1. In order to print the data sorted by product, you likely had to load all the data in memory. Imagine the csv file is over 1GB in size, 10GB, 1TB. What is the problem?

**When loading data of larger and larger sizes from a csv the biggest issue is your computer's volatile memory size. When dealing with data beyond the size that you can store in RAM you have to process data “slices” and once all of the slices are processed it will print and indicate that it’s processes are complete.**

2. How easy is it to change your code to sort by the Payment\_Type instead? How about allowing the user to determine to sort by any of the following: Product, Payment\_Type, Name, Country, or State? What about sorting by Product-Country-Name? This is not a request to make  
those changes, just describe the effort.

**My code already allows this. By giving the user a list of columns to choose from, I was able to pass that into my sort function for the desired result.**

3. What was the challenge with counting the number of users named “Amanda”? What issues did you have to handle?

**My biggest issue was making sure that my comparison was caseinsensitive and that it checked all substrings.**

4. Was there the problem calculating the average transaction amount? Although the data appears numeric, what did your program have to do to perform the necessary computation?

**I had to mutate the datatype of the price column to calculate the mean. Without doing so it just concatenated all the strings together then crashed.**

5. Changing "United States" to "USA" is one change. How many records in the data had to change?

**31 records for the first data set, 24 for the second, and 26 for the last.**

6. With the second CSV file, re-running your program caused some problems. What broke, and why? Why does your program have to change in order to work properly?

**Thanks to the formatting ability of the python tabulate package the table resized correctly and did not have any issues.**