Lab 02: Measurement Scales,  Statements and Role of Statistics

**Handout date:** Wednesday, September 2, 2020

**Due date:** Monday, September 14, 2020 via **SubmitLab02** link in eLearning

*This lab counts 4 % toward your total grade*

**Objectives:** In this lab   
[a] you will identify and justify the measurement scale of several variables.

[b] you will practice some data manipulation techniques as well as  statements, and  
[c] read two articles on the role of statistics.

**Format of answer:** Your answers (graphs and verbal description) should be handed in as ***hard-copy*** in ***one*** document. Add a running title into the header of the document with the following information: ***your name***, ***GISC6301-Lab02*** and ***page numbers***. Label each answer properly starting with its task number. Maintain the sequence of questions. Format any code and computer output properly before inserting it into the document with your answer. -code and text output need to be in a ***monospaced*** font (i.e., fixed-pitch font) such as Courier New so proper spacing and alignments are preserved. Excessive, but irrelevant, output will lead to a deduction of your accumulated points.

## Task 1: Identify and justify the measurement levels of several statistical variables (1 point)

Justify your selection of the measurement scale. You may want to look up Wikipedia for some of the variables

1. Give an example each for nominal, ordinal, interval and ratio scaled measurement. Do not use examples from the lecture or lab. (0.4 points)
2. The longitude and latitude in degrees on the earth’s spherical surface. Be cautious in your arguments with respect to the origins of the coordinate system. That is, are distances between two longitudes constant all over the globe? (0.2 points)
3. The wind direction in degrees. (0.1 points)
4. Grouping of census block groups into neighborhoods. (0.1 points)
5. Income brackets for taxation purposes (0.1 points)
6. Elevation above sea-level at a fixed point in time. (0.1 points)

## Task 2: Working with Data (2 points)

For all tasks below show your properly formatted code. You find the necessary code for the examples in Lander and the online help. Only if asked show also the output.

Import the SPSS data-file **Concord1.sav** in the **Week03** channel as ***data-frame*** into the  environment by using a function from the library **foreign**. Make sure to name your data-frame properly.

1. Discuss the summary statistics for the water consumption: How did the ***average*** water consumption change from 1979 to 1981? (0.1 points)
2. Discuss the summary statistics: Which variable has ***missing*** observations? (0.1 points)
3. Discuss the summary statistics: Which variable is a factor? (0.1 points)
4. List all ***case numbers*** (variable **case**), which have at least for one variable missing value in a variable. Show also the code. (0.2 points)
5. Which ***class*** are the following data selections: [a] **Concord$retire**, [b] **Concord["retire"]**, [c] **Concord[ , "retire"]**, and [d] **Concord[["retire"]]**? Show code and the output. (0.2 points)
6. Calculate the ***average*** water consumption over the 3 years for each household and save it the new variable **meanWater** into the data-frame. Caution: also include households, which have **NA**s in the water consumption. Hint: look at the documentation of the function **mean( )**. (0.2 points)
7. Use logical statements to identify those households (variable **case**), which have above average water consumption in 1981. Show the code and the household numbers (0.2 point)
8. Draw a sample of 10 households without repetitions. Show the household numbers (variable **case**) and the code. Hint: look at the documentation of the function **sample( )**. (0.2 points)
9. Add a new variable **seqID** by labeling each record by its record number ranging from 1 to the number of observations. Show the code. (0.2 points)
10. ***Bind*** the two variables **peop80** and **peop81** together into a ***matrix***. Show the code. (0.1 points)
11. Give a code example of the use of the **ifelse** statement. (0.1 points)
12. Give an example of the **while** statement (0.1 points)
13. What are [a] positional, [b] named and [c] default arguments of a function? (0.3 points)

## Task 3: Critical discussion of big data analyses (1 point)

Read the document **BigDataAndStatitics.pdf** and **MathMusicStatsLiterature.pdf** in the **Week03** channel.

[a] Give the reasons why large samples may not necessary be better than small samples. (0.5 points)

[b] What makes mathematics different from statistics. List the main differences. (0.5 points)