

Letter for future DAT-102 student

Experience in gathering data:

I highly recommend student to use the python to gather the data from the website instead of manually doing sampling. There is a tool in python called Beautiful Soup that could help you to retrieve the whole data that you want. It is much convenient for students who want to gather a bunch of data. For example, you could get more than 1000 data in one hour by using python.

Likely Sources of error:

Sampling with 100 restaurant's health score in Yelp might not completely concluding the whole health condition in the whole city. However, it could estimate the health condition in general. As the numbers of investigation increases, the conclusions of the data will become more accurate.

Revision:

I will gather the data from another website and compare it with data gathered from Yelp to make sure the accuracy.