Recitation 1

Exercise 1

Identify each of the following variables as categorical or quantitative.

1. Number of children in family
   Answer. Quantitative

2. Nationality
   Answer. Categorical

3. Final Grade on Statistics exam (Scale: A, B, C, D, F)
   Answer. Categorical

4. Distance in kilometers of commute to work
   Answer. Quantitative

5. Choice of diet (vegetarian, nonvegetarian)
   Answer. Categorical

6. Ownership of a personal computer (yes, no)
   Answer. Categorical

Identify each of the following variables as continuous or discrete.

1. The length of time to run a marathon
   Answer. Continuous

2. The number of people in line at a box office to purchase theater tickets
   Answer. Discrete

3. The weight of a baby
   Answer. Continuous

4. The number of people you have dated in the past five years
   Answer. Discrete

5. The distance between where you live and your statistics classroom
   Answer. Continuous

Exercise 2

A professor examined the results of the first exam given in her statistics class. The scores were

$$70,\ 84,\ 59,\ 73,\ 86,\ 35,\ 81,\ 75.$$

1. Find the mean and the median.
   Answer
   The mean is given by

   $$(70 + 84 + 59 + 73 + 86 + 35 + 81 + 75)/8 = 70.375;$$

   while to find the median, first you have to sort the data,

   $$35, 59, 70, 73, 75, 81, 84, 86.$$

   It follows that the median is given by the average of the two middle values, $(73+75)/2 = 74$

2. Would you guess that the distribution is skewed or roughly symmetric? Why?
   Answer
   Since the median is greater than the mean, the distribution is skewed to the left.

3. Find the standard deviation.
   Answer

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
35	-35.375	1251.39
59	-11.375	19.39
70	-0.375	0.14
73	2.625	6.89
75	4.625	21.39
81	10.625	112.89
84	13.625	185.64
86	15.625	244.14
Total	0	1951.875

   The variance is given by $1951.875/8 = 243.9844$. Thus, $s = \sqrt{243.9844} = 15.62$.

4. The instructor added the grade of a further student, that is 100. Describe the effect of this new observation.
   Answer
   The mean is given by

   $$(70 + 84 + 59 + 73 + 86 + 35 + 81 + 75 + 100)/9 = 73.77;$$

To find the median, we need to add 100 to the sorted data,

$$35, 59, 70, 73, 75, 81, 84, 86, 100.$$

The median is 75.

5. Suppose that 73 was incorrectly recorded and is supposed to be 3. Describe the effect of this on the mean and median.
Answer
The mean is given by

$$(70 + 84 + 59 + 3 + 86 + 35 + 81 + 75 + 100)/9 = 65.89;$$

Given the new sorted data,

$$3, 35, 59, 70, 75, 81, 84, 86, 100,$$

he median is still 75.
Thus, this new observation affects mean but not median.

Exercise 3

Consider the following two sets of observations:

Set 1: 2,3,3,3,4,4,4
Set 2: 2,3,3,3,3,3,4

1. Find the variance for each data set.
Answer.

   - Set 1: Mean=3.29

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
2	-1.2857143	1.65306122
3	-0.2857143	0.08163265
3	-0.2857143	0.08163265
3	-0.2857143	0.08163265
4	0.7142857	0.51020408
4	0.7142857	0.51020408
4	0.7142857	0.51020408
Total	0	3.29

   The variance is given by $3.29/7 = 0.49$.

   - Set 2: Mean=3.

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
2	-1	1
3	0	0
3	0	0
3	0	0
3	0	0
3	0	0
4	1	1
Total	0	2

The variance is given by $2/7 = 0.29$.

2. Which data set shows more variability?
Answer. Set 1.

Exercise 4

For the question "How many children have you ever had?", the results were

No.Children	0	1	2	3	4
Count	25	15	20	5	0

1. Provide a graphical representation of the distribution.
   Answer

2. Find the mode.
   Answer
   The mode is 0 (the value that occurs most often).

3. Find the variance and the standard deviation.
   Answer

x_i	n_i	$n_i * x_i$	$(x_i - \bar{x})^2$	$n_i * (x_i - \bar{x})^2$
0	25	0	$(-1.08)^2$	29.16
1	15	15	$(-0.08)^2$	0.096
2	20	40	$(0.92)^2$	16.928
3	5	15	$(1.92)^2$	18.432
4	0	0	$(2.92)^2$	0
Total	65	70		64.616

To compute the variance and the standard deviation, first we need to compute the mean, $\bar{x} = 70/65$. The variance is obtained as $s^2 = 64.616/65 = 0.99$; thus, $s = \sqrt{0.99} = 0.995$.

Exercise 5

The 2007 unemployment rates of countries in the European Union are shown in the table below.

Country	Unemployment rate	Country	Unemployment rate	Country	Unemployment rate
Belgium	7.8	France	8.4	Italy	6.7
Denmark	3.2	Portugal	7.2	Finland	7.0
Germany	7.7	Netherlands	3.6	Austria	4.5
Greece	8.7	Luxembourg	5.0	Sweden	6.0
Spain	8.6	Ireland	4.4	U.K.	5.4

1. Provide a graphical representation of the distribution.
   Answer
   First we need to construct the frequency table

Interval	Frequency
3.1 to 4	2
4.1 to 5	3
5.1 to 6	2
6.1 to 7	2
7.1 to 8	3
8.1 to 9	3

   The corresponding histogram is depicted in the following figure.

2. Find the mean and standard deviation.

   Answer
   The mean is 6.28; while the standard deviation is 1.84.

3. What unemployment value for a country would have a z-score equal to 0?
   Answer
   A z-score of 0 indicates that the country's unemployment rate is zero standard deviations from the mean; hence, the unemployment rate is equal to the mean. In this case, a country with an unemployment rate of 6.3 would have a z-score of 0.

Exercise 6

The mean and standard deviation of a sample may change if data are rescaled.

1. Scores on a difficult exam have a mean of 57 and a standard deviation of 20. The teacher boosts all the scores by 20 points before awarding grades. Report the mean and the variance of the boosted scores.
   Answer
   mean=77; s=20.
   $\bar{y} = a + b\bar{x}$. $a = 20$ and $b = 1$

2. Referring to point 1, what happens to the mean if the students get a grade rise of 3%?
   Answer
   mean=58.71; s=0.6;
   $\bar{y} = a + b\bar{x}$. $a = 0$ and $b = 1.03$
   Thus $s^2 = b^2 \times 20^2 = 20.6^2 = 424.36$, while $\bar{y} = 58.71$

3. Suppose that the annual income for some group has a mean of $ 39,000 and a standard deviation of $ 15,000. Values are converted to euros. If one euro equals $2.00, report the mean and standard deviation in European currency.
   Answer
   mean=19500; s=7500; $\bar{y} = a + b\bar{x}$. $a = 0$ and $b = 0.5$.
   The mean of the annual income converted to euros is 19500; while the standard deviation converted to euros is 7500.