**World Quant University**

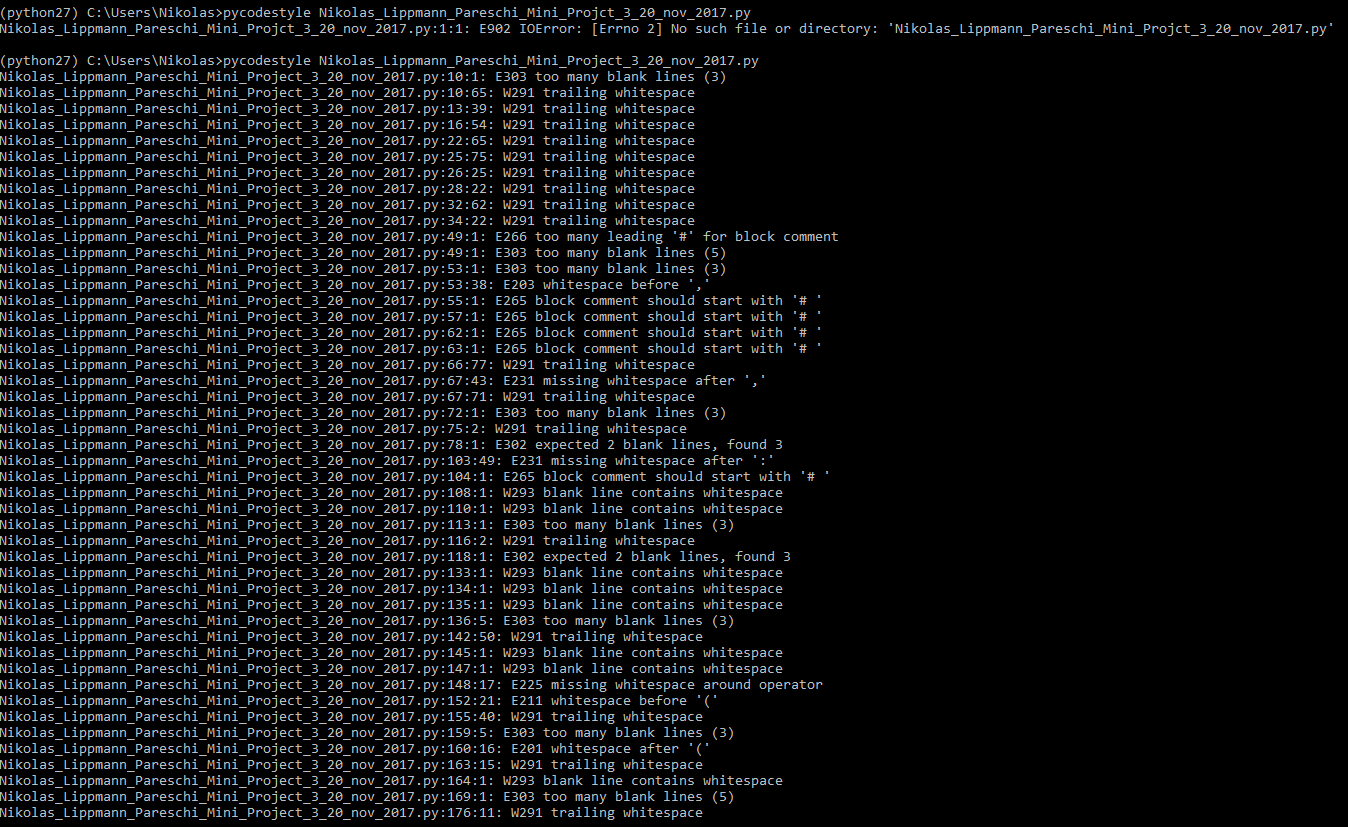
**Professor: Ritabrata Bhattacharyya**

**Python II**

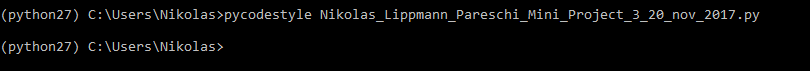
Nikolas Lippmann Pareschi - [nikolaslippmann@gmail.com](mailto:nikolaslippmann@gmail.com)

**Project 3:  USDA Food Database-Analyzing Nutrient Information**

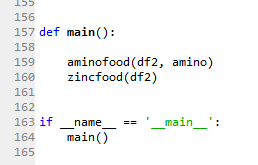
Introduction: I receive the feedback in my previous submission to use more modular code, to use a main method and to stick to PEP8 guidelines. I downloaded a PIP8 check and executed it from the prompt. In fact, I had to adjust several things:



After cleaning it:

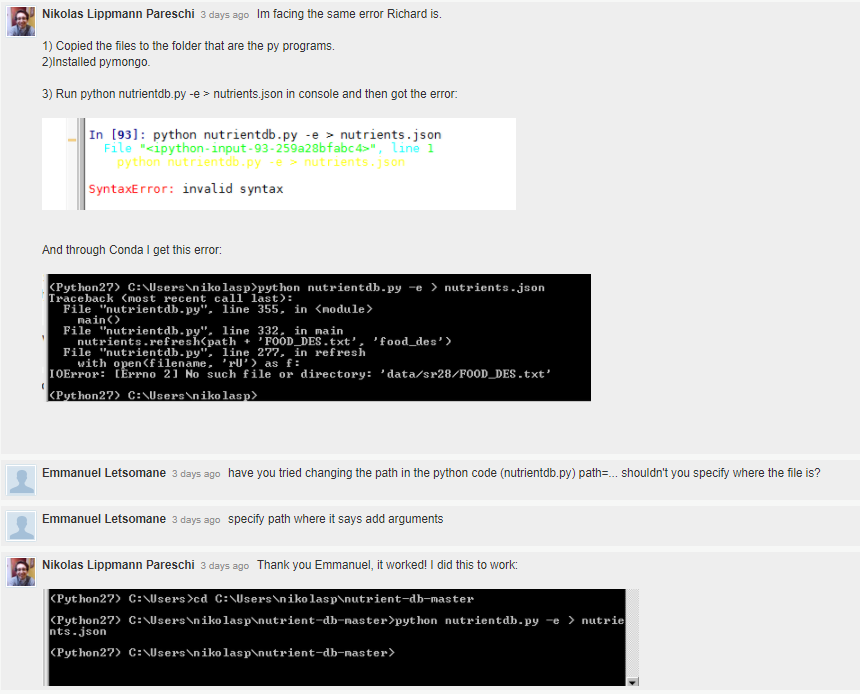


Regarding the modular code I use functions this time in the program and called the main method at the end:



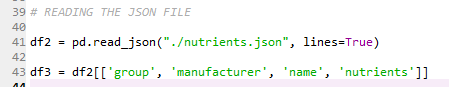
1. Check out the nutrient-db python utility from GitHub from <https://github.com/schirinos/nutrient-db.git>
2. Run the main program with python nutrientdb.py -e > nutrients.json to convert the USDA data to JSON format. For further details, check <https://github.com/schirinos/nutrient-db>. You might have to install the python utility for MongoDB interface via pip install pymongo

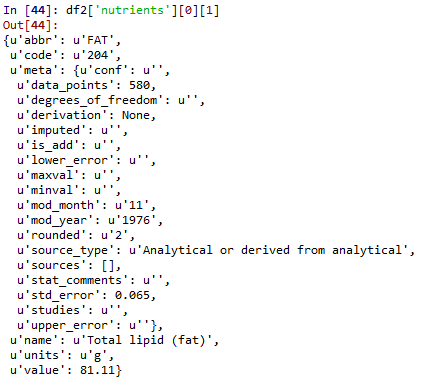
Steps 2 and 3 were done accordingly to piazza:



3.   Load the JSON dataset into Pandas dataframe using the built in python json class. Extract values of the following fields in to the dataframe - food names, group, id, and manufacturer

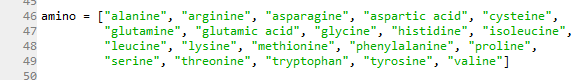
After exctracting the Jason we selected those columns. We got the nutrients because it contains several information, such as the ID (‘code’)



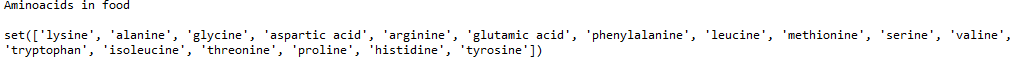


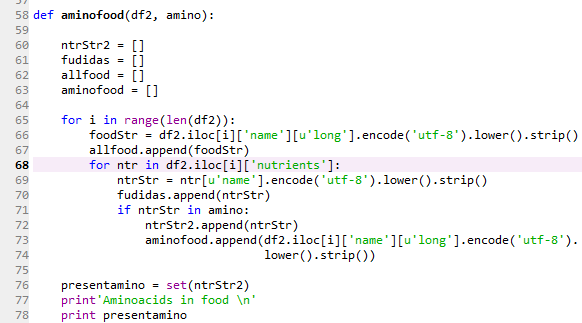
4. For the ‘Amino Acids’ nutrient group output a table showing the different constituents of the group (Alanine, Glycine, Histidine etc) and the foods in which they are present (Gelatins, dry powder, beluga, meat...etc)

We hard coded the following all the existent amino acids:

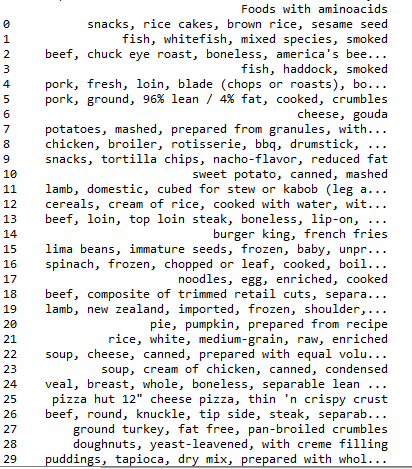


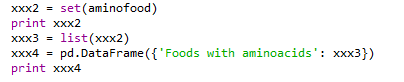
We checked to see which ones were present in the food listed:



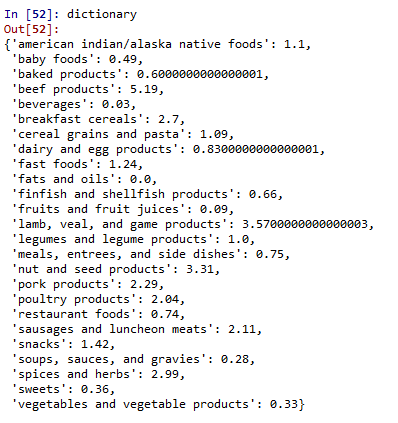


Foods in which the Aminos are present (it had several more items, it is truncated):

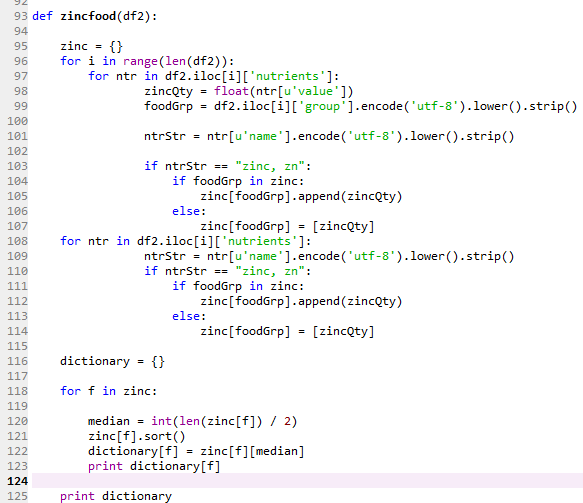




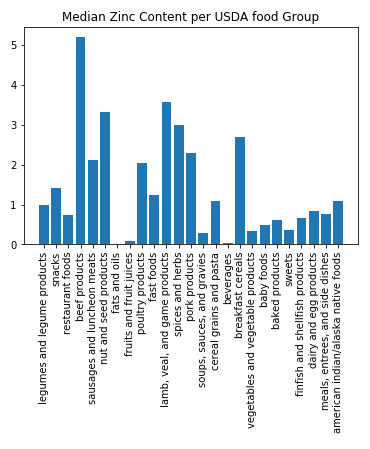
5.   For all the different nutrient group (beef Products, Pork Products, dairy and egg products etc.) calculate the median Zinc content (median of the zinc content in all the foods that constitute the nutrient group)



Which was done with the code below:



6.   Plot the distribution of median Zinc Content for different nutrient groups as a bar chart.



Code:

