

Homework #1 – Due Thursday, Jul. 12, before class

STAT-UB.0001 – Statistics for Business Control

Some of the homework assignments will involve data files. All data files will be available on NYU Classes under the “Datasets” section. You are permitted to work in teams, but each student must independently write up their own solutions (no direct copying).

Electronic submissions are not accepted, but you can write your solutions on a laptop and print it out. Print out and turn in your Minitab plots. See the “Minitab Tips” on NYU Classes for tips on opening files and printing or saving plots.

Problem 1

Much of our understanding of prehistoric peoples comes from caves, such as cave paintings, evidence of fire pits, etc. made nearly 40,000 years ago. Based on these, people tend to believe that most of prehistoric peoples lived in caves for most of their lives. Do you see any problem in this?

.....

Problem 2

Consider these values:

21 22 18 25 27 9 21 34 20 16

Find the mean, median, mode, and standard deviation for these. You need to compute these values by hand. This is for you to get familiar with those formulas.

.....

Problem 3

The file `HeightWeight.csv` contains data on 200 records of human heights and weights of 18 years old children. Here, we focus on the the *Weight* (in Pounds) column.

1. Use *Stat* \Rightarrow *Basic Statistics* \Rightarrow *Display Descriptive Statistics* to compute the mean, median and interquartile range, for the *Weight*. You will need to obtain the inter-quartile range by hand as the difference between the third quartile (Q_3 , the 75th percentile) and the first quartile (Q_1 , the 25th percentile).
2. Make a histogram of *Weight*. Does the data seem to have a reasonably bell-shaped distribution?
3. Make a boxplot of *Weight*. Do you see any outliers?
4. Use empirical rules to complete the statements:
 - For approximately 68% individuals, the weight is between $[?, ?]$
 - For approximately 95% individuals, the weight is between $[?, ?]$
 - For approximately 99.7% individuals, the weight is between $[?, ?]$

5. (Optional) Look at the data. What are the true percentages in those intervals you just computed? Do they agree with the empirical rules (roughly)?

.....

Problem 4

A World Cup 2018 quarter-final between Sweden and England will happen on Jul 7. Suppose Sweden will score 0-3 goals with equal probability, and same for England.

1. What is the sample space for this quarter-final?
2. What is the probability that Sweden scores more goals than England?

.....

Problem 5

Suppose among all soccer fans, 12% support Sweden, 23% support England, and 4% support both. A soccer fan is to be selected at random.

1. What is the probability that the soccer fan supports either Sweden or England (or both)?
2. What is the probability that the soccer fan supports neither team?
3. What is the probability that the soccer fan supports Sweden but not England?

.....

Problem 6

Suppose the probability that a World Cup game will have more than 3 goals scored is 30%, and the probability that the game will have more than 4 goals scored is 20%. For a World Cup game that has more than 3 goals scored, what is the probability of it having more than 4 goals scored?

.....