

## Lesson 5.5 Intro to basic features<sup>1</sup> of working with data files<sup>2</sup>

---

**Objectives:** You will be able to:

- Translate addresses into latitude/longitude
  - Sort files of data
  - Create subsets of data
  - Read location data from a file and plot points on maps
  - Create bubble plots
- 

**Journal Entry:**

**Consider the data that you have been collecting. How might seeing the data on a map help you analyze it?**

---

**NOTE:** Before you start this lesson, make sure you have installed and can use the data analysis tool (i.e. Google Sheets, MS Excel, or Apple's Numbers) which your 'research group has agreed beforehand to use for data analysis. Remember, to use Google Docs, Drive and related apps on campus, you must be logged into you account before coming on campus.

To prepare, for working on your own question(s), pre-collected data set and collected survey data, you and your 'research group' will complete a lesson that will require you to use all of the tools and complete all of the tasks that will be used in your 'research group's' Unit 5 Final Project.

In this **example lesson**, you will be using data collected in the vicinity of walking and biking routes within the City of Los Angeles California. Take a moment and find the boundaries for the City of Los Angeles and along with the GPS coordinates (latitude, -longitude) for the city's 'City Hall' building.

---

<sup>1</sup> The basic features of loading and saving files, sorting and creating subsets are explored.

<sup>2</sup> v.6 210

Here is a link to a Google Map made by your instructor using Google Maps, Google Earth and an available data source which includes the GPS coordinates outlining LA's City boundaries and the City Hall: <https://goo.gl/xx4LJC>

Next, find on *My Mentor*, under the *Unit 5 Resources*, the handout entitled '*Walking and Biking in LA*'. Open and read this document as an introduction to this lesson.

The 'Walking and Biking in LA' data collection 'campaign' is very similar to what you will be doing within your 'research group' for the final Unit 5 project— those that completed this 'Walking and Biking in LA' data collection 'campaign' project were people like yourself concerned about a topic so they collected data to inform their cause.

If you haven't already, find and load the Google Map using the link in the handout you just read, which contains the locations of the 56 intersections used in their survey [<https://goo.gl/QAMrBT>].

Next, go to <http://www.latlong.net> and figure out how to describe the location of your home, your school and your favorite restaurant using latitude, (-) longitude coordinates. Write your findings in your journal and then share your school and restaurant findings with an elbow partner.

**NOTES** regarding the use of <http://www.latlong.net>: The appropriate coordinate numbers are under the map inside the parenthesis (ex. this would be the coordinate set for Victorville expressed as 'latitude, -longitude) (34.536218, -117.292764) which means the location is 34.536218 north of the equator and 117.292764 west of the Prime Meridian). Why is the longitude coordinate (-) negative? Where in the world would you be if it were (+) positive? . The system of longitude and latitude will allow you to draw spatial objects (such as points houses, schools etc.) if you just think of longitude as the x-direction (it runs along the equator from east to west) and latitude as the y-direction (it runs from the North Pole to the South Pole along the Prime Meridian) when making a plot from these coordinates.

**Note:** Throughout this unit, you will be introduced to and expected to use a variety of strategies as you discover new ways/methods for data analysis. As you learn about and practice using these strategies, expect to provide your instructor with practical ‘demonstrations’ and ‘behavioral predictions’ as you work with other students in ‘driver/navigator’ pairings, or become an expert in a strategy and then share your expertise with others (known as ‘jig-sawing’).

Now, on **My Mentor** under the **Unit 5 Resources**, find and load the **labike data file** (labike.csv) using the data analysis tool that your group has chosen to use. As you look over the data in this file, consider for a moment that this is the format that your survey data will eventually take. Analyze features of the file’s data layout. Next, on **My Mentor**, under the **Unit 5 Resources**, find and open the handout entitled ‘**Exploring LACBC**’ and use this document to better understand the data layout in the **labike.csv** file. Make sure you find and understand the following file elements:

- Header, number of rows, categories
- Note that the file has many rows of data.
- The first line is the header and describes the names of each variable or column.
- Each **row** refers to a different intersection, and so there are 38 intersections represented in the data set.
- Each **column** refers to the various data that were collected for each intersection.
- Navigate through the survey and find the following variables in the data set:
  - **"name"** is the location of the intersection,
  - **"longitude"** is the longitude of the location
  - **"latitude"** is the latitude of the location
  - **"type"** is the type of bike transportation available at the intersection (bike lane, bike path, bike route, none)
  - **"bike\_count\_pm"** is the evening count of bike
  - **"ped\_count\_pm"** is the evening count of pedestrians

## NOW... For Your Consideration & Practice...

- How would you go about obtaining a table of frequencies for given transportation type?
  - Note that 20 of the intersections have nothing.
  - There will be more discussion of frequencies later in the unit, but note that this file is small enough that the counting could be done by hand, but that later data sets will be much larger.
- How would you go about doing a 'sort' by bike count and pedestrian count?
  - Which intersections have the most bike traffic/pedestrian traffic?
  - Are they the same?
- How would you create a subset of locations where the bike count is greater than or equal to the pedestrian count?
- What other subsets might be interesting to create? (i.e. locations with bike routes).
- Now, create a few subsets of your own and make a short list of questions you might want to ask about the subsets you created.
- How would you go about plotting the intersections on a map (including a title, axes and background).
- What are the various sizes and shapes of points and can they be 'zoomed in on'.
- Before you plot a feature, ask yourself 'Why am I adding that feature?'
  - Are there any outliers?
  - Are there 'clusters' of points?
  - Does the plot match the table?
  - The key here is that you 'reflect' on your analysis and justify your conclusions using the data.

---

## Activity: LA Bike Activity

On **My Mentor** under the **Unit 5 Resources**, find and open the LA Bike Activity. This project is to be completed on your own as an individual. After you have completed this activity gather with other members in your 'research group' and share your findings/responses with your group members. After 'debriefing' your findings with your 'research group' your instructor will lead a discussion to ensure that everyone has had an opportunity to become fluent in the use of each of the features of the data analysis tools thus far.

---

## Activity: LA Bus Stops Activity

**My Mentor** under the **Unit 5 Resources**, find and open the LA Bus Stops Activity. This project is to be completed on your own as an individual. After you have completed this activity gather with other members in your 'research group' and share your findings/responses with your group members. After 'debriefing' your findings with your 'research group' your instructor will lead a discussion regarding your analysis and the patterns that have been discovered and described thus far. In particular, expect to share and discuss your answers to questions 10-13.

---

## Discussion: Making Bubble Charts with LA Bike Count Data

What are bubble charts? On **My Mentor** under the **Unit 5 Resources**, find and open the handout on **Bubble Charts**.

1. What is being described when longitude and latitude is plotted on a map?
2. Is there a way to distinguish counts of pedestrians and bicyclists?
  - **Demo:** How to create a bubble chart with the pedestrian counts in the *labike file*.
  - **Demo:** How to change the size and color of the bubbles.

---

## **Activity: Bubble Chart Activity**

On *My Mentor* under the *Unit 5 Resources*, find and open the *Bubble Chart Activity*. This project is to be completed on your own as an individual. After you have completed this activity gather with the other members in your 'research group' and share your findings/responses with your group members.

---