MTH 145 SPSS Assignment

Instructions: You can work in pairs or group of 3 for this assignment. For each problem, use SPSS to do the calculations. Perform the instructions and answer the questions in the "**To be** *included in your work*" section. This is worth 45 points and is due on **Monday, May 12**.

I. (Test of Independence) Use SPSS on the data shown in the table below to test the claim that occupation is independent of whether the cause of death was homicide. The table is based on data from the U.S. Department of Labor, Bureau of Labor Statistics. Does any particular occupation appear to be most prone to homicides? If so, which one. This "occupation" SPSS data can be downloaded from my website.

	Police	Cashiers	Taxi Drivers	Guards
Homicide	82	107	70	59
Not Homicide	92	9	29	42

Follow the instructions below once you get the data set into SPSS.

- 1. First you have to tell SPSS which variable represents your cell counts. Do this by
 - Clicking on **Data** → **Weight Cases...**
 - Click in the circle next to Weight cases by
 - Select the column for the cell counts (Count) and click on 🕩 to make it the Frequency Variable
 - Click on **OK**
- 2. The chi-square test of independence can be accomplished by doing the following:
 - Analyze \rightarrow Descriptive Statistics \rightarrow Crosstabs
 - Put Homicide as the Row variable and Occupation as the Column variable.
 - Click on **Statistics** and check the box for **Chi-square**
 - Click **Continue**
 - Click on **Cells** and check boxes next to **Observed** and **Expected Counts**.
 - Click **Continue** and then **OK**

To be included in your work:

- 1. Formulate the appropriate null and alternative hypotheses.
- 2. What assumptions are you making? Did you satisfy these assumptions?
- 3. Copy and paste the relevant SPSS output.
- 4. Based on the SPSS output, write your conclusion in the context of this problem. Explain how you arrived at this conclusion.
- II. Do the same thing for problem 13.55 on page 816. The data set "*lifestyle*" for this problem can be found in the CD that comes with your textbook. You can also download it from my website.

III. (ANOVA) If a supermarket product is offered at a reduced price frequently, do customers expect the price of the product to be lower in the future? This question was examined by researchers in a study conducted on students enrolled in an introductory management course at a large Midwestern university. For 10 weeks 160 subjects received information about the products. The treatment conditions correspond to the number of promotions (1, 3, 5, or 7) that were described during this 10-week period. Students were randomly assigned to four groups. The table below gives the data. This "price_promo" SPSS data can be downloaded from my website.

	3.78	3.82	4.18	4.46	4.31	4.56	4.36	4.54	3.89	4.13
1	3.97	4.38	3.98	3.91	4.34	4.24	4.22	4.32	3.96	4.73
	3.62	4.27	4.79	4.58	4.46	4.18	4.40	4.36	4.37	4.23
	4.06	3.86	4.26	4.33	4.10	3.94	3.97	4.6	4.50	4.00
	4.12	3.91	3.96	4.22	3.88	4.14	4.17	4.07	4.16	4.12
3	3.84	4.01	4.42	4.01	3.84	3.95	4.26	3.95	4.30	4.33
	4.17	3.97	4.32	3.87	3.91	4.21	3.86	4.14	3.93	4.08
	4.07	4.08	3.95	3.92	4.36	4.05	3.96	4.29	3.60	4.11
	3.32	3.86	4.15	3.65	3.71	3.78	3.93	3.73	3.71	4.10
5	3.69	3.83	3.58	4.08	3.99	3.72	4.41	4.12	3.73	3.56
	3.25	3.76	3.56	3.48	3.47	3.58	3.76	3.57	3.87	3.92
	3.39	3.54	3.86	3.77	4.37	3.77	3.81	3.71	3.58	3.69
	3.45	3.64	3.37	3.27	3.58	4.01	3.67	3.74	3.5	3.60
7	3.97	3.57	3.50	3.81	3.55	3.08	3.78	3.86	3.29	3.77
	3.25	3.07	3.21	3.55	3.23	2.97	3.86	3.14	3.43	3.84
	3.65	3.45	3.73	3.12	3.82	3.70	3.46	3.73	3.79	3.94

To conduct the analysis using SPSS, perform the following steps.

- Analyze → Compare Means → One-Way ANOVA
- Put eprice in the Dependent List and promo as the Factor
- Click **Continue** and then **OK**

To be included in your work:

- a. Identify the following:
 - a. Response variable.
 - b. Factor(s) and levels of each factor.
 - c. Experimental units.
- b. Is this a designed experiment or an observational study? Explain.
- c. Formulate the appropriate null and alternative hypotheses.
- d. Construct a QQplot for the data in each of the four treatment groups. Copy and paste these qqplots into your document. Reduce their sizes so you can fit all four in one page. Based on these plots, draw a conclusion regarding the normality of these data.
- e. Obtain the standard deviations of each treatment group using SPSS. Using our rule of thumb (max. std < 2*min. std), do the data satisfy the assumption of equal standard deviations? Explain why or why not.
- f. Run the ANOVA procedure, then copy and paste the relevant SPSS output.
- g. Based on the SPSS output, write your conclusion. Explain how you arrived at this conclusion.

IV. (Linear Regression) Using the "*Health Exam Results*" data (the data you used for the previous SPSS project), use SPSS to fit a regression line to model the *weight* (y) of a person based on (a) *height* and (b) *waistline*.

To conduct the analysis using SPSS, perform the following steps.

- Analyze \rightarrow Regression \rightarrow Linear
- Put *weight* in the **Dependent** box and *height* as the **Independent**.
- Click **Continue** and then **OK**

Creating a Scatterplot

- 1. Click on Graph \rightarrow Legacy Dialogs \rightarrow Scatter/Dot <u>;</u> Choose Simple Scatter, then click Define.
- 2. Highlight the name of the response variable and click \mathbf{b} to be plotted on the *Y***-axis**.
- 3. Highlight the name of the explanatory variable and click ▶ to be plotted on the *X*-axis and click OK.
- 4. You can have the least squares regression line plotted on the scatterplot by following the steps outlined below:
 - a. Double Click on the scatterplot to bring up the **Chart Editor**.
 - b. Click on **Elements**→**Fit Line at Total**.
 - c. Click **Close** to close the **Properties Window**.
 - d. Close the Chart Editor.

To be included in your work:

- 1. Construct a scatterplot for the *height* variable versus the response variable.
- 2. Does the scatterplot indicate a linear relationship?
- 3. Add the appropriate regression line in your scatterplot. Copy and paste this into your word document.
- 4. Determine the regression line, correlation coefficient r, and coefficient of determination R^2 .
- 5. Explain the meaning of the value of R^2 in the context of this problem.
- 6. Use SPSS to test Ho: $\beta_1=0$ vs. H₁: $\beta_1\neq 0$. Copy and paste the relevant table from the SPSS output.
- 7. Repeat 1 6 for the other explanatory variable (*waistline*).
- 8. Which of the two explanatory variables is the better predictor of the response variable? Explain your answer.