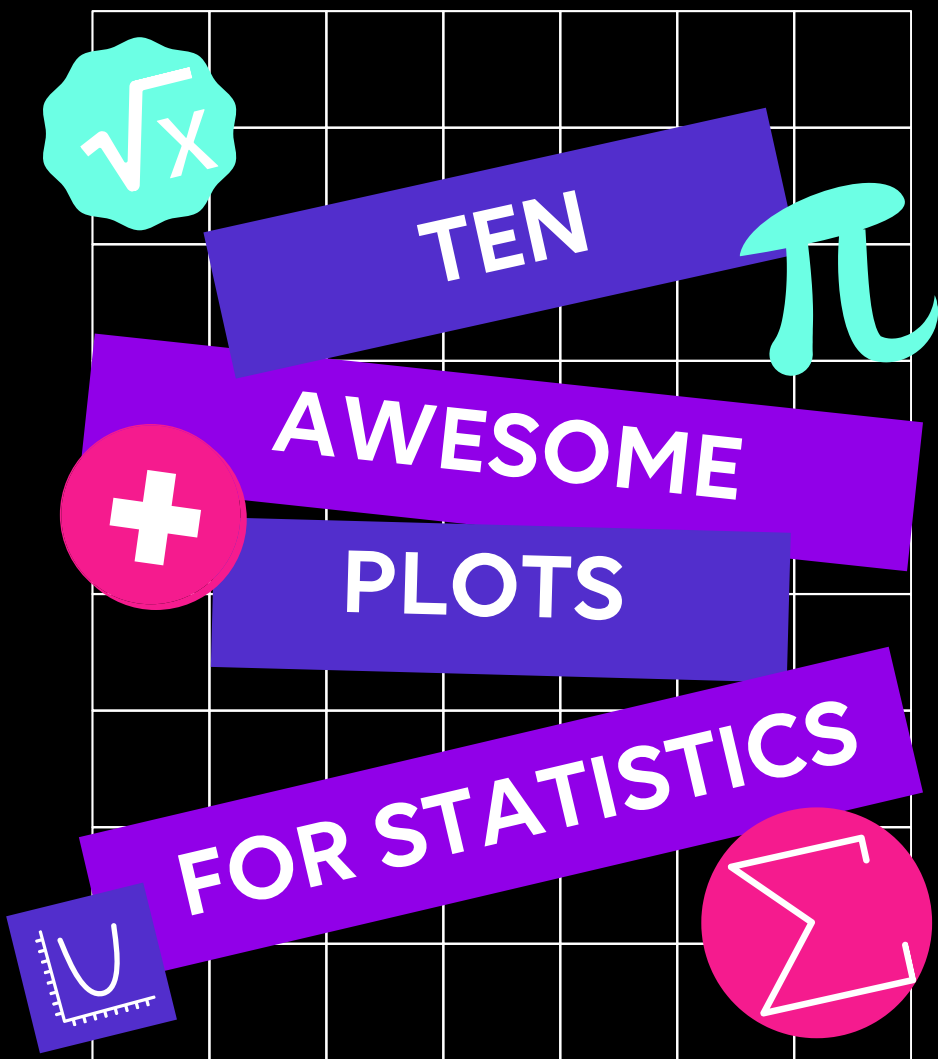




Jeffrey Michael Franc

STAT59

10 Awesome Plots For Statistics

Jeffrey Michael Franc

Copyright Jeffrey Michael Franc, 2021

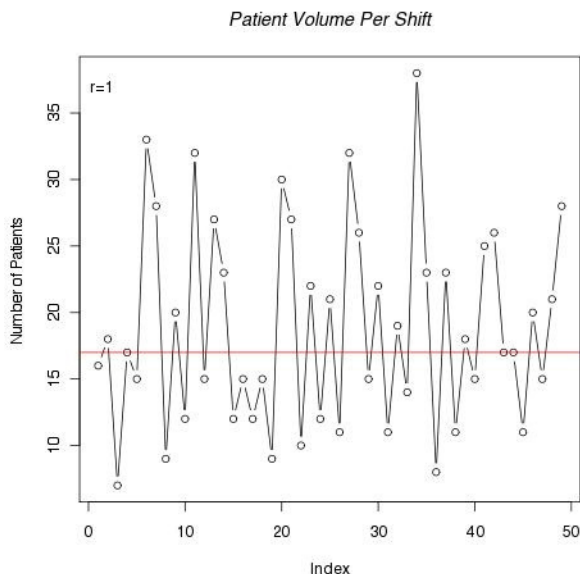
Published by

Stat59 Services Edmonton AB Canada

1

RUN CHART

DATA WITH TIME ORDER



EXAMPLE

A researcher suspects that the schedule of emergency shifts is poor. She feels that the volume of patients seen by each physician is unequal. She measures patient volume for each physician for one week and plots them in time order.

ADVANTAGES

- Preserves the order of the observations
- Makes it easy for readers to assess for patterns
- Simple to understand

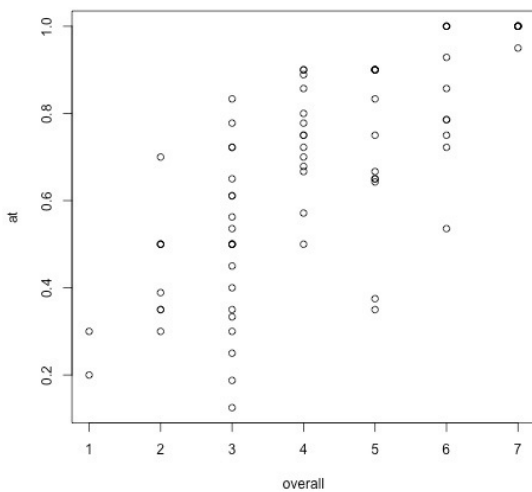
DISADVANTAGES

- Deceptive when order is not relevant

2

SCATTER PLOT

PLOTTING CORRELATION



EXAMPLE

A researcher is looking at the correlation between technical abilities score and global ratings score in a series of critical care simulations. He plots the technical abilities on the y-axis and overall global ratings score on the y-axis



ADVANTAGES

- Allows to quickly see relation between x and y variables



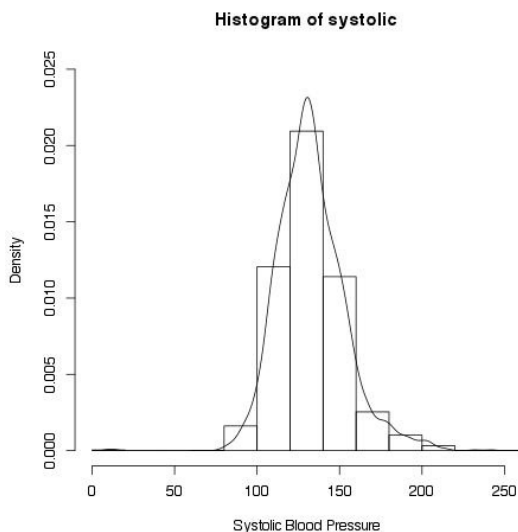
DISADVANTAGES

- Difficult to show association between more than two variables
- Overwhelming with large quantity of data

3

DENSITY HISTOGRAM

UNIVARIATE DATA



EXAMPLE

A researcher has recorded systolic blood pressure for approximately 2000 emergency department visits.



ADVANTAGES

- Good overall view of data
- Easily displays large quantity of data
- Mean and dispersion of data easily visualized



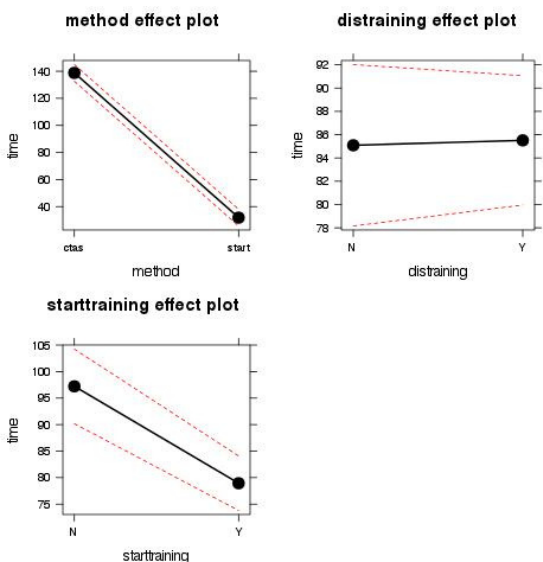
DISADVANTAGES

- Addition of density curve loses true units of measurement
- Fine detail of each observation is lost

4

EFFECT PLOTS

EFFECTS OF FACTORS



EXAMPLE

A researcher has recorded the amount of time it takes to triage patients. He is looking at the effect of triage method, experience disaster medicine training, and previous triage training and how they influence this time.



ADVANTAGES

- Graphical representation of important factors
- Easily rule out factors with little effect



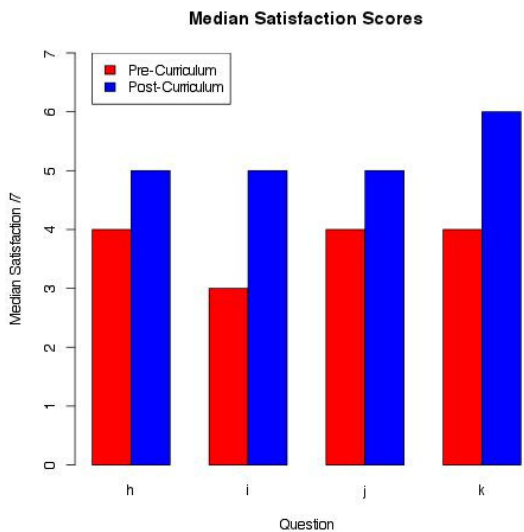
DISADVANTAGES

- Difficult to interpret for some readers

5

BAR PLOT

MULTIPLE RESPONSES



EXAMPLE

A researcher is looking at satisfaction with a disaster medicine curriculum. She measured a satisfaction score (1 – 7) for four different characteristics both before and after a new curriculum was introduced.



ADVANTAGES

- Small differences are easily seen
- Allows comparison of multiple groups



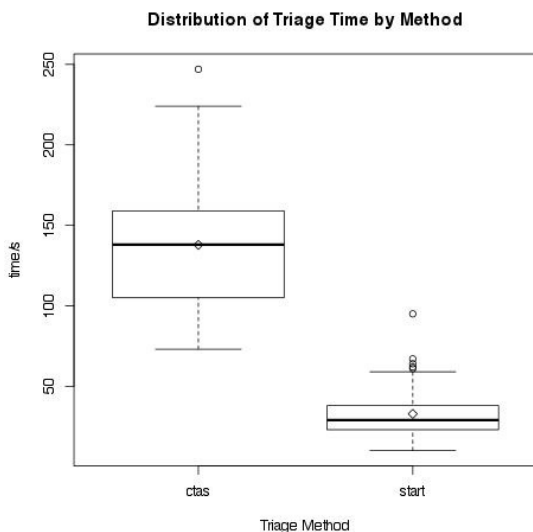
DISADVANTAGES

- Difficult to interpret as the number of factors increases

6

BAR PLOT

CONTINUOUS DISTRIBUTIONS



EXAMPLE

A researcher is looking at the difference between time for nurses to triage patients using two different tools: CTAS and START.

**ADVANTAGES**

- Excellent way to categorize distribution of sample
- Large amount of data in one plot
- Easily visualize median, shape, and outliers

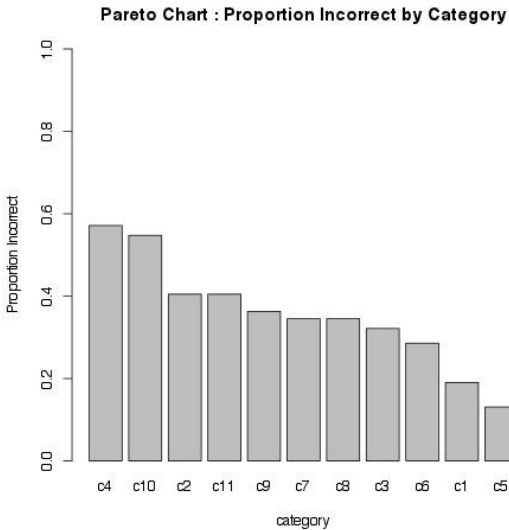
**DISADVANTAGES**

- May be difficult to understand for non-statisticians
- Uses median and interquartile range rather than mean and standard deviation

7

PARETO CHART

PRIORITIZATION OF DATA



EXAMPLE

A group of researchers has tested a group of residents in emergency medicine to see which areas of disaster medicine knowledge they are in most need of help. Pre-test scores as a percentage are available for 12 different knowledge areas.



ADVANTAGES

- Adds priority to the bar graph
- Easy to understand why priorities are chosen



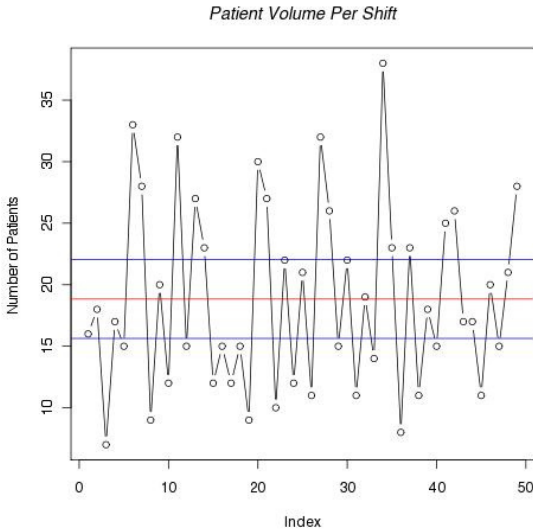
DISADVANTAGES

- Generally larger bars are considered more important: may need to reverse the data for clarity

8

CONTROL CHART

HYPOTHESIS TESTING



EXAMPLE

The same researcher who did the shift workload Run Chart has been asked by her colleagues to "Prove that the results were statistically significant."



ADVANTAGES

- Combines statistical hypothesis testing with data visualization
- More intuitive to understand than p-values



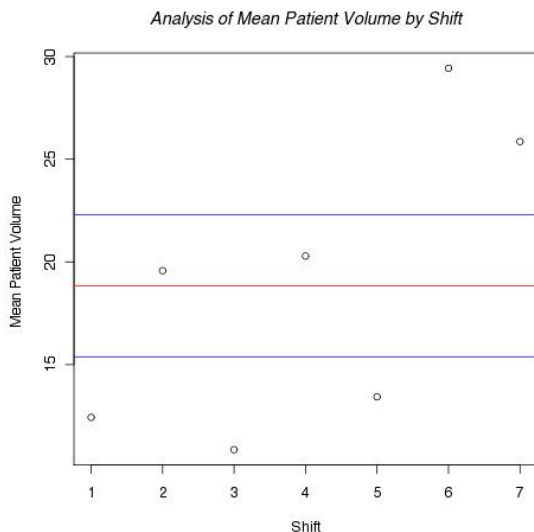
DISADVANTAGES

- Control limits often mistaken for specifications
- Need to understand principles of repeated testing

9

ANALYSIS OF MEANS

ANOVA RESULTS



EXAMPLE

In addition to the run chart, and control chart, the researcher also did ANOVA which showed that there was a significant difference between shift volumes. None of her colleagues understood what the ANOVA analysis meant.

**ADVANTAGES**

- Graphical method to compare several groups
- Corrects for multiple hypothesis testing
- Unlike ANOVA shows which groups are statistically different

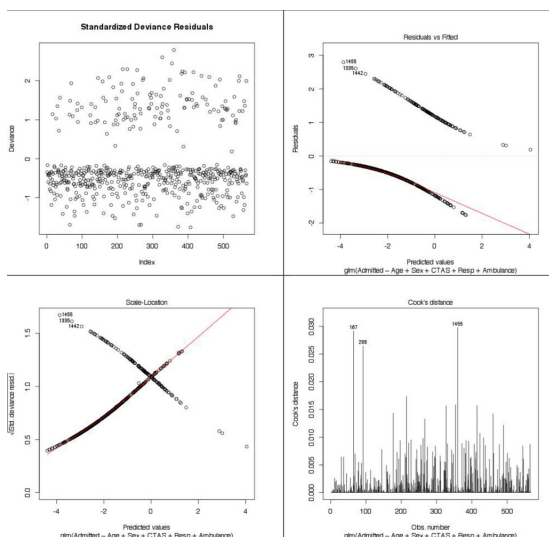
**DISADVANTAGES**

- Less well known than ANOVA
- May need some explanation for many readers to understand

10

RESIDUAL PLOTS

REGRESSION DIAGNOSTICS



EXAMPLE

A researcher used a linear regression equation to show a significant difference between triage methods and time to triage. A reviewer asked "How do we know that the linear assumption was correct?"



ADVANTAGES

- Mandatory for any study using regression
- Allows visual representation of utility of regression equation
- Useful for high-dimension multivariate data



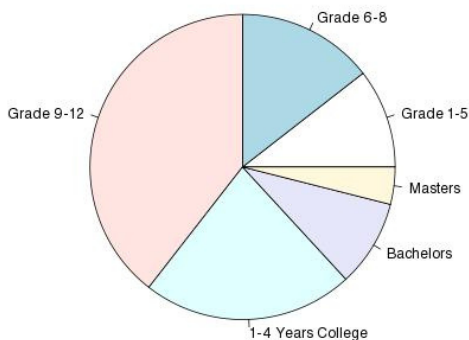
DISADVANTAGES

- Difficult to interpret (particularly logistic regression)
- May need explanation

PIE CHARTS

DISHONORABLE MENTION

Education Level



EXAMPLE

Although commonly used to represent many types of data, in most cases a much better choice exists. These should be reserved for cases where precise numbers are not important. They should never be used in academic research.



ADVANTAGES

- Easy to understand for almost all readers
- Gives good overview of the composition of the whole sample



DISADVANTAGES

- Perception of area is poor compared to height
- Easy to lose subtle differences
- Not advised if close comparison is needed
- Currently low popularity rating among statisticians



Awesome Plot Checklist

- ☐ The plot emphasizes or clarifies the overall conclusions of the study
- ☐ A typical reader from the intended audience understands the plot without reading the descriptive text
- ☐ All details needed to understand the plot are contained in the title and legend
- ☐ The plot emphasizes clarity over cleverness
- ☐ Grey-scale (not color) is used if the plot is likely to be printed
- ☐ Axes are labeled appropriately
- ☐ Units are added to the axes where appropriate
- ☐ If submitted for publication, the plot is formatted and saved in accordance with the "Instructions for Authors."