**Assignment 3**

**Data Analysis**

**Data coding**

For this assignment you will practice data analysis techniques using SPSS. Below you find instructions and examples on how to sort and code your data.

1. Check all questionnaires for any errors in recording data by hand.
2. Take a blank questionnaire and assign codes to each possible response. Here are some examples of how to assign codes:

**Example 1:**

What is your gender?

Male

1

Female

2

In example 1, create one field named “gender” and put “1” if the respondent is a male, “2” if a female.

# Example 2

Where do you do most of your grocery shopping? Please check one.

a. Hyvee 1

b. Baker’s 2

c. Walmart 3

d. Target 4

e. Whole Foods 5

In example 2, create one field named “store” (or whatever other name you choose), and put 1, 2, 3 etc. to reflect the appropriate response.

# Example 3

Where do you do your grocery shopping? Please check all that apply.

a. Hyvee

b. Baker’s

c. Walmart

d. Target

e. Whole Foods

 f. Other Please specify \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

Example 3 is a tricky question for coding, because unlike example 2, where the respondent is asked to choose just one, here we are asking to choose all that apply. Consequently, some respondents may check just one choice, some two, and so on. The approach we took in example 2 will not work here (think about this and try for yourself).

To code this right, you will have to create a variable (column) for each of the stores. In this case, there will be six fields for six choices, including one for “other”. Then, if someone chose *Baker’s* and *Super Target*, you put “1” in those two fields, and “0” in the rest of the fields. Similarly, if someone chose Albertson’s, you put “1” in that field and “0” in all others.

In reality, you will also need to create one extra variable in case someone chooses “other” and specifies the name of the store. that field, you will write down the name of the store ( say “Bag and Save”) and since it requires text entry, you will choose “string” as a type of variable.

# Example 4

From the following list of grocery store attributes, rank the three most important ones. Put “1” in front of the most important, “2” in front of the second most important and so on.

Convenient location

Wide assortment of products

Reasonable prices

Fresh produce

Open 24 hours

Has a pharmacy

Has a deli

Has a bank

Cart-to-the car service

In example 4, you need to find how to decide the rank of the above 9 attributes. For each attribute, you need to create a column and write “1”, “2”, or “3” as per the response. Now, suppose “Convenient Location” received 10 first ranks, 15 second ranks and 4 third ranks. You need to give a weight of 3, 2, and 1 to respectively the first, second, and third rank and calculate the composite score. Here, it will be (3\*10) + (15\*2) + (4\*1) = 64. Do similar calculations for every attribute in the list. The one with the highest score is ranked number 1 and so on. Also, some of the steps in this calculation may be carried out manually by you, as SPSS will be of little help there.

**Example 5**

On a scale of 1 to 5, please rate how satisfied you are with the service level of the store employees. Choose “1” for “most satisfied” and “5” for “most dissatisfied”, with the intermediate numbers for the intermediate levels of satisfaction:

**Most satisfied Satisfied Neutral Dissatisfied Most dissatisfied**

 1 2 3 4 5

In above, you do not need any coding, as the numbers “1” to “5” themselves could be used as codes.

# Example 6

Circle an appropriate range that suggests your annual household income:

1

2

3

4

5

1. Less than $15000
2. $15000 - $34999
3. $35000 - $49999
4. $50000 - $64999
5. $65000 and more

In example 6, you can code the income ranges as shown above. Then, you can report the % of respondents that came from each income range. You can also compare different respondents (for example, their satisfaction level with the employee service) among different income categories.

After some major sorting and coding you now most likely have a dataset that you can perform tests on!

**Data Analysis**

**Coding and descriptive statistics**

1. Code all data received from the survey, similar to how we coded data during class (also see the prior pages in this document for examples).
2. Calculate appropriate descriptive statistics for all demographic variables (e.g., nationality, gender, age).
3. Calculate the appropriate descriptive statistic to determine which brand of the chosen mobile IT device is most widely used.
	1. This requirement relates to objective #1 of the mobile IT devices survey.
4. Code and show the top three most used features.
	1. This requirement relates to objective #8 of the mobile IT devices survey.

**Dependent variable(s)**

1. Calculate descriptive statistics (mean, standard deviation) for usage of each feature (the dependent variable).
	1. This requirement relates to objective #2 of the mobile IT devices survey.

**Independent Variable(s)**

1. Conduct a reliability analysis in SPSS on each independent variable.

*Conduct a factor analysis for the independent variables [high difficulty level]*

* 1. Are there any survey items you would remove from a given scale? Why?
1. Calculate a scale for each independent variable, by averaging all questions measuring that independent variable.

*Or, extract factor scores from your factor analysis [high difficulty level]*

1. Calculate descriptive statistics (mean, standard deviation) for each independent variable.
	1. This requirement loosely relates to objective #3 of the survey in the sense that you can inspect the mean score for each independent variable as a proxy for why the chosen device are used.
2. Calculate descriptive statistics (e.g. mean, standard deviation) for each attitude question.
	1. This requirement relates to objective #4 of the survey.

**Tests**

1. Determine if perceptions of your chosen independent variables (e.g. usefulness, ease of use, subjective norm, attitude) varies by gender:
	1. Run a t-test comparing males to females for each independent variable.
	2. Are any of the t-tests significant?
	3. What is your interpretation of the results of the t-tests?
2. Conduct a correlation analysis to determine which independent variable is most highly correlated with usage.
	1. Indicate whether each correlation is significant and interpret the results.
	2. Are any of the independent variables highly correlated (e.g., correlated greater than .7)? Do you think this might be a problem in terms of establishing discriminant validity?
3. Run a regression using the scales calculated in step 7 to explain the use of one mobile IT device feature.
4. Is the regression significant? Interpret the regression and fit statistics.
5. Which independent variables explain usage?
6. Report on any additional tests you did that can be useful for your report.

**Important note on your submission**

DO NOT attach the SPSS output as spitted out by the computer with your report (you will get negative points if you do).

* Management wants to see the key numbers in a formatted fashion, not the entire output.
* Your deliverable for this project needs to be a formatted report. To that end, be sure to use appropriate plots and charts and explain the process and your results along the way. As such, use various chart types (e.g. bar-, pie-diagrams) whenever they are relevant.