# Topic for the class-Correlation

**Unit _3** : Title-Descriptive statistics

**Date & Time** : 5.9.24 10.00 AM – 10.50 AM

**Dr. Bhramaramba Ravi**

Professor

Department of Computer Science and Engineering

GITAM School of Technology (GST)

Visakhapatnam – 530045

Email: **bravi@gitam.edu**

# Unit3-syllabus

- **UNIT 3 Descriptive statistics 9 hours, P - 2 hours**
- Measures of Central Tendency – Measures of Variation – Quartiles and Percentiles –

  Moments – Skewness and Kurtosis. Exploratory Data Analytics Descriptive Statistics – Mean,

  Standard Deviation, Skewness and Kurtosis – Box Plots – Pivot Table – Heat Map – Correlation Statistics – ANOVA, Random variable, Variance, covariance, and correlation- Linear transformations of random variables, Regression.
- https://www.coursera.org/learn/data-visualization-r

# Correlation coefficients

- There are many ways to measure the strength of the relationship between two variables.
- For pairs of variables measured on an interval or ratio scale, a *correlation coefficient* (*r*) can be calculated.
- This value quantifies the *linear relationship* between the variables by generating values from −1.0 to +1.0.
- If the optimal straight line is drawn through the points on a scatterplot, the value of *r* reflects how closely the points lie to this line.
- Positive numbers for *r* indicate a positive correlation between the pair of variables, and negative numbers indicate a negative correlation.
- A value of *r* close to 0 indicates little or no relationship between the variables.

# Correlation coefficients contd.

- For example, the two scatterplots shown in Figure 4.15 illustrate different values for *r*.
- The first graph illustrates a strong positive correlation because the points lie relatively close to an imaginary line sloping upward from left to right through the center of the points; the second graph illustrates a weaker correlation.

The formula used to calculate *r* is shown here:

$$r = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{(n-1)s_x s_y}$$

# Correlation coefficients contd.

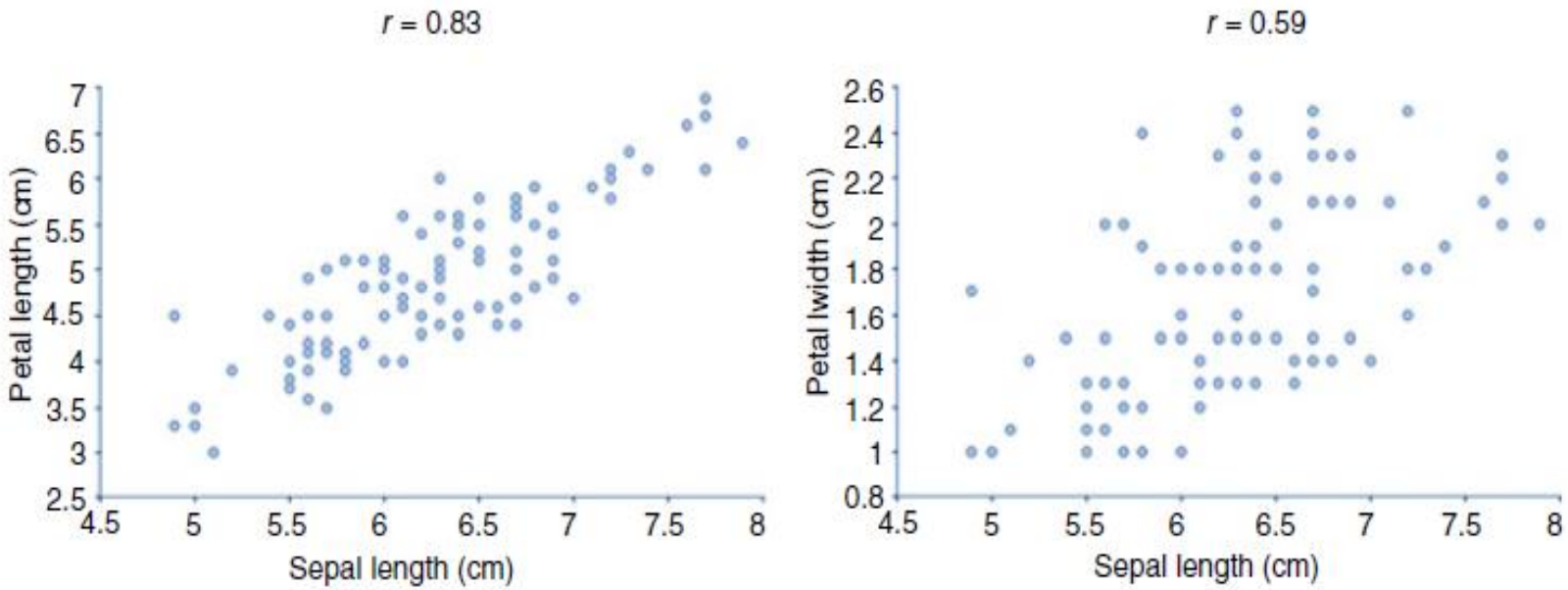


**FIGURE 4.15** Scatterplots illustrate values for the correlation coefficient ($r$).

THANK YOU