

Graphical Representation of Data

T.Krishnan

Strand Life Sciences, Bangalore

- graphs are visualization aids
- graphs are supposed to inform not mislead
- graphs are useful
 - before data analysis
 - during data analysis
 - after data analysis

Uses of Statistical Graphs

- Graphs before: visual representation of data and summaries: e.g. bar chart
- Graphs during: check assumptions and model fit : normal probability plot
- Graphs after:
 - check assumptions; e.g., normal probability plot of residuals
 - present results: scatter plot with regression line

Types of Statistical Graphs

- to examine a single variable:
 - categorical (e.g., bar graph)
 - continuous (e.g., histogram)
- interaction between two variables
 - categorical vs categorical—(e.g., subdivided bar chart)
 - continuous vs categorical—(e.g., conditioned box plot)
 - continuous vs continuous—(e.g., scatter plot)

- multidimensional plots—(e.g., SPLOM—scatter plot matrix)
- plot ensembles and statistical models—(e.g., interaction plots; residual plots)
- geographical data
- special graphs:
 - e.g., missing values
 - large data—summary-based plots

- In Statistics, graphs display relationship between variables or show the spread of a given variable or phenomenon.
- Basic Statistical Graphs:
 - Bar Graphs
 - Pie Charts
 - Line Graphs
 - Scatterplots
 - Histograms

- Plain Bar Chart
- Stacked Bar Chart
- Range Bar Chart
- Anchored Bar Chart
- Divided Bar Chart

- Histogram
- Frequency Polygon
- Cumulative Histogram
- Gap Histogram
- Stem-and-Leaf Plot

- Box Plot
- Box-Whisker Plot
- Notched Box Plot

- Quantile Plot
- Normal Probability Plot

- Scatter Plot
- Scatter Plot Matrix (SPLOM)
- Regression Line with Confidence Bands

- Dendrogram