SPSS Worksheet : Create a file SPSS_Worksheet_PUNetID.docx to contain your answers. Label each answer clearly. Place this worksheet in the CS130Drop box by March 29, 4:45pm.

Make sure you save your files to Turing!

1) Consider the following grade chart.

Name	ID#	Quiz1	Quiz2	Midterm	Final
Adams	0001	14	23	82	76
James	0002	12	21	76	68
Jones	0003	15	24	91	93
Mann	0004	14	19	88	73
Smith	0005	11	16	79	71
Tolls	0006	10	13	62	65
Wells	0007	5	10	43	55
Max Points	9999	15	25	100	100

You are working for a professor who has given you her grades. She wants you to compute the Average and Grades for each student and then produce a Bar Chart that she can display to the class. She has kept the grades in a table of a Word document. <u>Without typing in all of the information all over</u>, you are to copy the information from the table into an Excel document. Then compute the Average and Grade for each student in Excel. Since you like the charting in SPSS, import the Excel file into SPSS and create a properly labeled Bar Chart. In a Word document display each of the following:

a) All grade information including Average and Grade (letter grade) in a table (Copy and pasted from Excel)

b) The Bar Chart using SPSS of all A's, B's, C's, D's, and F's in the class.

2) The datafile CandyBars.xls exists in the CS130 Public\Datafiles folder.

a) Create a Scatterplot of the data Carbohydrates and Sugars.

b) Perform a linear regression using SPSS.

c) What is the regression equation?

3) Rework the AIDS problem using SPSS. Again, do not type in the information. Instead, find a way to import the data into SPSS and rework the problem. Paste in your results into a Word document.

Let us consider the following data which represents the number of deaths, N, from AIDS in the United States from 1981 to 1996, where t denotes the number of years after 1980.

t	N
1	159
2	622
3	2130
4	5635
5	12607
6	24717
7	41129
8	62248
9	90039
10	121577
11	158193
12	199287
13	243923
14	292586
15	340957
16	375904

After typing the data into an Excel spreadsheet, try to fit different types of nonlinear functions to the data. Which works the best? How do we know?

4) Let's work with some sports statistics.
The Pythagorean Theorem in the world of sports statistics is used to predict the number of wins a team should have based on the number of points that the team has scored and the number of points that team's opponents have scored against that team This algorithm was first applied to baseball and recently has been applied to football.

The website **http://waynewinston.com/wordpress/?page_id=13** contains a number of Excel files ("Mathletics book file") that apply various analyses to sports statistics. The file **Footballbasketballpythagoras.xls** contains the *predicted* winning percentage and *actual* winning percentage for two NFL seasons.   Our goal with this assignment is to determine the accuracy of the Pythagorean Football theorem.

Download the file **mathleticsjune09.zip** from the web page above. Unzip the file and copy **Footballbasketballpythagoras.xls** to your desktop. Rename this file **Football_PUNetID.xls**.   Open the file and examine the worksheet **NFL**. Note which row the table starts and ends in, as well as which column the table starts and ends in. What formula is used to predict the number of wins?

We want to import this file into SPSS. Close Excel and use the File | Open dialog to open the Football_PUNetID file. Rather than import the entire file, we want to import only the rows and columns in the table. In the import dialog box, list the range of cells to import (specify the top left of the table and the bottom right, like: C2:G100).

Once the data is imported, check the type and measure. Make any changes necessary. Save this SPSS file as **Football_PUNetID.sav**.

Run a linear regression to determine how closely "Predicted win Percentage" and "Actual win Percentage" are related. You will need to first build a scatter plot.

What is the dependent variable? Why? Give a detailed explanation.

What is the independent variable? Why? Give a detailed explanation.

What is the R2 value?
What is the equation of the regression line?

In your opinion, how well does this theorem predict actual wins? Justify this with your results above.

Open the Football_PUNetID Excel file again and run the same regression as you did in SPSS. Copy the scatter plot to the Word document under the heading Problem#2-Question5. Be sure the scatter plot contains the R2 value and the equation of the line.

Make sure you save your files to Turing!