**Assignment 1 Part 1**

**Data and Pivot Tables**

Analytical Methods Fall 2015

Professor César Zamudio

Download the dataset "Assignment 2 Dataset - Final.xls". This is a **real dataset** from a distribution company overseas. The names, locations and products in the dataset were masked for this reason.

To solve the following questions you will use Excel PivotTables. You will need to learn to use these tables. You may use sources such as these:

**Excel 2003:**

<http://www.edferrero.com/Pivot200301.aspx>

**Excel 2007:**

http://www.timeatlas.com/5\_Minute\_Tips/Chunkers/Learn\_to\_use\_Pivot\_Tables\_in\_Excel\_2007\_to\_Organize\_Data

**Excel 2010:**

http://www.gcflearnfree.org/excel2010/20

**Excel for Mac:**

* http://www.techonthenet.com/excel/pivottbls/create2011.php
* http://faculty.chicagobooth.edu/drew.creal/teaching/basicCourseMaterial/tutorials/MAC%20Tutorial%205%20-%20How%20to%20Create%20a%20Pivot%20Table.pdf
* http://mac2.microsoft.com/help/office/14/en-us/excel/item/df48adb7-b540-4bb0-aced-56160ecb6b77
* http://www.dummies.com/how-to/content/how-to-make-a-pivottable-in-excel-for-mac-2011.html

Deliver answers to the following questions in an Excel file. Organize your solution in various Sheets – for example, Question 1 in Sheet 1, Questions 2 and 3 in Sheet 2, etc. Feel free to organize in whichever way you deem best. Name Sheets appropriately!

**Note: Throughout the exercise, the important part is to be able to *extract the right data* from the dataset. If you find difficulties in plotting in PivotTables, you could copy the summarized data from the PivotTable to another Excel sheet and plot it there.**

*Top customers*

**Question 1.** The company is interested in knowing which of their customers are more valuable. Report the top 10 customers in terms of dollar sales...

* Overall
* By state
* By state and city (only Texas)

**Question 2.** What are the purchasing patterns of these customers? For the best 10 customers overall in terms of dollar sales, report their sales volume *by month.* You must show client names in rows and months in columns, producing a matrix (*Hint: you have to* ***group*** *the dates).*

**Question 3.** Plot overall sales by month. Given the sales pattern you observe, what kind of product do you think this is? Why do you think the sales spikes are observed?

*Regional analysis*

**Question 4.** Find the top 3 states based on dollar sales.

**Question 5.** Now we want to investigate deeper on which customers are the best, by city. Produce a table with the **top 5** clients based on dollar sales of every city within the top 3 states found in Question 4.

**Question 6.** Given the result of Question 6, what type of clients does the firm *seem* to be primarily serving in each **state**? Confirm this by adding the variable "Client type" to the analysis in Question 5. You can plot the proportion of clients by type in a stacked bar if you want. Also, suggest what type of Marketing efforts should be conducted (in general) in each of the top 3 states. Should the firm conduct the same type of effort in each state?

*Other questions*

**Question 7:** Which products are sold the most overall?

**Question 8:** Observing the pattern in Question 9, are some products being bought together? Why do you think this is the case?

**Question 9:** How many customers **repurchased**, overall? That is, which consumers bought from the company more than once? (For example, if Jane purchased *any* number of products in January, and nothing thereafter, she did not repurchase; if Joe purchases 1 product in January and 1 in February, he repurchased).

Hint: You want incidence in a month, not the number of times the client purchased in a given month. So if Jane purchases 10 times in Jan, you should count her as "purchased in Jan". You may need to move the data to a different Excel file to produce this data.

**Hint 2: Use conditional statements (IF(…))**

**Question 10:** Plot the monthly sales of repurchasers and the monthly sales of purchasers in a *single graph.* Are the patterns different?