If you want your 10% -- you need a clicker. They count the third week start practicing now!!!

Please send 1.234 using the clicker

Decimal

Chapter 1 reading

Assignment 1 is out! Get it from CANVAS

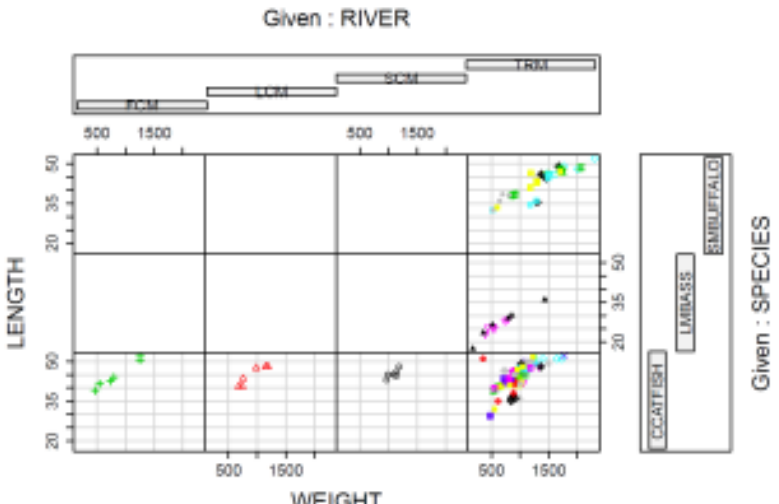
- **Get started asap. Qu1 and 2 can be done already**
- **5% of your grade for each Assignment**
- **Reminder – 10% quizzes, 15% Assignments, 20% Mid-term exams, 10% Labs, 10% Projects and 30% Final Exam**

Must understand code

2. A biologist wants to make a coplot of **LENGTH Vs WEIGHT** given **RIVER*SPECIES** for fish caught in the Tennessee river and recorded in the **DDT.csv** data set, so that each point is colored according to the variable **MILE** which is treated as a factor (Qualitative variable).

```
> head(ddt)
  RIVER MILE  SPECIES LENGTH WEIGHT DDT
1   FCM    5 CCATFISH   42.5    732  10
2   FCM    5 CCATFISH   44.0    795  16
3   FCM    5 CCATFISH   41.5    547  23
4   FCM    5 CCATFISH   39.0    465  21
5   FCM    5 CCATFISH   50.5   1252  50
6   FCM    5 CCATFISH   52.0   1255 150
# The following code may help
m=with(ddt, as.numeric(factor(MILE))) # A
length(unique(m)) #B
```

Make and interpret the plot



Qu 2

(f) What is the mean value of DDT found in the sample of CCATFISH caught in the FCM river?
Hint:

```
ddt=read.csv("..\\CSV\\DDT.csv")
head(ddt)
subset(ddt,RIVER=="FCM" & SPECIES=="CCATFISH",) #or
ddt[ddt$RIVER=="FCM" & ddt$SPECIES=="CCATFISH",]
```

Chapter 2

Descriptive statistics

What we will cover today!

- How to make a histogram!! (The most difficult plot)
- How to make a stem plot (easy)
- How to calculate measures of variation (important)
  - Range
  - Sd and variance
- Empirical and Chebyshev's rule.

Quick Review

Key Terms

Arithmetic mean 28
Bar graph 13
Box plots 42
Category frequency 13
Category relative frequency 13
Chebyshev's Rule 34
Class interval 21
Dot plot 19
Empirical Rule 34
Hinges 42
Histogram 21
Inner fences 42
Interquartile range (IQR) 42
Lower quartile 38
Mean 28
Measures of central tendency 27
Measures of relative standing 27
Measures of variation 27
Median 28
Midquartile 38
Mode 28
Mound-shaped distribution 30
100*p*thpercentile 38
Outer fences 43
Outlier 41
Parameter 28
Pareto diagram 14
Percentile 38
Pie chart 13
Population mean 28
Population standard deviation 33
Population variance 33
Range 32
Sample mean 28
Skewness 29
Standard deviation 32
Statistic 28
Stem-and-leaf display 20
Upper quartile 38
Variance 32
Whiskers 42
z-score 38

Key Formulas

$\frac{\text{Category frequency}}{n}$	Category relative frequency 44
$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}$	Sample mean 28
$s^2 = \frac{\sum_{i=1}^{n} (y_i - \bar{y})^2}{n-1} = \frac{\sum_{i=1}^{n} y_i^2 - \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n}}{n-1}$	Sample variance 33
$s = \sqrt{s^2}$	Sample standard deviation 33
$z = \frac{y - \bar{y}}{s}$	Sample z-score 40
$z = \frac{y - \mu}{\sigma}$	Population z-score 40
$\text{IQR} = Q_U - Q_L$	Interquartile range 42
$Q_L - 1.5(\text{IQR})$	Lower inner fence 43
$Q_U + 1.5(\text{IQR})$	Upper inner fence 43
$Q_L - 3(\text{IQR})$	Lower outer fence 43
$Q_U + 3(\text{IQR})$	Upper outer fence 43

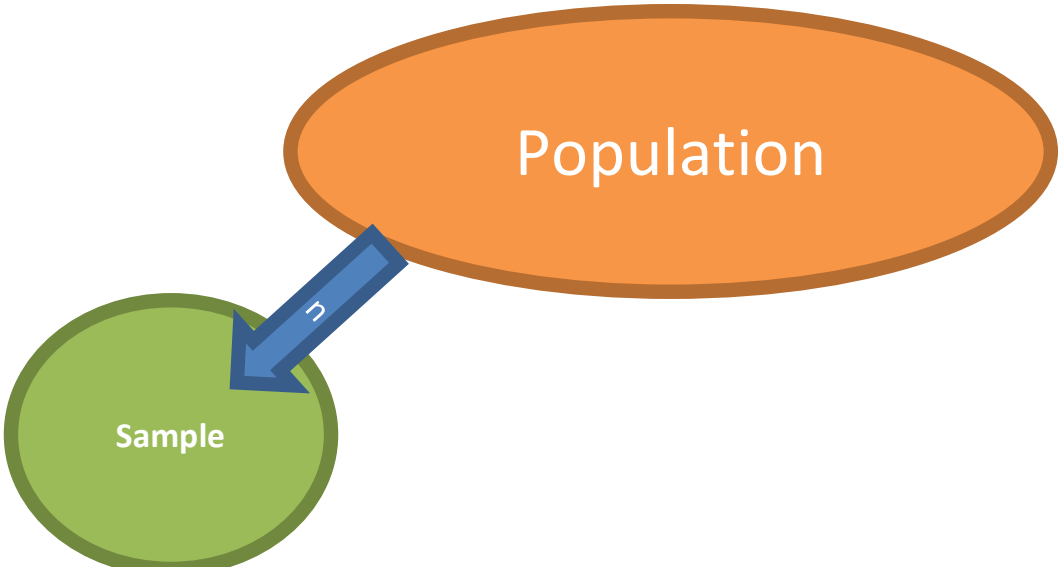
Chapter Summary Notes

- Graphical methods for qualitative data: **pie chart, bar graph**, and **Pareto diagram**
- Graphical methods for quantitative data: **dot plot, stem-and-leaf display**, and **histogram**
- Numerical measures of central tendency: **mean, median**, and **mode**
- Numerical measures of variation: **range, variance**, and **standard deviation**
- Sample numerical descriptive measures are called **statistics**.
- Population numerical descriptive measures are called **parameters**.
- Rules for determining the percentage of measurements in the interval (mean) $\pm$ 2 (std. dev.): **Chebyshev's Rule** (at least 75%) and **Empirical Rule** (approximately 95%)
- Measures of relative standing: **percentile score** and ***z*-score**
- Methods for detecting outliers: **box plots** and ***z*-scores**

Histogram

- The histogram comes in three flavors
  - Frequency
  - Relative Frequency
  - Density

GOBIANTS


Population
n
Sample

What is a parameter?

A) A description of a sample

B) A description of a population

C) Neither A or B

$$s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$$