AdTrans project , 2023-1-PL01-KA220-HED-000158917

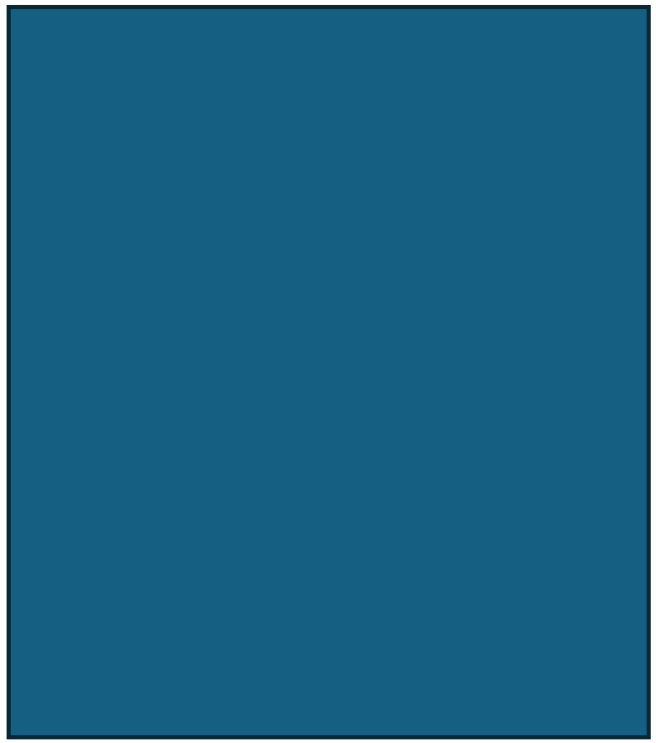
ADTRANS
ADVANCED DATA ANALYSIS FOR TRANSPORTATION AND MOBILITY ACADEMIC TEACHERS

Mathematical Statistics
Measures of Position

Anna Gobis

Co-funded by the European Union

Descriptive statistics



Measures of variability

- Variance
- Standard deviation
- Coefficient of variation

Asymmetry measures

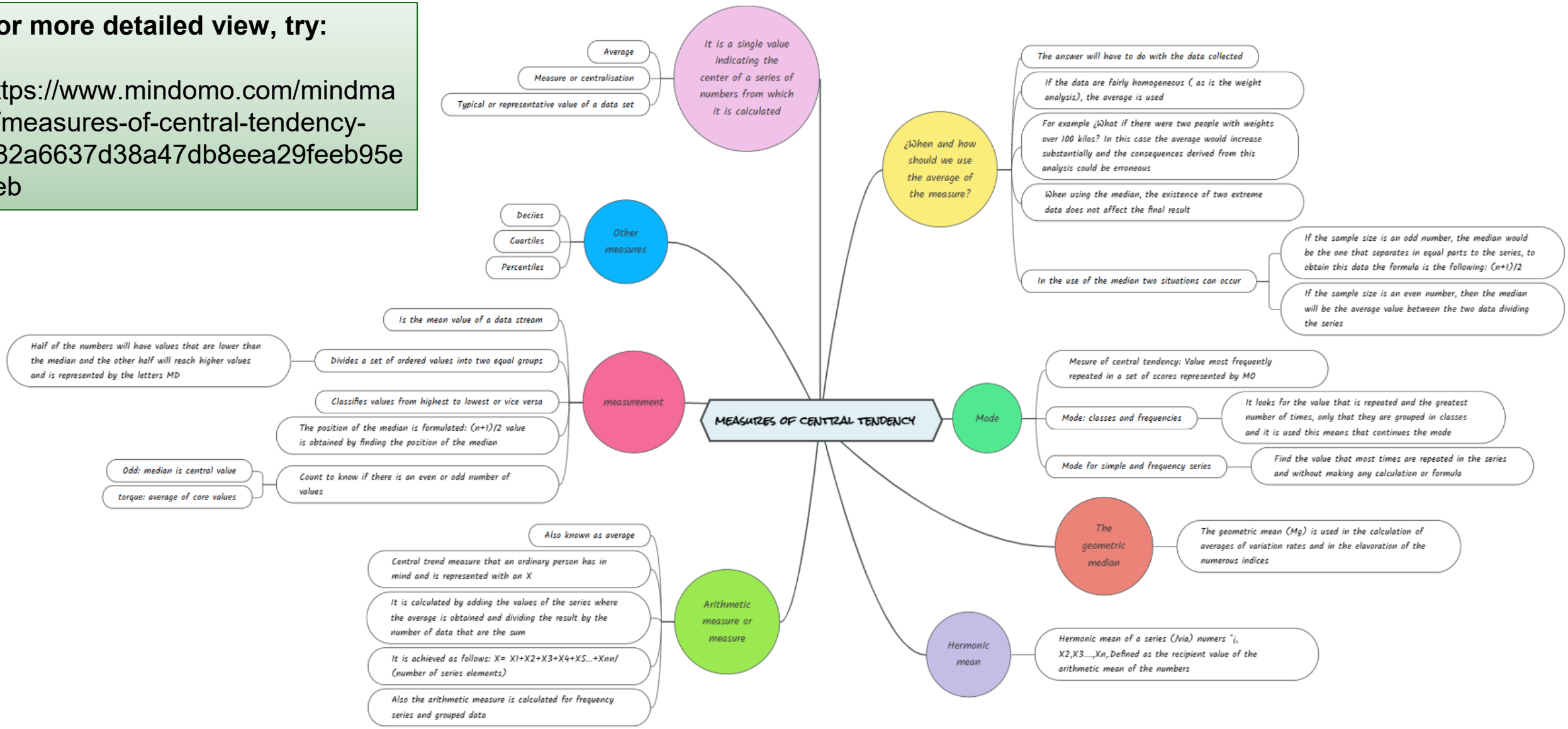
- Asymmetry coefficient
- Skewness coefficient

Concentration Measures

- Kurtosis

For more detailed view, try:

https://www.mindomo.com/mindmap/measures-of-central-tendency-482a6637d38a47db8eea29feeb95eceb



Measures of position - classic

Arithmetic Mean:

Detailed series: $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

Distribution series: $\bar{x} = \frac{1}{n}\sum_{i=1}^{k} \dot{x}_i n_i$

Measures of position - classic

Harmonic Mean

Detailed series:

$$= \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$$

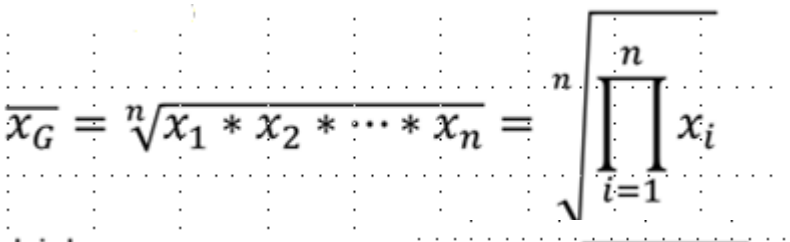
Distribution series:

$$= \frac{n}{\sum_{i=1}^{k} \frac{n_i}{x_i}}$$

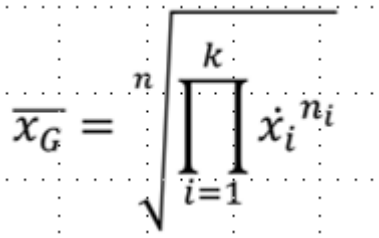
Measures of position - classic

Geometric Mean

Detailed series:



Distribution series:



Example 2.1

10 men were selected and asked about their height. The results were as follows: 182; 170; 171; 165; 179; 187; 169; 168; 181; 177.

Calculate the average height of men in the randomly selected sample.

Example 2.2

The cyclist travels at a speed of 20 km/h for the first 4 km. He travels at a speed of 10 km/h for the next 4 km. Calculate the cyclist's average speed.

Example 2.3

Prices of a certain product were recorded for several months (table).
Calculate the average relative increase in prices.

Month	Price
[redacted]	
2	18.77
[redacted]	
4	21.03
[redacted]	
6	22.86

Measures of position - positional

Median

Detailed series:

$$Me = \begin{cases} x_{\frac{n+1}{2}} \\ \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2} \end{cases}$$

⇦ when *n* is odd

⇦ when *n* is even

Distribution series: $Me = x_m^- + \frac{i}{n_m}\left[\frac{n}{2} - c_{m-1}\right]$

- x_m^- – lower boundary of the interval containing the median
- i – class interval
- n_m – frequency of the interval containing the median
- c_{m-1} – cumulative frequency up to the class preceding the one containing the median

Measures of position - positional

Mode (modal value, mode)

Detailed series:

Most frequently occurring value

Distribution series: $$Mo = x_D^- + \frac{n_D - n_{D-1}}{2n_D - n_{D-1} - n_{D+1}} * i$$

- x_D^- – lower boundary of the interval containing the mode
- n_D – frequency of the interval containing the mode
- n_{D-1} – frequency of the preceding interval
- n_{D+1} – frequency of the following interval

Measures of position - positional

Quantiles

Distribution series: $$q_p = x_q^- + \frac{i}{n_q}\left[p * n - c_{q-1}\right]$$

- x_Q^- – lower boundary of the interval containing the quantile
- i – class interval
- n_Q – frequency of the interval containing the quantile
- c_{Q-1} – cumulative frequency up to the class preceding the one containing the quantile
- q_p – quantile of order p, p – order of the quantile

Mean of Grouped Data

Grade	f	mid.	cf
40-49	3	44.5	3
50-59	5	54.5	8
60-69	6	64.5	14
70-79	9	74.5	23
80-89	8	84.5	31
90-100	7	95	38

$$\bar{x} = \frac{\text{Sum}}{n}$$

$$\text{Mean} = \frac{\sum f \cdot m}{\sum f}$$

Example 2.4

15 students were randomly selected and asked how much they spent on entertainment last month. The following values were obtained [PLN]: 109; 57; 79; 125; 122; 90; 109; 99; 87; 108; 99; 109; 118; 93; 109. Determine the detailed series. Calculate the measures of the location of the distribution of expenses on entertainment for this group of students

Example 2.5

10 men were randomly selected and asked about their height. The following results were obtained:

182; 170; 171; 165; 179; 187; 169; 168; 181; 177

Calculate quartiles .

Example 2.6

We have data for 50 employees regarding the time it takes to perform a certain task [min]. Calculate the average value. Calculate the modal and quartiles, indicate the median.

Time [min]	Number of employees
8-9	7
10-11	12
12-13	6

Example 2.7

Below are the results of a colloquium for a group of 30 students in the form of a detailed series. The maximum possible number of points was 10.

1	**3**	**3**	**3**	**3**	**4**	**4**	**4**	**4**	**4**	**4**	**5**	**5**	**5**	**5**
5	**6**	**6**	**6**	**6**	**7**	**7**	**7**	**7**	**8**	**8**	**9**	**9**	**10**	**10**

Build a grouped series. For both the detailed and grouped series, calculate: arithmetic mean, median, and quartiles. Determine the variability range of the characteristic ($Me \pm Q$). Present the distribution in the form of a box-and-whisker plot (median, quartiles, min, max).

Example 2.4

15 students were randomly selected and asked how much they spent on entertainment last month. The following values were obtained [PLN]: 109; 57; 79; 125; 122; 90; 109; 99; 87; 108; 99; 109; 118; 93; 109. Determine the detailed series. Calculate the measures of the location of the distribution of expenses on entertainment for this group of students

Example 2.5

10 men were randomly selected and asked about their height. The following results were obtained:

182; 170; 171; 165; 179; 187; 169; 168; 181; 177

Calculate quartiles .

Additional tasks

Task 2.1

Twelve temperature measurements (in degrees Fahrenheit) were taken and the following results were obtained: 440; 432; 412; 405; 415; 427; 409; 417; 422; 411; 439; 411

a) Calculate the average of the measurements

b) Recalculate the average of the measurements by first subtracting 400 from each value, and then adding 400 to the average of the results obtained in this way.

Task 2.2

Kowalski invested money in three accounts with different rates of return: PLN 5,000 at 3% per year, PLN 10,000 at 4% and PLN 10,000 at 4.5% per year. Determine the average percentage of profit received from these lectures

Reply 2.1

a) X=420

b) X=420

Reply 2.2

X=4%

Task 2.3

The following numbers refer to student absences for fifteen consecutive days: 39, 55, 46, 38, 41, 40, 33, 49, 42, 37, 49, 44, 51, 48, 44.

Calculate the mean, median and mid-range

Task 2.4

A fundraiser was organized for 12 people for a certain amount. The following numbers indicate the percentage of the actual amounts collected: 94, 110, 80, 98, 95, 108, 101, 460, 75, 105, 110, 85.

Calculate the median and mean of these numbers and indicate which of these measures is a better definition of the "average" collection results.

Reply 2.3

X=43.7, Me=44, R=44

Reply 2.4

X=126.75%, Me=99.5%, median is better

Sources:

- Kot, SM, Jakubowski, J., Sokołowski, A.: Statistics. Second revised edition. Difin Publishing House . Warsaw, 2011.
- Stanisz, A.: Accessible statistics course using STATISTICA PL on examples from medicine. Volume 1. Basic statistics. StatSoft Polska Publishing House. Cracow, 2006.
- Flisikowski , K.: Statistical methods in business. Materials of the AdvancedPhD project . Gdańsk University of Technology, 2015.
- Zimny, A.: Descriptive statistics. Supporting materials for exercises. Second revised edition. State Higher Vocational School in Konin, 2010.
- Freund, JE: Fundamentals of Modern Statistics, State Economic Publishing House, Warsaw 1968