Python Basics

Dr. Granger is interested in studying the relationship between the length of house-elves' ears and aspects of their DNA. This research is part of a larger project attempting to understand why house-elves possess such powerful magic. She has obtained DNA samples and ear measurements from a small group of house-elves to conduct a preliminary analysis (prior to submitting a grant application to the Ministry of Magic) and she would like you to conduct the analysis for her (she might know everything there is to know about magic, but she sure doesn't know much about computers). She has placed the file on the web for you to download.
You might be able to do this analysis by hand in Excel, but counting all of those bases would be a lot of work, and besides, Dr. Granger seems to always get funded, which means that you'll be doing this again soon with a much larger dataset. So, you decide to write a script so that it will be easy to do the analysis again.
Write a Python script that:

1. Imports the data into a data structure of your choice
2. Loops over the rows in the dataset
3. For each row in the dataset checks to see if the ear length is large (>10 cm) or small (<=10 cm) and determines the GC-content of the DNA sequence (i.e., the percentage of bases that are either G or C)
4. Stores this information in a table where the first column has the ID for the individual, the second column contains the string 'large' or the string 'small' depending on the size of the individuals ears, and the third column contains the GC content of the DNA sequence.
5. Prints the average GC-content for both large-eared elves and small-eared elves to the screen.
6. Exports the table of individual level GC values to a CSV (comma delimited text) file titled `grangers_analysis.csv`.

This code should use functions to break the code up into manageable pieces. For example, here's a function for importing the data from the web:

```
def get_data_from_web(url):
    webpage = urllib.urlopen(url)
    datareader = csv.reader(webpage)
    data = []
    for row in datareader:
        data.append(row)
    return data
```

This function imports the data as a list of lists. Another good option would be to use either a Pandas data frame or a Numpy array. An example function using Pandas looks like:

```
def get_data_from_web(url):
    data = pd.read_csv(url)
     return data
```

Throughout the assignment feel free to use whatever data structures you prefer. Ask your instructor if you have questions about the best choices.