Quantitative Methods – I

Practice 2

Lorenzo Cavallo

For any clarification:
lorenzo.cavallo.480084@uniroma2.eu

THEME #4

Measures of position

Measures of position

Mode
Value/category/class with the highest frequency

Median
Value of the observation(s) in the middle of the ranked data, where the middle position is $\frac{n+1}{2}$

Quartiles
Three values that divide the **ranked** data into four equal parts

Percentiles
Values that divide the **ranked** data into 100 equal parts

Arithmetic mean/average
Sum of all values divided by number of observations

POPULATION

$$\mu = \frac{\sum x}{N}$$

SAMPLE

$$\bar{x} = \frac{\sum x}{n}$$

Geometric mean
The *n*th root of the product of all observations

$$\left(\prod_{i=1}^{n} a_i\right)^{\frac{1}{n}} = \sqrt[n]{a_1 a_2 \cdots a_n}.$$

Harmonic mean
The reciprocal of the arithmetic mean

$$H = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_n}} = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$$

3.82 The following data give the 2014 profits (in millions of dollars) of the top 10 companies listed in the 2014 *Fortune 500*. Find the mean and median for these data. Do these data have a mode?

Company	2014 Profits (mil. of dollars)
Wal-Mart Stores	16022
Exxon Mobil	3258
Chevron	21423
Berkshire Hathaway	19476
Apple	37037
Phillips 66	3726
General Motors	5346
Ford Motor	7155
General Electric	13057
Valero Energy	2720

Step 1. Rank data

Company	2014 Profits (mil. of dollars)
Valero Energy	2720
Exxon Mobil	3258
Phillips 66	3726
General Motors	5346
Ford Motor	7155
General Electric	13057
Wal-Mart Stores	16022
Berkshire Hathaway	19476
Chevron	21423
Apple	37037
Total	129220

Step 2. Median $\frac{n+1}{2} = \frac{11}{2} = 5.5$ (between 5th and 6th values)

$$\text{Median} = \frac{7155+13057}{2} = 10106$$

Step 3. Mean $\mu = \frac{\sum x}{N} = \frac{129220}{10} = 12922.0$

Step 4. No Mode

Calculating the median for ungrouped data: odd number of data values.

Table 3.2 lists the 2014 compensations of female CEOs of 11 American companies (*USA TODAY*, May 1, 2015). (The compensation of Carol Meyrowitz of TJX is for the fiscal year ending in January 2015.)

Table 3.2 Compensations of 11 Female CEOs

Company & CEO	2014 Compensation (millions of dollars)
General Dynamics, Phebe Novakovic	19.3
GM, Mary Barra	16.2
Hewlett-Packard, Meg Whitman	19.6
IBM, Virginia Rometty	19.3
Lockheed Martin, Marillyn Hewson	33.7
Mondelez, Irene Rosenfeld	21.0
PepsiCo, Indra Nooyi	22.5
Sempra, Debra Reed	16.9
TJX, Carol Meyrowitz	28.7
Yahoo, Marissa Mayer	42.1
Xerox, Ursula Burns	22.2

Find the median for these data.

Calculating the median for ungrouped data: odd number of data values.

Table 3.2 lists the 2014 compensations of female CEOs of 11 American companies (*USA TODAY*, May 1, 2015). (The compensation of Carol Meyrowitz of TJX is for the fiscal year ending in January 2015.)

Table 3.2 Compensations of 11 Female CEOs

Company & CEO	2014 Compensation (millions of dollars)
General Dynamics, Phebe Novakovic	19.3
GM, Mary Barra	16.2
Hewlett-Packard, Meg Whitman	19.6
IBM, Virginia Rometty	19.3
Lockheed Martin, Marillyn Hewson	33.7
Mondelez, Irene Rosenfeld	21.0
PepsiCo, Indra Nooyi	22.5
Sempra, Debra Reed	16.9
TJX, Carol Meyrowitz	28.7
Yahoo, Marissa Mayer	42.1
Xerox, Ursula Burns	22.2

Find the median for these data.

Solution To calculate the median of this data set, we perform the following two steps.

Step 1: The first step is to rank the given data. We rank the given data in increasing order as follows:

16.2	16.9	19.3	19.3	19.6	21.0	22.2	22.5	28.7	33.7	42.1

Calculating the median for ungrouped data: odd number of data values.

Table 3.2 lists the 2014 compensations of female CEOs of 11 American companies (*USA TODAY*, May 1, 2015). (The compensation of Carol Meyrowitz of TJX is for the fiscal year ending in January 2015.)

Table 3.2 Compensations of 11 Female CEOs

Company & CEO	2014 Compensation (millions of dollars)
General Dynamics, Phebe Novakovic	19.3
GM, Mary Barra	16.2
Hewlett-Packard, Meg Whitman	19.6
IBM, Virginia Rometty	19.3
Lockheed Martin, Marillyn Hewson	33.7
Mondelez, Irene Rosenfeld	21.0
PepsiCo, Indra Nooyi	22.5
Sempra, Debra Reed	16.9
TJX, Carol Meyrowitz	28.7
Yahoo, Marissa Mayer	42.1
Xerox, Ursula Burns	22.2

Find the median for these data.

Solution To calculate the median of this data set, we perform the following two steps.

Step 1: The first step is to rank the given data. We rank the given data in increasing order as follows:

16.2	16.9	19.3	19.3	19.6	21.0	22.2	22.5	28.7	33.7	42.1

Step 2: The second step is to find the value that divides this ranked data set in two equal parts. Here there are 11 data values. The sixth value divides these 11 values in two equal parts. Hence, the sixth value gives the median as shown below.

16.2	16.9	19.3	19.3	19.6	21.0	22.2	22.5	28.7	33.7	42.1
					↑					
					Median					

$$\frac{(n+1)}{2} = \frac{12}{2} = 6$$

3.16 The following data give the annual salaries (in thousand dollars) of 20 randomly selected health care workers.

50	71	57	39	45	64	38	53	35	62
74	40	67	44	77	61	58	55	64	59

a. Calculate the mean, median, and mode for these data.

3.16 The following data give the annual salaries (in thousand dollars) of 20 randomly selected health care workers.

50	71	57	39	45	64	38	53	35	62
74	40	67	44	77	61	58	55	64	59

a. Calculate the mean, median, and mode for these data.

n=20

Rank data

35 38 39 40 44 45 50 53 55 57 58 59 61 62 64 64 67 71 74 77

3.16 The following data give the annual salaries (in thousand dollars) of 20 randomly selected health care workers.

50	71	57	39	45	64	38	53	35	62
74	40	67	44	77	61	58	55	64	59

a. Calculate the mean, median, and mode for these data.

n=20

Rank data

35 38 39 40 44 45 50 53 55 57 58 59 61 62 64 64 67 71 74 77

Mean

$$\bar{x} = \frac{\Sigma x}{n} = \frac{35+38+39+40+44+45+50+53+55+57+58+59+61+62+64+64+67+71+74+77}{20} = \frac{1113}{20} = 55{,}65$$

3.16 The following data give the annual salaries (in thousand dollars) of 20 randomly selected health care workers.

50	71	57	39	45	64	38	53	35	62
74	40	67	44	77	61	58	55	64	59

a. Calculate the mean, median, and mode for these data.

n=20

Rank data

35 38 39 40 44 45 50 53 55 57 58 59 61 62 64 64 67 71 74 77

Mean

$$\bar{x} = \frac{\Sigma x}{n} = \frac{35+38+39+40+44+45+50+53+55+57+58+59+61+62+64+64+67+71+74+77}{20} = \frac{1113}{20} = 55{,}65$$

Median

$$\frac{n+1}{2} = \frac{21}{2} = 10{,}5$$

35	38	39	40	44	45	50	53	55	57	58	59	61	62	64	64	67	71	74	77
1°	2°	3°	4°	5°	6°	7°	8°	9°	10°	11°	12°	13°	14°	15°	16°	17°	18°	19°	20°

$$Median = \frac{57+58}{2} = 57{,}5$$

3.16 The following data give the annual salaries (in thousand dollars) of 20 randomly selected health care workers.

50	71	57	39	45	64	38	53	35	62
74	40	67	44	77	61	58	55	64	59

a. Calculate the mean, median, and mode for these data.

n=20

Rank data

35 38 39 40 44 45 50 53 55 57 58 59 61 62 64 64 67 71 74 77

Mean

$$\bar{x} = \frac{\Sigma x}{n} = \frac{35+38+39+40+44+45+50+53+55+57+58+59+61+62+64+64+67+71+74+77}{20} = \frac{1113}{20} = 55{,}65$$

Median

$$\frac{n+1}{2} = \frac{21}{2} = 10{,}5$$

35	38	39	40	44	45	50	53	55	57	58	59	61	62	64	64	67	71	74	77
1°	2°	3°	4°	5°	6°	7°	8°	9°	10°	11°	12°	13°	14°	15°	16°	17°	18°	19°	20°

$$Median = \frac{57+58}{2} = 57{,}5$$

Mode

35 38 39 40 44 45 50 53 55 57 58 59 61 62 64 64 67 71 74 77

2 times

Mode = 64

Mean for Ungrouped Data

For ungrouped or raw data, the mean has the following formula.

$$\bar{x} = \frac{\sum x}{n}$$

where $\bar{x}$ = mean

$\sum x$ = sum of the measurements or values

n = number of measurements

Example 1:

Ms. Sulit collects the data on the ages of Mathematics teachers in Santa Rosa School, and her study yields the following:

38 35 28 36 35 33 40

Solution:

$$\bar{x} = \frac{38 + 35 + 28 + 36 + 35 + 33 + 40}{7}$$

$$= 35$$

Mean for Grouped data

$$\mu = \sum \frac{x_i \, n_i}{N}$$

where:

x_i is the value of each category

n_i is the frequency of each category (you can find the symbol of the frequency also as f_i)

N the total number of observation

Example 2: Given the following distribution: 4, 2, 4, 2, 6, 4, 0, 4, 0, 2, 4, 4.
The frequency distribution is:

x_i	n_i	$x_i n_i$
0	2	0
2	3	6
4	6	24
6	1	6
	12	**36**

To calculate the mean we have to multiply the value of each class for its frequency

The mean of the distribution is:

$$\mu = \sum \frac{x_i n_i}{N} = \frac{36}{12} = 3$$

Mean for Grouped data (in classes)

$$\text{Class midpoint or mark} = \frac{\text{Lower limit} + \text{Upper limit}}{2}$$

$$\mu = \sum \frac{m_i \, n_i}{N}$$

where:

m_i is the midpoint of each class

n_i is the frequency of each category (you can find the symbol of the frequency also as f_i)

N the total number of observation

Employees	Companies Frequency *f*	Midpoint *m*	*f* x *m*
1 - 5	1	3	3
6 - 10	3	8	24
11 - 15	16	13	208
16 - 20	4	18	72
21 - 25	1	23	23
TOTALS	25		330

To calculate the mean we have to find the midpoint of each class and multiply it for the frequency of each class

$$\mu = \sum \frac{m_i \, f_i}{N} = \frac{330}{25} = 13.2$$

3.53 For 50 airplanes that arrived late at an airport during a week, the time by which they were late was observed. In the following table, x denotes the time (in minutes) by which an airplane was late, and f denotes the number of airplanes.

Find the mean.

x_i	f_i
0 ├ 20	14
20 ├ 40	18
40 ├ 60	9
60 ├ 80	5
80 ├ 100	4

x_i	f_i	m_i	$m_i f_i$
0 ├ 20	14	10	140
20 ├ 40	18	30	540
40 ├ 60	9	50	450
60 ├ 80	5	70	350
80 ├ 100	4	90	360
	50		1840

Find the midpoint of the classes

Calculate the mean for grouped data

$$\mu = \frac{\sum m_i f_i}{N}$$

$$\mu = \frac{\sum m_i f_i}{N} = \frac{1840}{50} = 36.8$$

Properties of the Mean

1. $min(x_i) \leq \bar{x} \leq max(x_i)$ ➔ internality

2. $n\bar{x} = \sum x_i$ (or, for grouped data, $n\bar{x} = \sum x_i n_i$)

This derives directly from the definition of (simple or weighted) mean

3. $\sum(x_i - \bar{x}) = 0$ (or, for grouped data, $\sum(x_i - \bar{x})\, n_i = 0$)

Indeed, $\sum x_i - \sum \bar{x} = n\bar{x} - n\bar{x} = 0$

4. If $y = a + bx$, then $\bar{y} = a + b\bar{x}$

Proof $\bar{y} = \frac{1}{n}\sum y_i = \frac{1}{n}\sum(a + bx_i) = \frac{1}{n}\sum a + \frac{1}{n}\sum bx_i = \frac{1}{n}na + b\frac{1}{n}\sum x_i = a + b\bar{x}$

5. $\bar{x}$ is that value of c that minimizes $\sum(x_i - c)^2$

Proof: $\frac{\partial}{\partial c}\sum(x_i - c)^2 = 2\sum(x_i - c)(-1) = \sum(x_i - c) = 0$

$\sum x_i - \sum c = \sum x_i - nc = 0$ ➔ $c = \bar{x}$

5. Combined mean: $\bar{x} = \frac{\bar{x}_1 n_1 + \bar{x}_2 n_2 + \cdots + \bar{x}_k n_k}{n_1 + n_2 + \cdots + n_k}$

*3.25 The mean age of six persons is 52 years. The ages of five of these six persons are 56, 38, 43, 51, and 36 years, respectively. Find the age of the sixth person.

Solution

n=6

56 38 43 51 36 6^{th}?

The mean in 52: μ=52

From the property of the mean:

$$\Sigma(xi - \mu) = 0$$

Indeed,

$$\Sigma x_i = n\mu$$

$$n\mu = 6 \cdot 52 = 312$$

$$\Sigma x_i = 224 + (6^{th}\, person)$$

$$312 = 224 + (6^{th}\, person) \rightarrow \quad 312 - 224 = 88$$

The 6^{th} person is 88 years old

Exercise 3

Given the following distribution:

4	2	4	2	6	4	0	4	0	2	4	4

1. calculate the median and the arithmetic mean using the unitary distribution;
2. calculate the mean using the frequency distribution;
3. check the property of the mean: $\sum(x-\mu)=0$.

Solution

1. Calculate the Median

Rank the data

0	0	2	2	2	4	4	4	4	4	4	6
#1	*#2*	*#3*	*#4*	*#5*	*#6*	*#7*	*#8*	*#9*	*#10*	*#11*	*#12*

$$prof(med)=\frac{n+1}{2}=\frac{13}{2}=6.5, \qquad med=\frac{x_{(6)}+x_{(7)}}{2}=\frac{4+4}{2}=4.$$

Exercise 3

Given the following distribution:

4 2 4 2 6 4 0 4 0 2 4 4

1. calculate the median and the arithmetic mean using the unitary distribution;
2. calculate the mean using the frequency distribution;
3. check the property of the mean: $\sum(x-\mu)=0$.

Solution

1. Calculate the Mean

$$\sum_{i=1}^{12} x_i = 4+2+4+\cdots+4 = 36.$$

$$\mu = \sum \frac{x}{N} = \frac{36}{12} = 3$$

Exercise 3

Given the following distribution:

4 2 4 2 6 4 0 4 0 2 4 4

1. calculate the median and the arithmetic mean using the unitary distribution;
2. calculate the mean using the frequency distribution;
3. check the property of the mean: $\sum(x-\mu)=0$.

Solution

2. Calculate the Mean using the frequency distribution

x_i	n_i	$x_i n_i$
0	2	0
2	3	6
4	6	24
6	1	6
	12	36

$$\mu = \sum \frac{x_i n_i}{N} = \frac{36}{12} = 3$$

Exercise 3

Given the following distribution:

4 2 4 2 6 4 0 4 0 2 4 4

1. calculate the median and the arithmetic mean using the unitary distribution;
2. calculate the mean using the frequency distribution;
3. check the property of the mean: $\sum(x-\mu)=0$.

Solution

3. Verify the property of the mean $\sum(x-\mu)=0$

x_i	n_i	$x_i n_i$	$(x_i-\bar{x})$	$(x_i-\bar{x})n_i$
0	2	0	-3	-6
2	3	6	-1	-3
4	6	24	1	6
6	1	6	3	3
	12	**36**		**0**