Jayden Ma (C1)

Andrei Gabor (C2)

EK125 Final Project

Ma\_Jayden\_Gabor\_Andrei\_Project

Professor Attaway

1 May 2019

\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_

**Background**

Influenza and pneumonia are two common illnesses that can, if left untreated, claim many lives. While they are so common that we don’t generally see influenza and pneumonia as a serious threat that some Tylenol can’t solve, they can be dangerous to those who are vulnerable, such as children, elders, and people with compromised immune systems. Because of this, it should be important to note places where these two illnesses occur the most if one is vulnerable.

This data set provides 122 cities across the United States, logging deaths from influenza and pneumonia from 1962 to 2016, on a weekly interval. It also records the total deaths of people in those cities. Our team will be analyzing if there is a prevalence in influenza and pneumonia deaths in any of the cities, by graphing the data set out. Through this, we will be able to see which cities have a higher danger of someone catching these illnesses and potentially dying. The users can then figure out which cities to avoid if they are in risk of catching illnesses, so this project is important to us because we can potentially save the lives of people from influenza and pneumonia.

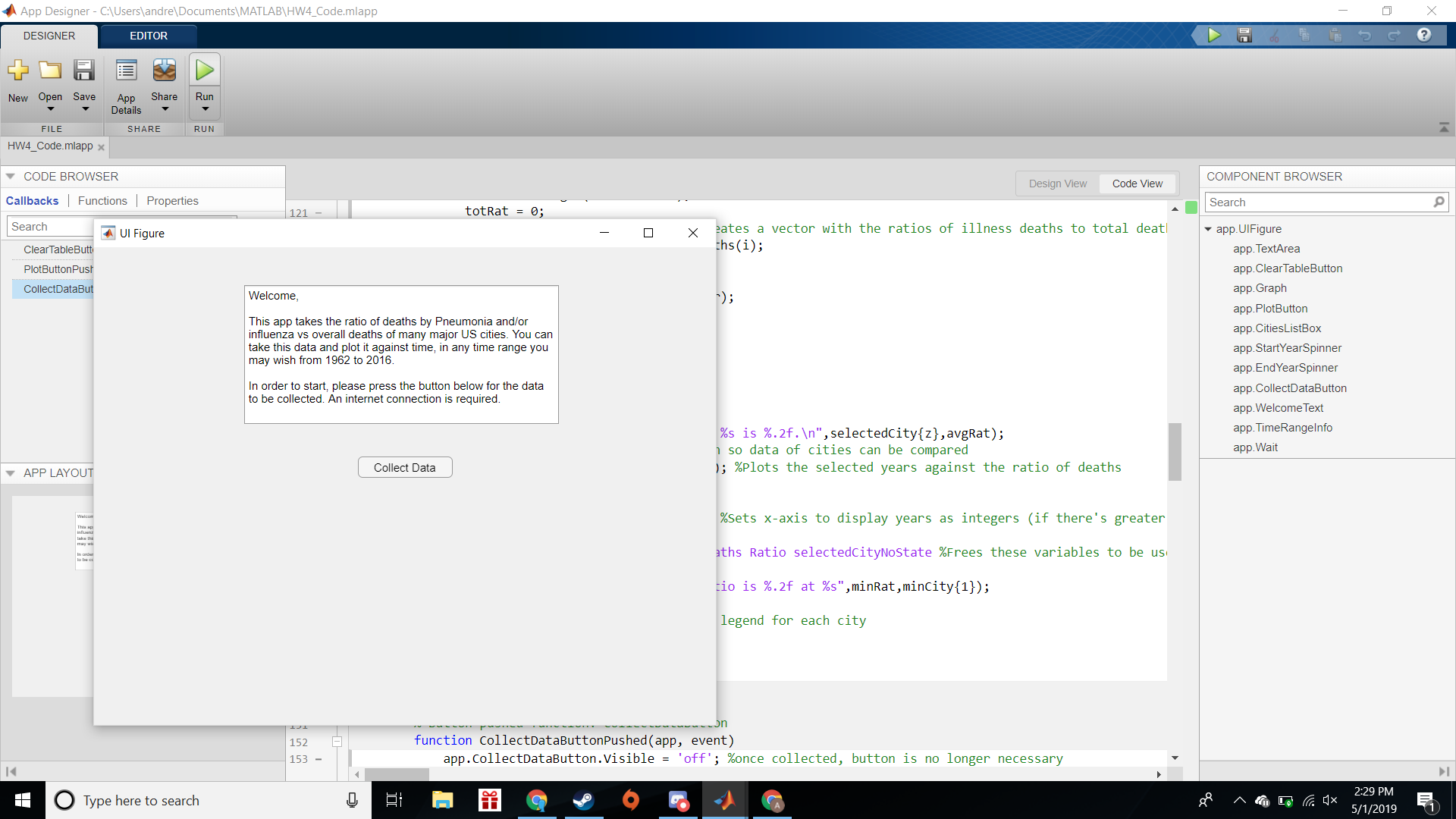
We made an interface that is easy to use, where users can select from 50 major cities and the time interval that they would like to observe, and be able to graph the influenza and pneumonia deaths against the total deaths. The time interval is yearly and can span from 1962 to 2016, where the user can choose the starting year and the ending year to get a range of every year between. Hopefully with this, people can avoid contracting illnesses that may be lethal to them.

**Objectives**

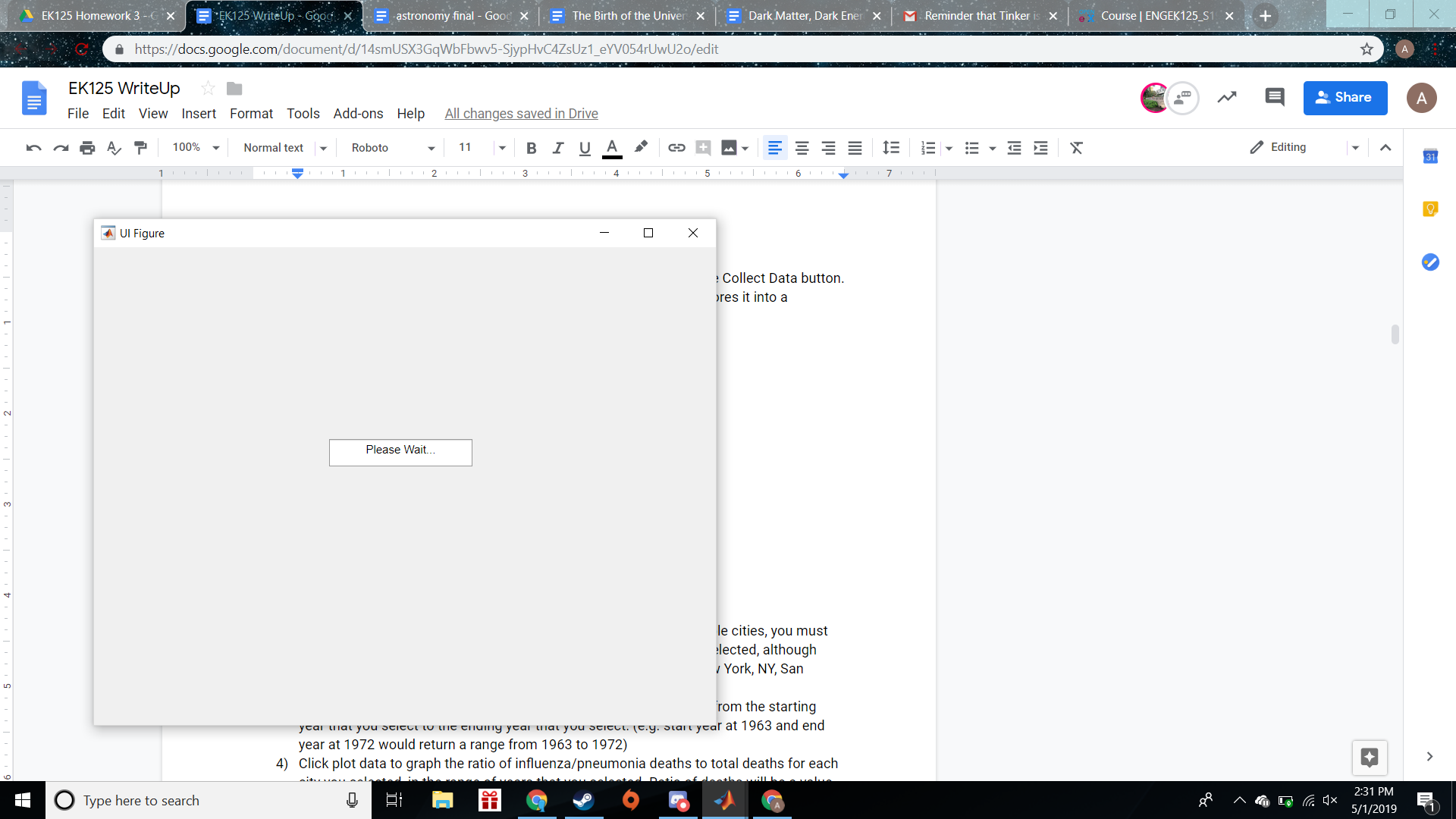
1. Reading into and analyzing a data set: Lines 157-158 take the data from the online database once the button “Collect Data” is pressed. The code accesses the dataset for through an API directly from the Internet. The dataset contains over 362764 data points with weekly data of the deaths that occurred in those cities. The deaths are split between deaths by influenza/pneumonia and total deaths. This dataset takes approximately 2 minutes to load into MATLAB, and is used for the rest of the code.
2. Plotting the data: Lines 66-146: The user is allowed to choose which cities they want and what year range they want as well. The code iterates through every year that matches each cities, and adds up the weekly deaths for each year. The sum of the total deaths is divided by the sum of the deaths for influenza/pneumonia, returning a ratio of deaths by influenza/pneumonia over total deaths. This data is then plotted, where the years are in the x-axis, the ratio is the y-axis from 0 to 1 (where 0 would be 0 deaths by influenza/pneumonia in that year and 1 would be 100% deaths by influenza/pneumonia in that, neither of which are possible) This is only completed when the user presses the button “Plot Data”.
3. Printing into a text file: Lines 67,69, 126-134, and 143 -144: Once the user hits plot, the code will open a file called project.txt. The program will then write the average ratio of each city, the name of those cities, and the minimum ratio of all the cities selected in the time range and the city in which that ratio is. The text written in the file gets rewritten every single time the user presses the plot button. Specifically, line 67 opens the file, and lines 134 and 143 write the data to the file. Line 144 closes the file.

**Instructions:**

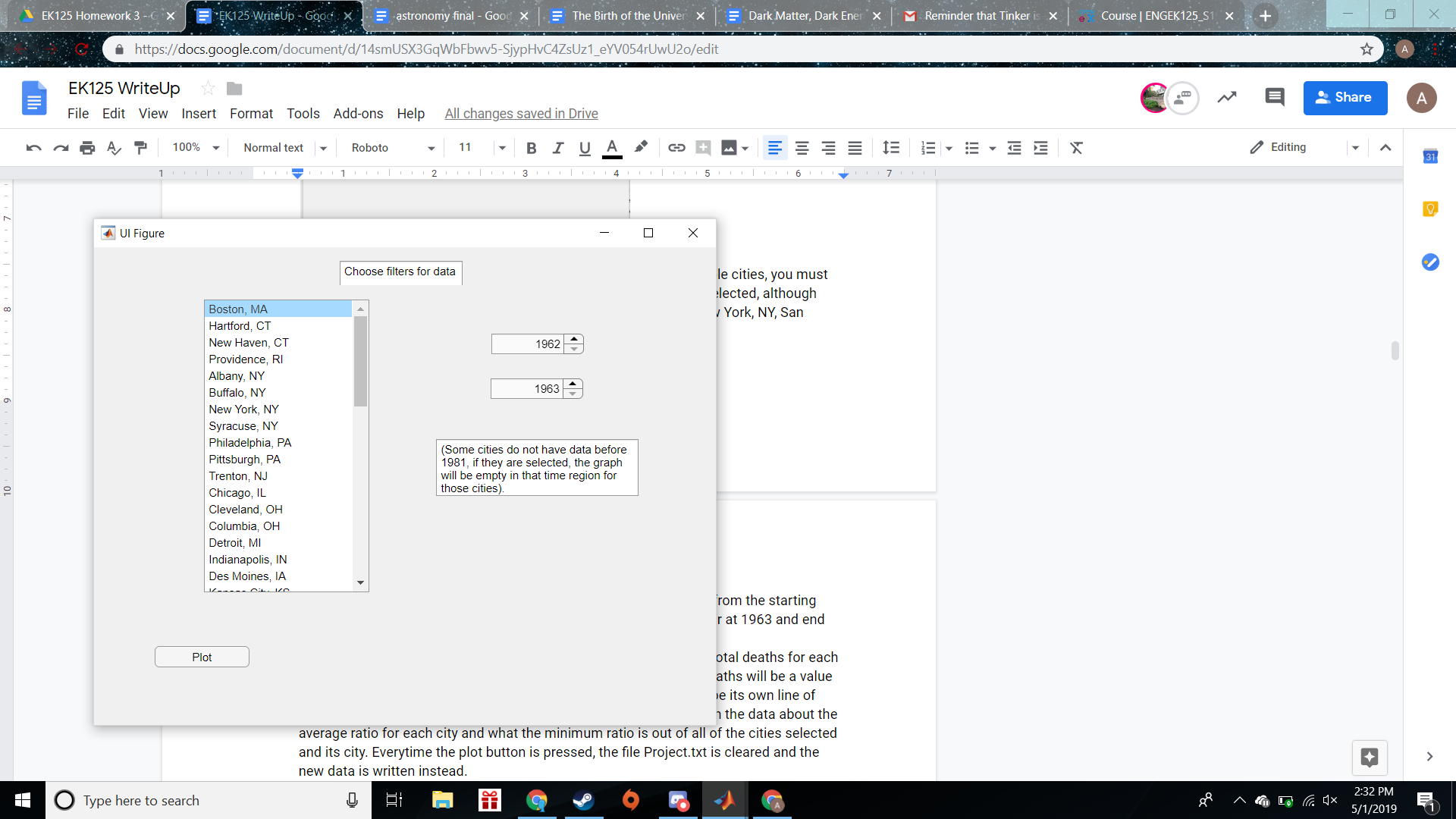
1. First the Welcome screen comes up, and the user needs to press the Collect Data button. This button grabs the data from the database on the internet and stores it into a structure the code later uses.



(after clicking wait screen appears)

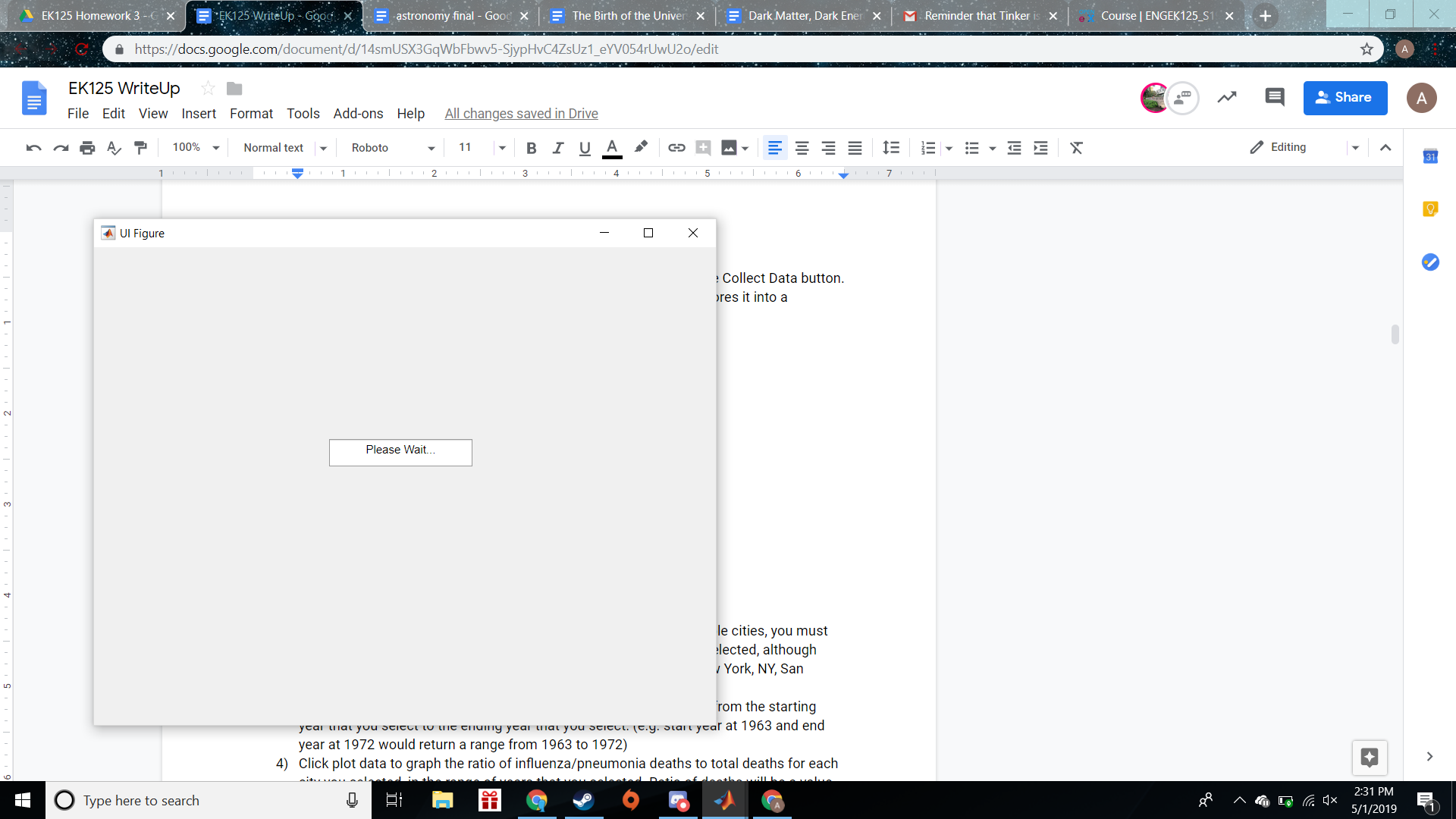


1. Select the cities that you want to plot under “Cities.” To select multiple cities, you must shift-click the next selections. As many cities as you want may be selected, although having too many may make the graph unclear. (e.g. Boston, MA, New York, NY, San Diego, CA)

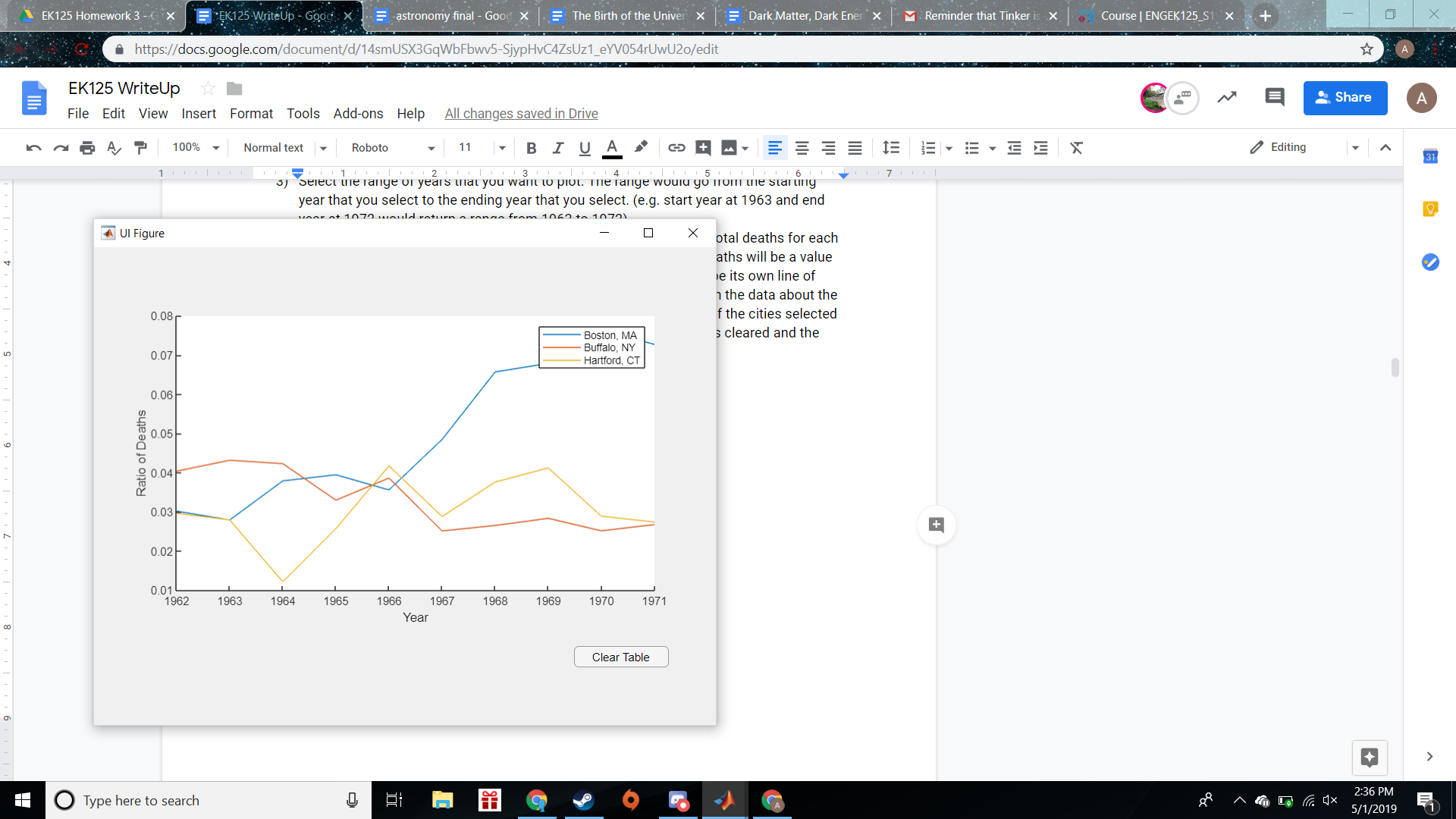


1. Select the range of years that you want to plot. The range would go from the starting year that you select to the ending year that you select. (e.g. start year at 1963 and end year at 1972 would return a range from 1963 to 1972)
2. Click plot data to graph the ratio of influenza/pneumonia deaths to total deaths for each city you selected, in the range of years that you selected. Ratio of deaths will be a value from 0 to 1 in the y-axis, the years is in the x-axis, and each city will be its own line of data with the legend. At the same time there will be a file written with the data about the average ratio for each city and what the minimum ratio is out of all of the cities selected and its city. Everytime the plot button is pressed, the file Project.txt is cleared and the new data is written instead.

(Another wait screen appears after the Plot button is pressed)



(Then the plot appears, below is an example)



1. Click Clear Graph to clear the graph and return to the select cities and Years page. You can then select a set of cities and a new range of years, and the data set will be loaded already.

