



**PES University, Bangalore**

(Established under Karnataka Act No. 16 of 2013)

**UE19CS203 – STATISTICS FOR DATA SCIENCE**

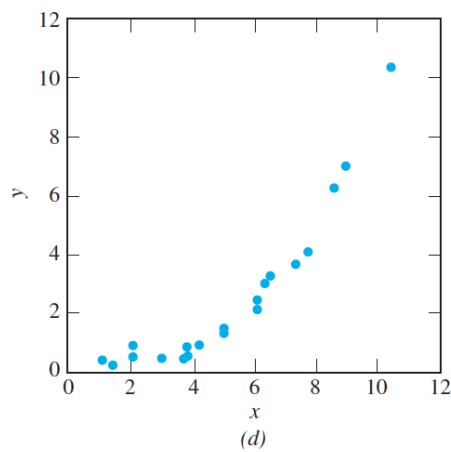
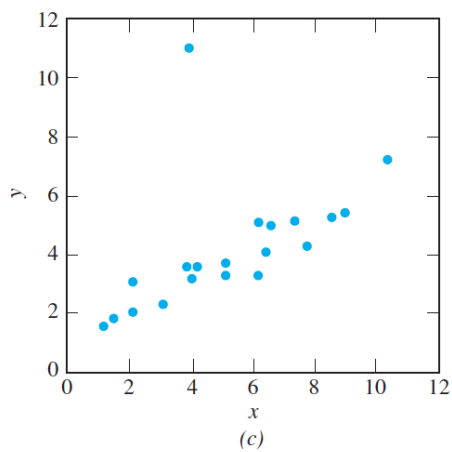
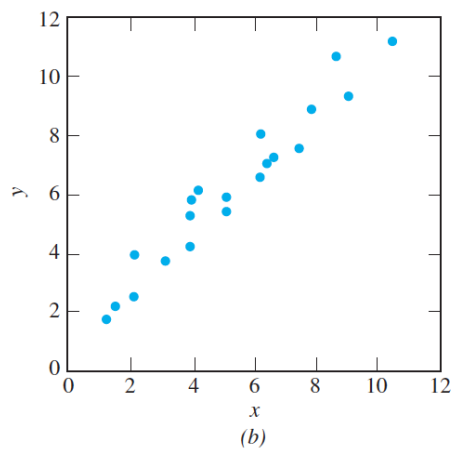
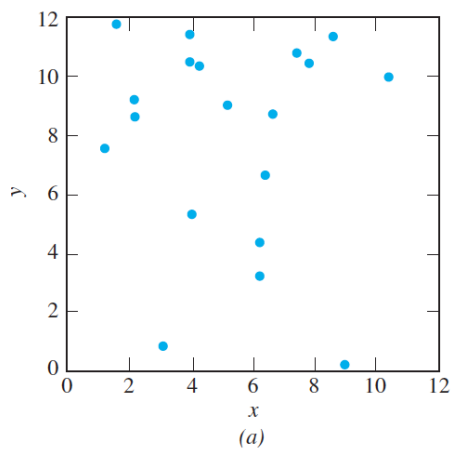
**Unit-1 - Introduction to Data Science**

**QUESTION BANK – SOLVED**

**Data Visualization Techniques – Scatter Plot, Scatter Plot, Bar Chart, Heat Map**

**Exercises for Section 1.3**

**1. Match each scatterplot to the statement that best describes it.**



- a. The relationship between x and y is approximately linear.
- b. The relationship between x and y is nonlinear.

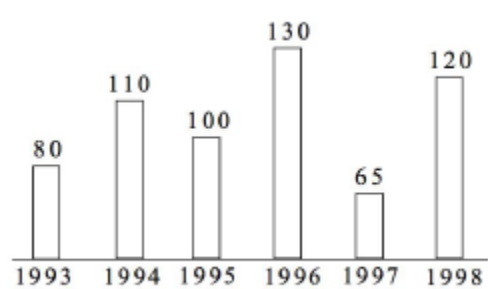
- c. There isn't much of any relationship between x and y.
- d. The relationship between x and y is approximately linear, except for an outlier.

**[Text Book Exercise – Section 1.3 – Q. No. 18 – Pg. No. 42]**

**Solution**

- a. There isn't much of any relationship between x and y.
- b. The relationship between x and y is approximately linear.
- c. The relationship between x and y is approximately linear, except for an outlier.
- d. The relationship between x and y is nonlinear.

- 2. The bar graph provided below gives the data of the production of food grains over the years which is (1000 tonnes). Study the bar chart and answer the questions.**



- a) In the case of how many years was the production below the average production of the given years?
- b) What was the percentage drop in production from 1996 to 1997?
- c) In which year was the production 50% of the total production in 1993 and 1998 together?
- d) If the production in 1999 will be above the average production of the given years, which of the following could be the minimum production for 1999?
- e) What was the approximate percentage increase in production from 1993 to 1994?

**Solution:**

- a) In the case of how many years was the production below the average production of the given years?

$$\text{Average production} = (80 + 110 + 100 + 130 + 65 + 120) / 6 = 605 / 6 = 100.833$$

Thousand tonnes = 100900 tonnes.

That is, production in 1993, 1995 and 1997 was below the average.

So it is for 3 years the production was below the average.

**b) What was the percentage drop in production from 1996 to 1997?**

Production in 1996 = 130 thousand tonnes.

Production in 1997 = 65 thousand tonnes

$$\text{decrease} = (130 - 65) / 130 * 100 = 65 / 130 * 100 = 50\%$$

**50%** is the percentage drop in production from 1996 to 1997

**c) In which year was the production 50% of the total production in 1993 and 1998 together?**

Total production in 1993 and 1998 =  $80 + 120 = 200$

$$50\% \text{ of } 200 = 100$$

ie, the production in 1995 was 50% of the total production in 1993 and 1998.

**d) If the production in 1999 will be above the average production of the given years, which of the following could be the minimum production for 1999?**

Average production = 100900 tonnes.

Therefore the minimum production in 1999 will be 100900 tonnes

**e) What was the approximate percentage increase in production from 1993 to 1994?**

Production in 1993 = 80 thousand tonnes.

Production in 1994 = 110 thousand tonnes

$$\text{Increase} = (110 - 80) / 80 * 100 = \text{approx } 40\%$$