**Lab 1: Introduction to Excel**

The goal of today’s class is to get more comfortable working with Excel and Python and manipulating data. There are other statistical software programs—such as SAS, STATA, and R—that can do the same things (and sometimes more quickly); still, Excel remains a staple of most planning offices. It is also often used for accounting and budgeting, so becoming familiar with looking at and manipulating data in Excel is an important professional planning skill. In addition, it can be useful to provide or present information in Excel if that better matches your audiences’ or coworkers’ skill level.

We will cover the following topics to help you get from the spreadsheet to these outputs:

• Navigating the Environment

• Basic Functions

**Part 2: Working with Census Data**

**Please download the file labeled “Lab1Data.xls” from Quercus and open the file**. Note that the data in the first two sheets has been sourced from the Canada Census using Statistics Canada. The data is from drawn from the 2021 Canada Census of Population, which provides statistical information about the population, age, sex, type of dwelling, household status, income, education, labor and more measured in the Census Programs.

**Exercise 1: Understanding What Options Are Available**

Try playing around with the “Lab\_1\_Data” worksheet – can you make the  
columns wider? Wrap the text within a cell to be able to see the whole variable  
description? See if you can sort the data so you can see which census tract in Toronto has the largest population. Can you make the text bigger?

Just as with Word, you can cut, copy, paste individual data, or a whole column or row (or multiple columns and rows at once).

See if you can delete all the columns having to do with households and dwellings. Uh oh! Now that data is gone. The “undo” button is one of the world’s greatest inventions – restore the columns and breathe a sigh of relief! See if you can cut and paste a row or column in a new location.

If you have the hang of it, think about using a shortcut for selecting columns and rows (PC: ctrl + arrow direction; Mac: shift + cmd + arrow direction).

**How many rows and columns are in the data set? Can you identify which census tract in Toronto has the largest population? Hint: Use the command, sort\_values(…).**

**Exercise 2: Basic Functions**

Let’s start analyzing our data. Make sure you are on the sheet that is labeled “Lab\_1\_Data.” **Calculate total population, total number of dwellings, and the total renters (the column is named v\_CA21\_4239: Renter) and owners (the column is named v\_CA21\_4238: Owner) of private households in Toronto.**

**Exercise:** Start by scrolling to the bottom of the worksheet, and adding a new row labeled “Census Subdivision Totals”. Darn – now you can’t see the variable labels along the top of the worksheet! Select cell A2, go to View (Mac: Window) – “Freeze Panes” and “Freeze” the top row. You can also freeze the left-most column by selecting cell B1 and pressing “Freeze panes”. To select both your first row and your left-most column, so that you can always see the labels associated with your data, select cell B2 and “Freeze panes.” This will work by freezing the row above your cell selection (Row 1, in this case) and the column to the left of your cell selection (Column A, in this case). Now, see if you can input a “formula” that calculates the total population for values in a selected column. Next, go to the “Formulas” menu and select “AutoSum”. Press Enter. Voilá!

**First,** did you get the same total? Then, look at the formula that Excel inserts. The important thing to understand is that the “=” sign in a cell tells Excel that you are going to give it instructions about what goes in that cell, rather than an actual number or text. The equation after the equal sign tells Excel to do an operation based on the Column and Row identification. “ =SUM(B2:C199)” tells Excel to Sum all the cells in column B, starting with the cell in row 2 and going all the way to row 199. You can always go in and change those reference cells. Try it, and see what happens!

**Second,** calculate the totals for the other columns. Try dragging the little black square in the first “total” cell over to the right. What does Excel put in each of the bottom cells? Excel tries to guess what you want to copy or duplicate, which is a great shortcut, but you have to be careful to check that that its guess is correct. If you ever want to “freeze” either a column or row reference, you need to put a dollar sign in front. Try typing “=SUM($B$2:$B$198)” in the cell and dragging that over – it should stay exactly the same.

**Third,** calculate the percent renters in Toronto. (Total Rental Units/ Total Private Households by Tenure\*100). First click in the cell where you want you answer simply as a ratio: Total Rental Units/Total Private Households by Tenure. Make the formula for this calculation using values from your Census Subdivision Totals. Once you have a decimal value, click your cell and notice that the cell type. It should read “General.” If you change the type to “Percentage,” what happens? You can alternatively get this Percentage manually by making your formula to multiply the ratio value by 100 (Total Rental Units/Total Private Households by Tenure\*100). Give this a try.

**Finally,** create a new worksheet called “summary\_table” and then cut and paste your totals into the new worksheet to start building a summary table. What went wrong? It is important to remember the difference between actual number values and “instructions” that you give Excel on what to put in a cell. When you copy the instructions, the “reference” disappears, giving you the dreaded “#REF!”. One solution is to copy and paste number “values” (“Paste Special” > “Values”). If you’re more advanced, see if you can figure out how to reference cells in another worksheet.

**Part 2: Pause for two Common Issues**

Click on the Part\_2 worksheet.

1. *Summarizing numbers* – Example 1 has two sets of numbers and a calculated total. Uh‐oh the totals are not the same. Why do you think that is? When you are adding numbers, make sure you understand what you are really adding.

2. *Format Issues* – In Excel, you will sometimes run into problems where your functions don’t work. This issue very frequently be related to how your data is formatted. If a cell is recorded as “Text” it will not be included when you conduct functions. What cell is the problem in Example 2?

**Exercise 3: Common Functions and Calculations**

The Part\_3 sheet, or lemonade\_sales.csv, includes a table summarizing sales at my lemonade stand made in the past two weeks. Let’s generate some summary statistics for my business using functions and basic calculations.

Note that there are no “standard” prices for my lemonade (which is why sales counts do not correspond to profits in any typical way). I’m wild.

**Calculate each of the summary statistics in bold at the bottom of the table**. You’ll want to think about what you need to calculate, and then become familiar with some of the most used functions: SUM(), AVERAGE(), MAX(), MIN(). You should also become familiar with syntax for addition (+), subtraction (-), division (/), and multiplication (\*).

What are the total sales throughout the day (AM and PM)? What are the total profits throughout the day?

What are the percentage of sales made in the morning? What about in the evening? What are the percentage of profits made in the morning? What about in the evening? Can you calculate the averages of the morning sales, morning profits, evening sales, and evening profits for week 1 and week 2 of the lemonade sale?

Now, instead of weekly averages, can you calculate the daily total sales and profits for each day that the lemonade stand was open? What are the daily average sales and profit?

What are the minimum sales in a day? What is the minimum profit in a day? What about maximum sales and profit?

Note that the Maximum and Minimum Sales in a Day calculations are extra tricky and require using data from multiple columns.