

Name	Ubaid El ullah Khan
ID	139994
BS MLT	
Semister	Six
Paper	Statistics

(1)

Q1

Answer

Part - A: ~

→ Overall mean consumption: ~

MAN

① Fresh-vegetables (Mean) .

$$204 + 259 + 266 + 317 = 1046$$

$$\text{Total mean} = \frac{1046}{4} \rightarrow 261.5$$

② Fruits

$$31 + 45 + 69 + 105 = \frac{250}{4}$$

$$\text{T. Consumption} = 62.5 \text{ mean}$$

③ Rice

$$367 + 337 + 269 + 246 = \frac{1219}{4}$$

$$\text{Total} = 304.7 \text{ mean}$$

④ Fish:

$$23 + 28 + 31 + 44$$

$$\text{total} = \frac{126}{4} = 31.5 \text{ mean}$$

⑤ Meats .

$$70 + 61 + 69 + 77$$

$$\rightarrow \frac{277}{4} \rightarrow 69.2 \text{ mean}$$

women

① Fresh-vegetable

$$\rightarrow 178 + 235 + 266 + 304 = \frac{983}{4}$$

$$\text{Total consumption} = 245.2 \text{ mean.}$$

② Fruits

$$28 + 46 + 70 + 121 = \frac{265}{4}$$

$$\text{Total} = 66.25 \text{ mean.}$$

③ Rice

$$315 + 276 + 243 + 220$$

$$\text{Total} = \frac{1054}{4} = 263.5 \text{ mean.}$$

④ Fish

$$19 + 21 + 28 + 46$$

$$\rightarrow \frac{114}{4} = 28.5 \text{ mean}$$

⑤ Meats

$$48 + 43 + 54 + 63$$

$$\text{Total } \frac{218}{4} \rightarrow 52 \text{ mean.}$$

Combine mean consumption of man and woman Per day ?

① Fruits & vegetables → combined mean = $\frac{261.5 + 245.2}{2} = \frac{506.7}{2} \rightarrow 253.35 \text{ mean}$

② Fruits $62.5 + 66.25 \rightarrow \frac{128.75}{2} \rightarrow 64.36 \text{ mean.}$

③ Rice $304.7 + 263.5 \rightarrow \frac{568.2}{2} \rightarrow 284.1 \text{ mean.}$

④ Fish $31.5 + 28.5 = 60/2 \rightarrow 30 \text{ mean}$


⑤ Meats $69.2 + 52 \rightarrow \frac{121.2}{2} \rightarrow 60.6 \text{ mean}$

(2)

Part B:-

This part is a theoretical one as the figures mentioned in the table for consumption of milk, root vegetables and wheat flour indicate that those men and women which are in category (1) $\rightarrow Q_1$ on both sides have more consumption of these ingredients than the rest groups.

They are ingesting more food than the rest groups.

\rightarrow If we calculate the overall consumption of milks, root vegetables and flour b/w 1308  1540 female, then the females despite males are less in number.

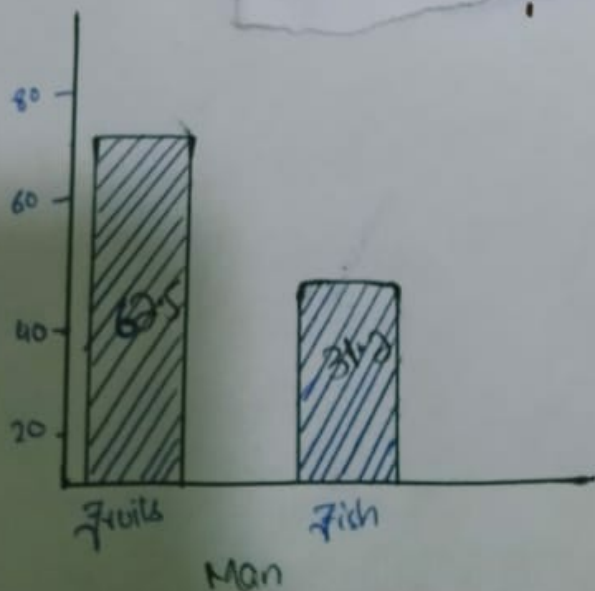
Part C:- The Pathem is quite interesting and logical among the ingredients.

for Man:- for man from Q_4 to Q_1 its descending order as the consumption of rice decreases gradually.

while for fruits and fish ascending order as consumption increases.

for women:- For women from Q_4 to Q_1 its same in rice for men its also decrease consumption so its descending order. while for fruits and fish the consumption is an ascending order. it increases gradually.

Part D



- (E) If we compare the main food ingredients of man category 4th with female of Q₃ consume more fresh vegetables than men Q₄ female of Q₁ consume more fruits than males Q₄ category while Q₄ males have consume more Rice than female Q₂.
- females of Q₁ have consumed more wheat flour than Q₄ men.
 - Q₄ males have consumed more meat than Q₁ females
 - Q₁ females consumed more fish than Q₄ males.

	Males (Q ₄)	Females (Q ₁)
Fresh vegetables	204	304
Fruits	31	121
Rice	367	220
wheat flour	79	180
Meat	70	63

(F) Explain in your own words. (3)

what the standard deviation is better for men for women.

Man

$$0.9 \times \sqrt{30.8} = 32.5$$

$$0.5 \times \sqrt{1308} = 18.0$$

$$1.0 \times \sqrt{1308} = 36.1$$

$$0.1 \times \sqrt{1308} = 3.61$$

$$0.1 \times \sqrt{1308} = 36.1$$

$$0.4 \times \sqrt{1308} = 14.0$$

$$0.2 \times \sqrt{1308} = 7.2$$

$$0.3 \times \sqrt{1308} = 16.8$$

~~161.92~~

women

$$0.8 \times \sqrt{1540} = 31.4$$

$$0.4 \times \sqrt{1540} = 15.69$$

$$0.4 \times \sqrt{1540} = 15.69$$

$$0.8 \times \sqrt{1540} = 31.4$$

$$0.9 \times \sqrt{1540} = 31.4$$

$$0.3 \times \sqrt{1540} = 11.77$$

$$0.2 \times \sqrt{1540} = 7.84$$

$$0.3 \times \sqrt{1540} = 11.77$$

~~235-84~~

So the standard deviation of the men of all commodities are best from the standard deviation of the women.

Q2:-

(01)

Answer:-

Part a:-

Purpose of a census:-

A census is like a survey asking information about your age, marital status, place of residence and home.

The purpose of census is to get sense about the population in general. Modern census data are used for research, business marketing and planning also, the census ensures that each community gets the right number of representation in government. census helps out to distribute public funds on educational setup, health care, law enforcement and highway allocated in parts based on population. In short, it helps us see how our country is changing.

Part B :-

The census is different from sample surveys and routine collection of data by many ways.

- * The census happens throughout the countries after decades while sample surveys are common everywhere.
- * Census needs big fund as its a long and hectic process and need more efforts as compared to the other means of surveys.
- * A sample is not the actual picture of the whole population of the country. A sample - survey cannot be generalized to all population. Its result are quite confined in many ways, while census is a reliable, source of the actual information about the population and the things related to it like dispersion of religion, ethnicity members of the family and many - more.

(2)(2)

Part. C in

As mentioned in the question that that census of 2011 in UK attracted 94% of the population its a positive sign for a country if that much population get interest in a census and co-operates with government authorities for a better outcome in interest of a common man. High response rate doesn't effect the validity of a census. In contrast its a good sign and it alleviate the uncertainty in data ~~and also~~ enhances the accuracy of a census. It helps to get a better picture of the whole-population.

Part. D:~

In 2011 Census of UK, almost 170,000 people stated their ~~of~~ religion as Jedi-Knigh which is a fiction-character nothing real. such responses can increase uncertainty. The reason behind such attitude of these people might be illiteracy, non-serious behaviours, unawareness about the sensitivity of census and the consequences of ambiguous statements.

such responses does not invalidat the question of religion in the survey form of census. If we compare the number of these people with the whole population its quiet-less. That means huge-population dont have problem from the portion of religion except some people.

(03)

Part - E :-

Census is a hectic process manually but if you plan for online census, that you might get some potential-problems.

- * The problems might be like this need for authentication from each household, lack of coverage of household, and along with the fear that hackers could compromise the integrity of the census. Moreover, data collected via internet would have to be integrated mail-back questionnaire and telephone responses. Privacy of people may also be disclosed.
- * Such problems might be overcome by assuring the safety of all types of information of the population. The method of the census through internet must be easy and keen that everyone should access it easily.
- * Proper protocols must be the need for hacking prevention and losing of data.

Part - F :-

The additional data held by government agencies might cause some issues too. The potential or chances of problems might be confidentiality and privacy disclosure of the common-people.

- (*) They may also know about the income of many people which was not on the record and they were not paying-tax. So the fear of such things will spread uncertainty among general-population.
- (*) Hackers may also get that extra information and use that data in illegal-things.

Q.2

(1)

(a) Find G.M, A.M, H.M, Median, mode, Quartiles, Deciles, Percentiles
 Rang, M.D, Q.D, variance and standard deviation, coefficient
 of variation, skewness for the following data.

Sol: -

Class	F	x	freq	f/x	C.B	C.F	$F(x_i - \bar{x})$	$F(x_i - \bar{x})^2$
20-24	1	22	1.34	0.0455	10.5-24.5	1	13.6	184.96
25-29	3	51	4.29	0.1111	24.5-29.5	4	25.8	281.32
30-34	5	160	7.53	0.1563	29.5-34.5	9	18	64.8
35-39	8	746	12.55	0.2168	34.5-39.5	17	11.2	15.69
40-44	5	210	8.12	0.1190	39.5-44.5	22	32	204.8
45-49	2	94	3.34	0.0426	44.5-49.5	24	22.8	259.92
50-54	0	0	0	0	49.5-54.5	24	0	0
55-59	1	57	1.76	0.0175	54.5-59.5	25	21.4	457.96
Total	25	890	38.98	0.7082			144.8	1416

(2)

i) A.M = ?

$$\bar{x} = \frac{\sum f x}{\sum f} = \frac{890}{25} = 35.6$$

$$\bar{x} = 35.6$$

ii) G.M = ?

$$G.M = A \log \frac{1}{\sum f} \cdot \sum f \log a$$

$$= A \log \frac{1}{25} (38.93)$$

$$= A \log (1.5572)$$

$$= 36.07$$

iii) H.M = ?

$$H.M = \frac{\sum f}{\sum f/x}$$

$$= \frac{25}{0.708}$$

$$= 35.30$$

(5) Median = Median is the mid-value of a data set. (3)

$$\text{Median} = \left[\left(\frac{n}{2} \right) + 1 \right] \text{th}$$

$$= \left[\left(\frac{25}{2} \right) + 1 \right] \text{th}$$

$$= (12 + 1) \text{th}$$

Median = 13th, integer in a data set.

$$\boxed{\text{Median} = 37}$$

(6) Quartiles.

$$Q_1 = \left[\left(\frac{n}{4} \right) + 1 \right] \text{th}$$

$$= \left[\left(\frac{25}{4} \right) + 1 \right] \text{th}$$

$$= (6.25 + 1) \text{th}$$

$$(7.25) \text{th}$$

$$= 7 \text{th}$$

$$\boxed{Q_1 = 32}$$

$$Q_3 = \left[\left(\frac{3n}{4} \right) + 1 \right] \text{th}$$

$$= \left[\left(\frac{3 \times 25}{4} \right) + 1 \right] \text{th} \rightarrow 19.75 \text{th}$$

$$(20) \text{th}, \text{ integer}$$

$$\boxed{Q_3 = 42}$$

(7) Deciles:

$$\text{As, } D_5 = \left[\left(\frac{2n}{10} \right) + 1 \right] \text{th} \rightarrow \left[\left(\frac{25 \times 2}{10} \right) + 1 \right] \text{th}$$

$$= (5 + 1) \text{th} \rightarrow 6 \text{th}, \text{ integer}$$

$$\boxed{D_5 = 38}$$

(8) Percentile:

$$\text{As, } P_{30} = \left[\left(\frac{30n}{100} \right) + 1 \right] \text{th}$$

$$\rightarrow = \left[\left(\frac{30 \times 25}{100} \right) + 1 \right] \text{th}$$

$$= (7.50 + 1) \text{th} \rightarrow (8.50) \text{th}$$

$$P_{30} = 9 \text{th}, \text{ integer}$$

$$\boxed{P_{30} = 32}$$

(4)

Quartiles = ?

$$Q_1 = l + h/f (n/4 - c)$$

$$n/4 = 25/4 = 6.25$$

$$\begin{aligned} Q_1 &= 29.5 + 5/5 (6.25 - 4) \\ &= 29.5 + (1)(2.25) \end{aligned}$$

$$Q_1 = 31.75$$

$$Q_2 = \text{Med} = 36.69$$

$$Q_3 = l + h/f (3n/4 - c)$$

$$\text{Now } 3n/4 = \frac{(3)(25)}{4} = 18.75$$

$$\begin{aligned} Q_3 &= 39.5 + 5/5 (18.75 - 17) \\ &= 39.5 + (1)(1.75) \\ &= 39.5 + 1.75 \end{aligned}$$

$$Q_3 = 41.25$$

(5)

Deciles = ?

$$Df = l + h/f \left(\frac{7n}{10} - c \right)$$

$$\text{Now } \rightarrow \frac{7n}{10} = \frac{(7)(25)}{10} = 17.5$$

$$Df = 29.5 + 5/5 (17.5 - 17)$$

$$29.5 + (1)(0.5)$$

$$Df = 40$$

Percentiles = ?

$$P_{33} = l + h/f \left(\frac{33n}{100} - c \right)$$

$$\text{Now } \frac{33n}{100} = \frac{(33)(25)}{100} = 8.25$$

$$P_{33} = 29.5 + 5/5 - (8.25 - 4)$$

$$= 29.5 + (1)(14.25)$$

$$= 29.5 + 14.25$$

$$P_{33} = 43.75$$

(6)

Rang = ?

$$\begin{aligned} R &= x_m - x_a \\ &= 59.5 - 19.5 \\ &= 59.5 - 19.5 \\ &= 40 \end{aligned}$$

Mean deviation = (M.D) = ?

$$\begin{aligned} M.D &= \frac{\sum f |x_i - \bar{x}|}{\sum f} \\ &= \frac{144.8}{25} \\ &= 5.79 \end{aligned}$$

Variance = ?

$$\begin{aligned} \text{Var} &= \frac{\sum f (x_i - \bar{x})^2}{\sum f} \\ &= \frac{1410}{25} \\ &= 56.4 \end{aligned}$$

(7)

Standard deviation = ?

$$SD \sqrt{\frac{\sum f(x_i - \bar{x})^2}{\sum f}}$$

$$= \sqrt{\frac{141}{28}} = \sqrt{5.0357}$$

$$S.D = 7.51$$

Quartiles Deviation (Q.D) = ?

$$Q.D = \frac{Q_3 - Q_1}{2}$$

$$= \frac{41.25 - 31.75}{2}$$

$$= \frac{9.5}{2}$$

$$Q.D = 4.75$$

(8)

Coefficient of variation (C.V) = ?

$$C.V = \frac{S.D}{\text{Mean}} \times 100$$

$$= \frac{7.51}{35.6} \times 100$$

$$= \frac{751}{35.6}$$

$$= 21.09\%$$

Skewness = ?

$$SK = \frac{\text{Mean} - \text{Mode}}{S.D}$$

$$= \frac{35.6 - 36.69}{7.51}$$

$$= \frac{-1.09}{7.51}$$

$$= -0.145$$

$$SK = 0.145$$

$$SK = 0.145$$

(1) ~~RTD~~

x	$1/x$	$\log(x)$	$(x - \bar{x})$	$(x - \bar{x})^2$	$1x \cdot \bar{x}$
22	0.045	1.34	-14.8	219.04	14.8
27	0.037	1.4313	-9.8	96.04	9.8
27	0.037	1.4313	-9.8	96.04	9.8
32	0.0312	1.4313	-9.8	23.04	4.8
32	0.0312	1.5051	-4.8	23.04	4.8
32	0.0312	1.5051	-4.8	23.04	4.8
32	0.0312	1.5051	-4.8	23.04	4.8
37	0.027	1.5051	0.2	0.04	0.2
37	0.027	1.5051	0.2	0.04	0.2
37	0.027	1.5051	0.2	0.04	0.2
37	0.027	1.5682	0.2	0.04	0.2
37	0.027	1.5682	0.2	0.04	0.2
37	0.027	1.5682	0.2	0.04	0.2
37	0.027	1.5682	0.2	0.04	0.2
37	0.027	1.5682	0.2	0.04	0.2
42	0.027	1.5682	5.2	27.04	5.2
42	0.0238	1.5682	5.2	27.04	5.2
42	0.0238	1.5682	5.2	27.04	5.2
42	0.0238	1.5682	5.2	27.04	5.2
42	0.0238	1.6232	5.2	27.04	5.2
47	0.0212	1.6720	1.76	3.10	1.76
47	0.0212	1.6720	1.76	3.10	1.76
57	0.019	1.7558	20.2	408.04	20.2
Total	10.73 10.73	38.92		1172.12	136.40

Q3 part (B)

un-group data.

$$(1) A.M = \frac{\sum_{i=1}^n x_i}{n}$$

$$\rightarrow \frac{920}{25}$$

$$= A.M = \boxed{36.8}$$

$$(2) H.M = \frac{n}{\sum_{i=1}^n (1/x_i)}$$

$$= \frac{25}{0.73}$$

$$\rightarrow \boxed{34.25}$$

$$(3) G.M = \text{Antilog} \left[\frac{\sum_{i=1}^n \log(x_i)}{n} \right]$$

$$= \text{Antilog} \left[\frac{38.92}{25} \right]$$

$$= \text{Antilog} (1.56)$$

$$= \boxed{G.M = 36.04}$$

(4) Modes (Most repeated figure)

So Most of it in (37)

$$\text{Mode} = \boxed{37}$$

(5) Median Mid value of data (set)

$$M = \left[\left(\frac{n}{2} + 1 \right)^{\text{th}} \right]$$

$$= \left[\left(\frac{25}{2} + 1 \right)^{\text{th}} \right]$$

$$= (12 + 1)^{\text{th}}$$

= Median = 13th integrent of data set.

(6) Quartiles

$$Q_2 = \left[\left(\frac{n}{4} + 1 \right)^{\text{th}} \right]$$

$$\left(\frac{25}{4} + 1 \right)^{\text{th}}$$

$$= (6.25 + 1)^{\text{th}}$$

$$= (7.25)^{\text{th}}$$

= 7th integor

$$\boxed{Q_1 = 32}$$

$$Q_3 = \left[\left(\frac{3n}{4} + 1 \right)^{\text{th}} \right]$$

$$= \left[\left(\frac{3 \times 25}{4} + 1 \right)^{\text{th}} \right]$$

$$= (19.75)^{\text{th}}$$

= 20th → integor

$$\boxed{Q_3 = 42}$$

⑦ Deciles

$$\begin{aligned} \text{As } D_2 &= \left[\left(\frac{2n}{10} \right) + 1 \right]^{\text{th}} \\ &\rightarrow \left[\left(\frac{2 \times 25}{10} \right) + 1 \right]^{\text{th}} \\ &\rightarrow (5+1)^{\text{th}} = 6^{\text{th}} \text{ integer} \\ &= \boxed{D_2 = 32} \end{aligned}$$

⑧ Percentile

$$\begin{aligned} \text{As } P_{30} &= \left[\left(\frac{30n}{100} \right) + 1 \right]^{\text{th}} \\ &\rightarrow \left[\left(\frac{30 \times 25}{100} \right) + 1 \right]^{\text{th}} \\ &= (7.50+1)^{\text{th}} \rightarrow 8.50 = 9^{\text{th}} \\ &= P_{30} = 9^{\text{th}} \text{ integer} \\ &= \boxed{P_{30} = 32} \end{aligned}$$

⑨ Range

Largest value - smallest value

$$57 - 22$$

$$\boxed{R = 35}$$

⑩ Q.D

$$\begin{aligned} &\frac{Q_3 - Q_1}{2} \\ &= \frac{42 - 32}{2} \\ &= \boxed{Q.D = 5} \end{aligned}$$

⑪ skewness

$$\begin{aligned} &\frac{\text{Mean} - \text{Mode}}{S.D} \\ &= \frac{36.8 - 37}{7.41} \\ &= \frac{-0.20}{7.41} \rightarrow \boxed{-0.03} \end{aligned}$$

⑫ Coefficient of variation

$$\begin{aligned} C.V &= \frac{S.D}{\bar{x}} \times 100 \\ &= \frac{7.41}{36.8} \times 100 \\ &= \boxed{C.V = 20.14} \end{aligned}$$

(13) coefficient of variation (C.V)

$$\frac{S.D}{\bar{x}} \times 100$$
$$= \frac{7.41}{36.80} \times 100$$

$$C.V = 20.14$$

(14) Quartile Deviation (Q.D) = $\frac{Q_3 - Q_1}{2}$

where $Q_3 = 41.25$ and $Q_1 = 31.75$

so

$$Q.D = \frac{41.25 - 31.75}{2}$$

$$Q.D = 4.75$$

(15) Skewness (S.K) = $\frac{\text{Mean} - \text{Mode}}{S.D}$

where, Mean = 36.80, Mode = 37

and S.D = 7.41

$$SK = \frac{36.80 - 37}{7.41}$$

$$= \frac{-0.20}{7.41}$$

$$Skewness = 0.03$$