FUNDAMENTALS OF STATISTICS (4ECON013C)

WEEK 1 TUTORIAL

1. Decide whether the statement makes sense (or is clearly true) or does not make sense (or is clearly false). Explain clearly; not all of these statements have definitive answers, so your explanation is more important than your chosen answer.

a) A data set of incomes has modes of \$50,000 and \$80,000. (makes sense)

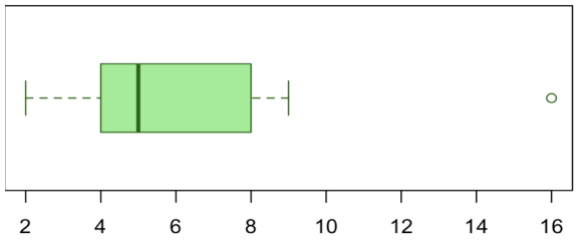
b) A researcher studying an income distribution obtains the same value of \$75,000 for the mean, median, and mode. (makes sense)

c) Jennifer received an SAT score that was equal to the first quartile and the 35th percentile. (does not make sense)

d) The house key lengths of 15 statistics students are measured and rounded to the nearest centimeter, and all 15 values are the same, so the standard deviation is 0 cm. (makes sense)

e) When there are outliers in the dataset, the median is a better measure of central tendency compared to the mean (true).

f) The following boxplot shows a negatively skewed data (false).



2. The monthly incomes of a sample of middle-management employees at KPMG Uzbekistan are given below:

Monthly salary (in USD)	# of employees
\$2000	5
\$1900	4
\$3000	3
\$2500	6

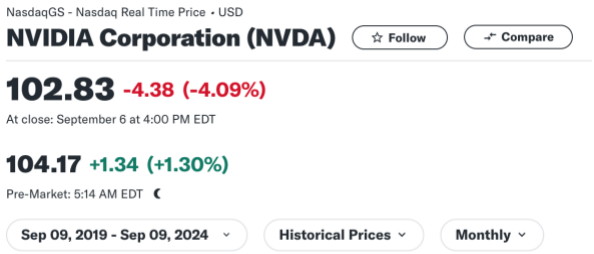
a) Compute the sample mean (\$2311), median (\$2250) and mode (\$2500).

b) If $S_{xx} = \sum(x_i - \bar{x})^2 = 2{,}797{,}778$,

then find the sample variance ($= S_{xx}/(n-1) = 164{,}575.18$) and standard deviation (\$405.68).

3. The following data provides the monthly closing stock price of NVIDIA Corporation (NVDA) rounded to the nearest integer:

Month	Close Price ($)
Sep 1, 2024	$103
Aug 1, 2024	119
Jul 1, 2024	117
Jun 1, 2024	124
May 1, 2024	110
Apr 1, 2024	86
Mar 1, 2024	90
Feb 1, 2024	79
Jan 1, 2024	62
Dec 1, 2023	50
Nov 1, 2023	47
Oct 1, 2023	41
Sep 1, 2023	44



a. Find the range (124 – 41 = $83).
b. Find the IQR. $Q_1 = 50$, $Q_3 = 110$, IQR = 60
c. Provide the five-number summary: $41, 50, 86, 110, 124

4. Annual incomes (defined as X, in thousand $) of randomly surveyed households are given in the following table:

52	23	66	41	93	27	98	84
22	77	80	45	30	99	42	46
88	26	53	58	78	47	44	66

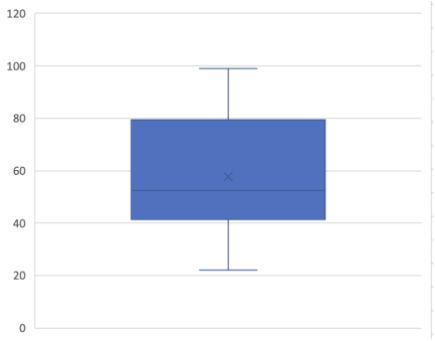
If $\sum x = 1{,}385$ and $\sum x^2 = 93{,}905$, do the following:
a. Find the mean (57.71) and median (52.5).
b. Compute sample variance and standard deviation.

$s^2 = \frac{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}{n-1} = \frac{93905 - \frac{1385^2}{24}}{24-1} = 607.78$, s = 24.65

c. Plot the box plot and comment on the presence of outliers.

Q1 = 41.75
Q3= 78.5
IQR = 36.75

(Q1 - 1.5IQR, Q3+1.5IQR) => (-13.38, 133.63)
So, no outliers in the dataset.



5. Suppose two-bedroom apartment rental prices in Tashkent have an average of $350 with a standard deviation of $60. If nothing is known about the shape of the distribution, at least what percentage of rental prices lie between $260 and $440?

260 = 350 – k*60 => k = 1.5
1 – 1/2.25 = 0.556 or 55.6%

6. The mean score on an accounting test is 80, with a standard deviation of 10. Suppose the scores follow a bimodal distribution. Between which two scores must this mean lie to represent at least 96% of the data set? Note that the student scores can range between 0 and 100.

$1 - \frac{1}{k^2} = 0.96$
k = 5
80 – 5*10 = 30
80 + 5*10 = 130, but maximum score is 100, so the interval is (30, 100).

7. Work in groups of 3-4 students to discuss this task.
All 100 first-year students at a small college take three modules in the Core Studies program. Two modules are taught in large lectures, with all 100 students in a single class. The third module is taught in 10 classes of 10 students each. Students and college administrators get into an argument about whether classes are too large. The students claim that the mean size of their Core Studies classes is 70. The administrators claim that the mean class size is only 25. Can both sides be right? Explain.

Solution. Yes, both can be right.
The students calculated the mean class size per student:

$$\frac{\text{total enrollment in student's classes}}{\text{number of classes student is taking}} = \frac{100 + 100 + 10}{3} = 70$$

While the administrators calculated the mean number of students per class. There are two classes with 100 students and 10 classes with 10 students, making a total enrollment of 300 students in 12 classes.

$$\frac{\text{total enrollment}}{\text{number of classes}} = \frac{300}{12} = 25$$

HOMEWORK

8. For the distribution drawn below comment on the skewness of the plots. Identify the points for mean, median and mode.

A, F – mode
B, E – median
C, D – mean

9. The manager of a local RV sales lot has collected data on the number of RVs sold per month for the last five years. That data is summarized below:

# of Sales	0	1	2	3	4	5	6
# of Months	2	6	9	13	21	7	2

What is the weighted mean number of sales per month?
$\bar{x}$ = (3.23)

10.
The data is given regarding the distance in miles between exits on I-75 in Kentucky.
Find the modal distance.

11	4	10	4	9	3	8	10	3	14	1	10	3	5
2	2	5	6	1	2	2	3	7	1	3	7	8	10
1	4	7	5	2	2	5	1	1	3	3	1	2	1

What is the modal distance?

Answer: 1

11. A bored carpenter counts the actual number of nails in 10 boxes of nails and records his findings as: 230, 235, 302, 287, 312, 323, 265, 319, 342, and 298. What can we say about the shape of the distribution of the number of nails?
Answer: skewed to the left, because mean = 291.3 < median = 300

12. Suppose Adam has assessment components midterm exam, group project and final exam with 30%, 30% and 40% weights respectively. He received 58 from midterm and 72 from the group project. What is the minimum score he needs from final exam to receive at least 70 overall mark?
Overall=0.30×58+0.30×72+0.40×F
F ≥ 78 (rounded to integer)

13.
The following data is given.

1	11.5	6	7.2	4	8	9
10	6.8	8.3	2	2	10	1

a. Find the range.
b. Find the IQR.
c. Provide five-number summary.

Answers.
a. Range = 11.5 – 1 = 10.5
b. Sorted data: 1 1 2 2 4 6 6.8 7.2 8 8.3 9 10 10 11.5
Location of Q1: (14-1)/4 + 1 = 4.25
Location of Q3: (14-1)*0.75+1= 10.75
Q1 = 2.5, Q3 = 8.825
IQR = 6.325
c. Five-number summary: 1, 2.5, 7, 8.825, 11.5