

# Math 4030 Assignment 1 (Problems)

Due: Sept. 21, 4:00PM

## Study Guide:

- Identify the types of a variable or data set being
  - qualitative or quantitative;
  - categorical or numerical;
  - discrete or continuous;
  - at nominal, ordinal, or scale level of measurement.
- Given a data set from a population or a sample, construct
  - a frequency table (grouped or ungrouped; with frequency, relative frequency, and/or cumulative frequency columns);
  - bar chart, ogive, or histogram;
  - pie chart and/or Pareto chart if appropriate;
  - dot diagram and/or stem-and-leaf display if suitable;
  - boxplot if suitable.
- Given a data set from a population or a sample, calculate
  - mean and median;
  - sample or population variance and standard deviation;
  - coefficient of variation (meaning and application);
  - quartiles and percentiles;
  - range and interquartile range;
  - outliers, if any.
- Given a particular value in a (numerical) data set, find is
  - percentile rank;
  - z-score (meaning and application).
- Given a grouped or ungrouped frequency table, calculate (or estimate)

the mean and variance.

- Calculate weighted means.

**Assignment Problems:**

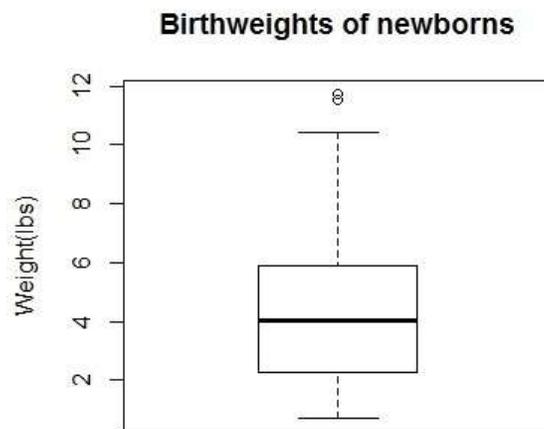
1. Determine whether the following possible responses should be classified as variables at level of nominal, ordinal, or scale.
  - a. The age of each of your classmates;
  - b. The size of fries (small, medium, large) ordered by a sample of Burger King customers;
  - c. Your hometown;
  - d. The type of car you currently drive;
  - e. Lily's travel time from her dorm to the student union;
  - f. The availability of parking on campus: Insufficient, Moderate, or Abundant;
  - g. Your marital status;
  - h. Heidi's favorite brand of tennis balls.
2. Below is a list of numerical variables from an annual survey of university students. For each of variables, decide whether it should be treated as discrete or continuous in the study.
  - a. Commuting distance to university (in kilometers);
  - b. Annual tuition fees;
  - c. Number of textbooks purchased;
  - d. Total cost of textbooks;
  - e. Monthly rent for student accommodation;
  - f. The number of students shared an off-campus apartment;
  - g. Starting time of the first class of a day;
  - h. The length of time needed for a student to complete a homework assignment.
3. Consider the following data set:

## Math 4030 Assignment 1 (Problems)

39, 24, 22, 54, 12, 53, 12, 13, 49, 39, 22, 14,  
44, 38, 36, 24, 54, 53, 13, 39, 14, 52, 22, 20.

Find mean, median, the 1st quartile, and the 84th percentiles.

- Find the percentile rank of values 36 and 49 in the data set given in Problem 3.
- The boxplot below displays the birth-weights of 100 randomly selected infants born at a hospital. Use this to answer the following questions (using visual estimation).



- Find the Median.
  - Find the interquartile range.
  - Suppose we removed the two outliers from the data set, the standard deviation would increase, decrease, or remain approximately the same?
  - Suppose we removed the two outliers from the data set, the interquartile range would increase, decrease, or remain approximately the same?
- Ted took 4 courses last semester: Physics, Calculus, Biology, and History. The means and standard deviations for the final exams, and Ted's scores are given in the table below. Convert Ted's score into z scores.

Math 4030 Assignment 1 (Problems)

Subject	Mean	S.D.	Ted's Mark	Ted's z-score
Physics	60	14	77.5	
Calculus	70	12	88	
Biology	77	10	77	
History	53	16	45	

On what exam did Ted perform the best comparing to his classmates?

7. A student conducted a survey on some people's weekly income:

\$1900, \$1700, \$300, \$185000, \$1000, \$1500, \$600, \$2000.

Find the mean and the median. Which number, mean or median, is a better way to represent these people's weekly income?

8. Given three different data sets.

Data set A : 66, 68, 69, 71, 72, 75

Data set B : 15, 15, 19, 21, 26, 29

Data set C : 5, 6, 7, 7, 7, 9

Which data set is the most variable? (Use sample standard deviations.)

9. Use the given data values (number of automobile accidents)

0	2	4	4
5	0	1	3
1	4	0	2
3	2	0	1
0	2	3	1
2	1	6	5

- Make the ungrouped frequency table with columns of frequency/count and cumulative frequency/count.
  - Make bar chart.
10. A civil engineer monitors water quality by measuring the amount suspended solids in a sample of river water. Over the period of 24 days, she observed

Math 4030 Assignment 1 (Problems)

10	21	30	38	60	29	55
21	19	43	30	26	18	19

- a. Construct a dot diagram.
  - b. Construct a grouped frequency table using 5 groups.
  - c. Draw the histogram based on the above frequency table.
  - d. Calculate the mean and median.
  - e. Calculate the sample standard deviation.
  - f. Calculate the coefficient of variations.
  - g. Find the first and the third quartiles.
  - h. Draw the boxplot.
  - i. Find the interquatile range.
  - j. Use the interquatile range to identify the outliers.
11. For the following ungrouped frequency table, calculate the mean and (sample)variance.

Number of siblings	Frequency
0	7
1	15
2	20
3	0
4	2

12. For the following grouped frequency table, estimate the mean and (population) variance.

Math 4030 Assignment 1 (Problems)

Sale Price (\$000)	Relative Frequencies (%)
0 - 100.0	6.6
100.1 - 150.0	11.2
150.1 - 200.0	24.6
200.1 - 250.0	19.7
250.1 - 300.0	23.5
300.1 - 350.0	14.4

13. 291 engineering students (from 3 classes) took the exam for Math 4030. Class results are listed in the following table.

	Class Size	Class Average	Class SD
Class A	78	76.2	5.4
Class B	120	61.4	6.0
Class C	93	78.9	6.5

- a. Find the combined average of all 291 students.
- b. Marks from which class is more variable? and explain why.