

Problem 1

The following data set y_1, y_2, \dots, y_{39} consists of the annual income for 39 people selected at random from people living in Conestogo. The data are in thousands dollars.

23	26	28	28	31	34	35	37	39	41
42	44	44	48	49	51	51	53	53	55
56	56	56	58	58	59	59	62	65	66
69	73	76	77	81	86	92	95	101	

$$\sum_{i=1}^{39} y_i = 2157 \quad \text{and} \quad \sum_{i=1}^{39} y_i^2 = 133837$$

Problem 1(a)

Determine the following:

sample mean, median, mode

sample variance, sample standard deviation, IQR, range

Give your answers to 3 decimal places.

Numerical Answers

sample mean = 55.308

sample median = 55

mode = 56

sample variance = 382.587

sample standard deviation = 19.560

IQR = 25

range = 78

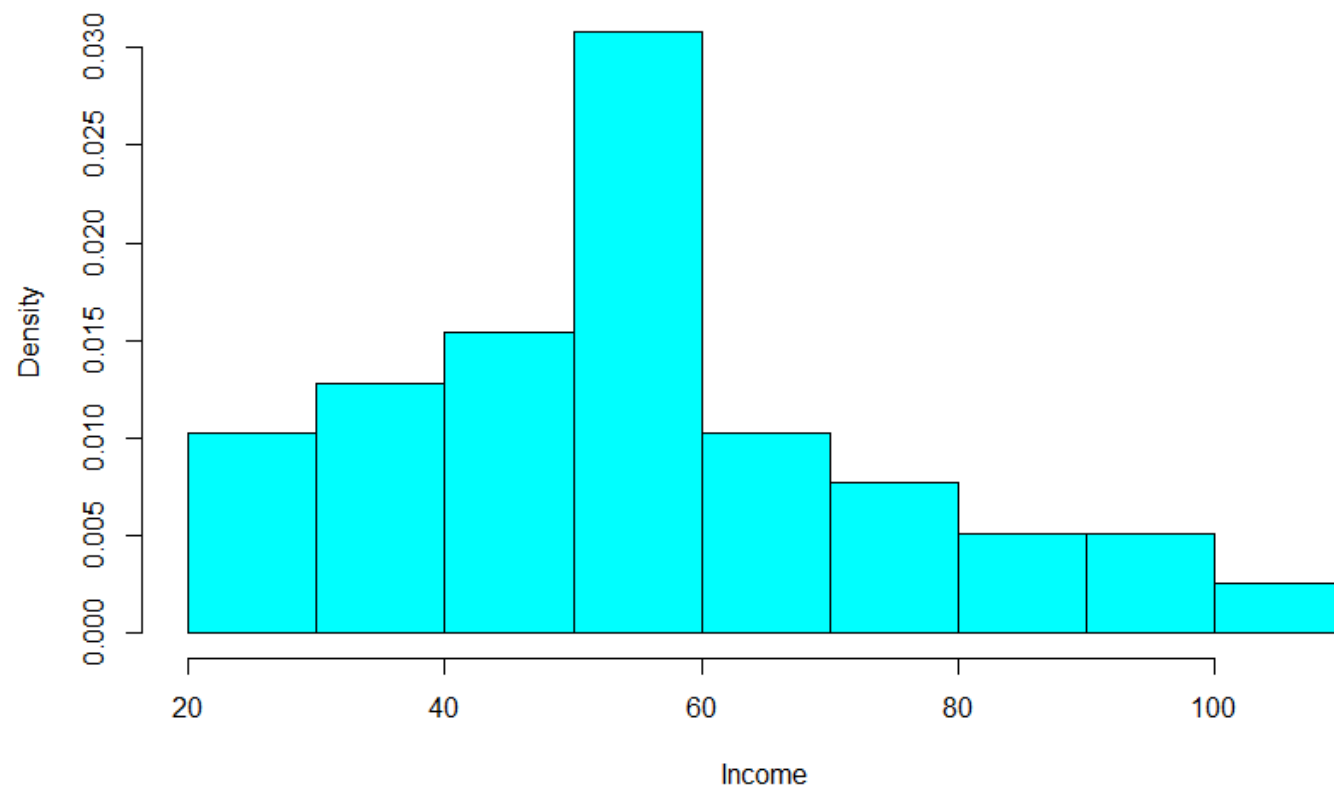
Problem 1(b)

From these values is it possible to determine any information about the sample skewness and sample kurtosis for these data?

The sample skewness is 0.468.

The sample kurtosis is 2.675.

Histogram of Incomes



Problem 1(c)

To celebrate Canada Day, the government adds one thousand dollar to each person's annual income.

Find the sample mean, median, mode, sample variance, sample standard deviation, IQR, range of the new data set.

Numerical Answers

sample mean = 56.308

sample median = 56

mode = 57

sample variance = 382.587

sample standard deviation = 19.560

IQR = 25

range = 78

Problem 1(d)

If the sample skewness for the original data is 0.468 what is the sample skewness for the new data?

If the sample kurtosis for the original data is 2.675 what is the sample skewness for the new data?

Sample skewness and sample kurtosis are unchanged.

Problem 1(e)

To compare the original incomes with incomes in 2006 these data should be multiplied by 0.8.

Find the sample mean, median, mode, sample variance, sample standard deviation, IQR, range, sample skewness, and sample kurtosis for the new data set.

Numerical Answers

sample mean = 44.246

sample median = 44

mode = 44.8

sample variance = 244.856

sample standard deviation = 15.648

IQR = 20

range = 62.4

Sample skewness and sample kurtosis are unchanged.

Problem 1(f)

For the original data set, the data point $y_2 = 26$ is unreliable and is discarded from the data set.

Determine the following for the data set of 38 observations:

sample mean, median, mode

sample variance, IQR, range

New data

23	26	28	28	31	34	35	37	39	41
42	44	44	48	49	51	51	53	53	55
56	56	56	58	58	59	59	62	65	66
69	73	76	77	81	86	92	95	101	

Numerical Answers

sample mean = 56.079

sample median = 55.5

mode = 56

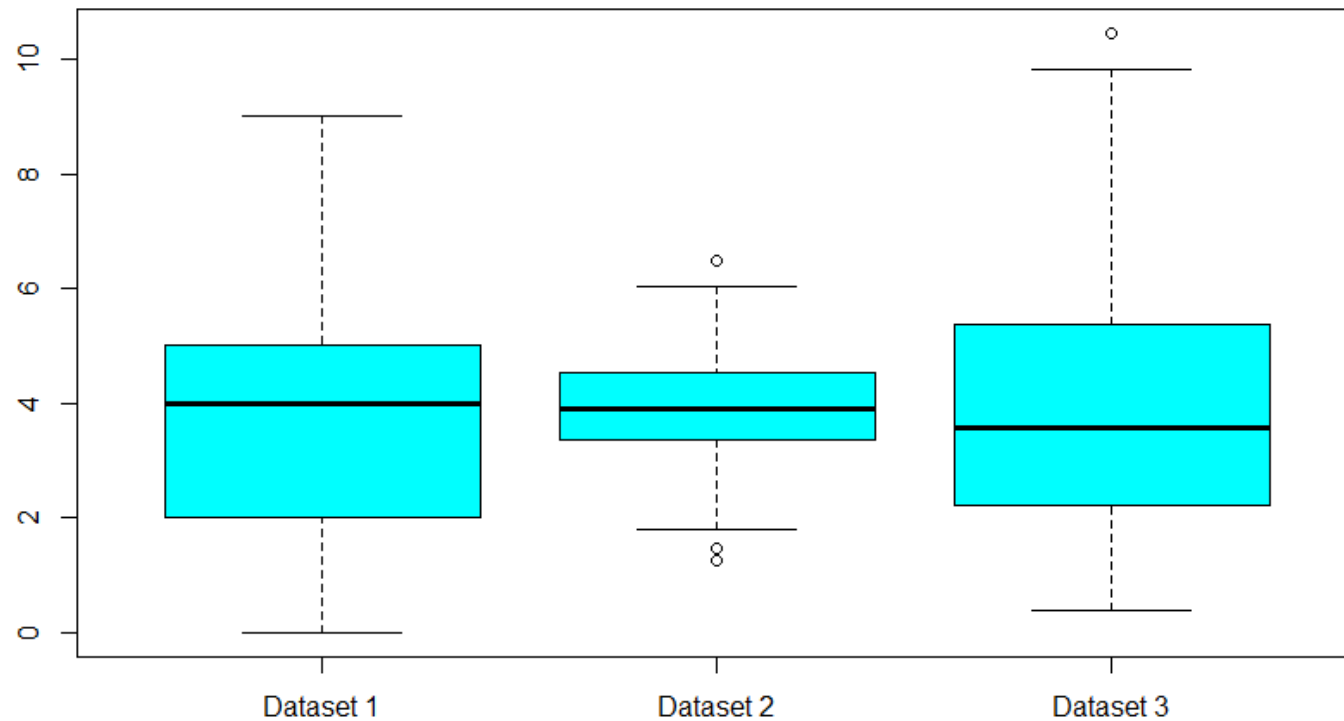
sample variance = 369.102

sample standard deviation = 19.212

IQR = 26

range = 78

Problem 2



Problem 2

For each data set determine the approximate values of the sample median, IQR, and range.

Determine for each data set whether the sample skewness is approximately zero, positive, or negative.

Which data set is the most skewed?

Numerical Answers (Approximate)

Data set 1:

sample median = 4, IQR = 3, range = 9

skewness is positive

Data set 2:

sample median = 4, IQR = 1, range = 5.2

skewness is approximately zero

Data set 3:

sample median = 3.6, IQR = 3.1, range = 10

skewness is positive

Data set 3 is the most skewed.

Problem 2

Which data set has the smallest variability?

Which data set is most bell-shaped (Gaussian)?

Data set 2 has the least variability.

Data set 2 is the most bell-shaped.