

Recitation 1

Carlo Cavicchia

carlo.cavicchia@uniroma1.it 

Exercise 1

Identify each of the following variables as categorical or quantitative.

1. Number of children in family

Exercise 1

Identify each of the following variables as categorical or quantitative.

1. Number of children in family
2. Nationality

Exercise 1

Identify each of the following variables as categorical or quantitative.

1. Number of children in family
2. Nationality
3. Final Grade on Statistics exam (Scale: A, B, C, D, F)

Exercise 1

Identify each of the following variables as categorical or quantitative.

1. Number of children in family
2. Nationality
3. Final Grade on Statistics exam (Scale: A, B, C, D, F)
4. Distance in kilometers of commute to work

Exercise 1

Identify each of the following variables as categorical or quantitative.

1. Number of children in family
2. Nationality
3. Final Grade on Statistics exam (Scale: A, B, C, D, F)
4. Distance in kilometers of commute to work
5. Choice of diet (vegetarian, nonvegetarian)

Exercise 1

Identify each of the following variables as categorical or quantitative.

1. Number of children in family
2. Nationality
3. Final Grade on Statistics exam (Scale: A, B, C, D, F)
4. Distance in kilometers of commute to work
5. Choice of diet (vegetarian, nonvegetarian)
6. Ownership of a personal computer (yes, no)

Exercise 1

Identify each of the following variables as categorical or quantitative.

1. Number of children in family
2. Nationality
3. Final Grade on Statistics exam (Scale: A, B, C, D, F)
4. Distance in kilometers of commute to work
5. Choice of diet (vegetarian, nonvegetarian)
6. Ownership of a personal computer (yes, no)

Identify each of the following variables as continuous or discrete.

1. The length of time to run a marathon

Exercise 1

Identify each of the following variables as categorical or quantitative.

1. Number of children in family
2. Nationality
3. Final Grade on Statistics exam (Scale: A, B, C, D, F)
4. Distance in kilometers of commute to work
5. Choice of diet (vegetarian, nonvegetarian)
6. Ownership of a personal computer (yes, no)

Identify each of the following variables as continuous or discrete.

1. The length of time to run a marathon
2. The number of people in line at a box office to purchase theater tickets

Exercise 1

Identify each of the following variables as categorical or quantitative.

1. Number of children in family
2. Nationality
3. Final Grade on Statistics exam (Scale: A, B, C, D, F)
4. Distance in kilometers of commute to work
5. Choice of diet (vegetarian, nonvegetarian)
6. Ownership of a personal computer (yes, no)

Identify each of the following variables as continuous or discrete.

1. The length of time to run a marathon
2. The number of people in line at a box office to purchase theater tickets
3. The weight of a baby

Exercise 1

Identify each of the following variables as categorical or quantitative.

1. Number of children in family
2. Nationality
3. Final Grade on Statistics exam (Scale: A, B, C, D, F)
4. Distance in kilometers of commute to work
5. Choice of diet (vegetarian, nonvegetarian)
6. Ownership of a personal computer (yes, no)

Identify each of the following variables as continuous or discrete.

1. The length of time to run a marathon
2. The number of people in line at a box office to purchase theater tickets
3. The weight of a baby
4. The number of people you have dated in the past five years

Exercise 1

Identify each of the following variables as categorical or quantitative.

1. Number of children in family
2. Nationality
3. Final Grade on Statistics exam (Scale: A, B, C, D, F)
4. Distance in kilometers of commute to work
5. Choice of diet (vegetarian, nonvegetarian)
6. Ownership of a personal computer (yes, no)

Identify each of the following variables as continuous or discrete.

1. The length of time to run a marathon
2. The number of people in line at a box office to purchase theater tickets
3. The weight of a baby
4. The number of people you have dated in the past five years
5. The distance between where you live and your statistics classroom

Exercise 1

Identify each of the following variables as categorical or quantitative.

1. Number of children in family - **quantitative**
2. Nationality - **categorical**
3. Final Grade on Statistics exam (Scale: A, B, C, D, F) - **categorical**
4. Distance in kilometers of commute to work - **quantitative**
5. Choice of diet (vegetarian, nonvegetarian) - **categorical**
6. Ownership of a personal computer (yes, no) - **categorical**

Identify each of the following variables as continuous or discrete.

1. The length of time to run a marathon - **continuous**
2. The number of people in line at a box office to purchase theater tickets - **discrete**
3. The weight of a baby - **continuous**
4. The number of people you have dated in the past five years - **discrete**
5. The distance between where you live and your statistics classroom - **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.

- › Find the mean and the median.

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.

- › Find the mean and the median.

up to you -> <https://forms.gle/Z9AUsyRrcpkAedhy9>

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.

- › Find the mean and the median.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{70 + 84 + 59 + 73 + 86 + 35 + 81 + 75}{8} = 70.375$$

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.

- › Find the mean and the median.

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

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

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.

- › Find the mean and the median.

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

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

$$x_{Med} = \frac{73 + 75}{2} = 74$$

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.

- › Would you guess that the distribution is skewed or roughly symmetric? Why?

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.

- › Would you guess that the distribution is skewed or roughly symmetric? Why?

Since the median is greater than the mean, the distribution is skewed to the left.

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.

- › Find the standard deviation.

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.

- › Find the standard deviation.

up to you -> <https://forms.gle/tLNwL7EEwiA5NKR19>

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.

- › Find the standard deviation.

$$s = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$$

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.

- > Find the standard deviation.

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

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.

- › Find the standard deviation.

$$s = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} = \sqrt{\frac{1951.875}{8}} = \sqrt{243.9844} = 15.62$$

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.

- › The instructor added the grade of a further student, that is 100. Describe the effect of this new observation.

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.

- > The instructor added the grade of a further student, that is 100. Describe the effect of this new observation.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{70 + 84 + 59 + 73 + 86 + 35 + 81 + 75 + 100}{9} = 73.77$$

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.

- > The instructor added the grade of a further student, that is 100. Describe the effect of this new observation.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{70 + 84 + 59 + 73 + 86 + 35 + 81 + 75 + 100}{9} = 73.77$$

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

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

$$x_{Med} = 75$$

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, 100.

- › Suppose that 73 was incorrectly recorded and is supposed to be 3. Describe the effect of this on the mean and median.

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, 100.

- › Suppose that 73 was incorrectly recorded and is supposed to be 3. Describe the effect of this on the mean and median.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{70 + 84 + 59 + 3 + 86 + 35 + 81 + 75 + 100}{9} = 65.89$$

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, 100.

- Suppose that 73 was incorrectly recorded and is supposed to be 3. Describe the effect of this on the mean and median.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{70 + 84 + 59 + 3 + 86 + 35 + 81 + 75 + 100}{9} = 65.89$$

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

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

$$x_{Med} = 75$$

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

- › Find the variance for each data set.

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

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

- > Find the variance for each data set.

Set 1: $\bar{x} = 3.29$

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

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

- › Find the variance for each data set.

Set 1: $\bar{x} = 3.29$ and $s^2 = \frac{3.29}{7} = 0.49$

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

- › Find the variance for each data set.

Set 2: $\bar{x} = 3$

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

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

- › Find the variance for each data set.

Set 2: $\bar{x} = 3$ and $s^2 = \frac{2}{7} = 0.29$

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

- › Which data set shows more variability?

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

- › Which data set shows more variability?

Set 1

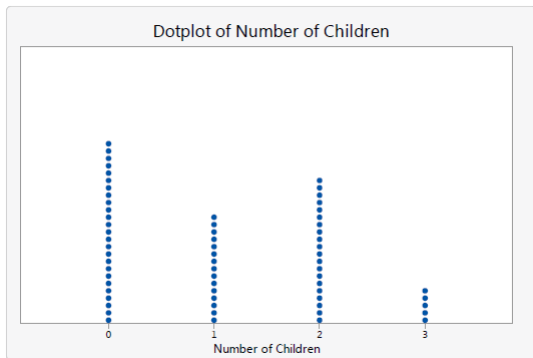
Exercise 4

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

- › Provide a graphical representation of the distribution.

Exercise 4



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

- › Find the mode.

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

- › Find the mode.

up to you -> <https://forms.gle/9CbddeS5PcFY8HvR9>

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

- › Find the mode.

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

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

- › Find the variance and the standard deviation.

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

- › Find the variance and the standard deviation.

up to you -> <https://forms.gle/4j4n2ut8Rbc5wArUA>

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

- › Find the variance and the standard deviation.

To compute the variance and the standard deviation, first we need to compute the mean, $\bar{x} = 70/65$.

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

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

- › Find the variance and the standard deviation.

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	Unempl. rate	Country	Unempl. rate	Country	Unempl. 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

- › Provide a graphical representation of the distribution.

Exercise 5

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

Country	Unempl. rate	Country	Unempl. rate	Country	Unempl. 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

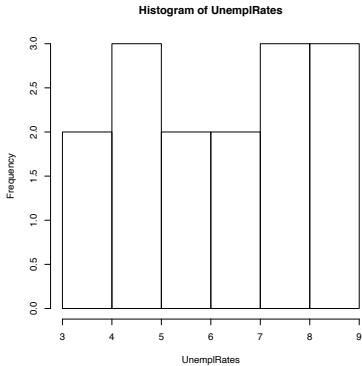
- › Provide a graphical representation of the distribution.

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

Exercise 5

The corresponding histogram is depicted in the following figure.



Exercise 5

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

Country	Unempl. rate	Country	Unempl. rate	Country	Unempl. 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

- › Find the mean and standard deviation.

$$\bar{x} = \frac{7.8 + \dots + 5.4}{15} = 6.28$$

Exercise 5

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

Country	Unempl. rate	Country	Unempl. rate	Country	Unempl. 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

- › Find the mean and standard deviation.

$$s = \dots = 1.84$$

Exercise 5

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

Country	Unempl. rate	Country	Unempl. rate	Country	Unempl. 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

- › What unemployment value for a country would have a z-score equal to 0?

Exercise 5

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

Country	Unempl. rate	Country	Unempl. rate	Country	Unempl. 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

- › What unemployment value for a country would have a z-score equal to 0?

up to you -> <https://forms.gle/UrtL736myMUQDpxo9>

Exercise 6

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

- › 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 standard deviation of the boosted scores.

Exercise 6

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

- › 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 standard deviation of the boosted scores.

Applying the linear transformation $y = a + bx$ we have: $\bar{y} = a + b\bar{x}$ and $s_y = bs_x$

In our case, $a = 20$ and $b = 1$. Thus, $\bar{y} = 77$ and $s_y = 20$.

Exercise 6

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

- › Referring to the previous point, what happens to the mean if the students get a grade rise of 3%?

Exercise 6

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

- > Referring to the previous point, what happens to the mean if the students get a grade rise of 3%?

Applying the linear transformation $y = a + bx$ we have: $\bar{y} = a + b\bar{x}$ and $s_y = bs_x$

In our case, $a = 0$ and $b = 1.03$. Thus, $\bar{y} = 58.71$ and $s_y = 20.6$.

Exercise 6

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

- › 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.

Exercise 6

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

- › 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.

Applying the linear transformation $y = a + bx$ we have: $\bar{y} = a + b\bar{x}$ and $s_y = bs_x$

In our case, $a = 0$ and $b = 0.5$. Thus, $\bar{y} = \text{€}19,500$ and $s_y = \text{€}7,500$.