NAME: Ibrahim Shalash

ID: UD53394SEN62351

Course Title:

Statistics I: Frequency Distribution & Data Description

.

SCHOOL: Science & Engineering

Major: Environmental Science

Atlantic International University

July 10, 2019

Contents

Introduction

Statistics is the branch of mathematics that collects, tabulates, analyzes and displays data in order to provide scientists, politicians, scholars, and decision- makers with information and data on all matters related to people's lives and livelihoods.

Statistics are included in all areas of life. In the field of health, statistical data collection and analysis of data on chronic diseases and genetic diseases in order to develop appropriate plans and strategies to combat these diseases and reduce their spread.

In sports, the science of statistics analyzes the data on athletes in terms of speed, performance and endurance and follow-up effect of stimulant drugs on the performance of players.

In pharmacology, the science of statistics collects and analyzes data on the uses of drugs and their effects on human health, which helps in making appropriate decisions about the use of certain drugs before putting them on the market.

In agriculture and food, the science of statistics is applied to the evaluation of agricultural products and the follow-up and evaluation of food products in terms of their impact on the fight against famine in order to reduce health risks.

Statistics is one of the most important sciences and tools in the field of economics, finance and business at the personal level, and the global level, where data is collected on projects related to these areas around the clock in order to follow the success or failure of those projects.

In education, statistics play an important role in monitoring, and developing methods of education, monitoring student achievement and teacher performance. Monitoring curricula and

educational tools and their effects on students and teachers. Developing plans for measuring, evaluating and developing the educational process

The science of statistics is also included in other fields and sciences of interest to humans, such as weather, climate change, heat stress and space science in order to increase human awareness towards environmental and societal issues, so as to reduce and illuminate adverse effects on environment, which result from the wrong human practices.

In this exam, three topics are of concern, the first topic is about the nature of probability and statistics, which concerns about the types of data, statistical terms, variables and their nature, sampling techniques and methods of study. The second topic is about data analysis, frequency distribution, presentation and interpretation of data. The third topic concerns about data description by means of variance, standard deviation, percentile and quartile.

Chapter 1

The Nature of Probability and Statistics

For Exercises 1–3, state whether descriptive or inferential statistics have been used.

1. By 2040 at least 3.5 billion people will run short of water (World Future Society).

   Answer:

   This is inference statistics because it predicts the number of people going into a shortage of water.

2. In a sample of 100 on-the-job fatalities, 90% of the victims were men.

Answer:

Descriptive statistics have been used, as the value, 90% is taken from the 100 samples.

3. In a survey of 1000 adults, 34% said that they posted notes on social media websites.

   Answer:

   Descriptive statistics have been used because 34% of adults who posted notes on social media websites was taken from the 1000 sample.

   For Exercises 4–7, classify each as nominal-level, ordinal level, interval-level, or ratio-level measurement.

4. Pages in the 25 best-selling mystery novels.

   Answer:

Ratio level has been used, because the count of pages of the novels can be done in ratios as the first novel is twice the third one or the last novel has half the pages of number 15 novel.

5. Rankings of golfers in a tournament.

   Answer:

   Ordinal scale level, because it describes compatible differences between golfers.

6. Temperatures inside 10 pizza ovens.

   Answer:

   Interval level because temperature measurement can be measured over a scale interval.

7. Weights of selected cell phones.

   Answer:

   Ratio level because cell phone may have different weights so there can be ratios of weights between one cell phone and another.

   For Exercises 8–11, classify each variable as qualitative or quantitative.

8. Marital status of nurses in a hospital.

   Answer:

   Qualitative variable because it describes a social status and can't be measured or given a value.

9. The time it takes to run a marathon.

   Answer:

Quantitative variable, because time can be measured accurately and can have a numerical value.

10. Weights of lobsters in a tank in a restaurant.

Answer:

Quantitative variable because the weights of lobsters can be measured accurately by balances.

11. Colors of automobiles in a shopping center parking lot.

Answer:

Qualitative variable because it gives a description of the cars and not a number or a count number.

For Exercises 12–15, classify each sample as random, systematic, stratified, cluster, or other.

12. In a large school district, all teachers from the two buildings are interviewed to determine whether they believe the students have less homework to do now than in previous years.

Answer:

Cluster method of data collection had been used. As the whole group of teachers has participated in the interview.

13. . Every seventh customer entering a shopping mall is asked to select her or his favorite store.

Answer:

A systematic method of data collection had been used. Because the choice of customers was sequential and takes a specified interval

14. Nursing supervisors are selected using random numbers to determine annual salaries.

Answer:

Random way of data collection had been used. The choice of nurses was random and was not limited to a certain group or individuals.

15. Every 100th hamburger manufactured is checked to determine its fat content.

Answer:

A systematic method of data collection had been used. Because the choice of samples was sequential.

For Exercises 16–17, identify the independent and dependent variables for each study.

16. Various types of coffees are selected from local coffee shops, and the number of milligrams of caffeine per ounce is determined.

Answer:

The independent variable is the type of coffee,

The dependent variable is the concentration of caffeine in each type. Because caffeine content depends on the type of coffee.

17. People who walk at least 3 miles a day are randomly selected, and their blood triglyceride levels are measured in order to determine if the number of miles that they walk has any influence on these levels.

The independent variables are people who walk, and the number of miles they walk.

The dependent variable is the concentration of triglyceride in their blood.

Chapter 2

Frequency Distributions and Graphs

18. BUN Count The blood urea nitrogen (BUN) count of 20 randomly selected patients is given here in milligrams per deciliter (mg/dl). Construct an ungrouped frequency distribution for the data.

17	18	13	14
12	17	11	20
13	18	19	17
14	16	17	12
16	15	19	22

Answer:

The lower limit of Blood Urea Nitrogen (BUN) is = 11 mg/dl

The upper Limit of BUN is = 22 mg/dl

The range is R= 22-11 = 11 mg/dl, since the range is small, so classes of single number data can be used.

The lower boundary of data is calculated by subtracting 0.5 from the lower limit, or 11-0.5 = 10.5 mg/dl

The upper boundary is calculated by adding 0.5 to the upper limit, or:

22+ 0.5 = 22.5 mg/dl

The values of BUN are tallied, and the frequencies are calculated as in table 1 below. From table 1, it is obvious that there is a peak on BUN count 17 mg/dl.

Table 1: Tallied data and ungrouped frequency distribution of BUN

Class limits	Class boundaries	Tally	Frequency
11	**10.5-11.5**	/	**1**
12	**11.5-12.5**	//	**2**
13	**12.5-13.5**	//	**2**
14	**13.5-14.5**	//	**2**
15	**14.5-15.5**	/	**1**
16	**15.5-16.5**	//	**2**
17	**16.5-17.5**	////	**4**
18	**17.5-18.5**	//	**2**
19	**18.5-19.5**	//	**2**
20	**19.5-20.5**	/	**1**
21	**20.5-21.5**		**0**
22	**21.5-22.5**	/	**1**
			20

The cumulative frequency distribution is shown in table 2

Table 2: Cumulative frequency distribution for BUN

	Cumulative frequencies
less than 10.5	0
less than 11.5	1
less than 12.5	3
less than 13.5	5
less than 14.5	7
less than 15.5	8
less than 16.5	10
less than 17.5	14
less than 18.5	16
less than 19.5	18
less than 20.5	19
less than 21.5	19
less than 22.5	20

19. Wind Speed The data show the average wind speed for 30 days in a large city. Construct an ungrouped frequency distribution for the data.

8	15	9	8	9	10
8	10	14	9	8	8
12	9	8	8	14	9
9	13	13	10	12	9
13	8	11	11	9	8
9	13	9	8	8	10

Answer:

(Note: There is no specified unit for wind speed in the question, I will assume it as km/h)

The lower limit of wind speed is = 8 km/h

The upper limit of wind speed is =15 km/h

The range R = Upper limit – Lower limit = 15 - 8 = 7 km/h

Because R is small, then single numbers classes are used

The lower boundary is = 8 - 0.5 = 7.5 km/h

The upper boundary is = 15 + 0.5 = 15.5 km/h

The wind speed data is tallied in table 3, which shows that about two thirds of wind speed values lie between 7.5 km/h and 9.5 km/h.

Table3: Tally and frequency distribution of grouped Wind speed frequency

Class limits of wind speed	**Class boundaries of wind speed(**km/h)	**Tally**	**Frequency**
8	**7.5-8.5**	/// /// /	**11**
9	**8.5-9.5**	/// ///	**10**
10	**9.5-10.5**	////	**4**
11	**10.5-11.5**	//	**2**
12	**11.5-12.4**	//	**2**
13	**12.5-13.5**	////	**4**
14	**13.5-14.5**	//	**2**
15	**14.5-15.5**	/	**1**
			36

Table 4: Cumulative frequency distribution of wind speed

	Cumulative frequencies
less than 7.5	0
less than 8.5	11
less than 9.5	21
less than 10.5	25
less than 11.5	27
less than 12.5	29
less than 13.5	33
less than 14.5	35
less than 15.5	36

20. Spending of College Freshmen, the average amounts spent by college freshmen for school items are shown. Construct a pie graph for the data.

Electronics/computers	$728
Dorm items	344
Clothing	141
Shoes	72

Answer:

The pie chart for spending in freshman level is shown in figure 1

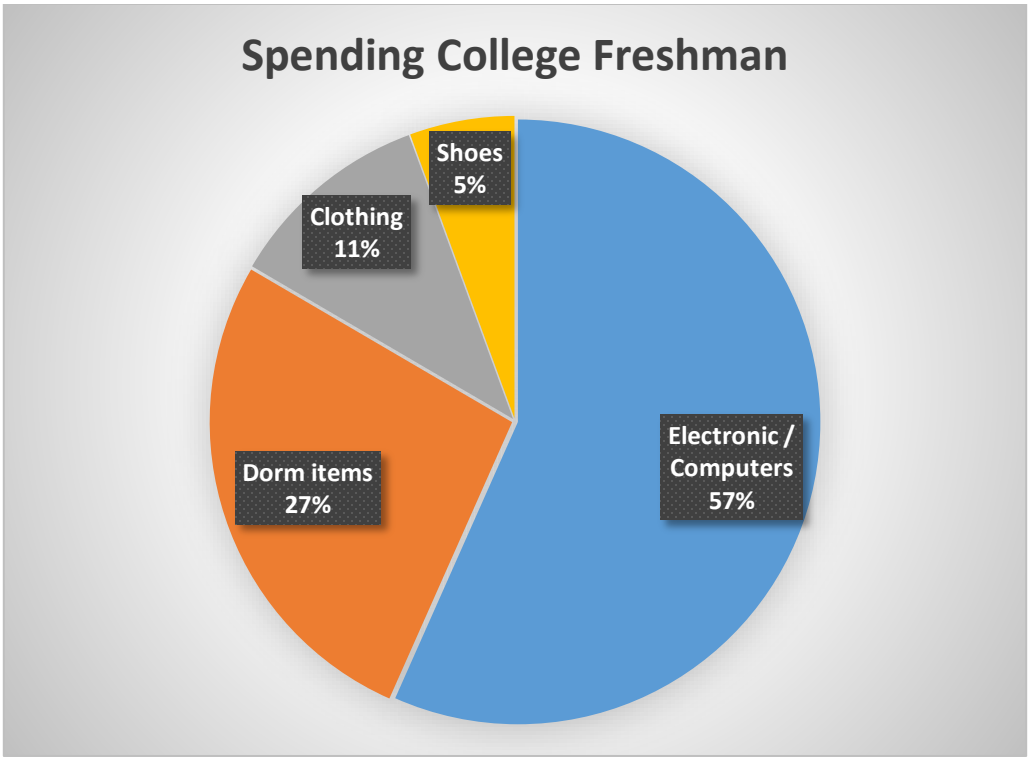

Figure 1: Spending in College Freshman

The chart shows that about 57% of expenses in Freshman College, go to electronic and computers, and dormitory expenses take 27%. While clothing and shoes take about 16%.

Chapter 3

Data Description

21. Net Worth of Wealthy People The net worth (in billions of dollars) of a sample of the richest people in the United States is shown. Find the mean, median, mode, and midrange for the data.

59	52	28	26	19
19	18	17	17	17

Answer:

The data is arranged in ascending order

17, 17, 17, 18, 19, 19, 26, 28, 52, 59

Mean $\overline{X} = \frac{\Sigma x}{N} = \frac{(17+17+17+18+19+19+26+28+52+59)}{\mathbf{10}} = \frac{\mathbf{272}}{\mathbf{10}}$ = \$27.2 billion

The median: Because the data set count is even number,

The median = $\frac{19+19}{2}$ = \$19 billion

The mode: the number 17 is highest frequency number, so mode= \$17 billion

The midrange (MR) = $\frac{(\text{highest value} + \text{lowest Value})}{2}$

MR = $\frac{59+17}{2} = \frac{\mathbf{76}}{\mathbf{2}}$ = \$38 billion

22. Shark Attacks The number of shark attacks and deaths over a recent 5-year period is shown. Find the mean, median, mode, and midrange for the data.

Attacks	71	64	61	65	57
Deaths	1	4	4	7	4

Answer:

This question has two data sets

1- The shark attacks, 71,64 .61.65.57
2- The deaths, 1,4,4,7,4

- For the first data set:

Arranging data set into ascending order, x_i = 57,61,64,65, 71

The Mean $\overline{X} = \frac{\Sigma xi}{n} = \frac{318}{5}$ = 63.5 attacks

The median; since the data set count is odd number, so the median = 64 attacks

This data set has no mode.

The Midrange (MR) = $\frac{\text{Upper limit} - \text{Lower limit}}{2} = \frac{71-57}{2}$ = 7 attacks

- For the second data set

Arranging data into ascending order, x_i = 1, 4, 4, 4, 7

The Mean $\overline{X} = \frac{\Sigma xi}{n} = \frac{20}{5}$ = 4 deaths

The median; since the data set count is odd number, so the median = 4 attacks

This data set multimodal.

The means of both data sets, shark attacks and deaths resulted from the attacks, indicate that for 63.5 attacks there are 4 deaths only.

23. Tornado Occurrences The data show the number of tornados recorded for each month of a specific year. Find the range, variance, and standard deviation for the data.

33	10	62	132	123	316	123	133	18	150
26	138								

Answer:

The data set are sorted in ascending order:

10, 18, 26, 33, 62, 123, 123, 132, 133, 138, 150, 316

Range = Upper limit – Lower limit = 316 – 10= 306 tornados

The procedure to find variance and standard deviation is illustrated in table 5

Table 5: Calculations for variance and standard deviation for tornados

No.	# of Tornados	Mean $\overline{X}$	$x_i-\overline{X}$	$(x_i-\overline{X})^{\wedge 2}$
1	10	105.33	-95.33	9088.44
2	18	105.33	-87.33	7627.11
3	26	105.33	-79.33	6293.78
4	33	105.33	-72.33	5232.11
5	62	105.33	-43.33	1877.78
6	123	105.33	17.67	312.11
7	123	105.33	17.67	312.11
8	132	105.33	26.67	711.11
9	133	105.33	27.67	765.44
10	138	105.33	32.67	1067.11
11	150	105.33	44.67	1995.11
12	316	105.33	210.67	44380.44
Total	1264		0.00	79662.67

The Mean $\overline{X} = \frac{\Sigma\, xi}{n} = \frac{1264}{12}$ = 105.33 tornados

$\Sigma\,(x_i-\overline{X})^{\wedge 2}$ = 79662.67 tornados$^{\wedge 2}$

Sample Variance $S^{\wedge 2} = \frac{\Sigma\,(xi-\overline{X})^{\wedge 2}}{n-1} = \frac{79662.67}{11}$ = 7,242.06 tornados$^{\wedge 2}$

Standard deviation $s = \sqrt{\frac{\Sigma\,(xi-\overline{X})^{\wedge 2}}{n-1}} = \sqrt{7{,}242.06}$ = 85.10 tornados

The high value of standard deviation indicates that the values are widely distributed around the mean.

24. Tallest Buildings The number of stories in the 13 tallest buildings in Houston are shown.

Find the range, variance, and standard deviation for the data.

75 71 64 56 53 55 47 55 52 50 50
50 47

Answer:

The raw data set are sorted in ascending order:

47,47,50,50,50,52,53,55,55,56,64,61,75

Range = Upper limit – Lower limit = 75 – 47= 28 stories

The procedure to find variance and standard deviation is illustrated in table 6

Table 6: Calculations for variance and standard deviation for Building stories

No.	# of Stories	Mean $\overline{X}$	$x_i-\overline{X}$	$(x_i-\overline{X})^{\wedge 2}$
1	47	55.77	-8.77	76.90
2	47	55.77	-8.77	76.90
3	50	55.77	-5.77	33.28
4	50	55.77	-5.77	33.28
5	50	55.77	-5.77	33.28
6	52	55.77	-3.77	14.21
7	53	55.77	-2.77	7.67
8	55	55.77	-0.77	0.59
9	55	55.77	-0.77	0.59
10	56	55.77	0.23	0.05
11	64	55.77	8.23	67.75
12	71	55.77	15.23	231.98
13	75	55.77	19.23	369.82
Total	725		0.00	946.31

The Mean $\overline{X} = \frac{\Sigma \text{ xi}}{n}$

$= \frac{725}{13}$ = 55.77 stories

$\Sigma (x_i-\overline{X})^{\wedge 2}$ = 946.31 stories$^{\wedge 2}$

Sample Variance $S^{\wedge 2} = \frac{\Sigma (\text{xi}-\overline{X})\text{^}2}{n-1} = \frac{946.31}{12}$ = 78.88 stories$^{\wedge 2}$

Standard deviation $S = \sqrt{\frac{\Sigma (\text{xi}-\overline{X})\text{^}2}{n-1}} = \sqrt{78.88}$ = 8.88 stories

25. Labor Charges The average labor charge for automobile mechanics is $54 per hour. The standard deviation is $4. Find the minimum percentage of data values that will fall within the range of $48 to $60. Use Chebyshev's theorem.

    Answer:

    - Determining the larger value of labor charge = $60
    - Subtracting the mean from the upper value $60- $54 = $ 6
    - Determining k value by dividing the deference on Standard deviation

      K = $\frac{6}{4}$ = 1.5

    - Applying Chebyshev's theorem:

      0r 1- $\frac{1}{k^2}$ = 1 - $\frac{1}{(1.5)^2}$ = 1- $\frac{1}{2.25}$ = 1-0.44 = 0.56 = 56%

      Which means that the 56% of the labor charge for mechanics will fall between $48 and $60.

26. Costs to Train Employees, for a certain type of job, it costs a company an average of $231 to train an employee to perform the task. The standard deviation is $5. Find the minimum percentage of data values that will fall in the range of $219 to $243. Use Chebyshev's theorem

    Answer:

    - Determine the larger value of training cost = $243
    - Subtract the mean from the upper value $243- $231 = $12
    - Determine k value by dividing the deference on Standard deviation

K= $\frac{12}{5}$ = 2.4

- Applying Chebyshev's theorem:

0r 1- $\frac{1}{k^2}$ = 1 - $\frac{1}{(2.4)^2}$ = 1- $\frac{1}{5.88}$ = 1-0.17 = 0.83 = 83%

Which means that the 83% of the data values will fall between $219 and $243.

27. Commuter Times The mean of the times it takes a commuter to get to work in your hometown is 29.7 minutes. If the standard deviation is 6 minutes, within what limits would you expect approximately 68% of the times to fall? Assume the distribution is approximately bell-shaped.

Answer:

Since the data are normally distributed, the 68% of data set will fall between

By applying the Empirical (Normal) Rule:

$\overline{X}$ –s and $\overline{X} + s$ for 68% distribution on normal curve

So, the limits at which the Commuters will get to work are:

($29.7_{(min)} - 6_{(min)}$) to ($29.7_{(min)} + 6_{(min)}$) = $23.7_{(min)}$ to $35.7_{(min)}$

This means that 68% of Commuters will get to work within 23.7 -35.7 minutes

28. High Temperatures The reported high temperatures of 23 cities in Africa in October are shown. Find the z values for

a) A temperature of 80º

b) A temperature of 56º

62 72 66 79 83 61 62 85 72 64 74 71
42 38 91 66 77 90 74 63 64 68 42

Answer:

The sum of data ΣX = 1566°F

Number of data set = n = 23

As shown in table 7:

Table 7: Calculations of the mean and standard deviation

No.	Temperature °F	Mean $\overline{X}$	$(X - \overline{X})$	$(X - \overline{X})^{\wedge 2}$
1	62	68.09	-6.09	37.05
2	72	68.09	3.91	15.31
3	66	68.09	-2.09	4.36
4	79	68.09	10.91	119.09
5	83	68.09	14.91	222.40
6	61	68.09	-7.09	50.22
7	62	68.09	-6.09	37.05
8	85	68.09	16.91	286.05
9	72	68.09	3.91	15.31
10	64	68.09	-4.09	16.70
11	74	68.09	5.91	34.96
12	71	68.09	2.91	8.49
13	42	68.09	-26.09	680.53
14	38	68.09	-30.09	905.22
15	91	68.09	22.91	525.01
16	66	68.09	-2.09	4.36
17	77	68.09	8.91	79.44
18	90	68.09	21.91	480.18
19	74	68.09	5.91	34.96
20	63	68.09	-5.09	25.88
21	64	68.09	-4.09	16.70
22	68	68.09	-0.09	0.01
23	42	68.09	-26.09	680.53
Total	1566		0.00	4279.83

Finding the Mean $\overline{X} = \frac{\Sigma X}{n} = \frac{1566}{23}$ = 68.086957 °F , rounded to 68.09°F

$\Sigma (X - \overline{X})^{\wedge 2}$ = 4279.83

Finding the standard deviation (s) of the temperature data set

S = $\sqrt{\frac{\Sigma (xi-\overline{X})^2}{n-1}}$; s = $\sqrt{\frac{4279.83}{22}}$ = 13.94767, and rounded to 13.95

Finding the z value for each temperature in the

1- For 80° z = $\frac{(X-\overline{X})}{s} = \frac{80-68.09}{13.95}$ = 0.84

2- For 56° z = $\frac{(X-\overline{X})}{s} = \frac{56-68.09}{13.95}$ = 0.67

From the z value, the temperature 80° is in a better relevant position.

29. Exam Grades Which of these exam grades has a better relative position?

a) A grade of 82 on a test with $\overline{X}$ = 85 and s = 6.

b) A grade of 56 on a test with $\overline{X}$ = 60 and s = 5.

Answer:

The z value is an indicative of the relative position

1- For the grade 82, z = $\frac{(X-\overline{X})}{s} = \frac{82-85}{6}$ = -0.5

2- For the grade 56, z = $\frac{(X-\overline{X})}{s} = \frac{56-60}{5}$ = -0.8

From the z value, it looks that grade 82 have a higher score than grade 56 and have a better relative position.

30- Check each data set for outliers.

a) 506, 511, 517, 514, 400, 521

b) 3, 7, 9, 6, 8, 10, 14, 16, 20, 12

Answer:

Outliers are those values that are extremely high or extremely low, which they have adverse effect on the mean and standard deviation. The outliers for a data set are calculated as follows

- Sorting data set in ascending order from lowest to the highest

  So for data series (a) will be: 400, 506, 511, 514, 517, 521

- Finding the median or Q_2; because the data set count is even,

  $Q_2 = \frac{511+514}{2} = 512.5$

- Finding the median or Q_1 for numbers below Q_2

  $Q_1 = \frac{400+506}{2} = 453$

- Finding the median or Q_3 for numbers above Q_2

  $Q_3 = \frac{517+521}{2} = 519$

- Finding the Interquartile Rang (IQR), which is $Q_3 - Q_1$

  IQR= 519 - 453 = 66

- Multiplying IQR by 1.5

- To check for lower outliers, subtracting the value of (1.5 IQR) from Q_1

  Q_1- 1.5 (IQR) = 453 -1.5 x66 = 453 – 99 =354

- To check for the upper outliers, adding the value of (1.5 IQR) to Q3

  Q_3 +1.5 IQR = 519 + 99 = 618

  In the data set (a), there are no values below or above the limits 354 -618

  So, there are no outliers.

  - For data set (b), the same procedure is done as in data set (a)

  - Sorting data set in ascending order: 3, 6, 7, 8, 9, 10, 12, 14, 16, 20

$Q_2 = \frac{9+10}{2} = 9.5$

$Q_1 = \frac{6+7}{2} = 6.5$

$Q_3 = \frac{14+16}{2} = 15$

IQR = $Q_3 - Q_1 = 15 - 9.5 = 5.5$

1.5(IQR) = 1.5x 5.5 = 8.25

- Check for lower outliers $Q_1 - 1.5$ (IQR) = 9.5 -8.25 = 1.25
- Check for upper outliers Q_3+1.5(IQR) = 15 +8.25 = 23.25

There are no values below the lower limits or above the upper limits, so there are no outlies in data set (b)

31- Check each data set for outliers.

a) 14, 18, 27, 26, 19, 13, 5, 25

b) 112, 157, 192, 116, 153, 129, 131

Answer:

Sorting data set (a) into ascending order

Data set (a) will be: 5, 13, 14, 18, 19, 25, 26, 27

- Finding the median or Q_2; because the data set count is even,
- $Q_2 = \frac{18+19}{2} = 18.5$
- Finding the median or Q_1 for numbers below Q_2

$Q_1 = \frac{13+14}{2} = 13.5$

- Finding the median or Q_3 for numbers above Q_2

$Q_3 = \frac{25+26}{2} = 25.5$

- Finding the Interquartile Range (IQR), which is $Q_3 - Q_1$

  IQR= 25.5 – 13.5 = 12

- Multiplying IQR by 1.5
- To check for lower outliers, subtracting the value of (1.5 IQR) from Q_1

  Q_1 - 1.5 (IQR) = 13.5 -1.5 x12 = 13.5 – 18 = - 4.5

- To check for the upper outliers, adding the value of (1.5 IQR) from Q_3

  Q_3 +1.5 IQR = 25.5 +18 = 43.5

  There are no values in data set (a) outside the lower and upper limits

  (-0.45- 43.5), so, there are no outliers recorded.

- Sorting data set (b) into ascending order from lowest to highest values.

  Data set (a) will be: 112, 116, 129, 131, 153, 157, 192

- Finding the median or Q_2; because the data set count is odd number, the middle number will be Q_2
- $Q_2 = 131$
- Finding the median or Q_1 for numbers below Q_2

  $Q_1 = 116$

- Finding the median or Q_3 for numbers above Q_2

  $Q_3 = 157$

- Finding the Interquartile Rang (IQR), which is $Q_3 - Q_1$

  IQR= 157 -116 = 41

- multiplying IQR by 1.5
- To check for lower outliers, subtracting the value of 1.5 IQR from Q_1

Q_1 - 1.5 IQR = 116 -1.5 x41 = 116 – 61.5 = 54.5

- To check for the upper outliers, adding the value of 1.5 IQR to Q_3

  Q_3 +1.5(IQR) = 157 + 1.5x 41 = 157 + 61.5 = 218.5

  There are no values in data set (b) outside the lower and upper limits

  54.5 – 218.5, so, no outliers are recorded.

Conclusion

Statistics is of great help and valuable for researchers and student. It provides students with the required knowledge on collecting, analyzing and interpretation of data. The statistics course, forms the firm step in starting research on a scientific basis. It replenishes knowledge of statistical principles, which will be of great help to start research.

The course reinforced knowledge and practice the terms of data description and interpretation, such as variance, standard deviation, and quartile. Also, the course reinforces knowledge of data collection tools and methods of data sorting.

The book recommended by AIU staff on “Elementary statistics: a step by step approach” by Bluman, is a very rich and helpful source for such courses. There is no need for other references on the subject.

I would recommend all students in all majors to enroll in such courses before they start their research.

Bibliography

Bluman, A. G. (2014). *Elementary statistics: A step by step approach* (9th ed.). New York, NY: McGraw-Hill Education.

L. (Director). (2-110, May 20). *MAT 110 Basic Statistics Lesson 1* [Video file]. Retrieved July 8, 2019, from tps://www.youtube.com/watch?v=rFNWDLSeaiU