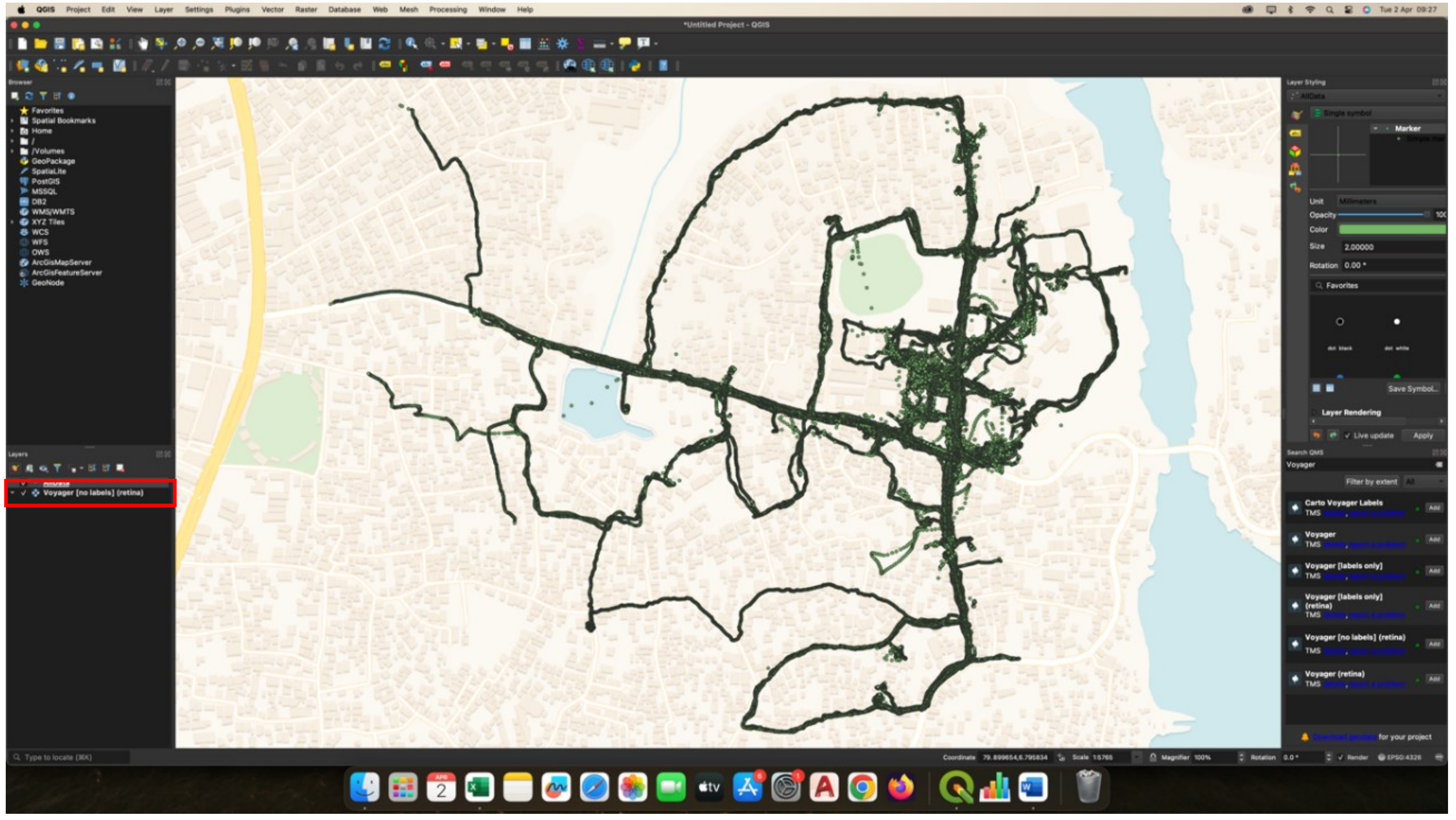


Figure 28 -Edit base map

Step 4.6

Symbology (To make the more visible points, we can go to the property of the All-Data file and change its symbology)

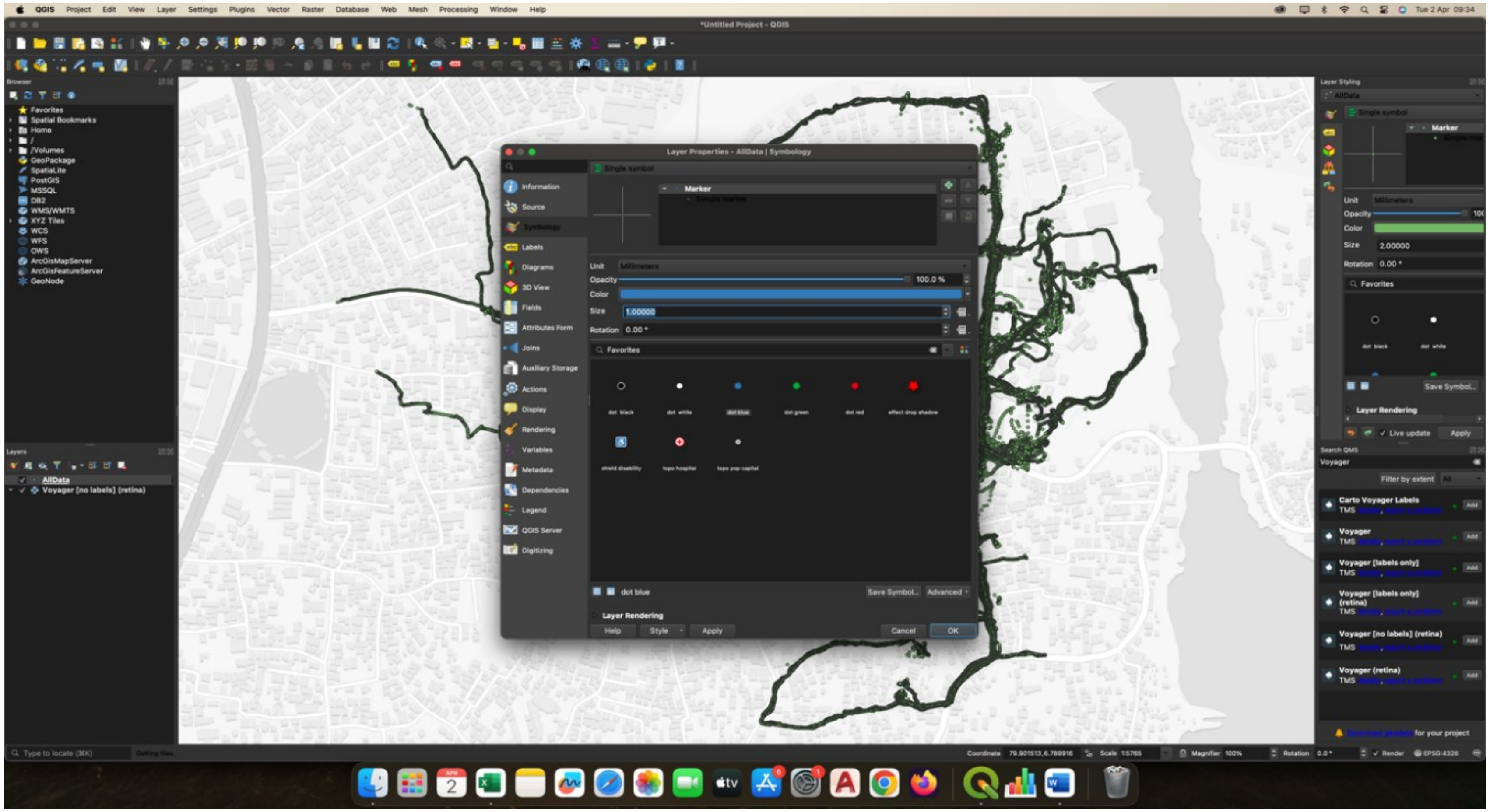


Figure 29 -Symbology

Step 4.7

Choose any icon or symbol you prefer to represent the data points. For this demonstration, let us select blue dots. Adjust the size of the dots to 1pt for a clear and concise display on the map.

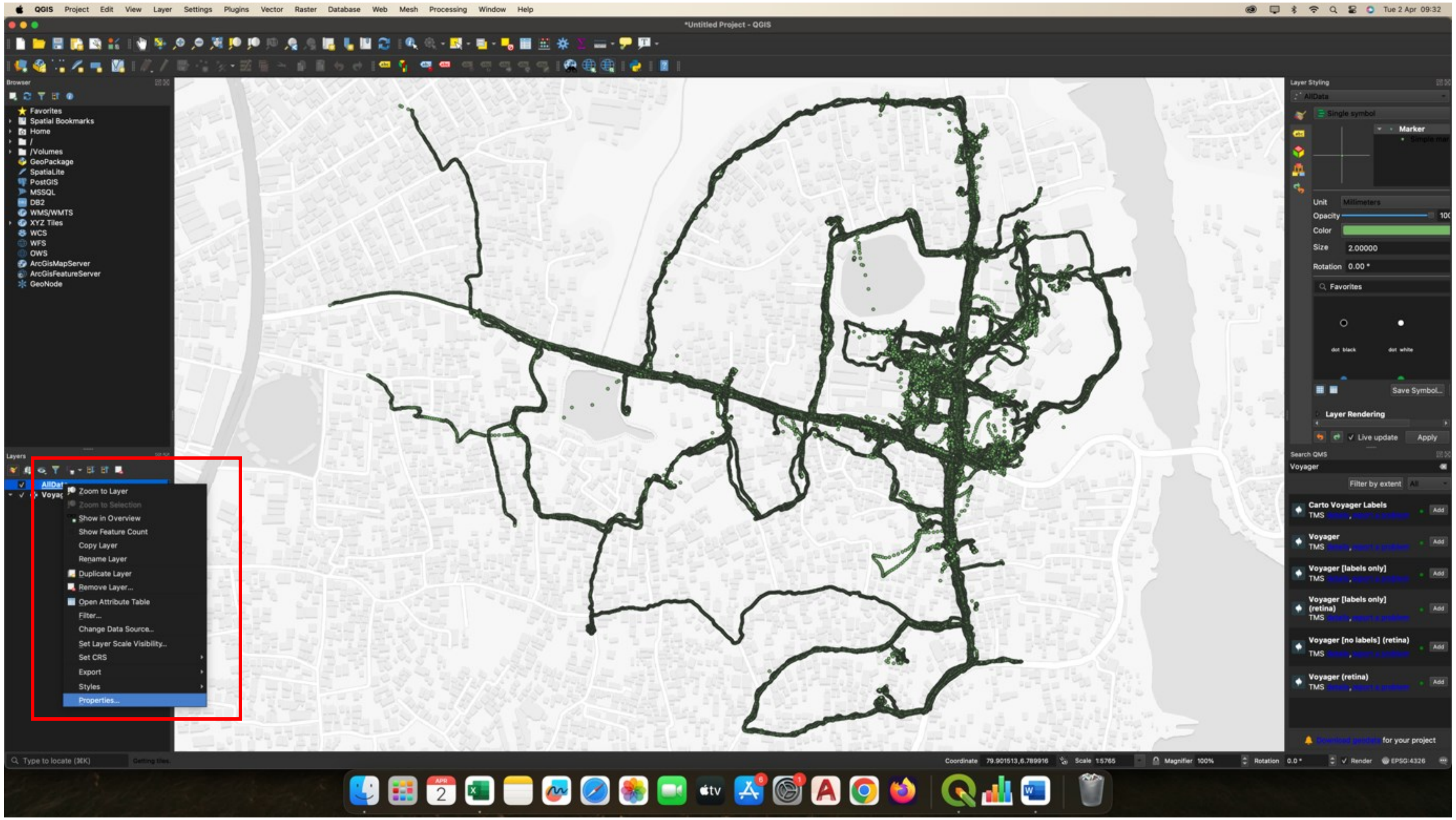


Figure 30 - Symbology

Step 4.8

To categorize the points based on walking speed, go to the "Symbology" tab for the layer. Choose the "Graduated" option from the dropdown menu.

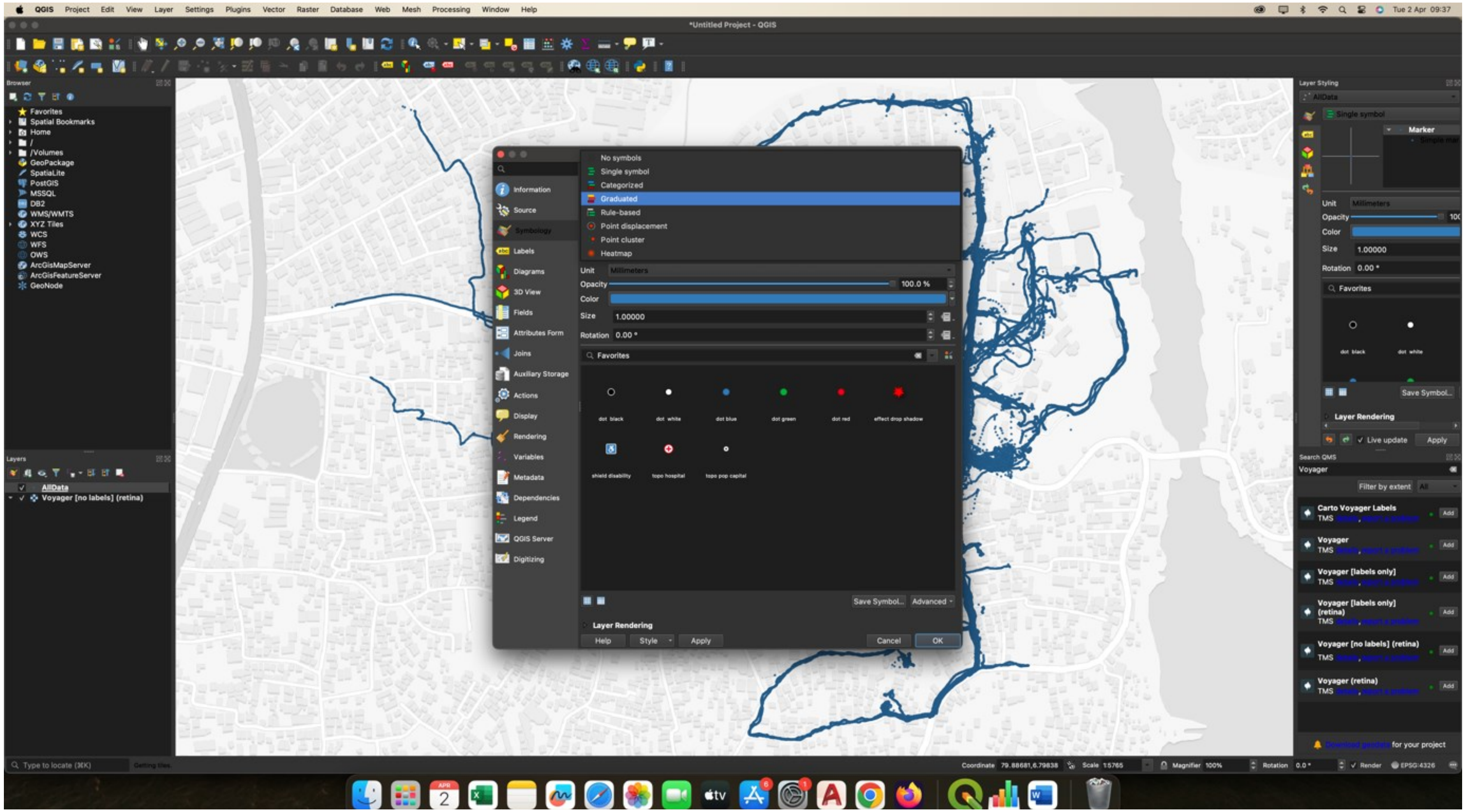
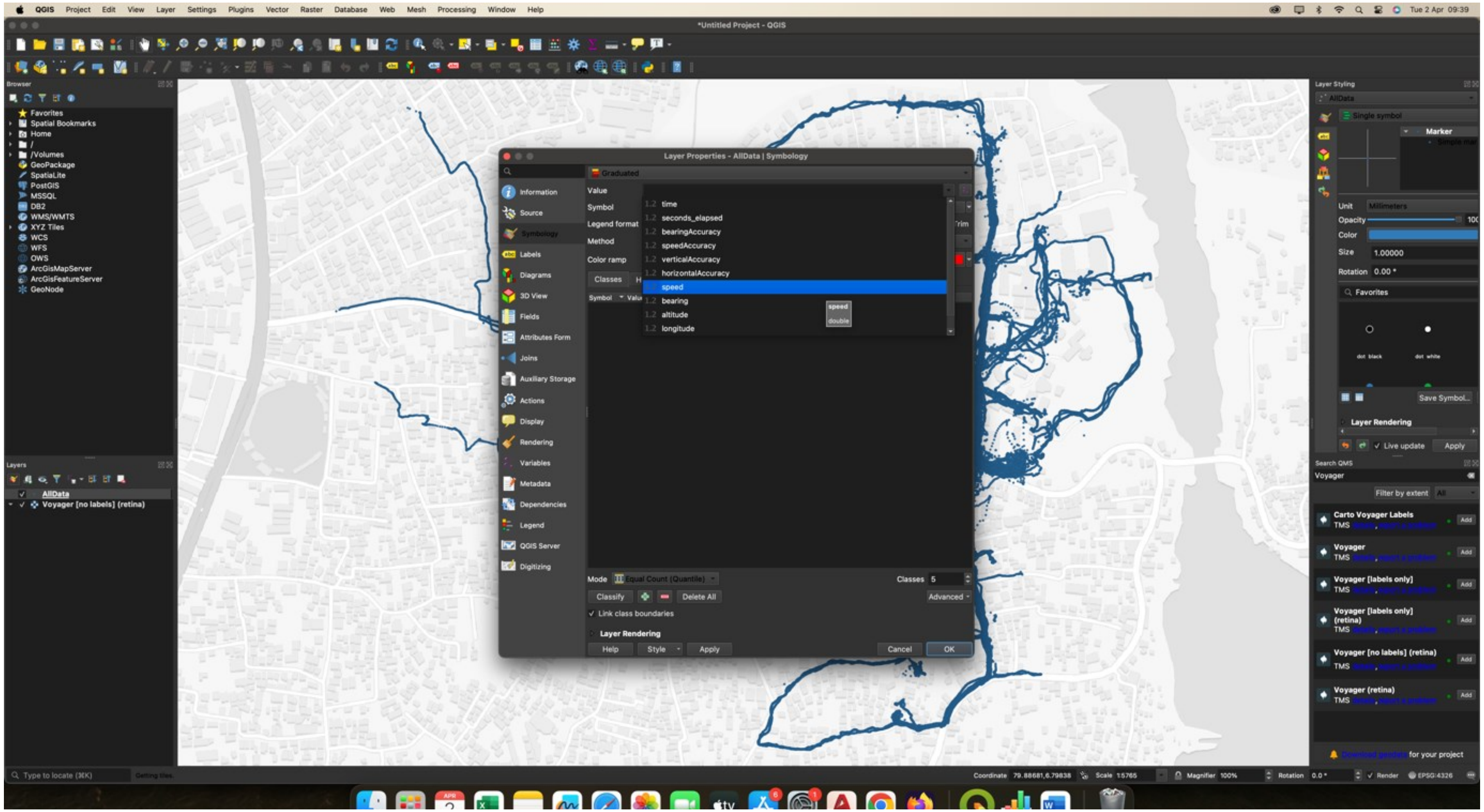


Figure 31- Categorize the points

Step 4.9

Then select Speed for a Value option.



4.10

Figure 32 - Value option

Step 4.10

Select a color ramp to visually represent different speed categories. For example, use warmer colors like red for higher speeds and cooler colors like blue for slower speeds. You can also customize the size of the symbols for each speed category to enhance visibility and clarity. Larger symbols can indicate higher speeds, while smaller symbols represent lower speeds.

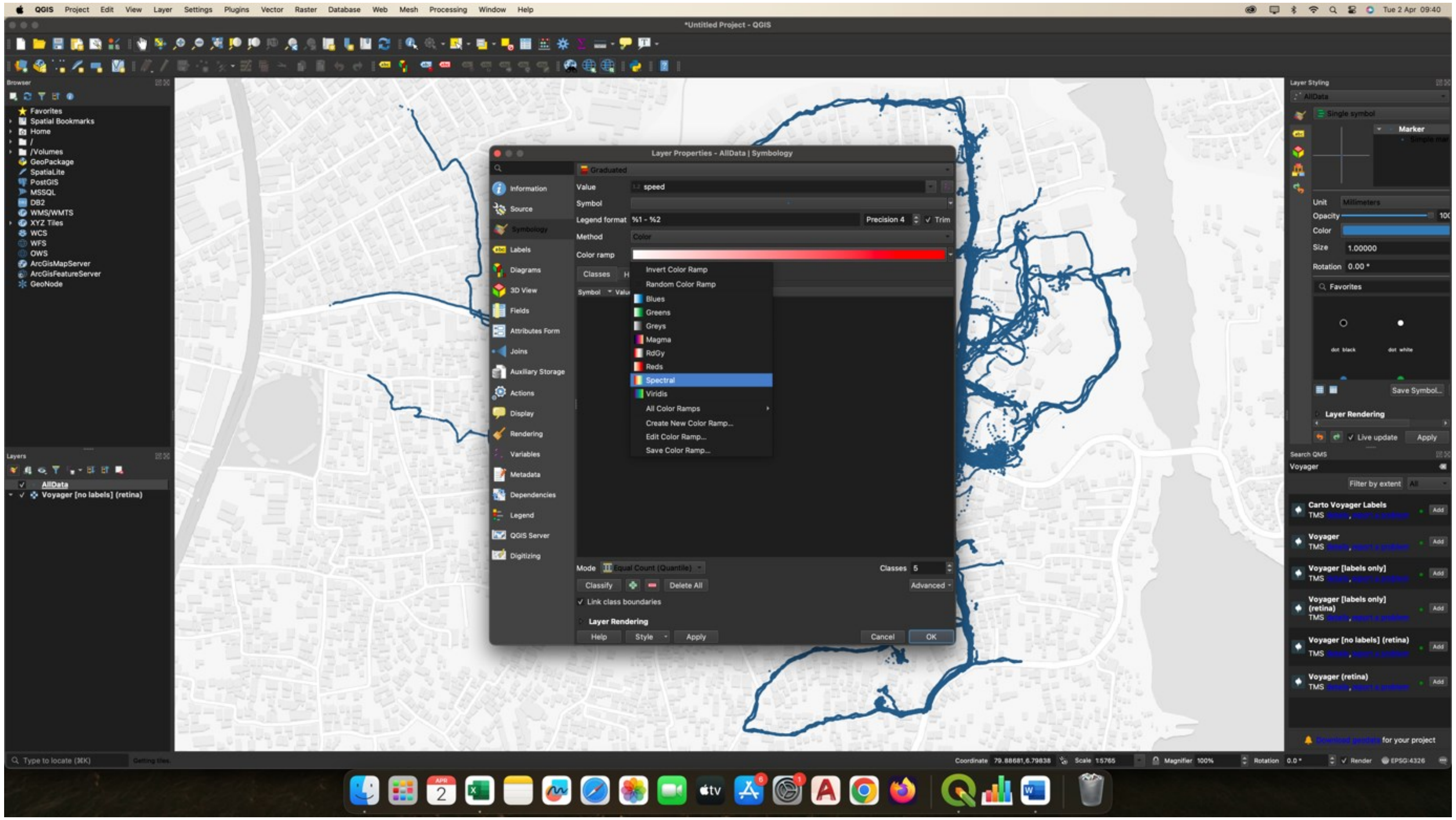


Figure 33 - Color ramp

Step 4.11

In the "Graduated" symbology options, choose the classification mode that best suits your dataset. For example, you can select "Natural Jenks", or "Equal Interval" based on the purpose of your analysis. The classification mode determines how the data will be grouped into categories or classes. If you choose "Natural Jenks," the algorithm will find natural breaks in the data to create meaningful categories.

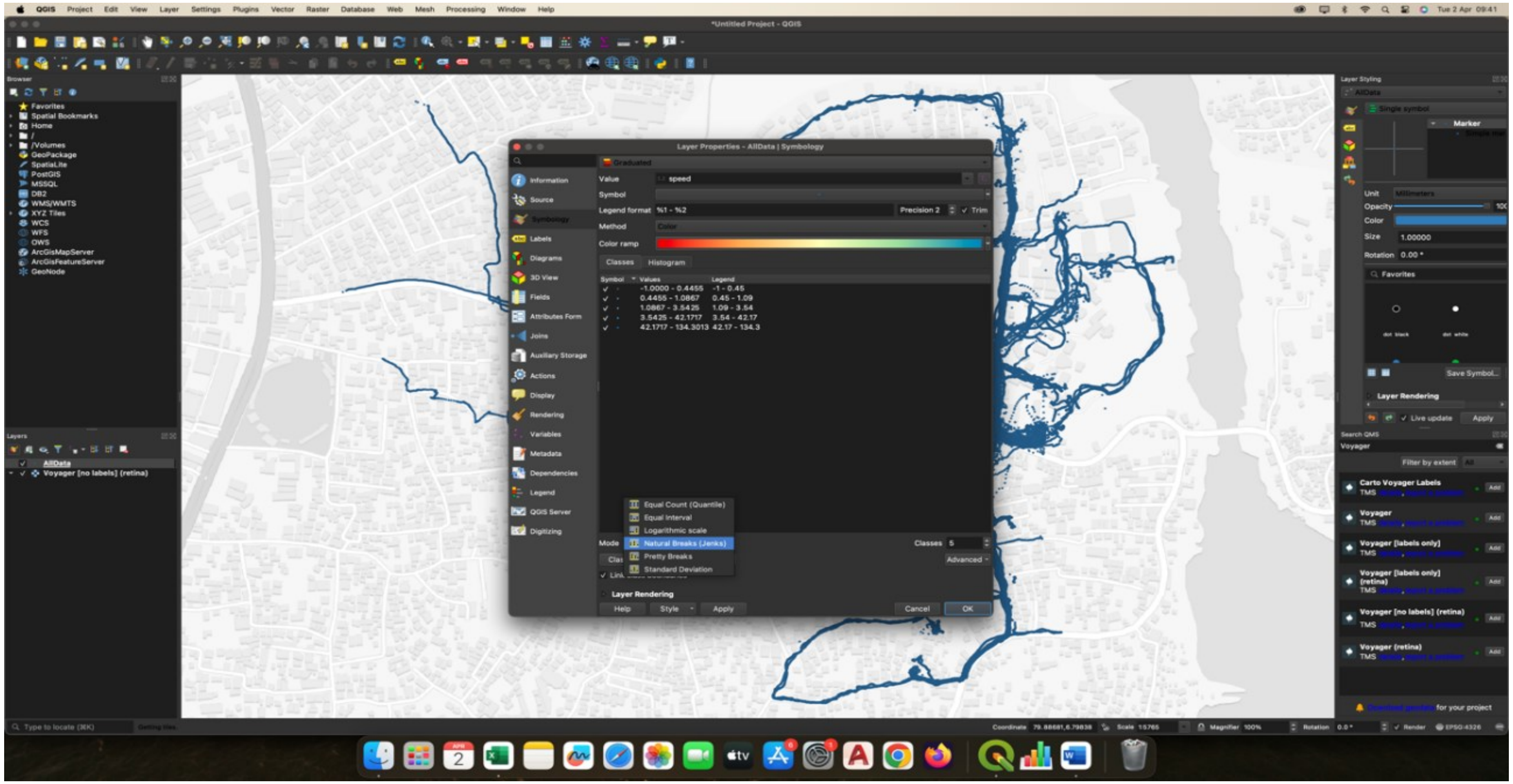


Figure 34 -"Graduated" symbology options

After following the steps outlined above, you will observe the output displayed on the screen, as depicted in Figure 35. This visualization illustrates the variation in pedestrian walking speeds across different areas, represented by the graduated symbols on the map.

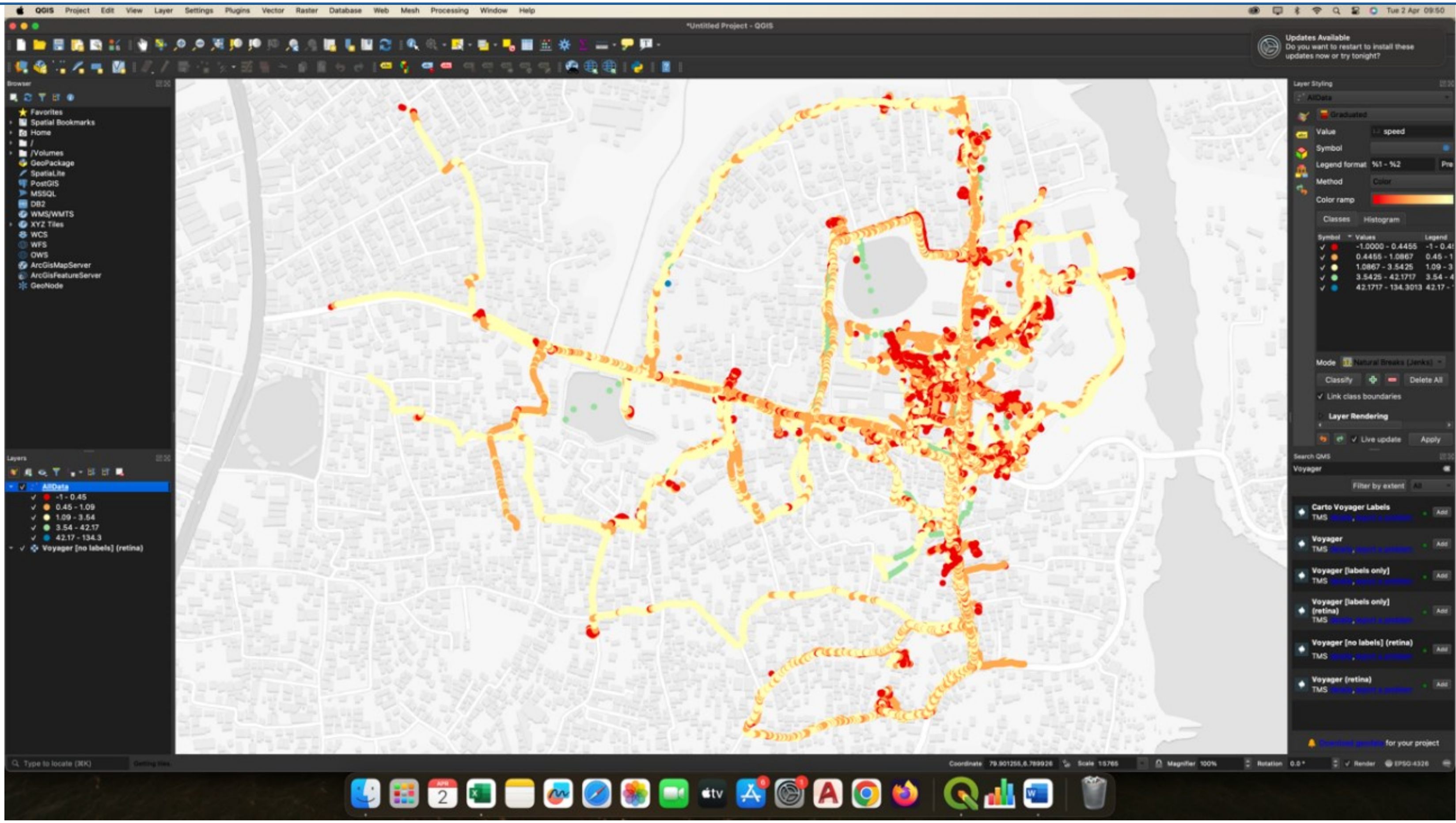


Figure 35 - Pedestrian walking speeds across different areas

Step 4.12

The next step involves clustering the study area into segments, which is an optional step. In this study, the area was manually segmented for ease of analysis in the upcoming stages. To manually segment the area, Here uses the "Select Features by Freehand Tool" in the QGIS software.

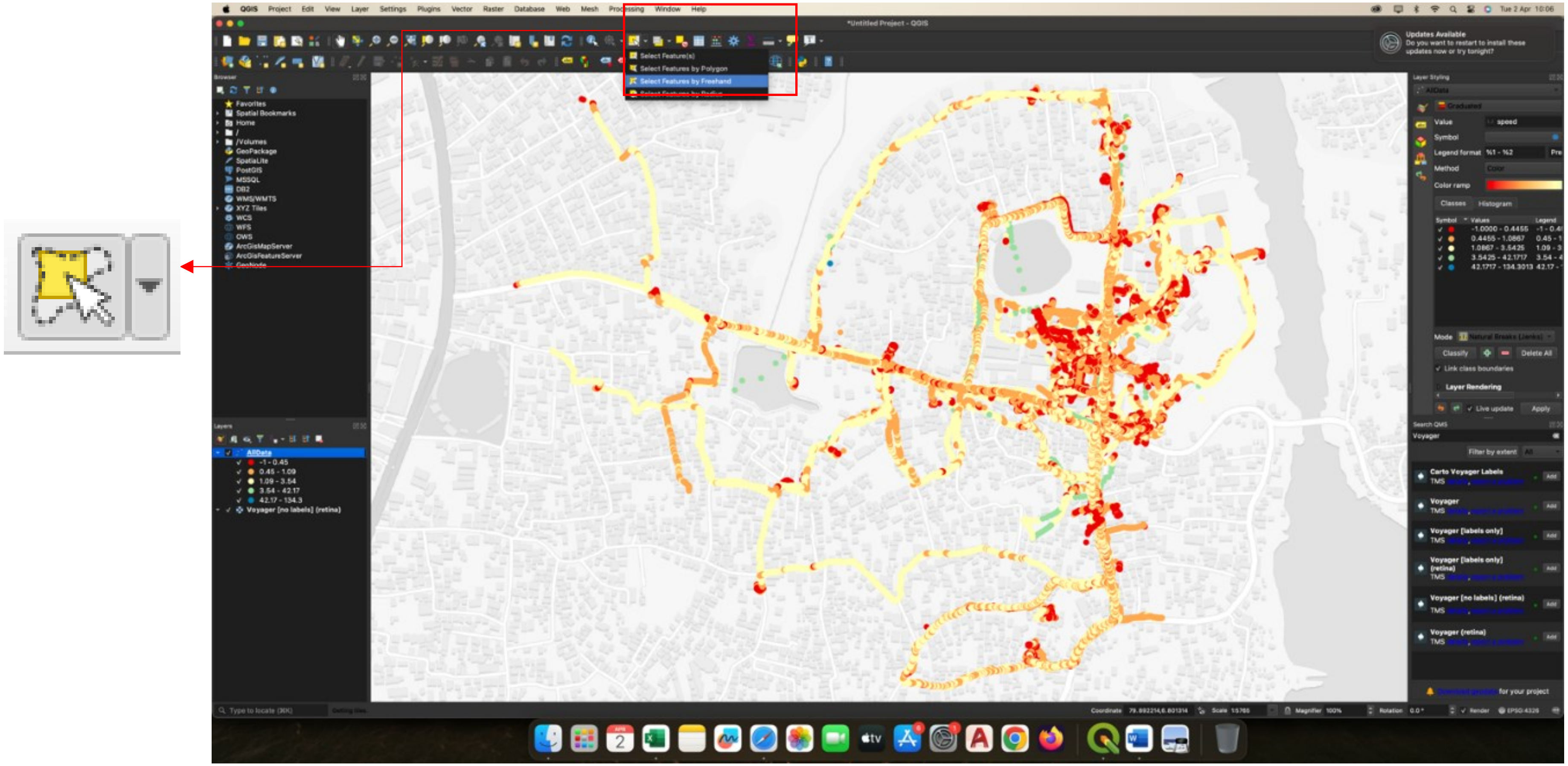


Figure 36 - Select Features by Freehand Tool

Step 4.13

In this step, we will cluster some segments of the study area based on nodes and intersections. We will start by selecting the TCP building as the first segment, which will represent the library in our analysis.

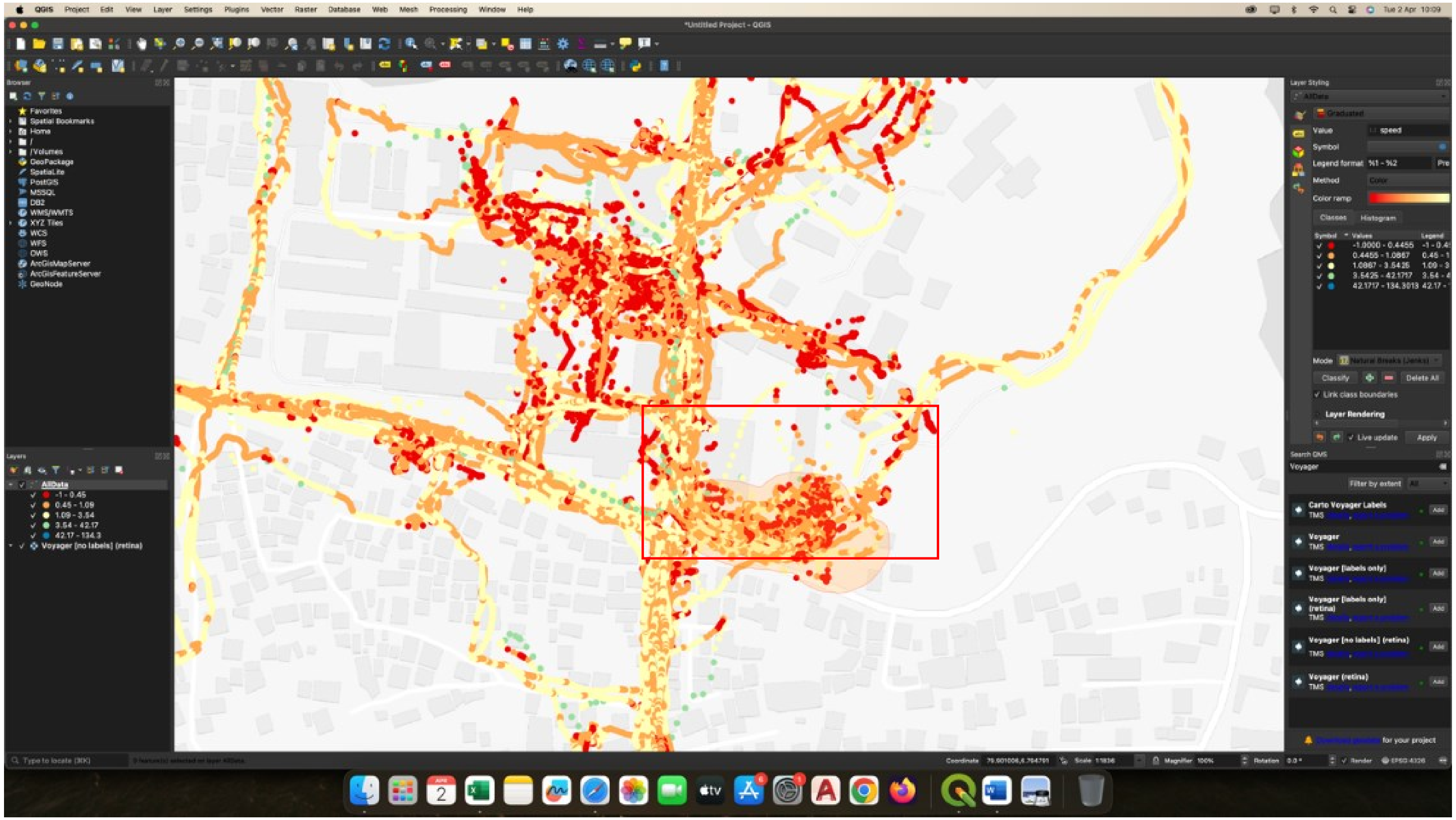


Figure 37- Cluster the segments

Step 4.14

If the segment is carefully selected, it will appear highlighted in yellow color. This visual indication helps in identifying the specific area or segment that has been clustered based on the chosen node or intersection.

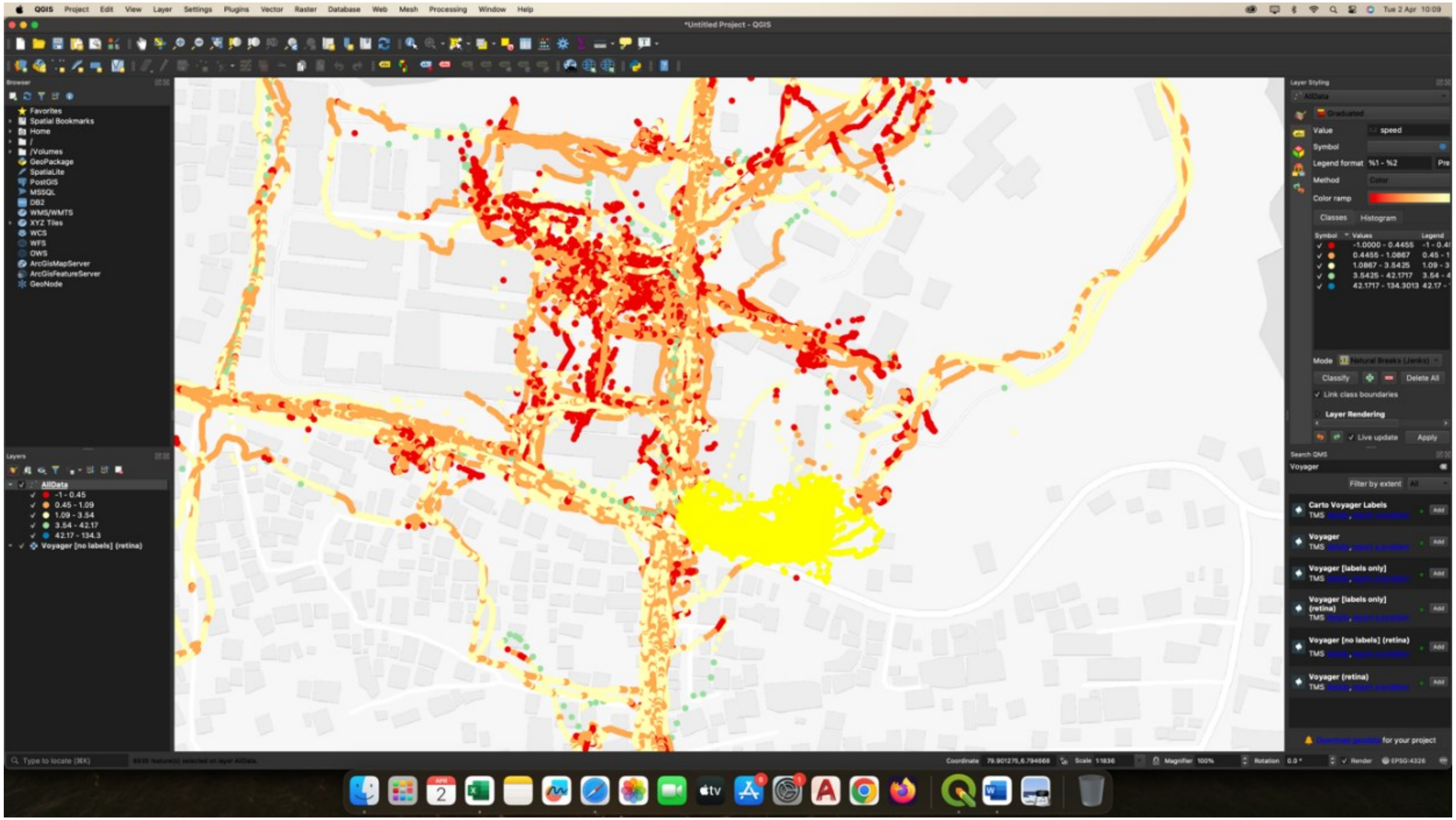


Figure 38 - Identifying the segment

Step 4.15

To export the selected segment as a separate shapefile, follow these steps:

Right-click on the data file

Click on "Export" from the menu options.

Choose the "Selected Features as..." option.

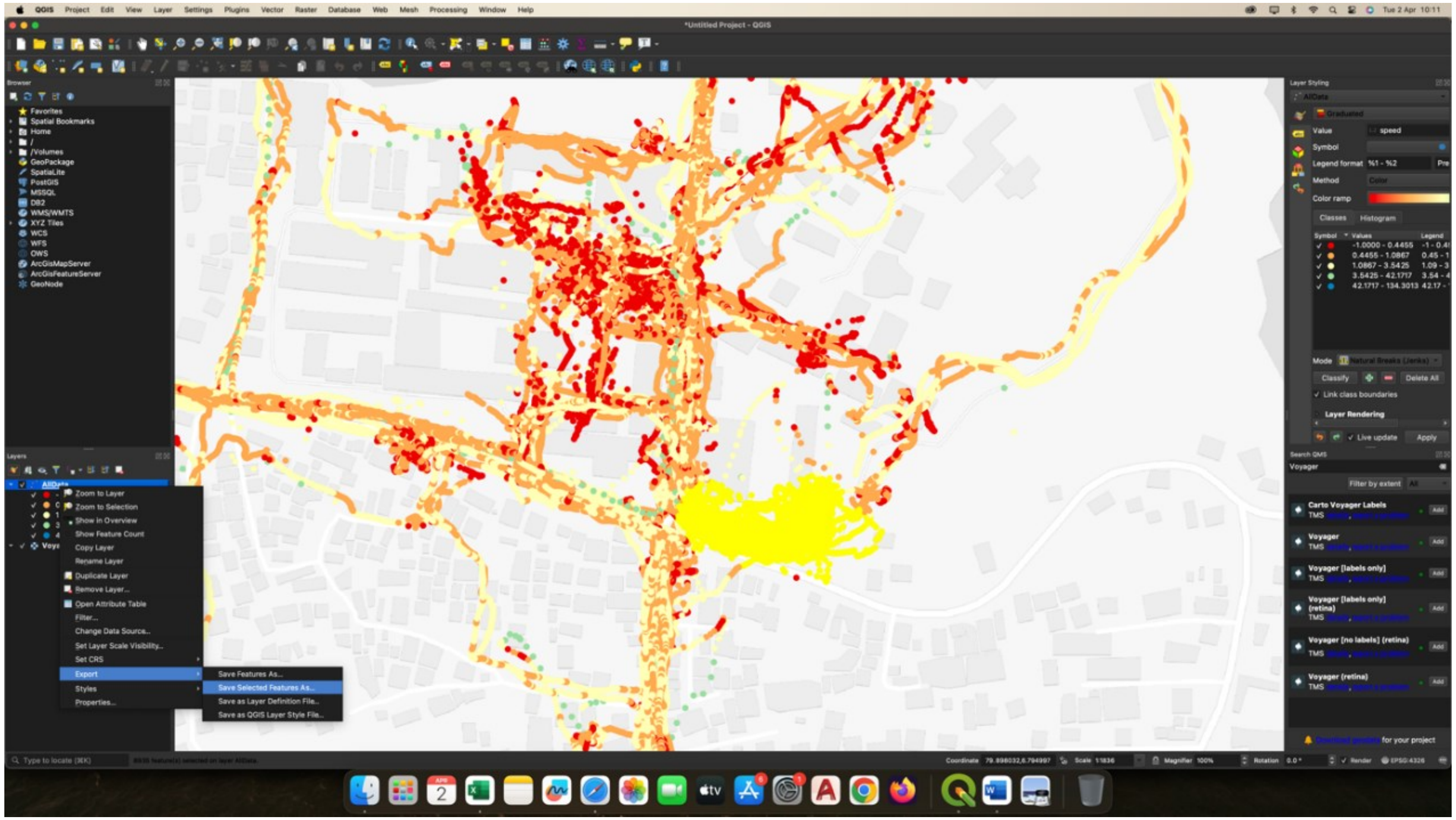


Figure 39 - Export the selected segment

Step 4.16

For the format, choose "ESRI Shapefile" and provide a location and file name for the exported shapefile.

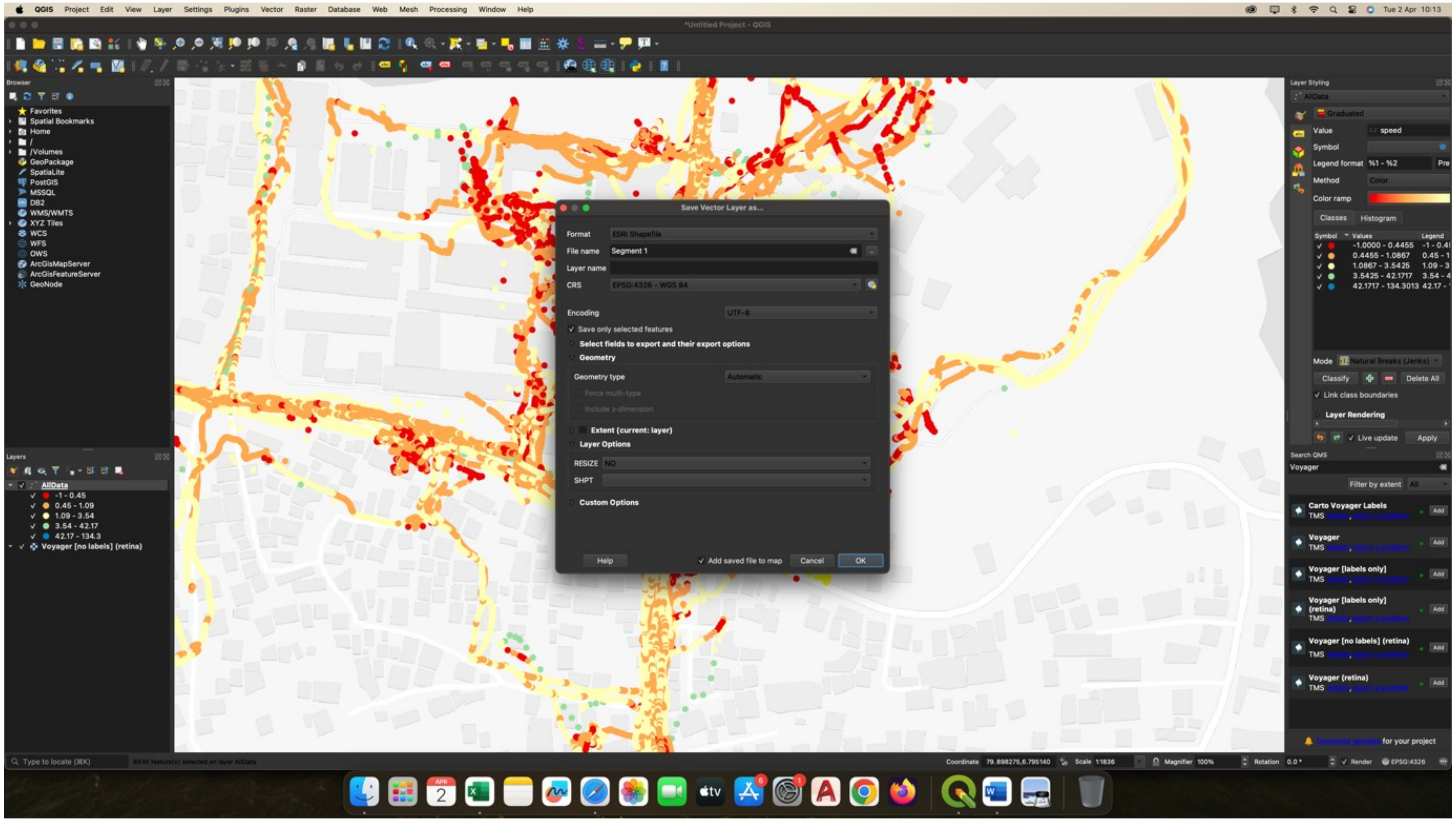


Figure 40 - ESRI Shapefile

Figure 41 show the output of the exported segments. In this study, a total of 40 segments were exported separately.

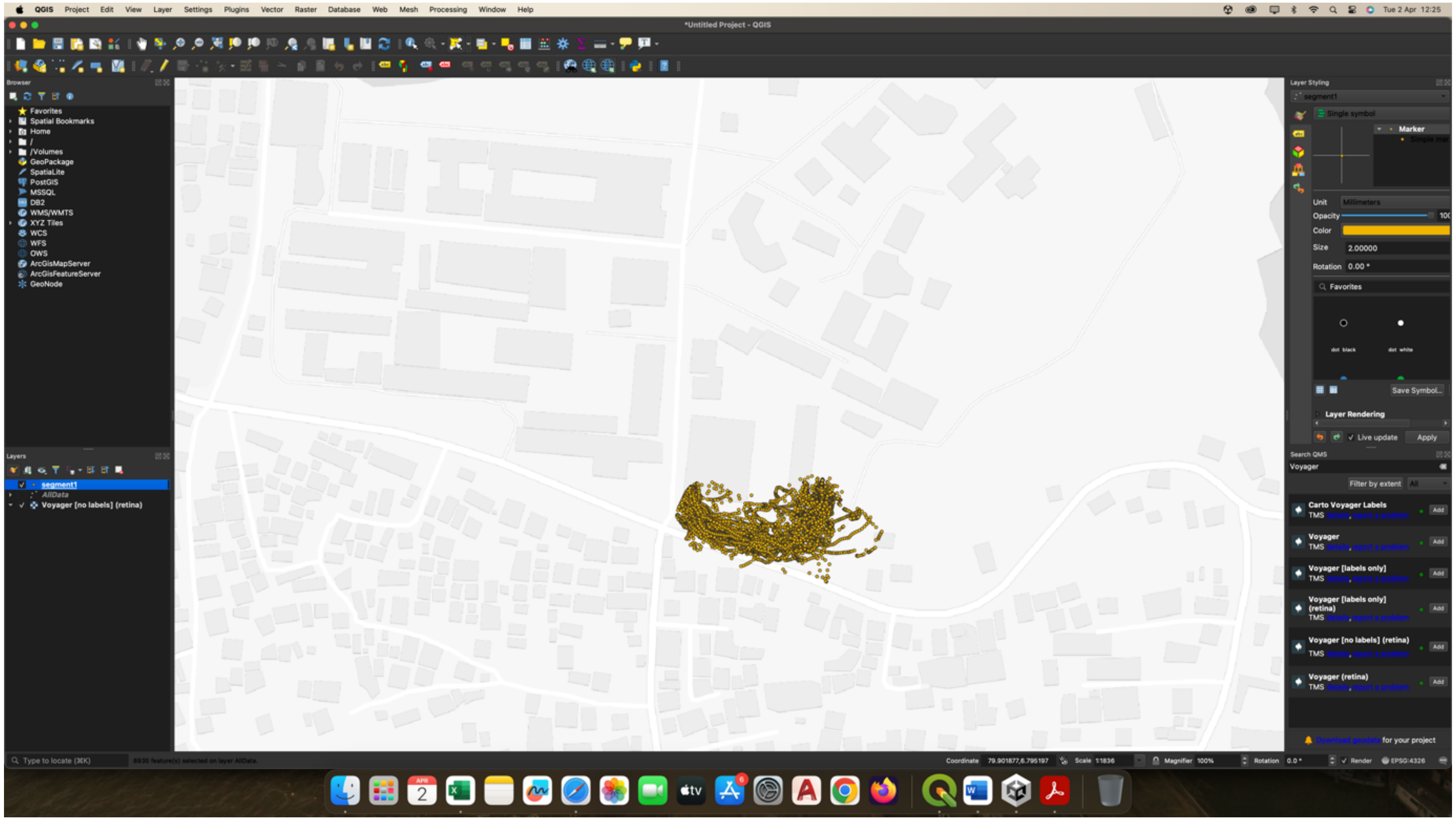


Figure 41- Exported segments.

3. 3 Data Clustering

Clustering people by walking speed allows accurate walking behavior categorization. First, categorize walking behavior to understand urban walking patterns. Mapping finds pedestrian-heavy areas and patterns. Also, clustering detects high-foot traffic zones or congestion hotspots where analyzing spatiotemporal factors in those hotspots can inform better interventions.

In QGIS, you have identified different segments within your data. After identifying these segments, you export each segment into its own CSV file. Each CSV file should represent one segment. This ensures that you have separate data files for each area of interest.

Now, you want to perform clustering analysis on each segment individually. This could be based on your study purposes and in this study, it is the location and speed.

K mean clustering,

This study used k mean clustering method for cluster the walking dynamics. KMeans clustering is a popular unsupervised machine learning technique used for partitioning a dataset into a predetermined number of clusters. The goal of KMeans is to group data points into clusters based on similarities in their features.

Why K mean clustering,

- **K-means allows us to efficiently group participants into clusters based on their location.**
- **Simple and Intuitive**
- **The algorithm is scalable to large datasets, allowing us to analyze a substantial amount of GPS and speed data collected from numerous participants.**

https://www.youtube.com/watch?v=4b5d3muPQmA

https://www.youtube.com/watch?v=YEwt6BJROug

Google Colab,

Google Colab is a cloud-based platform provided by Google that allows for the creation and sharing of Jupyter notebook files. It offers free access to GPU (Graphical Processing Unit) and TPU (Tensor Processing Unit) resources, making it ideal for running machine learning algorithms and data analysis tasks. And you need to have a google account.

Visual Studio Code (VS Code) with Python Extension, PyCharm, RStudio also can use.

You can access to the google colab through the website https://colab.research.google.com/

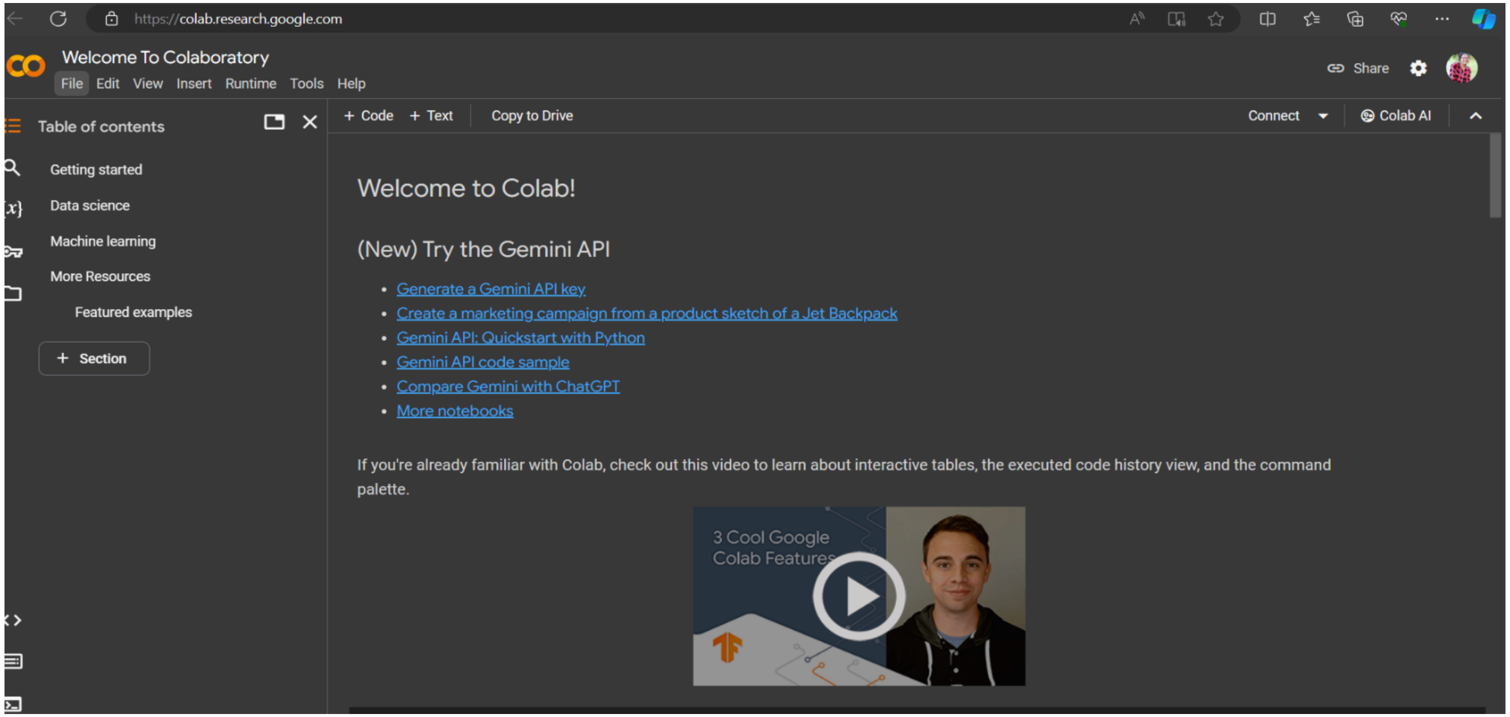


Figure 42 - Interface of the Google Colab

Editable Code (As per the purpose of the study, user need to change the code

Fixed\Common Code (User can use this code as it is)

Step 1

Import Required Libraries.

```
import pandas as pd.

import NumPy as np.

from sklearn. cluster import KMeans.

import os
```

```
import matplotlib.pyplot as plt

import seaborn as sns

%matplotlib inline
```

Step 2

Load the dataset.

Step 2.1. To load the dataset First you need to import the file into Google Collab. For that go to the upload tab and browse your file as shown in Figure 43.

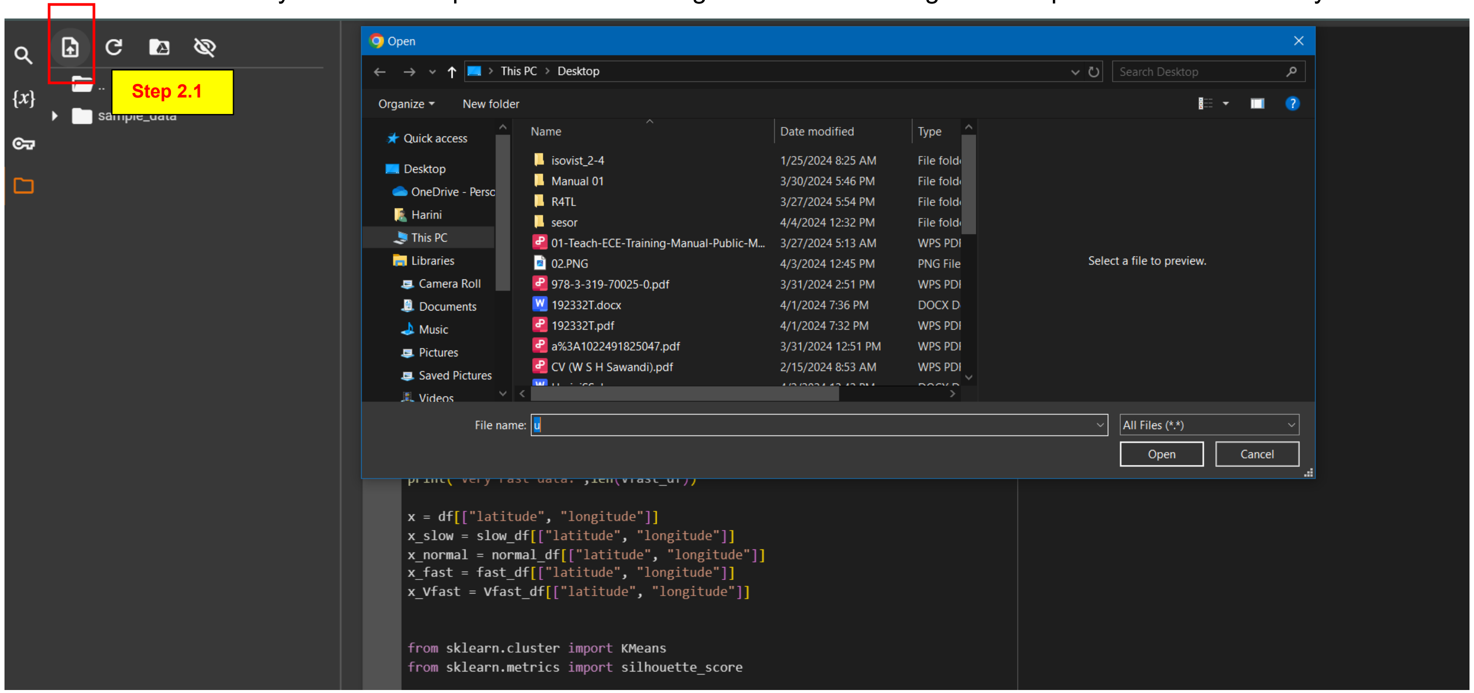


Figure 43 - Load the dataset.

2.2. Next click on the uploaded file copy path and paste it into the code.

```
file_path = "/content/Evening_without_outliers.csv" (copy path

df = pd.read_csv(file_path)

df.head()
```

Step 3

Trim the data to remove outliers. As per the purpose of the study, you can adjust the class ranges.

```
slow_df = df.loc[(df['speed'] <= 0.86) & (df['speed'] > 0.44)]

normal_df = df.loc[(df['speed'] <= 1.28) & (df['speed'] > 0.94)]

fast_df = df.loc[(df['speed'] <= 1.68) & (df['speed'] > 1.28)]

Vfast_df = df.loc[(df['speed'] <= 2.08) & (df['speed'] > 1.68)]
```

How to categorize the speed?

(Schimpl et al., 2011)

Mean speed (Sd)

- ***Slow : 0.65 (0.21)***
- ***Normal :1.11 (0.17)***
- ***Fast : 1.47 (0.20)***
- ***Very Fast: 1.83 (0.46***

Step 4

Print the counts of data points in each category.

```
print ('All data:', len(df))

print ('Slow data:',len(slow_df))

print ('Normal data:',len(normal_df))

print ('Fast data:',len(fast_df))

print ('Very Fast data:',len(Vfast_df))
```

Step 5

Prepare data for clustering.

```
x = df[["latitude", "longitude"]]

x_slow = slow_df[["latitude", "longitude"]]

x_normal = normal_df[["latitude", "longitude"]]

x_fast = fast_df[["latitude", "longitude"]]

x_Vfast = Vfast_df[["latitude", "longitude"]]
```

Step 6

Perform clustering for slow-walking data.

```
slow_clusters = []
slow_optimal_clusters = 50  #

if (slow_optimal_clusters > 1):
    model = KMeans(n_clusters=slow_optimal_clusters)
    y_kmeans = model.fit_predict(x_slow)
    slow_df['y1'] = y_kmeans

    for i in range(slow_optimal_clusters):
        cluster_i = slow_df.loc[slow_df['y1'] == i].copy()
```

```
        file_name = os.path.basename(file_path).split('.')[0]
        cluster_name = f"{file_name}_slow_{i}"
        # Calculate cluster statistics
        mean_speed = cluster_i["speed"].mean()
        std_dev_speed = cluster_i["speed"].std()
        min_speed = cluster_i["speed"].min()
        max_speed = cluster_i["speed"].max()
        range_speed = max_speed - min_speed
        count = cluster_i.shape[0]  # Number of data points in the cluster
        centroid_latitude = cluster_i["latitude"].mean()
        centroid_longitude = cluster_i["longitude"].mean()
        slow_clusters.append({
            "Cluster Name": cluster_name,
            "Count": count,
            "Mean Speed": mean_speed,
            "Std Dev Speed": std_dev_speed,
            "Min Speed": min_speed,
            "Max Speed": max_speed,
            "Range Speed": range_speed,
            "Centroid Latitude": centroid_latitude,
            "Centroid Longitude": centroid_longitude
print(slow_clusters_df)
```

```
    })
slow_clusters_df = pd.DataFrame(slow_clusters)
# Display the cluster statistics
print("Slow Walking Clusters:")
```

Step 7

Save the Clustered file as CSVs.:After running this step, all the CSV files are automatically saved in the section shown in Figure 44 below. You may need to refresh the area to see the updated files. From there, you can download these CSV files by clicking on the respective links. This process will generate four separate speed files for each segment

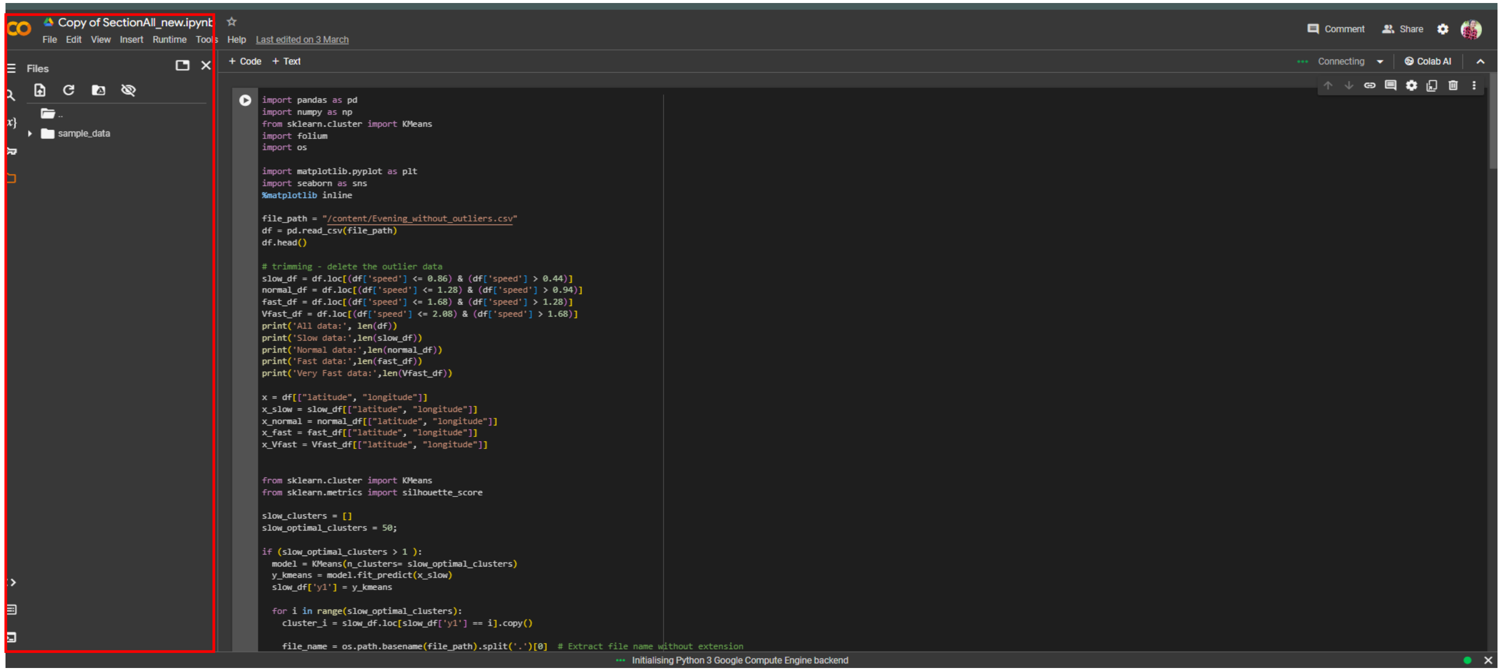


Figure 44 - Save the Clustered file as CSVs

Step 8

Display the Data Frame

```
slow_df.to_csv(f"{file_name}_slow.csv", index=False)

slow_clusters_df.to_csv(f"{file_name}_slow_clusters.csv", index=False)
```

Repeat the above steps for normal, fast, and very fast walking data (Users can follow the same structure and modify parameters as needed)

Link to the code –

https://colab.research.google.com/drive/1n_6Zfd388IW4690b8iULk4FcFbd8QMeC?authuser=0

4. MAP OUTPUTS

In the outputs section, you will find the outputs of the maps. These visualizations are based on the study's objectives, and you have the flexibility to adjust parameters as needed for your analysis.

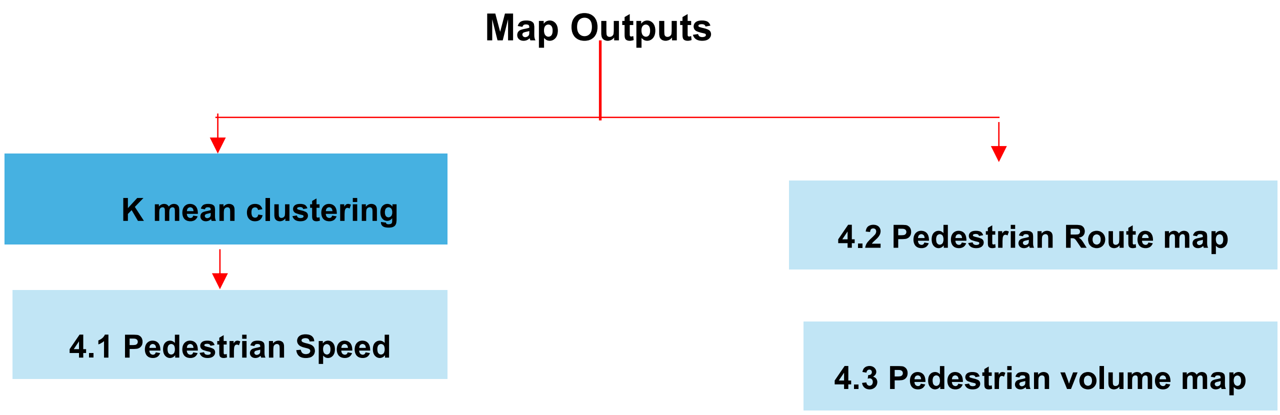


Figure 45 – Map outputs

4.1. Pedestrian Speed Map

After running the k-means clustering based on the study's purpose, you can map the data. In this study, k-means clustering resulted in 622 speed clusters (centroid points based on longitude and latitude). It provides separate files for each segment, and using these CSV files, you can map the data.

Step 01

In this step, we will use a sample CSV file, 'Normal clusters of Evening data,' for analysis. The process to import the data is the same as shown in Figure 23.

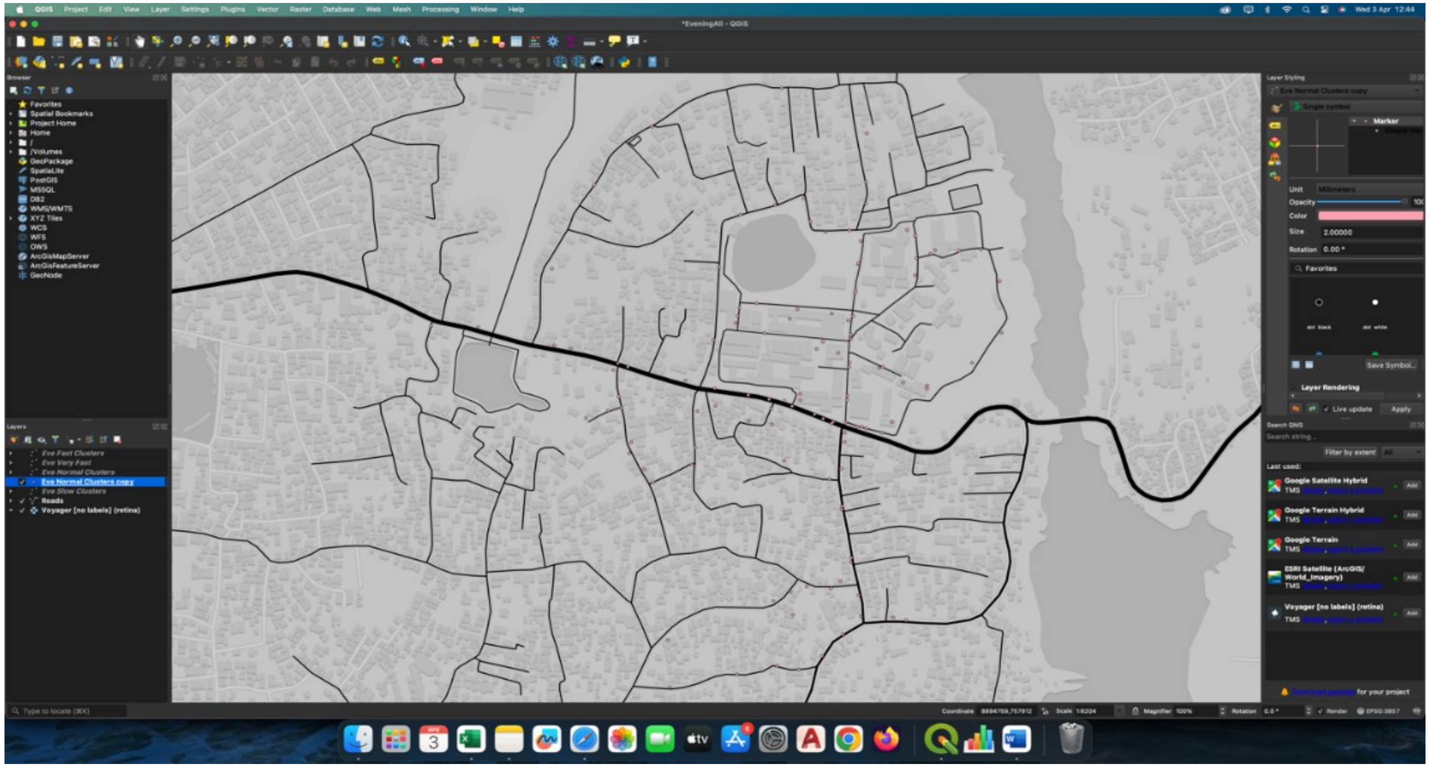


Figure 46 – 'Normal clusters of Evening data, for analysis

Step 02

Next, we will analyze areas that have a higher concentration of normal speeds. To do this, we will utilize the properties function of the layer

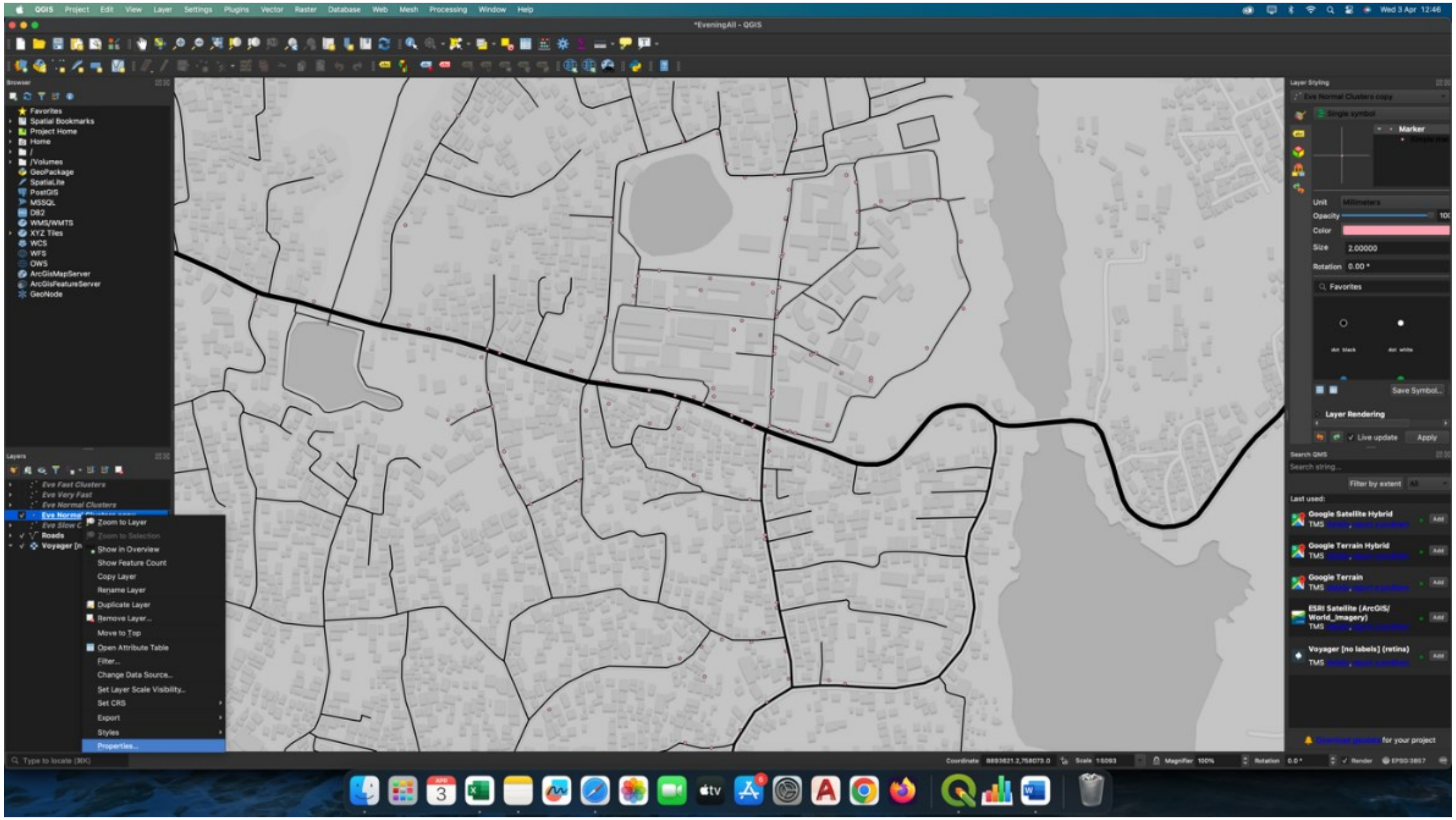


Figure 47- Higher concentration of normal speeds

Step 03

Select the Graduated Symbol

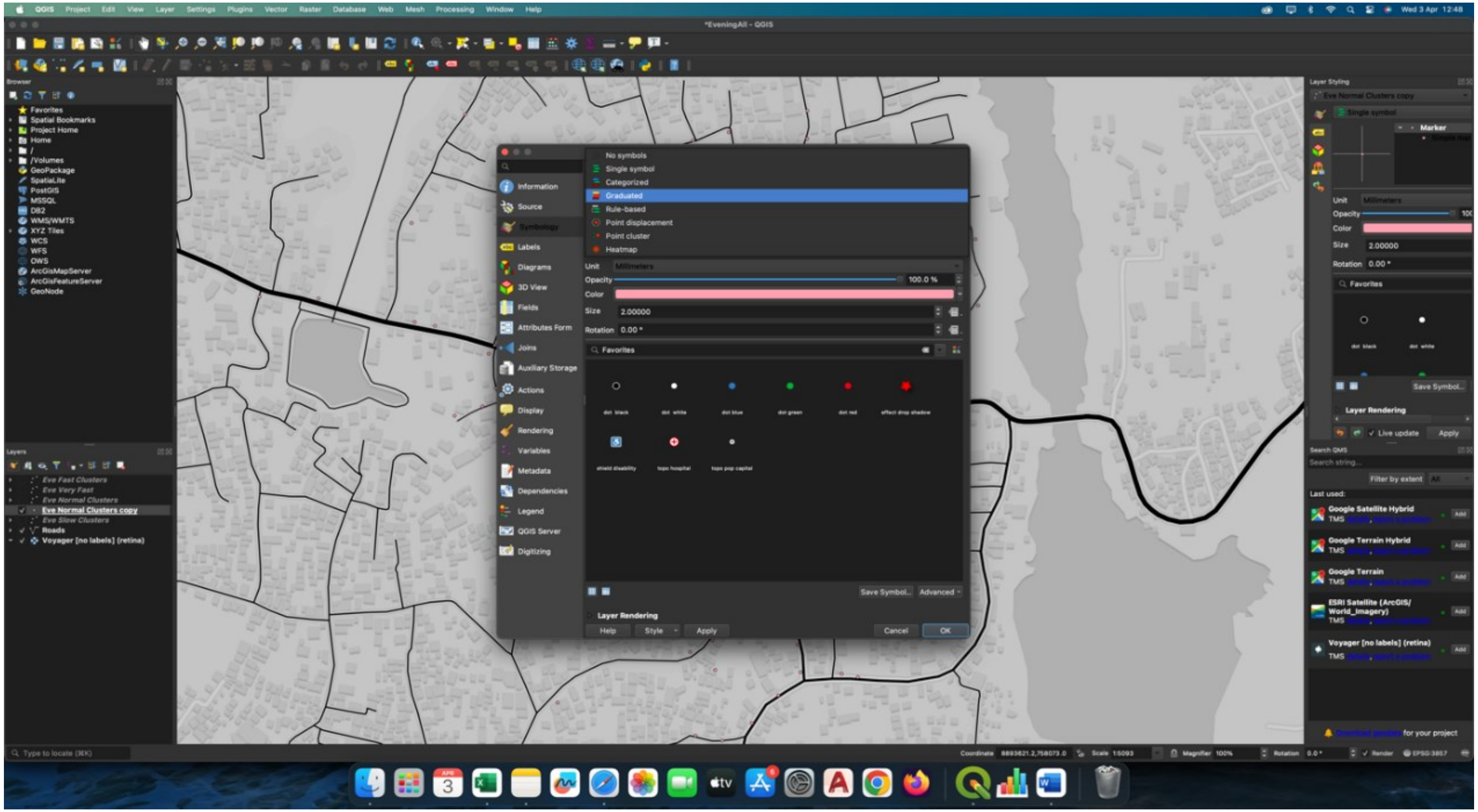


Figure 48 -Symbology

Step 04

Select the value as Count.

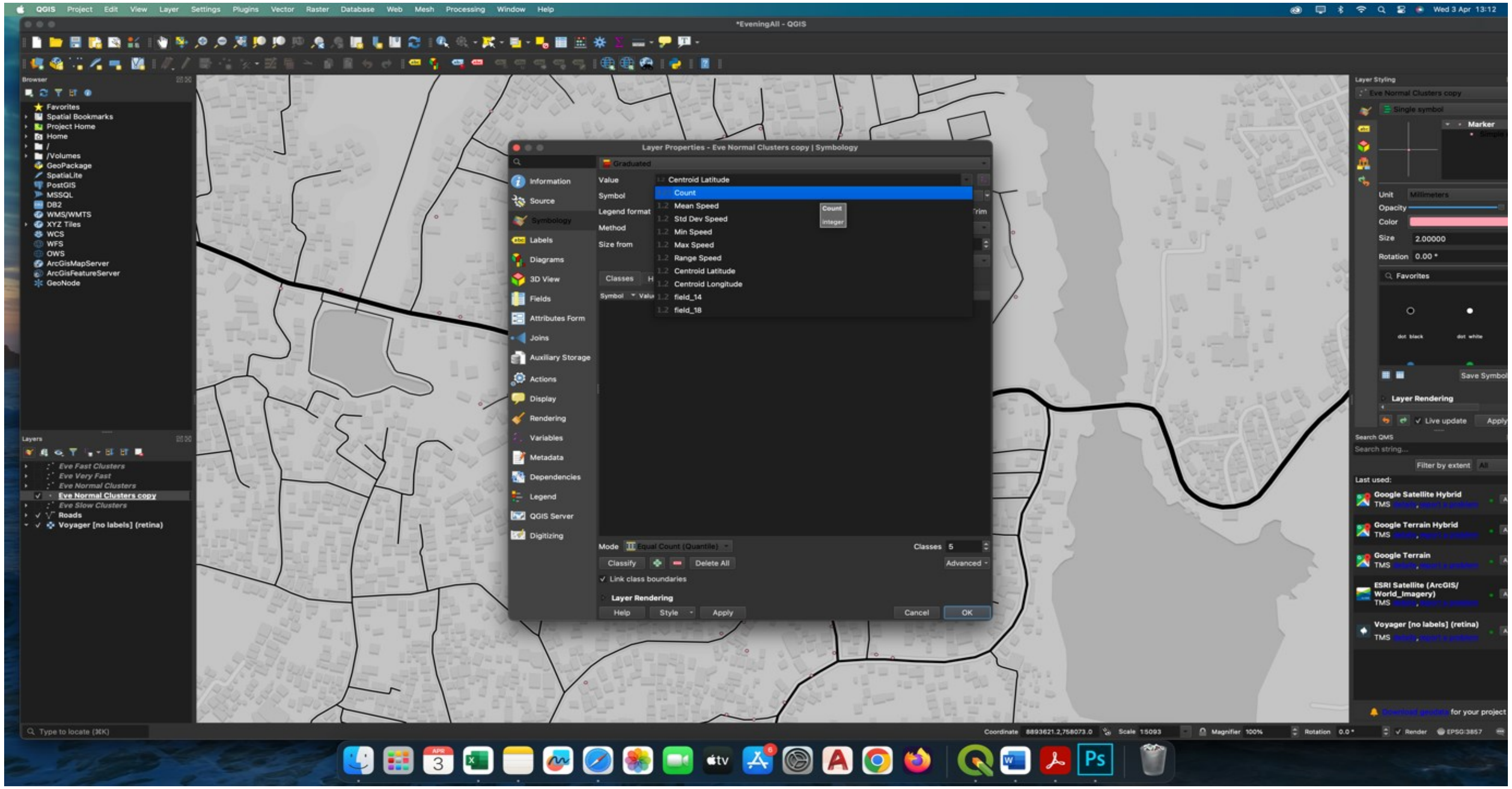


Figure 49 – Select value as Count.

Step 05

Select the Method as Size

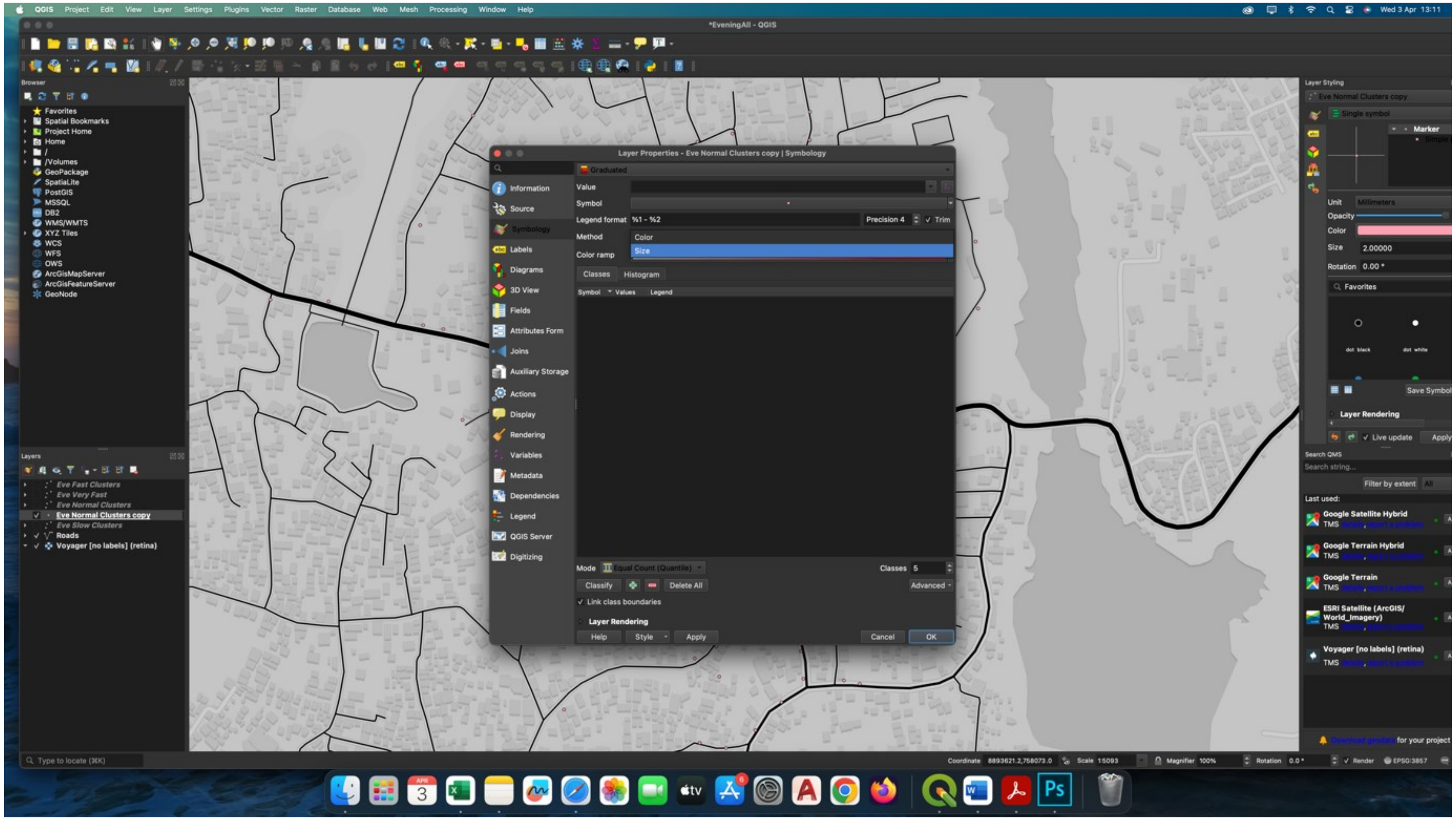


Figure 50 - Select the Method

Step 06

Then Classify according to equal intervals.

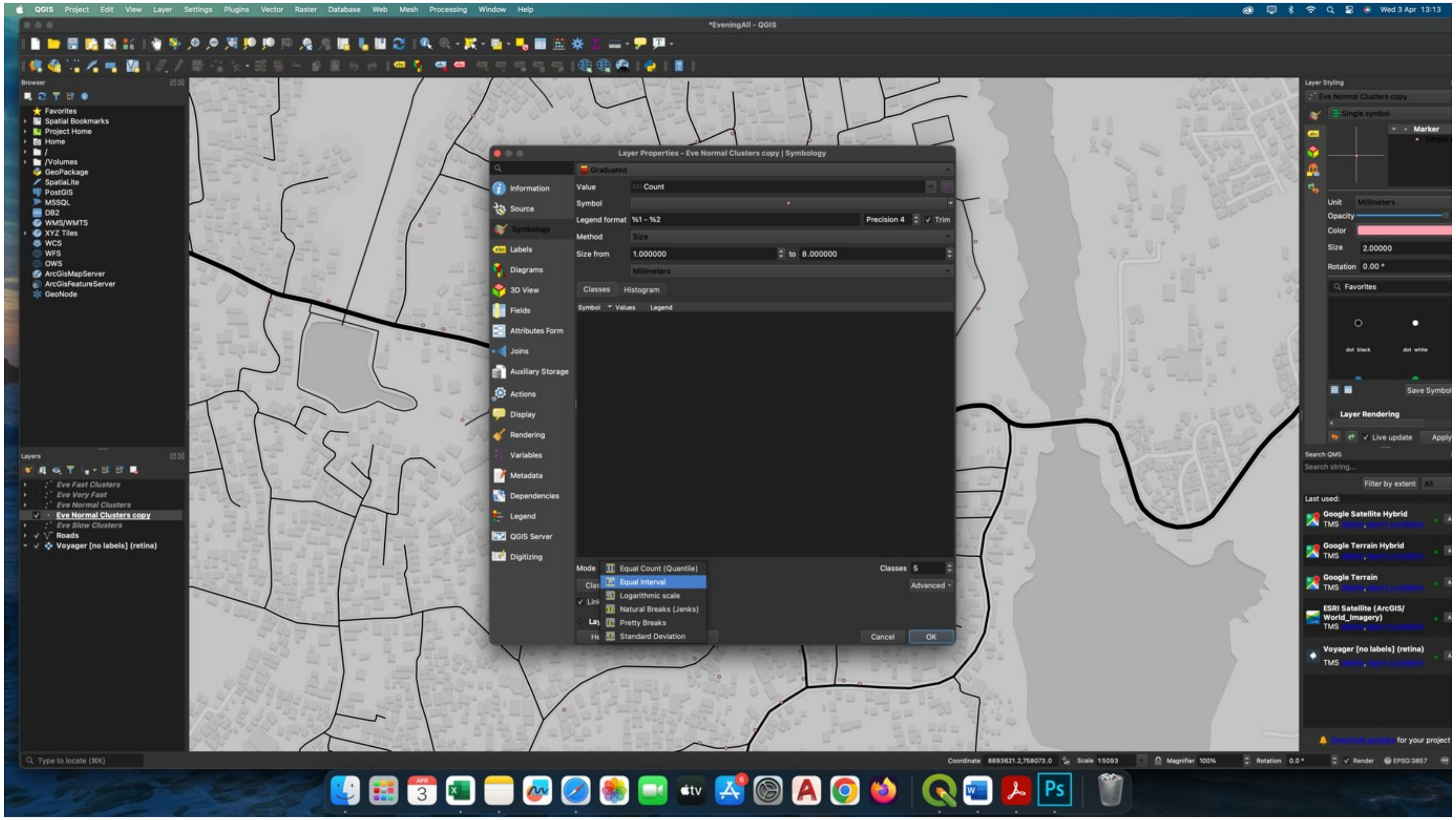


Figure 51 - Classify to equal intervals.

Step 07

Then it will appear without any color difference.

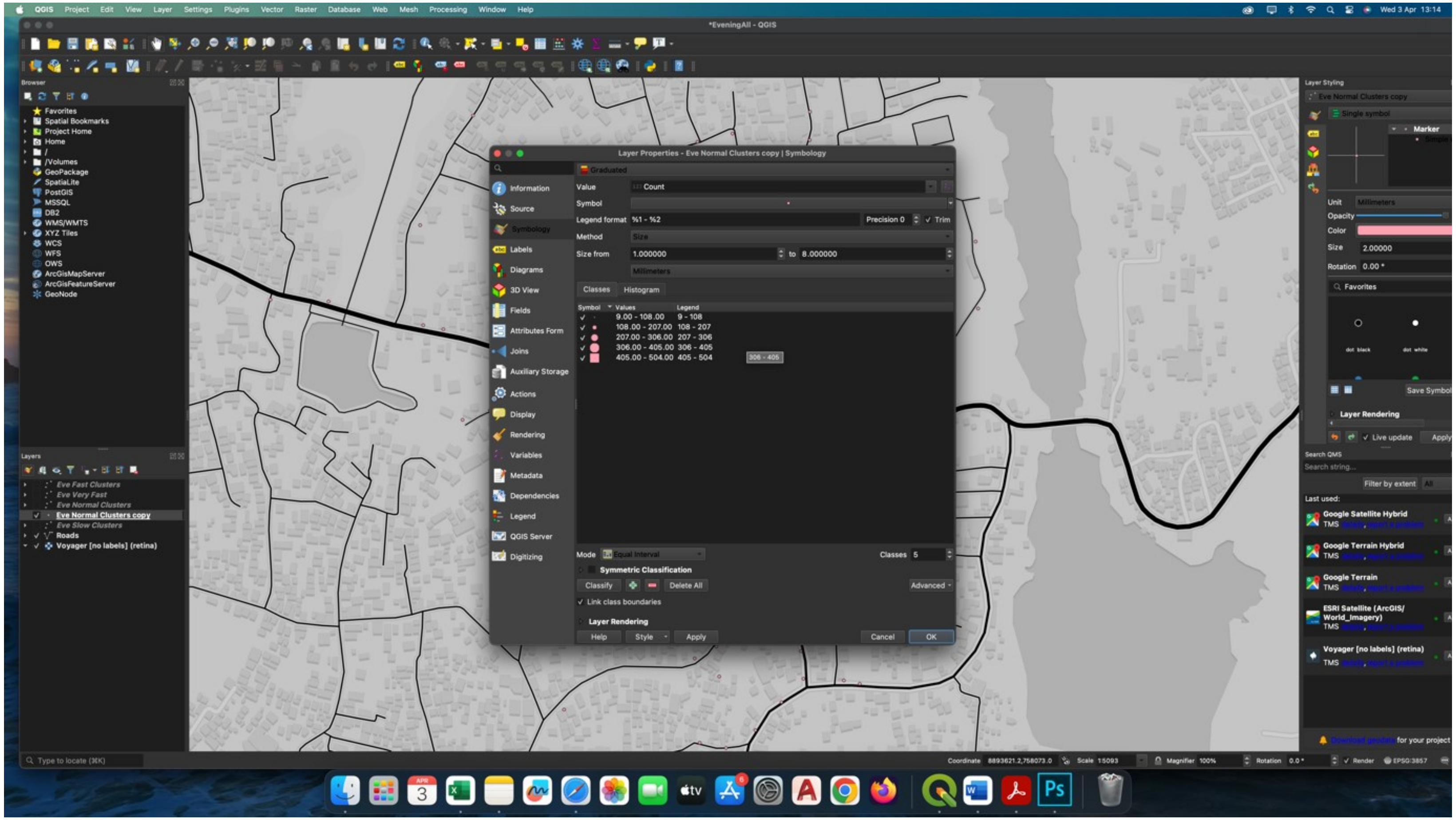


Figure 52 – Symbology

Step 08

Then select each symbol and manually change colors and sizes according.

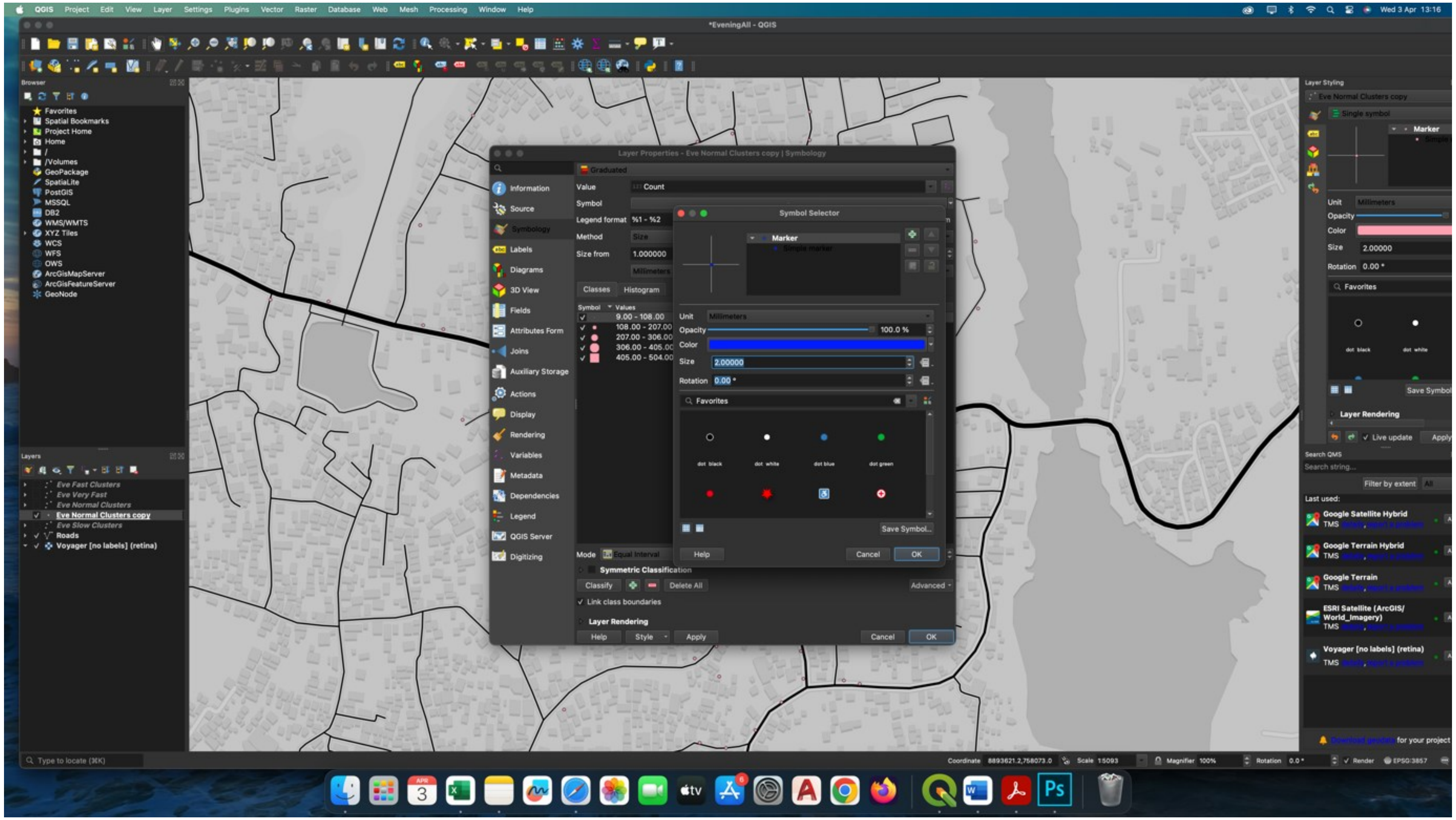


Figure 53 – Symbology

Step 09

The final categories will be as follows taking a blue to red color ramp with size differences of 2pt.

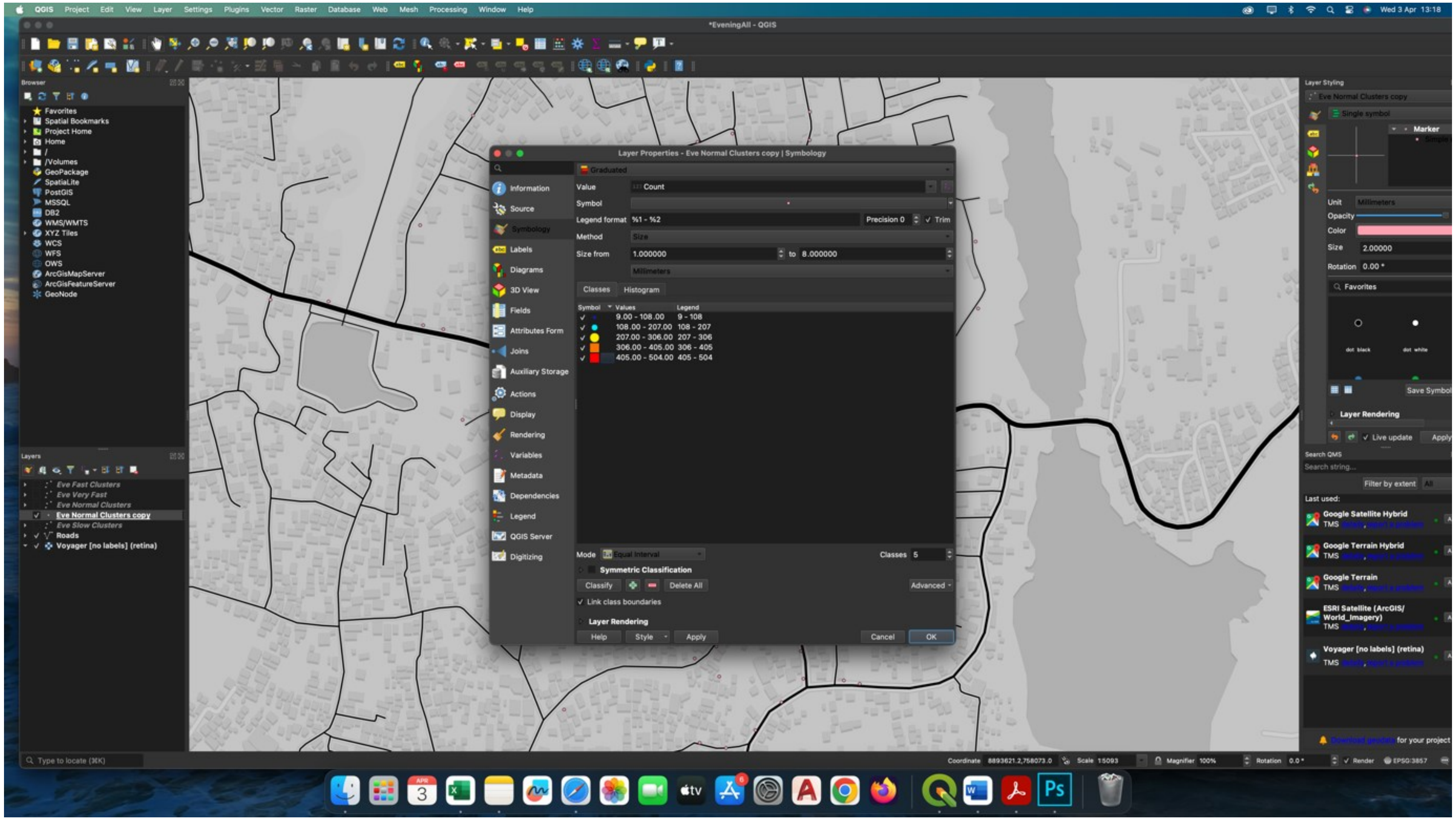


Figure 54- Final categories

Figure 55 show the output of the clusters.

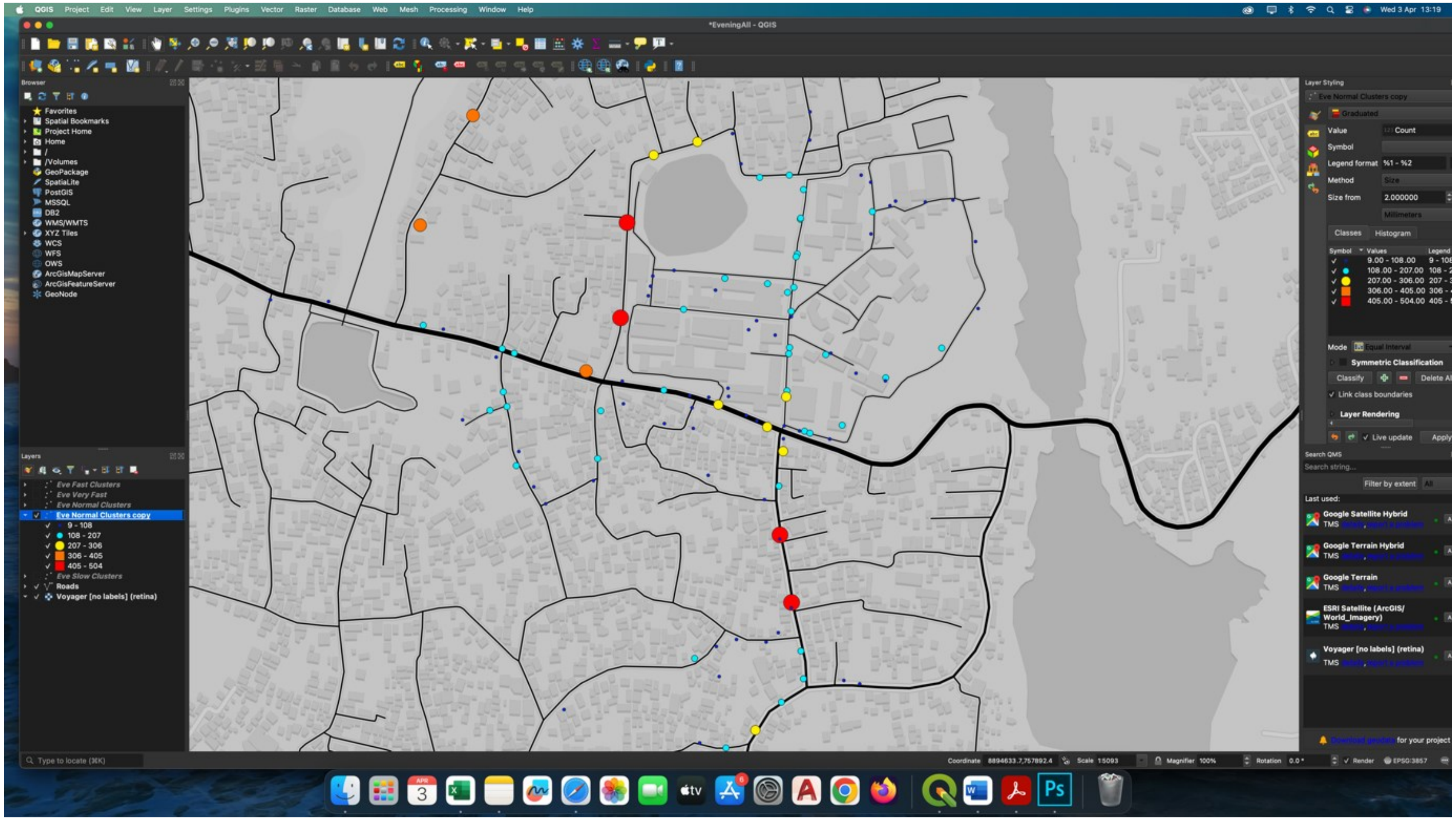


Figure 55 - Output of the clusters.

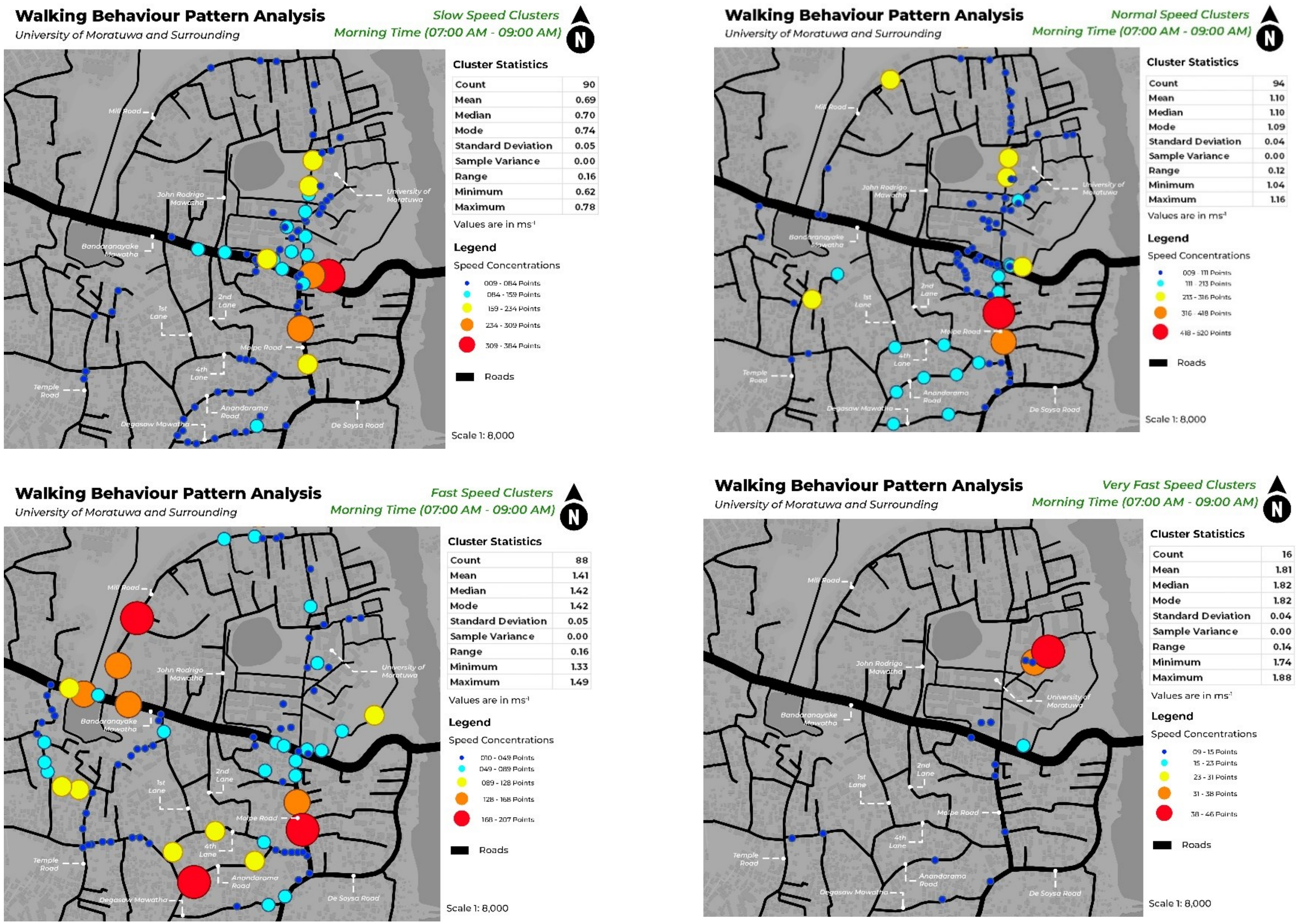


Figure 56 -Pedestrian Concentration of Different Speed clusters in Morning Hours

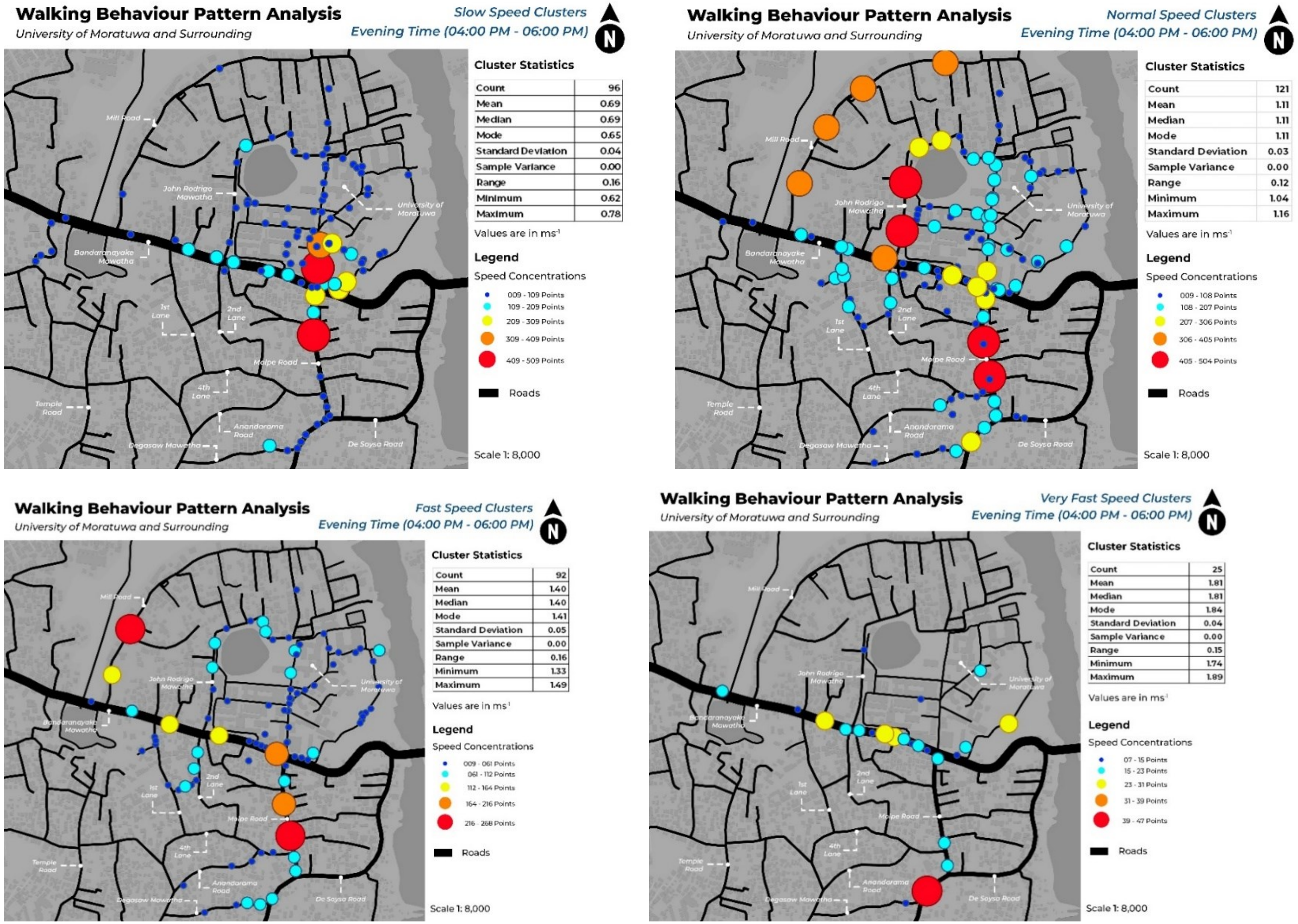


Figure 57 -Pedestrian Concentration of Different Speed clusters in Evening Hours

4.2. Pedestrian route map

This study specifically focuses on walking speed, time, and direction. However, in addition to these parameters, we can create pedestrian route maps by using each participant's CSV file, which includes cleaned outlier data. This allows us to gain insights into the most commonly used roads by students in the case study area. By mapping these routes, we can identify areas that might benefit from targeted intervention.

Step 01

In this step, first import the csv file same as shown in Figure 23. Then go to the processing tab and under that click on Toolbox.

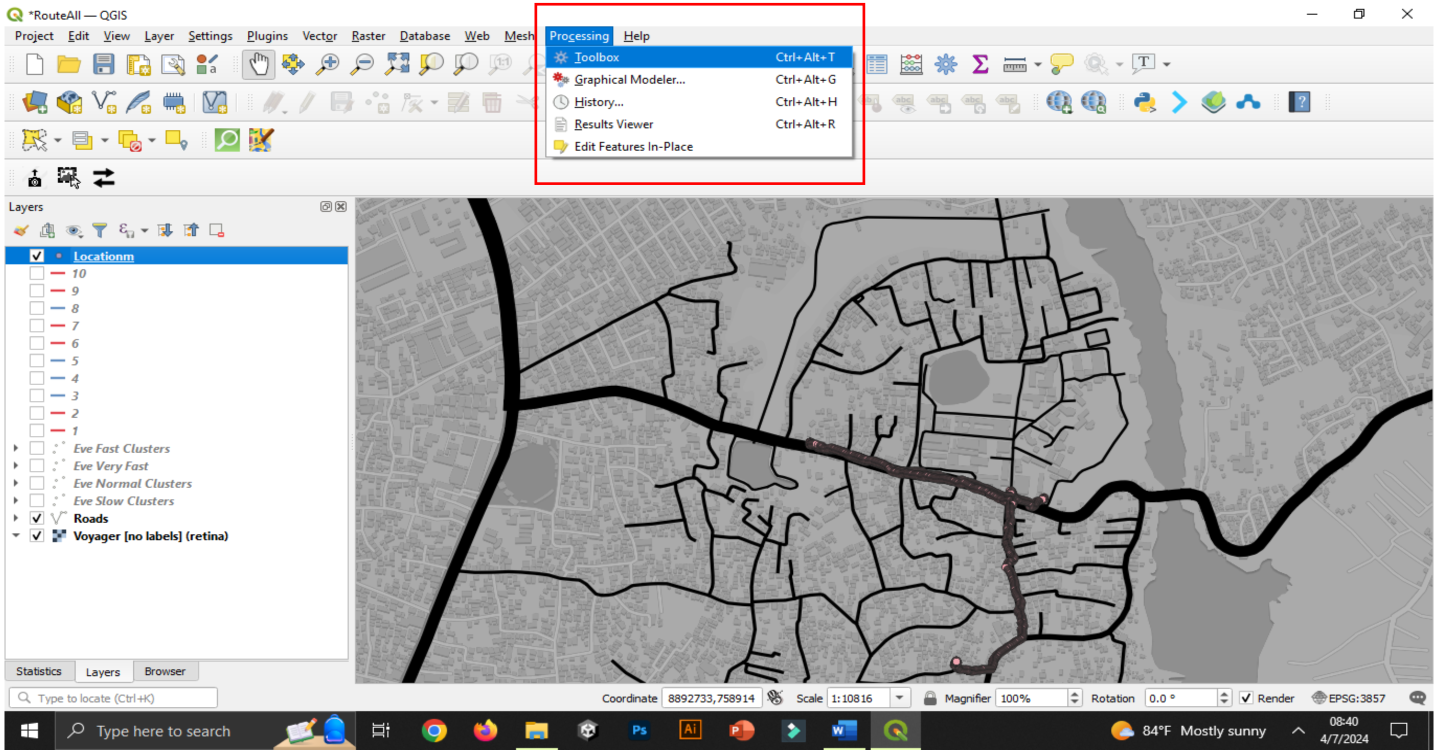


Figure 58 - Import the CSV file

Step 02

Under "Vector Creation," select the "Point to Path" tool.

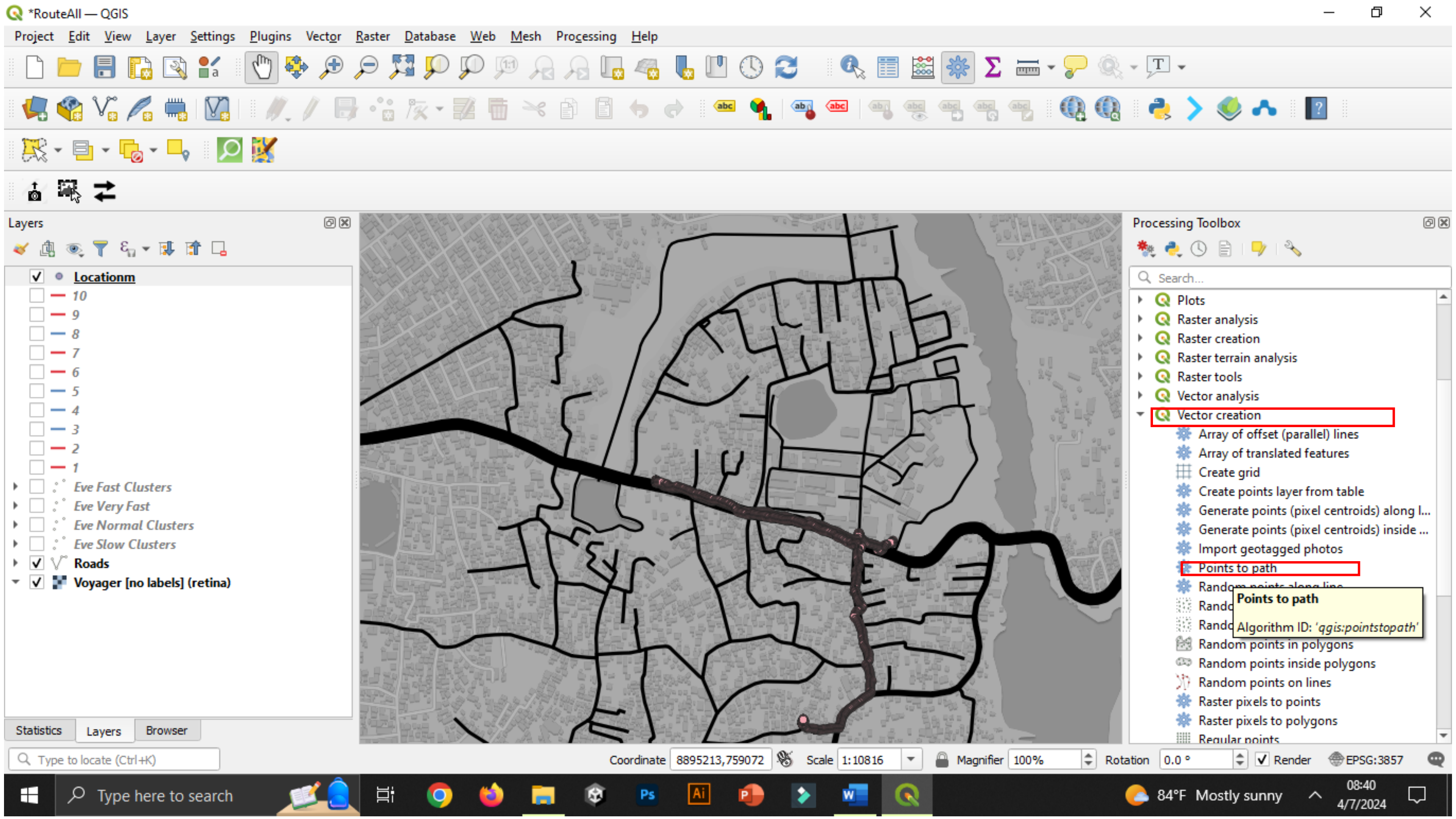


Figure 59 - Vector Creation

Step 03

Select the point file and select time as the order.

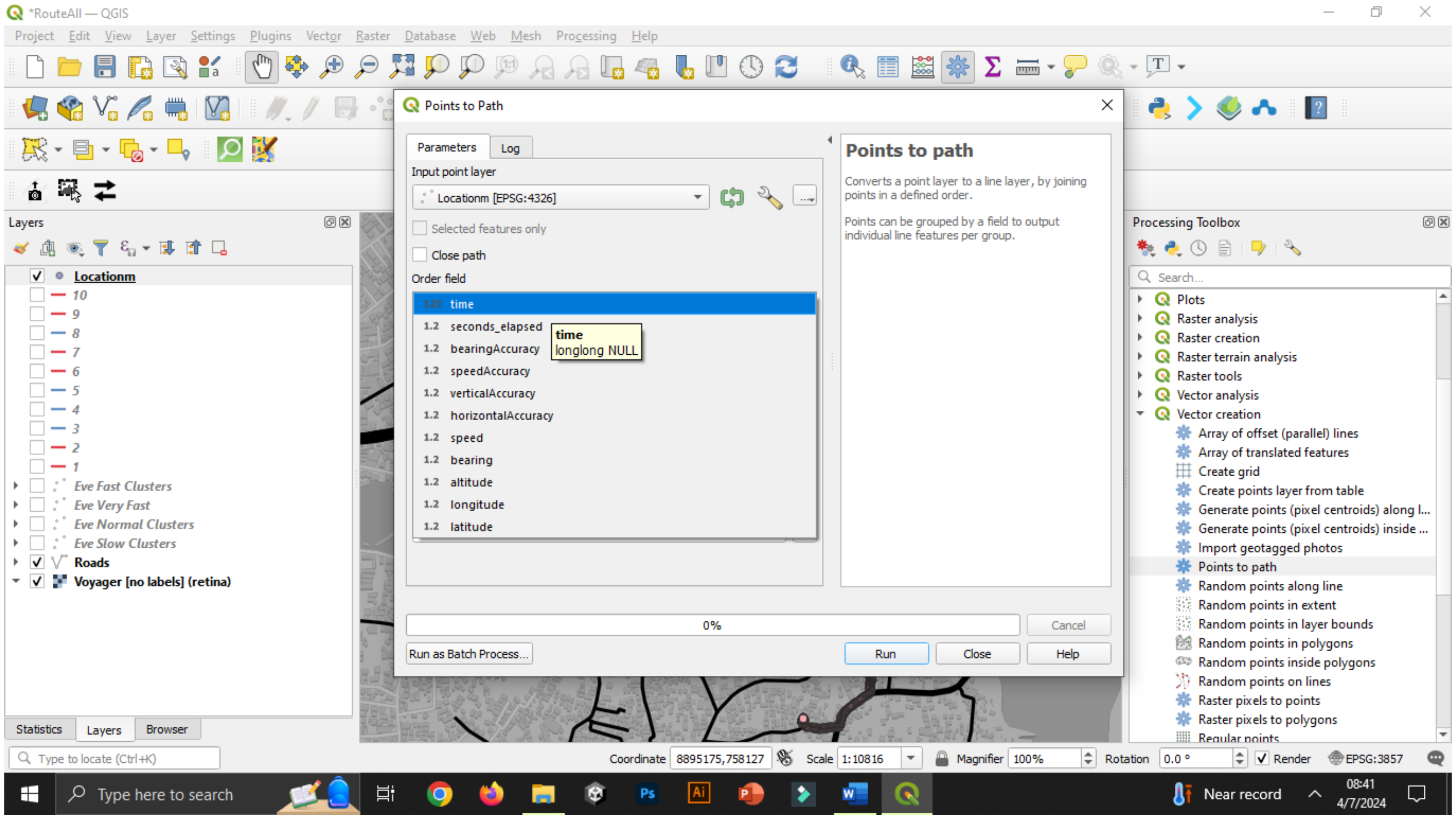


Figure 60 - Select the point file

Step 04

Give an output location as a save file.

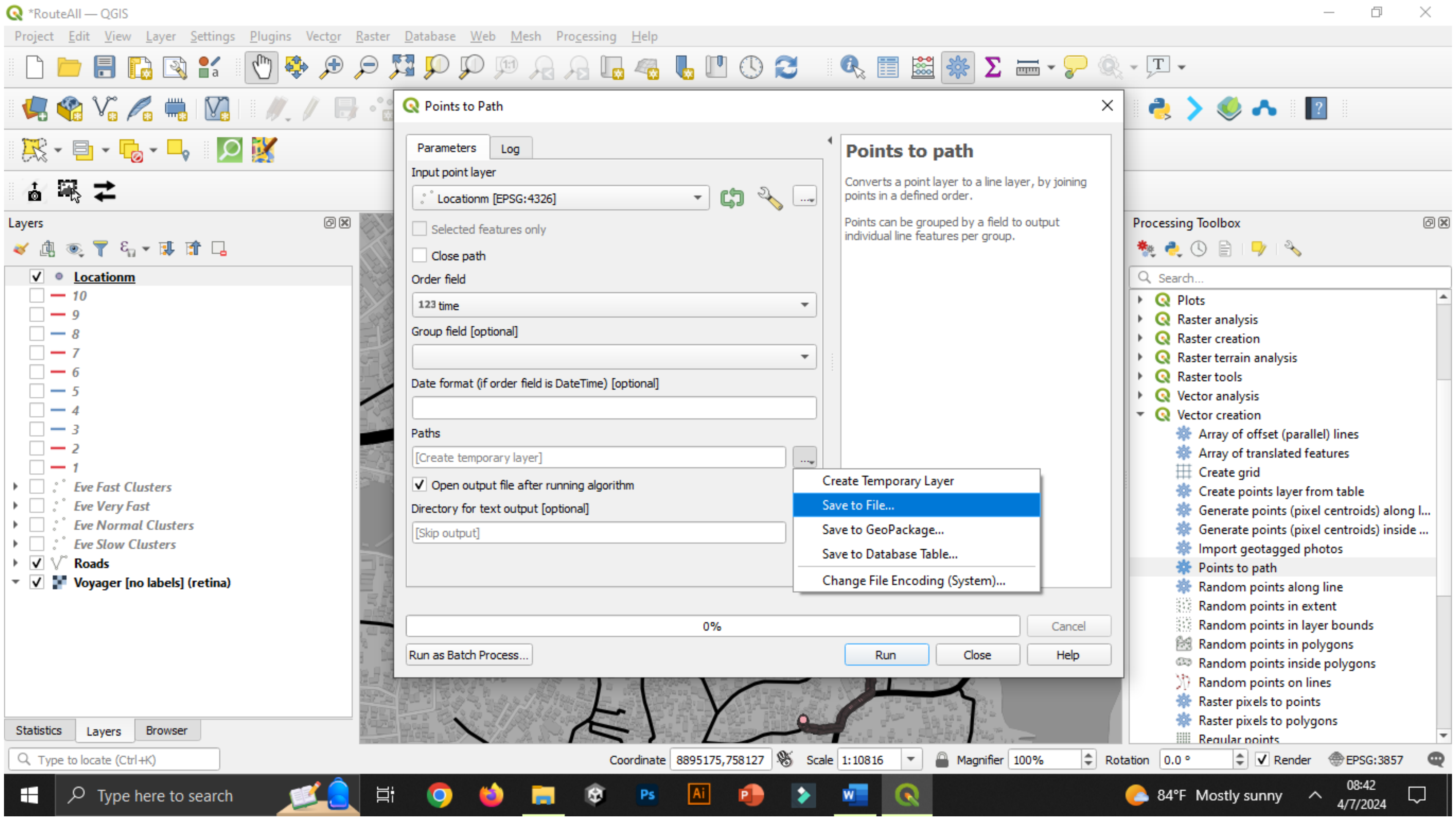


Figure 61 – Save to File

Step 05

Then Run the Process

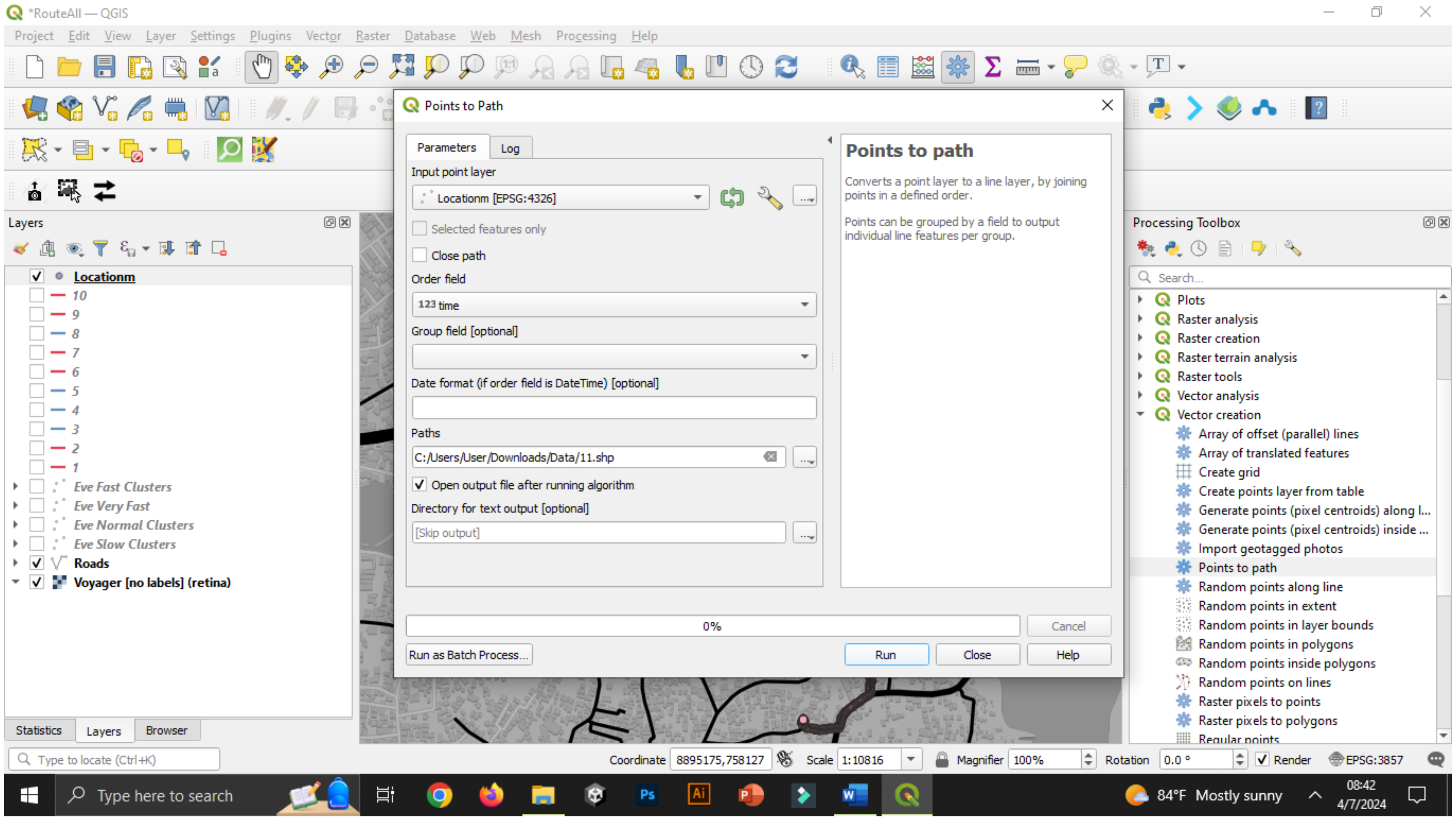


Figure 62 - Run the Process

Figure 63 show the output of the Routes. This can be repeat with all participant data to get the overall image of the route map.

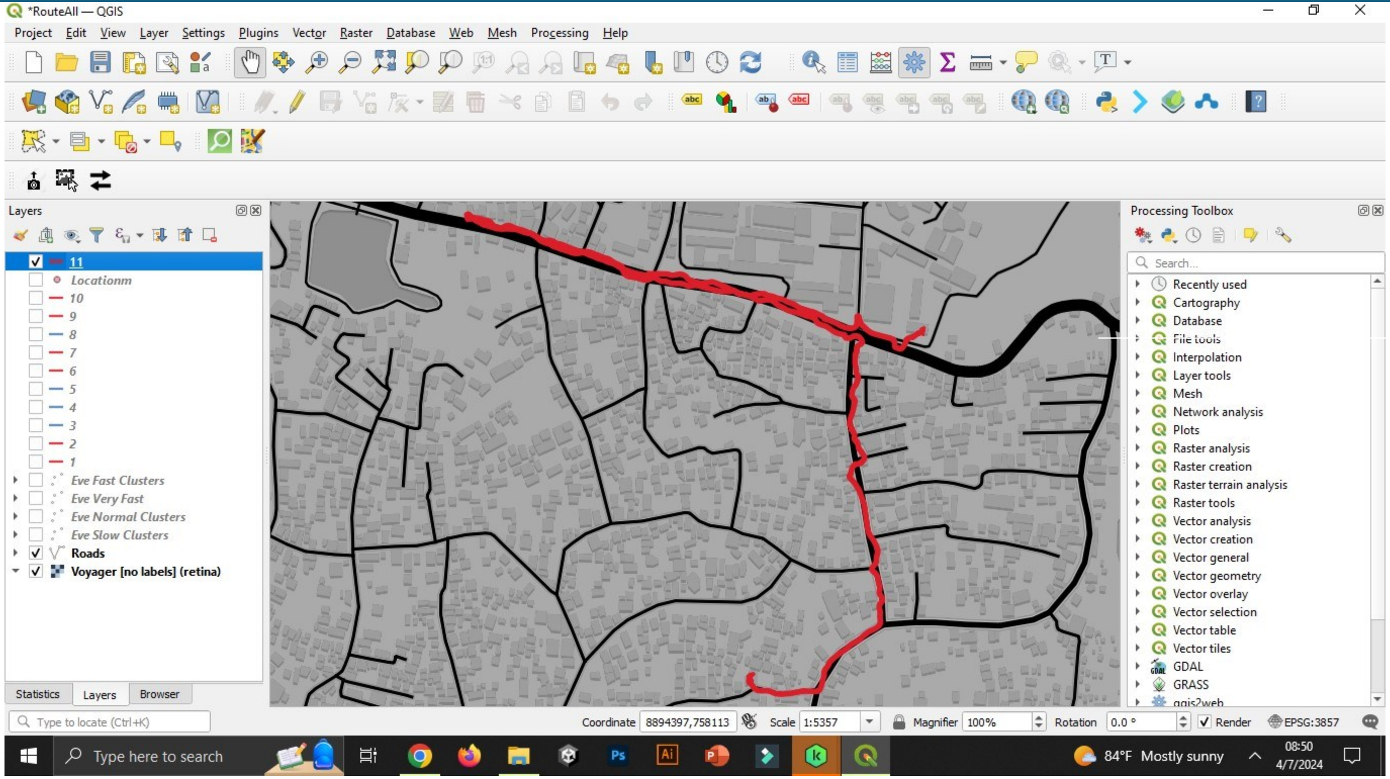


Figure 63 - Output of the Routes

Input

Data (Points)

→

Transforms a series of input points into continuous paths, facilitating the analysis of spatial movement patterns.

→

Output

Pedetrian Route Map

4.3. Pedestrian Volume Map

This study aims to analyze the patterns of pedestrian movement, with a particular focus on parameters such as walking speed, time, and direction. Although our focus is on these characteristics, we also explore the pedestrian volume of the area by collecting and analyzing the collected data. Using a CSV file that includes all participant walking data that was used in the previous analysis, which has been carefully filtered to remove any outliers, we analyze the pedestrian volume. By utilizing kernel density estimation methods in QGIS, we can reveal a detailed comprehension of the specific locations that pedestrians visit most frequently within and around the chosen research area. This approach allows us to clearly define areas with heavy traffic, which helps with making well-informed decisions on urban planning and allocation of resources.

What is Kernal Density?

Kernel density estimation (KDE) is a statistical technique used to estimate the probability density function of a random variable. In simpler terms, it helps us understand where data points are concentrated within a given space. In our case, we use KDE in QGIS to create a visual representation of where pedestrians tend to gather the most. By analyzing this density map, we can pinpoint hotspots of pedestrian activity, helping urban planners make informed decisions about infrastructure, traffic flow, and public space design.

Step 01

In this step, first, import the CSV file same as shown in Figure 23. Then go to the processing tab and under that click on Toolbox. Then Search “heat map”

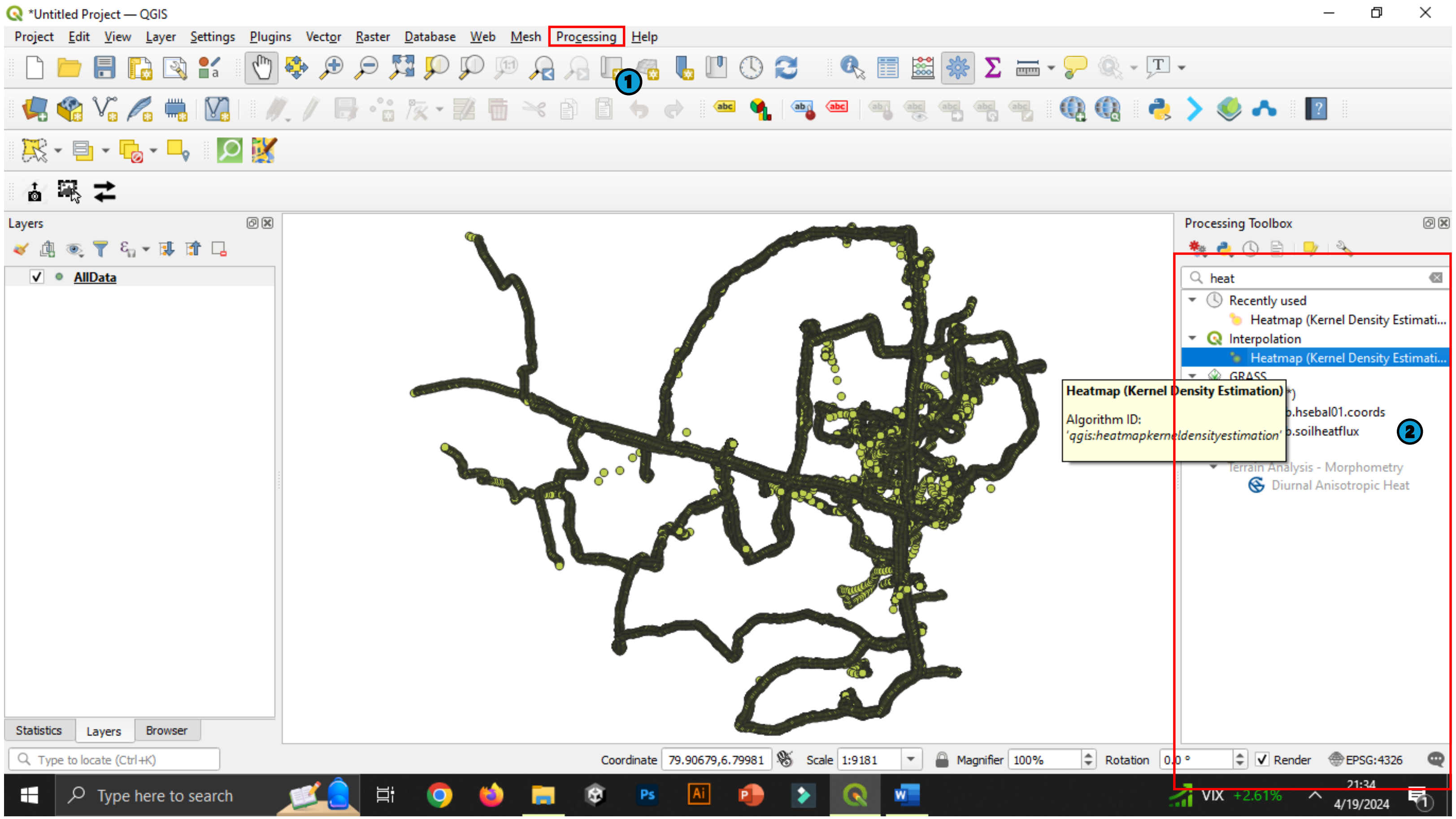


Figure 64 -Import the CSV file

Step 02

In this step, first, the CSV file sets the radius and necessary parameters and saves the file., then click "Run" to generate the density map.

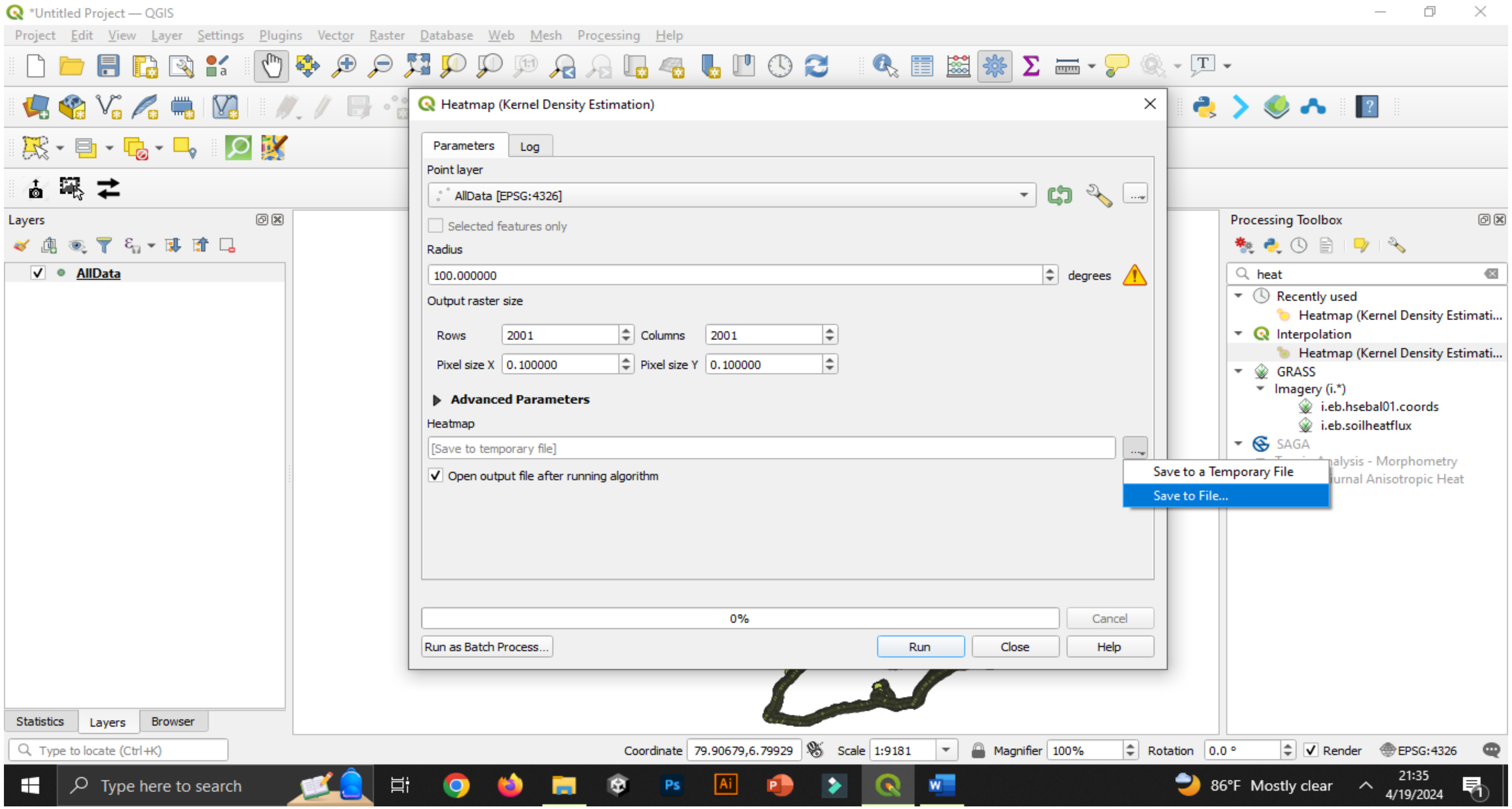


Figure 65- Set the Parameters

Step 03

Then as shown below it will take time to run the map.

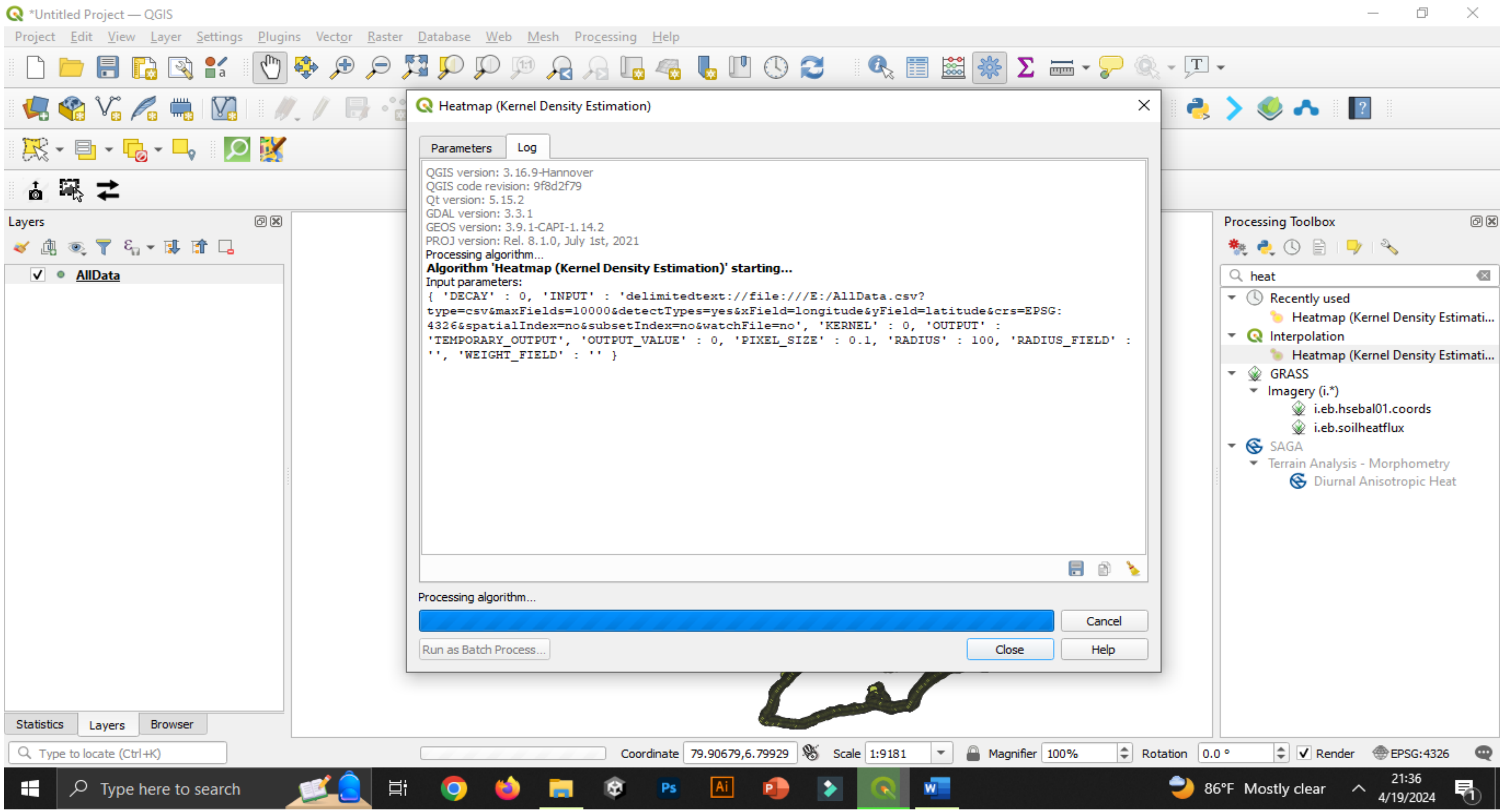


Figure 66 -Run the map

Step 04

Go to the data source manager, click on the "Raster" option, and proceed to open the previously saved files.

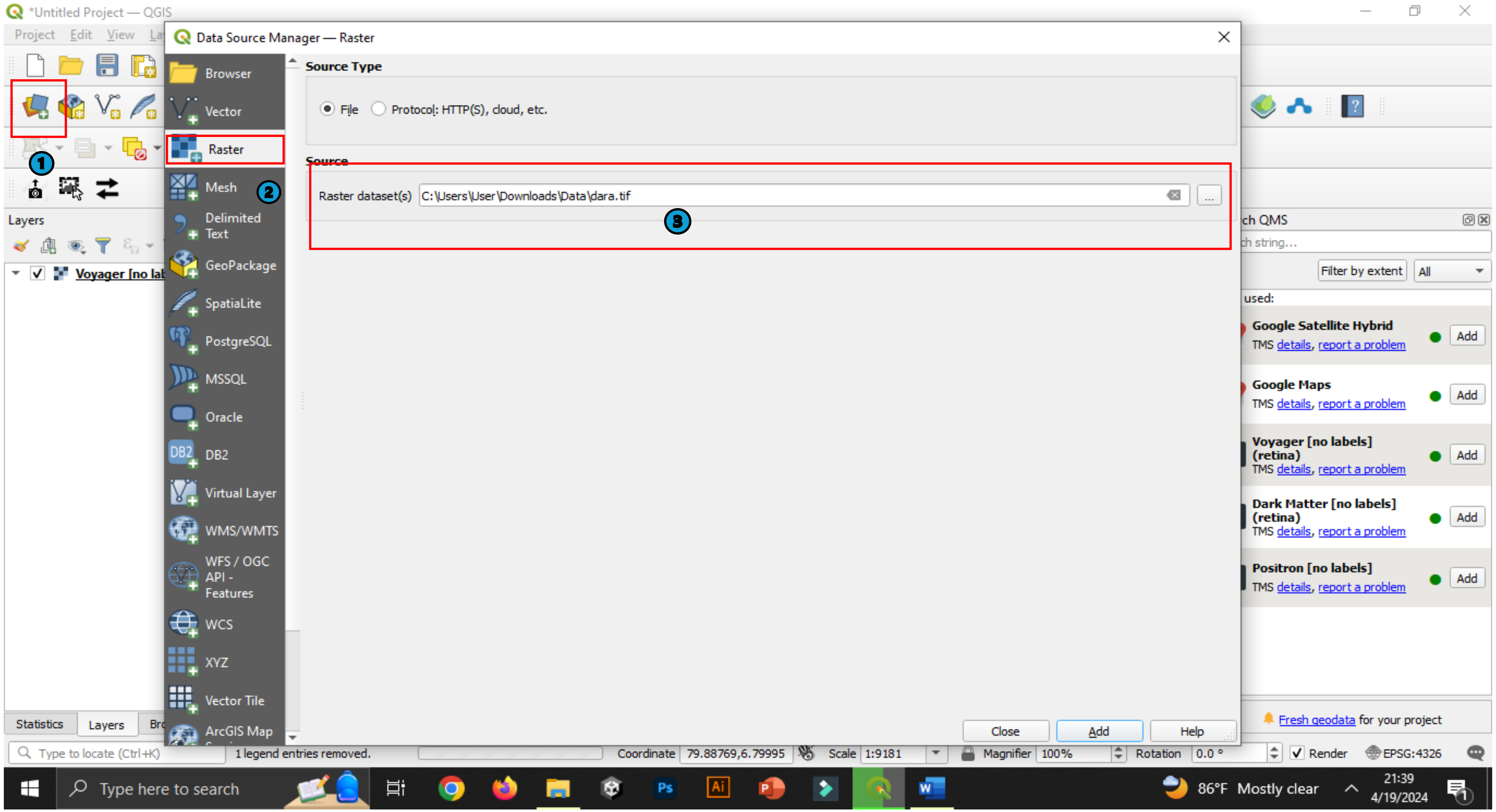


Figure 67 - Data source manager

Step 05

Then the map will display as follows. Then go to the properties of the file.

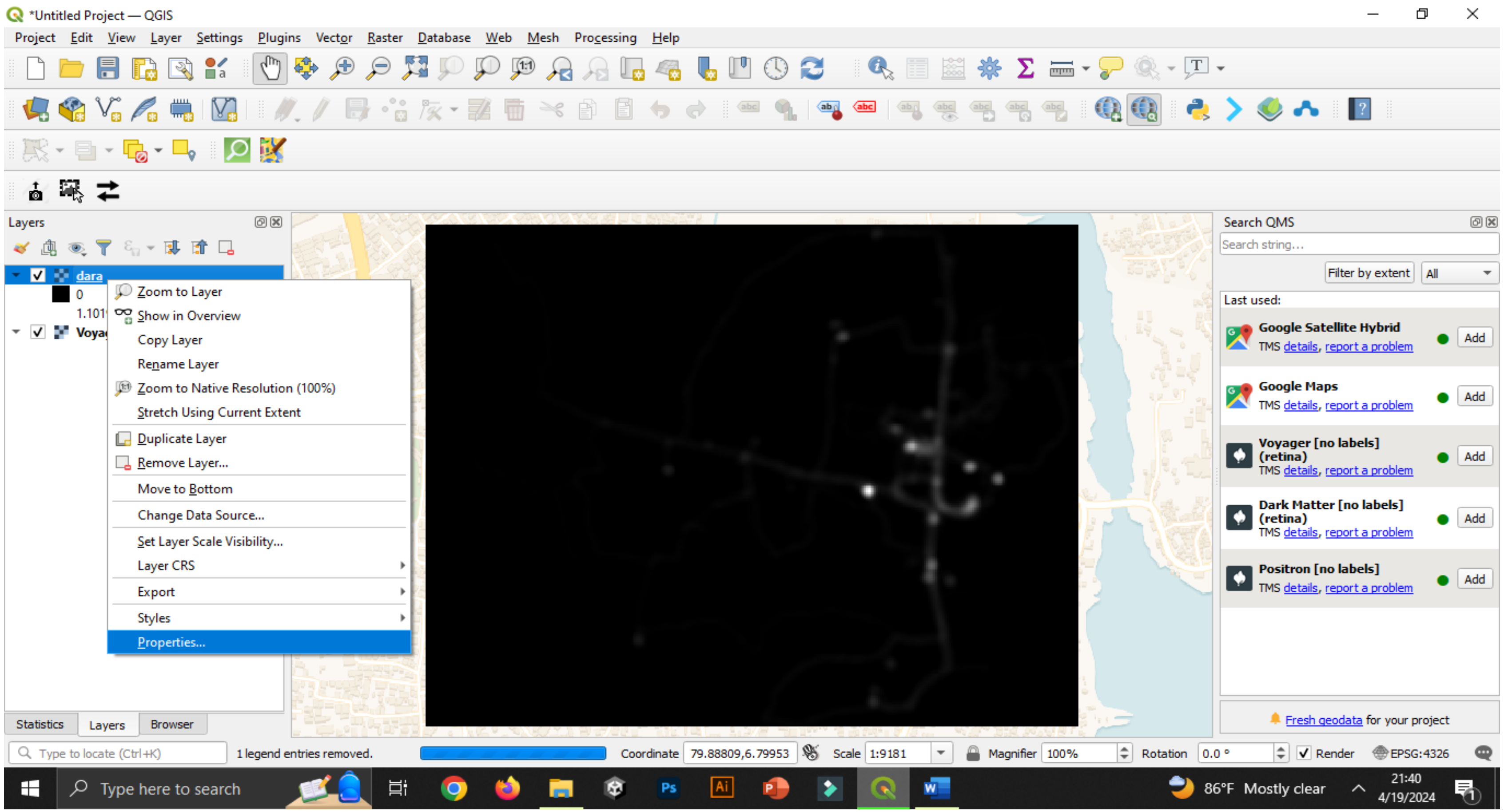


Figure 68 -Properties

Step 06

Then under the symbology tab select the “single band pseudo color.”

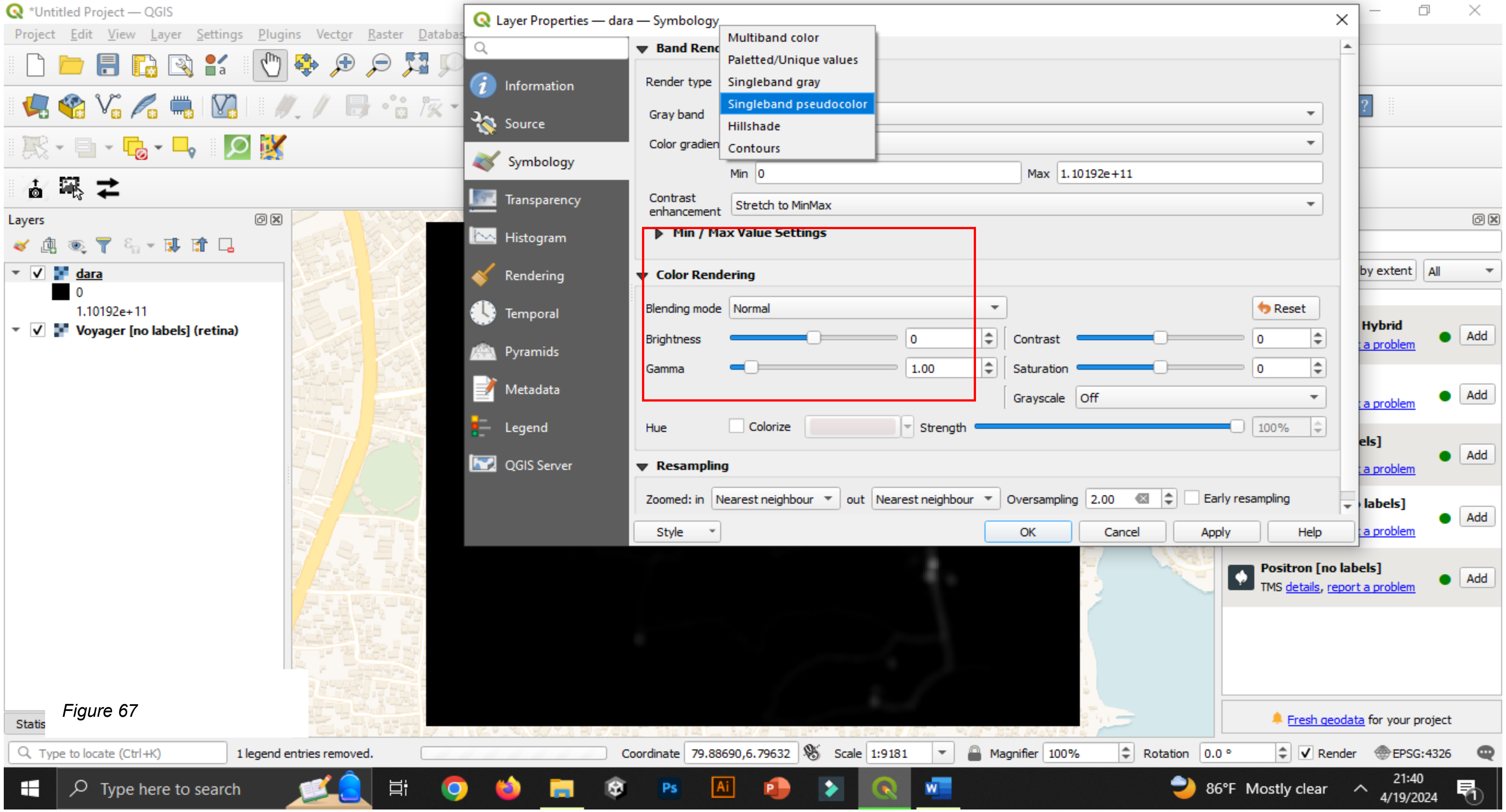


Figure 69- Single band pseudo color

Step 07

Then select equal interval and based on your preference select the color ramps and set other parameters.

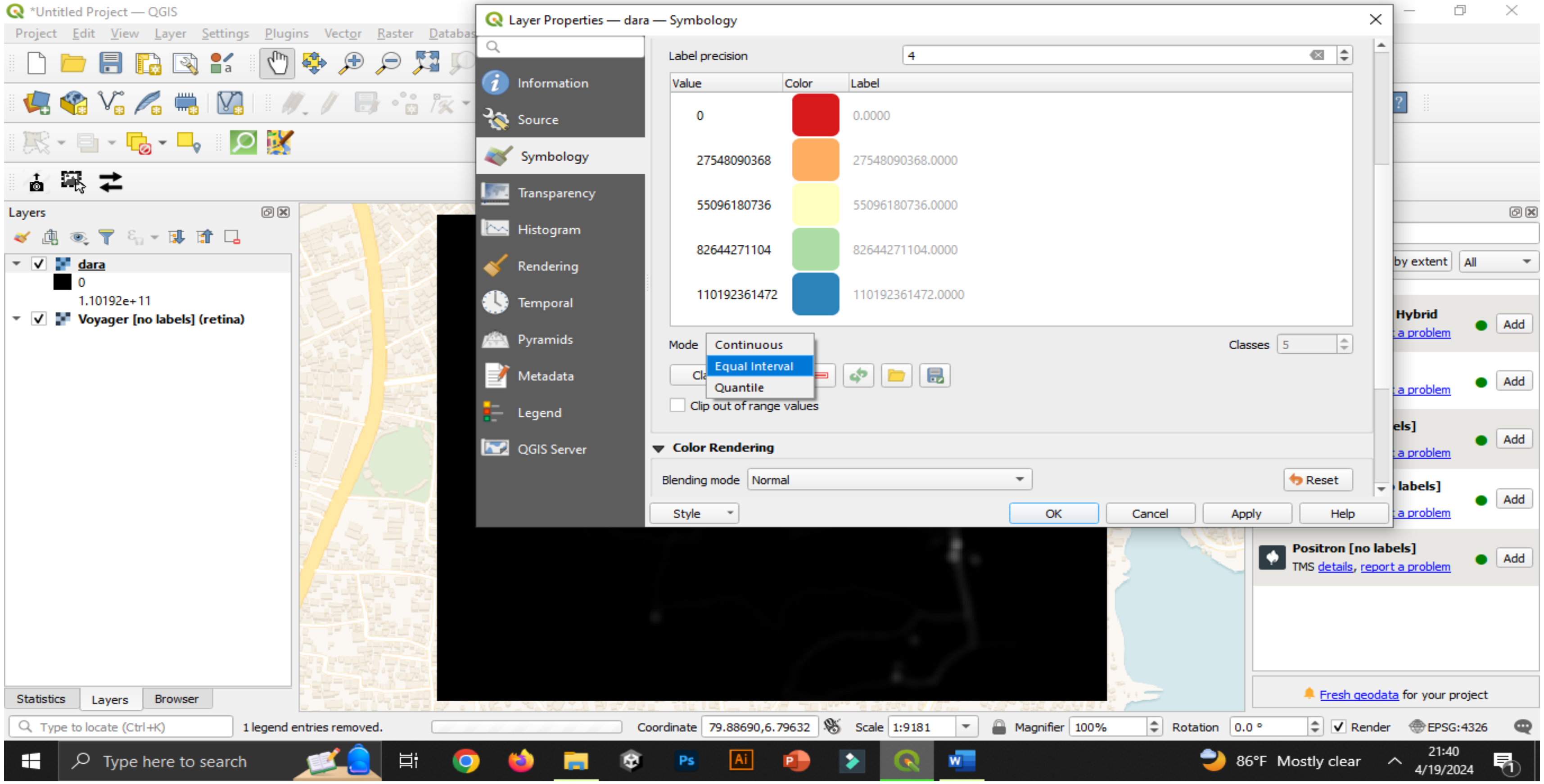


Figure 70 -Colour ramps

Figure 71 show the output of the pedestrian volume map.

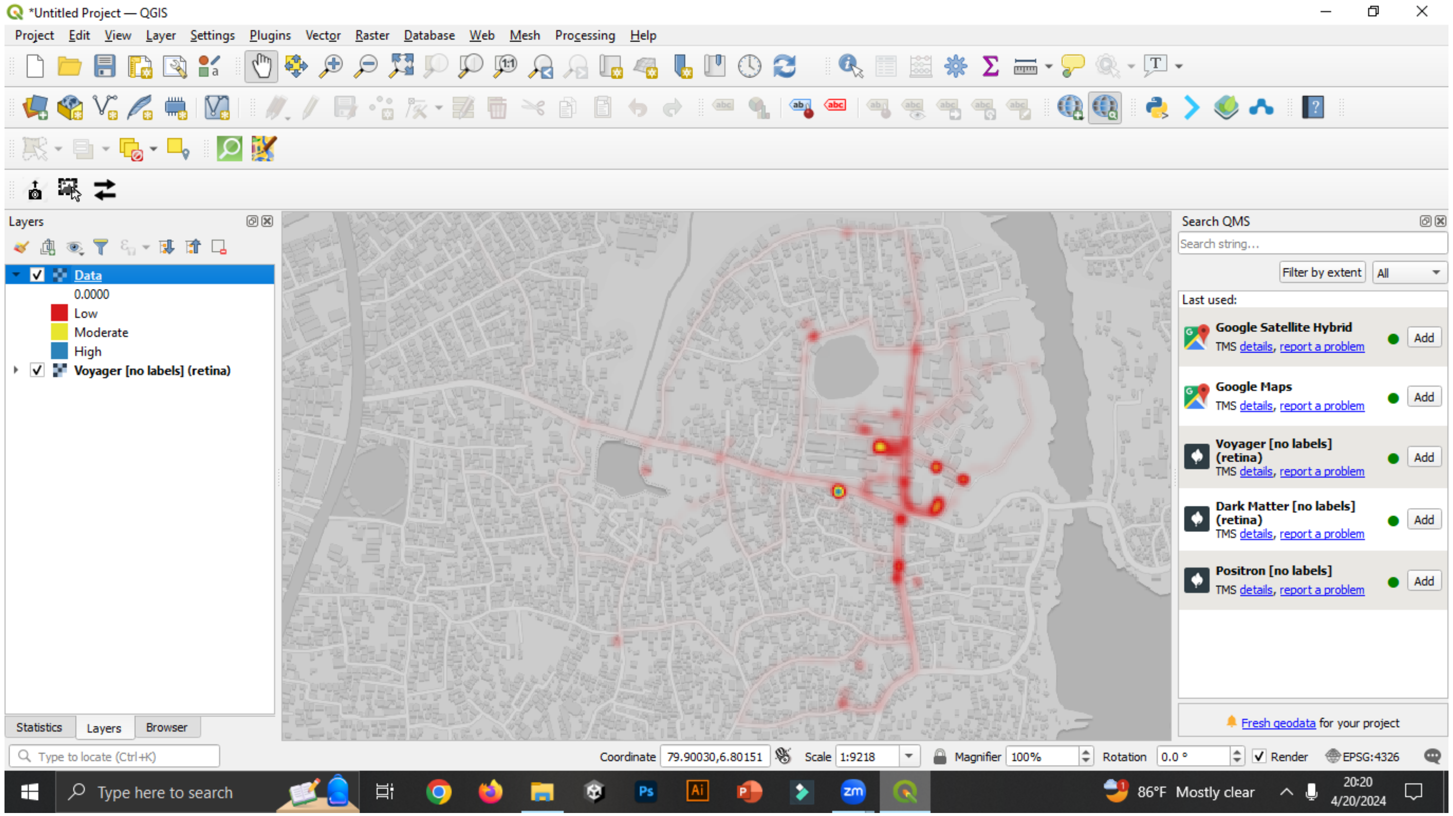


Figure 71 –Pedestrian volum map

Input

Data (Points)

Processed to generate a smoothed surface representing the density of occurrences, providing spatial patterns and concentration.

Output

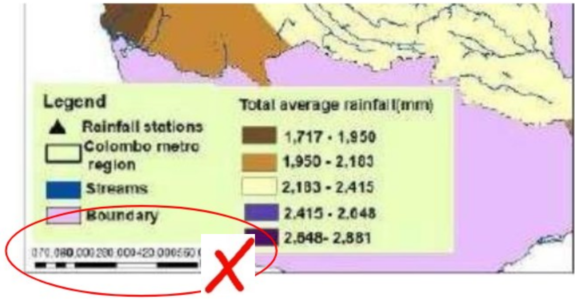
Pedetrian Volume Map

5. PREPARATION OF MAPS

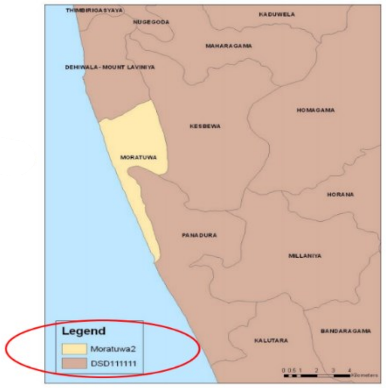
1. Use Suitable Colors & Symbols
2. Show adjacent administration boundaries.
3. Maximum utilization of map space
4. Show the Graticule Network with appropriate grid size.
5. Show the basic elements in the map.
    a. Transportation Networks
    b. Water Bodies
6. Prepare a descriptive map.
    a. Label notable features.
        i. Major Road Types
        ii. Major Rivers
        iii. GN Boundaries
        iv. If needed, show the location of the area in a different data frame.
7. Check the units of the scale bar.
8. Mention correct units at the legend.
9. Check the text given in the legend.
10. Mention the correct sources.

Examples

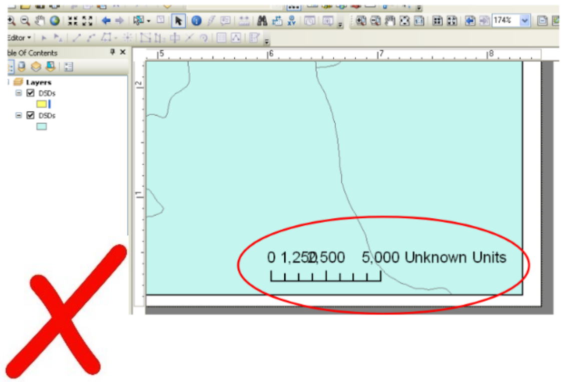
Maximum utilization of map space



Check the units of the scale bar.



Mention the correct sources.



Check the text given in the legend.

Figure 72 -Incorrect formats of maps

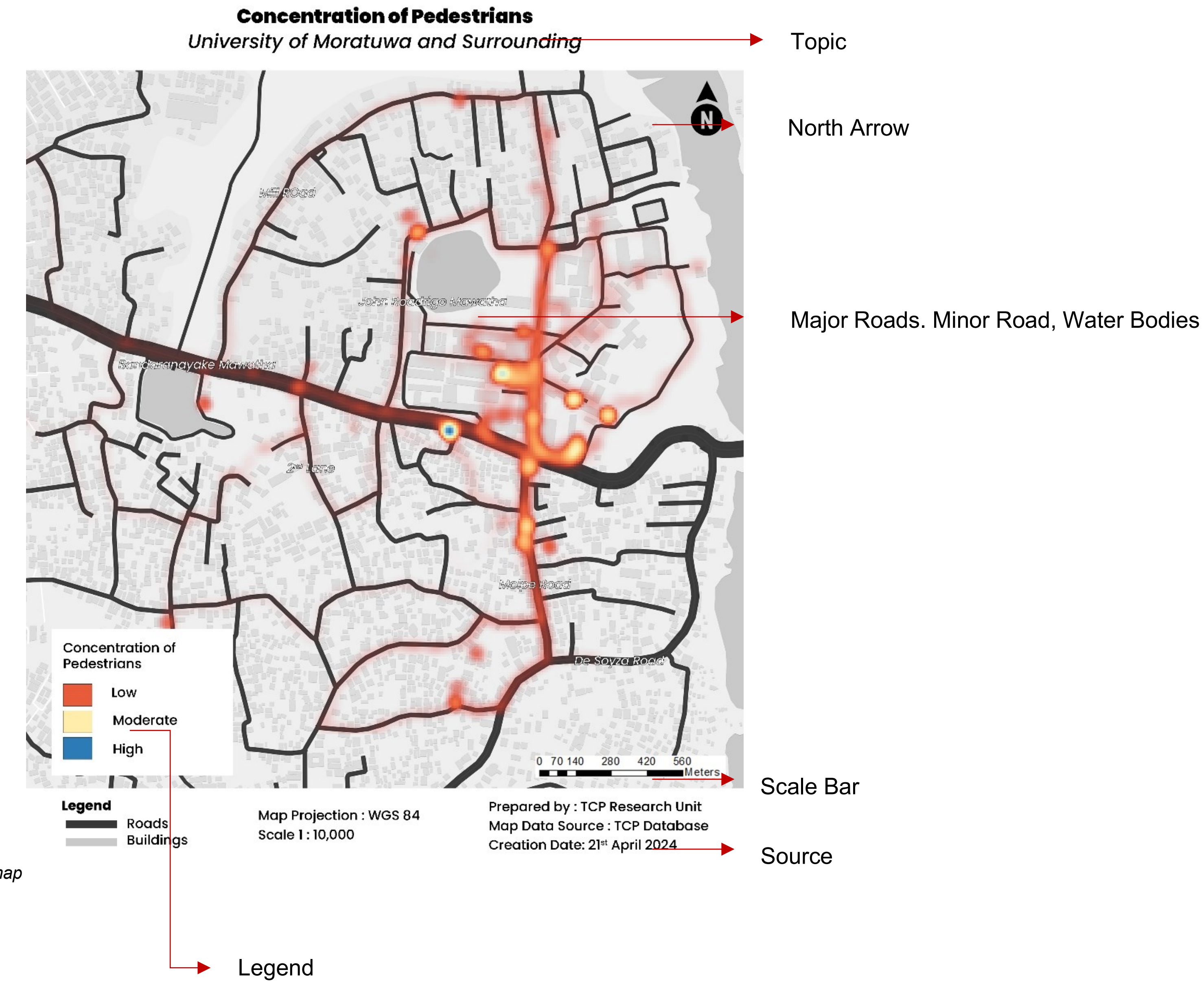


Figure 73 -Correct format of the map

AFTERWORD

As cities strive to become more walkable and human-centered, understanding pedestrian movement through data-driven methods becomes increasingly important. This guide equips readers with practical tools to capture and analyze walking patterns using accessible technologies like smartphones and GIS.

We hope this book inspires students, educators, and professionals to embrace spatial thinking and apply these insights toward creating healthier, more sustainable urban environments. Let this resource serve as a foundation for continuous learning, thoughtful analysis, and impactful planning.

lbs2its.net

ISBN 978-955-9027-93-5

618657-EPP-1-2020-1-AT-EPPKA2-CBHE-JP