

PROBLEM SET 1 – DATA UNDERSTANDING

Problem 1. Suppose that the data for analysis includes the attribute age. The age values for the data tuples are (in increasing order) {13, 15, 16, 16, 19, 20, 20, 21, 22, 22, 25, 25, 25, 25, 30, 33, 33, 35, 35, 35, 35, 36, 40, 45, 46, 52, 70}.

- What is the *mean* of the data? What is the *median*?
- What is the *mode* of the data? Comment on the data's modality (i.e., bimodal, trimodal, etc.).
- What is the *midrange* of the data?
- Can you find (roughly) the first quartile (Q_1) and the third quartile (Q_3) of the data?
- Give the *five-number summary* of the data.
- Show a *boxplot* of the data.
- What is the *variance* and *standard deviation* of the data?

Problem 2. Suppose that a hospital tested the age and body fat data for 18 randomly selected adults with the following results:

<i>age</i>	23	23	27	27	39	41	47	49	50
<i>%fat</i>	9.5	26.5	7.8	17.8	31.4	25.9	27.4	27.2	31.2
<i>age</i>	52	54	54	56	57	58	58	60	61
<i>%fat</i>	34.6	42.5	28.8	33.4	30.2	34.1	32.9	41.2	35.7

- Calculate the mean, median, variance, and standard deviation of *age* and *%fat*.
- Draw the boxplots for *age* and *%fat*.
- Draw a *scatter plot* for these two variables.
- Calculate the Pearson and Spearman correlation coefficients between the two attributes.

Problem 3. Let D (in the previous slide) be the number of appearances, goals, and assists of Cristiano Ronaldo in Champions League from season 2003/2004 to 2021/2022.

- Calculate the *mean*, *median*, *mode*, *variance*, *standard deviation* of three attributes appearances, goals, and assists.
- Find the *first quartile* (Q_1) and the *third quartile* (Q_3) of three attributes appearances, goals, and assists.
- Show the *boxplots* of the three attributes appearances, goals, and assists.
- Show the *scatter plots* of appearances and goals; of appearances and assists; of goals and assists.
- Compute the *covariance matrix* of the three attributes appearances, goals, and assists.

$D =$

<i>Club</i>	<i>Season</i>	<i>Appearances</i>	<i>Goals</i>	<i>Assists</i>
Man United	03/04	5	0	1
Man United	04/05	7	0	2
Man United	05/06	6	0	0
Man United	06/07	11	3	5
Man United	07/08	11	8	1
Man United	08/09	12	4	3
Real Madrid	09/10	6	7	1
Real Madrid	10/11	12	6	4
Real Madrid	11/12	10	10	3
Real Madrid	12/13	12	12	1
Real Madrid	13/14	11	17	5
Real Madrid	14/15	12	10	4
Real Madrid	15/16	12	16	4
Real Madrid	16/17	13	12	6
Real Madrid	17/18	13	15	3
Juventus	18/19	9	6	2
Juventus	19/20	8	4	1
Juventus	20/21	6	4	2
Man United	21/22	7	6	0