Question #1 [07 marks]

Construct a frequency distribution for the energy consumption of natural gas (in billions of Btu) by states and the district of Columbia. Use 9 classes and starting first class from the minimum observation of the data.

474	475	205 –	639	197 –	344 –	3 –	409	247 –	66 –
377 –	87 –	747	1166	223 –	248 –	958	406	251 –	3462
2391	514	371 –	58 –	224 –	530	317–	267 –	769	9 –
188 –	289 –	76 –	678	331 –	52 –	214 –	165 –	255 –	319 –
34 –	1300	284 –	834	114 –	1082	73 –	62 –	95 –	393
146 –									

Range = 3462 - 3 = 3459

No. of observations = 51

No of classes = 9

Width of class = 3459/9 = 385

Class Boundaries	Tallies	Frequency
3 — 388	卌 卌 卌 卌 卌 卌 |||	33
388 — 773	卌 卌		11
773 — 1158	|||	3
1158 — 1543	||	2
1543 — 1928		0
1928 — 2313		0
2313 — 2698			1
2698 — 3083		0
3083 — 3468			1
		51

Question #2 [04 marks+04 marks = 08 marks]

A sociologist has been studying the yearly changes in the number of convicts assigned to the largest correctional facility in the state. His data are expressed in terms of the percentage increase in the number of prisoners (a negative number indicates a percentage decrease). The sociologist's most recent data are as follows:

1999	2000	2001	2002	2003	2004
– 5%	6 %	9 %	4 %	7 %	– 6 %

(a) Calculate the average percentage increase using only the 1999-2002 data.

(b) A new penal code was passed in 1998. Previously, the prison population grew at a rate of about 2 percent per year. What seems to be the effect of the new penal code?

a)   From   1999 – 2002

$$= -5, 6, 9, 4$$

$$= \frac{-5 + 6 + 9 + 4}{4}$$

$$= 3.5\%$$

The average percentage increase from 1999 – 2002
                    is   3.5 %.

Question #3 [03 marks +05 marks +05 marks = 13 marks]

Define the following terms with detailed examples.

Discrete and Continuous Variable.

Discrete variable:-

Discrete variable that is complete or a whole number integer.

Example: 5, 6, House number, street #, etc

Continuous variable:

Continuous variable is a variable with decimal part such as it measured.

Example:- 4.3, GPA, Temperature
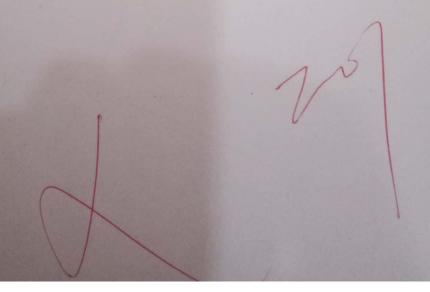
Data Collection through Questionnaire

Data collection through Questionnaire is an important method of data collection from Sample. This data plays an important role in statistical calculations and decision making but sometimes the person who is Questione can give wrong data for being biased that could output of calculations.

Example: ~~course fe~~ The customer of a specific company can be questioned about customer service

Shape of the Distribution

Shape of Distribution means when statistical tools are applied such as frequency is calculated and Graph can represent the shape of distribution **such as** symmetric, etc.

The shape of distribution can play an important role in understanding the whole data or for decision making.

Example:-
  Histogram can be made for data and the shape of Histogram can represent the shap of distribution e.g. J shape, Reverse J, Bell Graph

Question #4 [08 marks +04 marks = 12 marks]

Starting teacher salaries (in equivalent U.S. dollars) for upper secondary education in selected countries are listed below. Which set of data is more variable? (The U.S. average starting salary at this time was $29,641). Also sketch a suitable graphical display.

Europe	Salary	Asia	Salary
Sweden	$48,704	Korea	$26,852
Germany	41,441	Japan	23,493
Spain	32,679	India	18,247
Finland	32,136	Malaysia	13,647
Denmark	30,384	Philippines	9,857
Netherlands	29,326	Thailand	5,862
Scotland	27,789		

Question #5 [08 marks +05 marks + 07 marks = 20 marks]

The data show the number of murders in 25 selected cities

Class limits	Cumulative Frequency	Mid Point	f	fx_i	$\lvert x-\bar{x} \rvert$	$f\lvert x-\bar{x}\rvert$
34-96	13	65	13	385	105.08	1366.04
97-159	15	128	2	256	42.08	84.16
160-222	15	191	0	0	20.92	0
223-285	20	254	5	1270	83.92	419.6
286-348	21	317	1	317	146.92	146.92
349-411	22	380	1	380	209.92	209.92
412-474	22	443	0	0	272.92	0
475-537	23	506	1	506	335.92	1335.92
538-600	25	569	2	1138	398.92	797.84
			25	4252		3360.4

a. Find Quartile deviation and mean absolute deviation.
b. Show that sum of the square of the deviations of observations taken from arithmetic mean is minimum as compared to any other measure of central tendency.

a) 1)

$$Q_3 = \frac{3(9)}{4} = 6.75 \Rightarrow 7^{th}$$

$$Q_3 = l + \frac{h}{f}\left(\frac{3n}{4} - c\right)$$

$$= l + \frac{h}{f}\left(\frac{3(25)}{4} - \dot{c}\right)$$

$$= l + \frac{h}{f}\left(18.75 - c\right)$$

$$= 223 + \frac{62}{5}\left(18.75 - 15\right)$$

$$= 223 + \frac{62}{5}\left(3.75\right)$$

$$= 223 + 46.5$$

$$= 269.5$$

$$Q_1 = 34 + \frac{62}{13}\left(6.25 - 0\right)$$

$$= 34 + 29.80$$

$$Q_1 = 63.80$$

$$= \frac{Q_3 - Q_1}{2}$$

a) $$\bar{x} = \frac{\sum f_i x_i}{n}$$

$$= \frac{4252}{25}$$

$$\bar{x} = 170.08$$

$$\frac{\sum |x - \bar{x}|}{n} = \frac{3360.4}{25}$$

$$\text{Mean Abs Devi} = 134.416$$

b)   Square of A.M

$$= \cancel{18067.66}$$

$$= 170.08$$

$$= 28927.2064$$

09

Good Luck