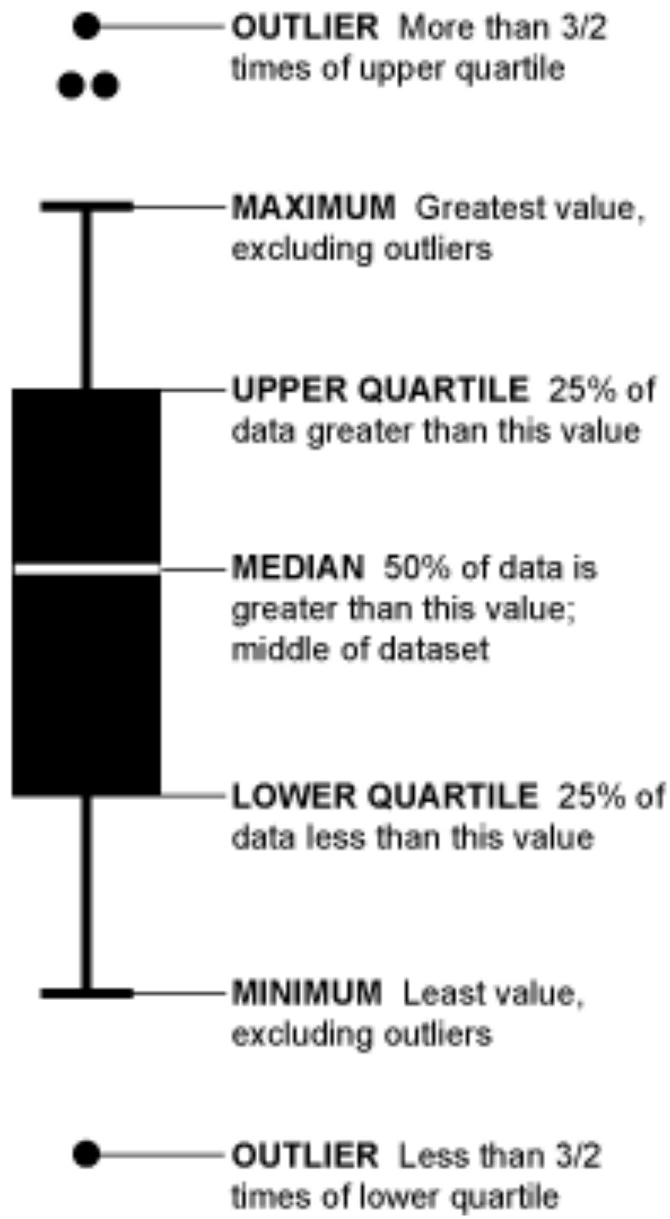
# Review Midterm and Discuss Grades

# box plot


OUTLIER More than 3/2 times of upper quartile
MAXIMUM Greatest value, excluding outliers
UPPER QUARTILE 25% of data greater than this value
MEDIAN 50% of data is greater than this value; middle of dataset
LOWER QUARTILE 25% of data less than this value
MINIMUM Least value, excluding outliers
OUTLIER Less than 3/2 times of lower quartile

# Hypothesis Testing

Beyond 2 ‘Groups’

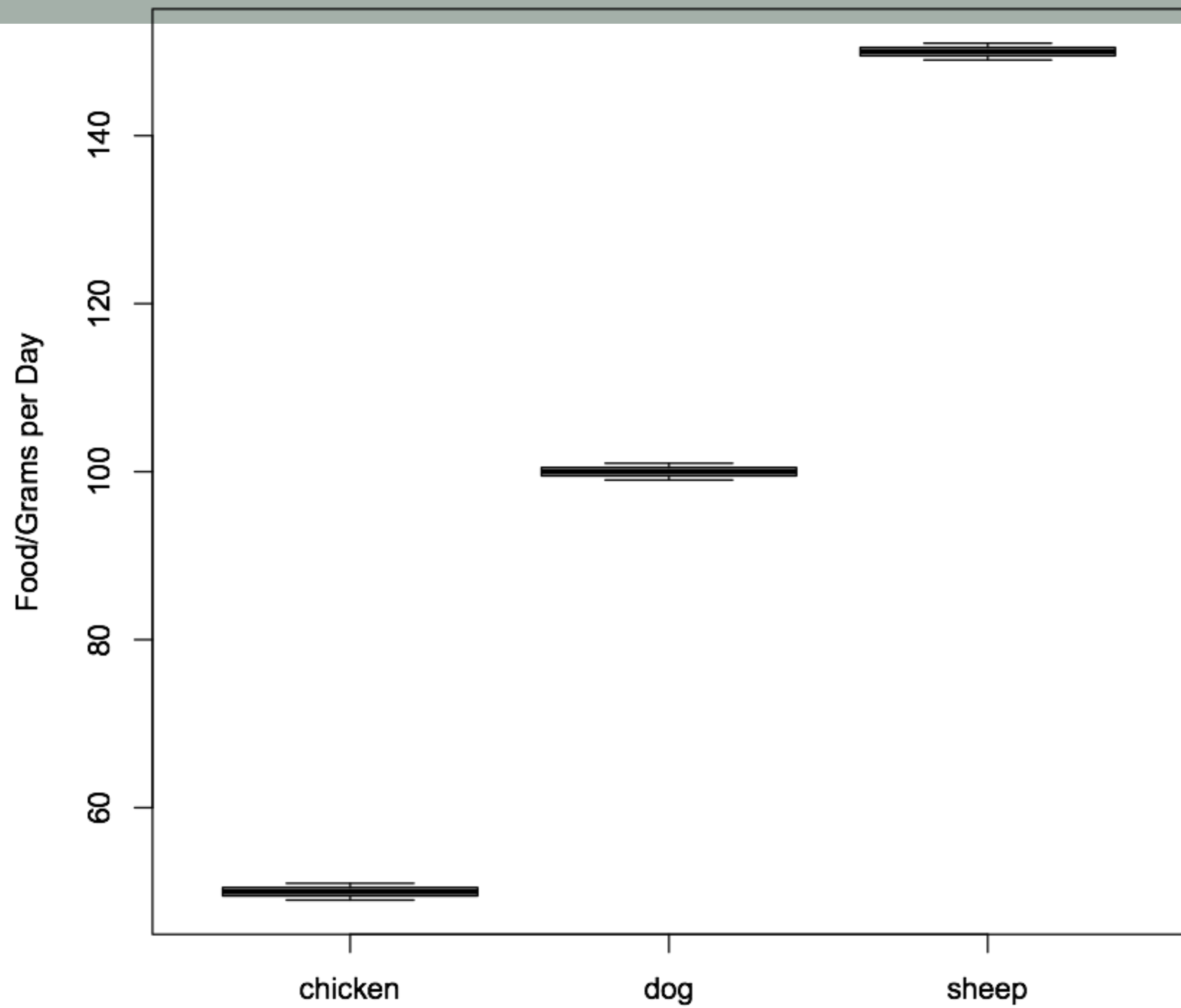
# Animal Food Consumption

| animal | grams/day |
|---|---|
| chicken | 49 |
| chicken | 50 |
| chicken | 51 |
| dog | 99 |
| dog | 100 |
| dog | 101 |
| sheep | 149 |
| sheep | 150 |
| sheep | 151 |
| **std** | **43.3** |
| **XBAR** | **100** |

# Animal Food Consumption

| animal | grams/day |
|---|---|
| chicken | 49 |
| chicken | 50 |
| chicken | 51 |
| dog | 99 |
| dog | 100 |
| dog | 101 |
| sheep | 149 |
| sheep | 150 |
| sheep | 151 |
| **std** | **43.3** |
| **XBAR** | **100** |

| | chicken | dog | sheep |
|---|---|---|---|
| | 49 | 99 | 149 |
| | 50 | 100 | 150 |
| | 51 | 101 | 151 |
| **std** | **1** | **1** | **1** |
| **XBAR** | **50** | **100** | **150** |


Food/Grams per Day
60
80
100
120
140
chicken
dog
sheep

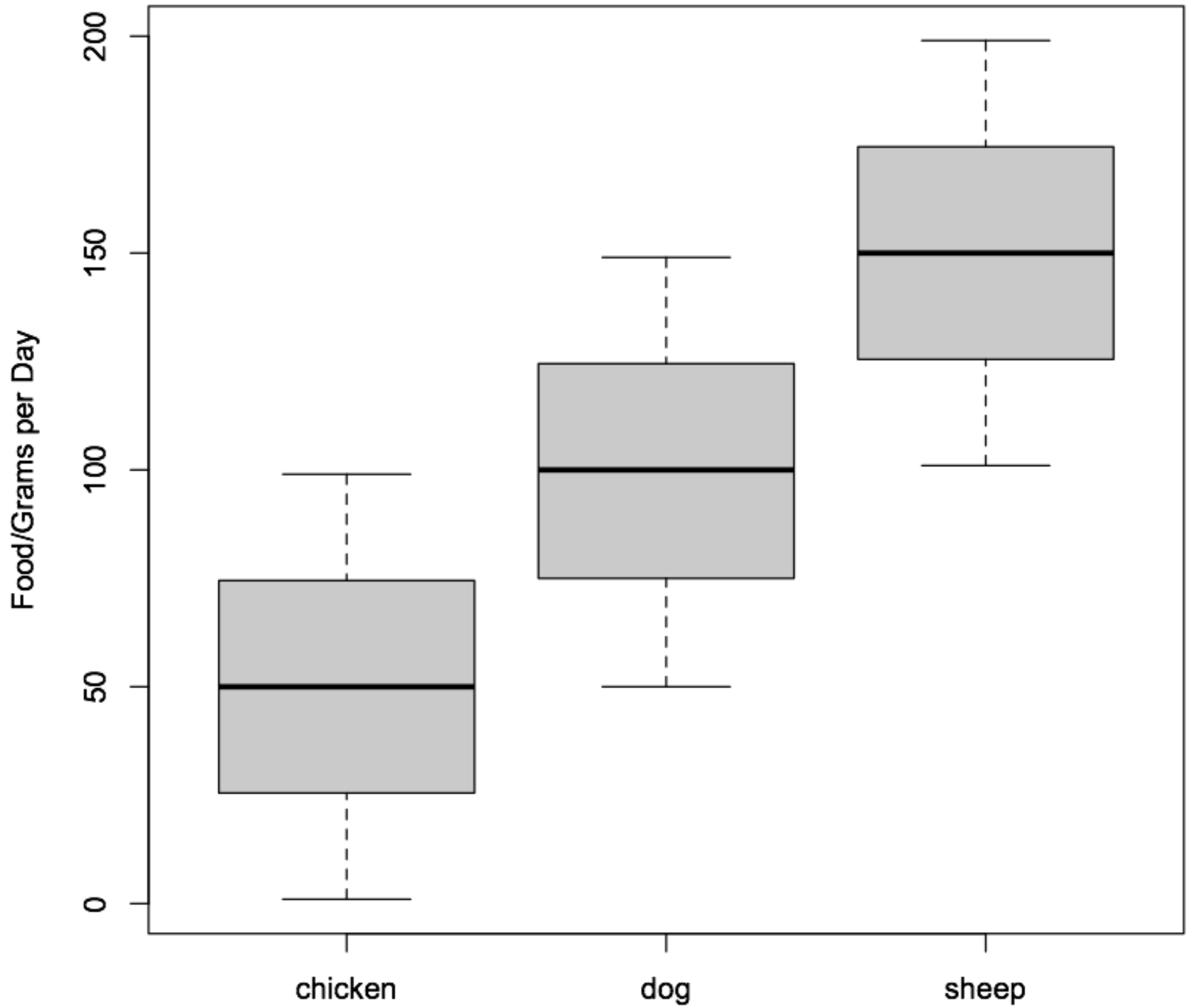
# Another Scenario

| animal | grams/day |
|---|---|
| chicken | 1 |
| chicken | 50 |
| chicken | 99 |
| dog | 50 |
| dog | 100 |
| dog | 149 |
| sheep | 101 |
| sheep | 150 |
| sheep | 199 |
| std | 60.7 |
| XBAR | 99.9 |

# Another Scenario

| animal | grams/day |
|---|---|
| chicken | 1 |
| chicken | 50 |
| chicken | 99 |
| dog | 50 |
| dog | 100 |
| dog | 149 |
| sheep | 101 |
| sheep | 150 |
| sheep | 199 |
| **std** | **60.7** |
| **XBAR** | **99.9** |

| | chicken | dog | sheep |
|---|---|---|---|
| | 1 | 51 | 101 |
| | 50 | 100 | 150 |
| | 99 | 149 | 199 |
| **std** | **49** | **49** | **49** |
| **XBAR** | **50** | **100** | **150** |


Food/Grams per Day
0
50
100
150
200
chicken
dog
sheep

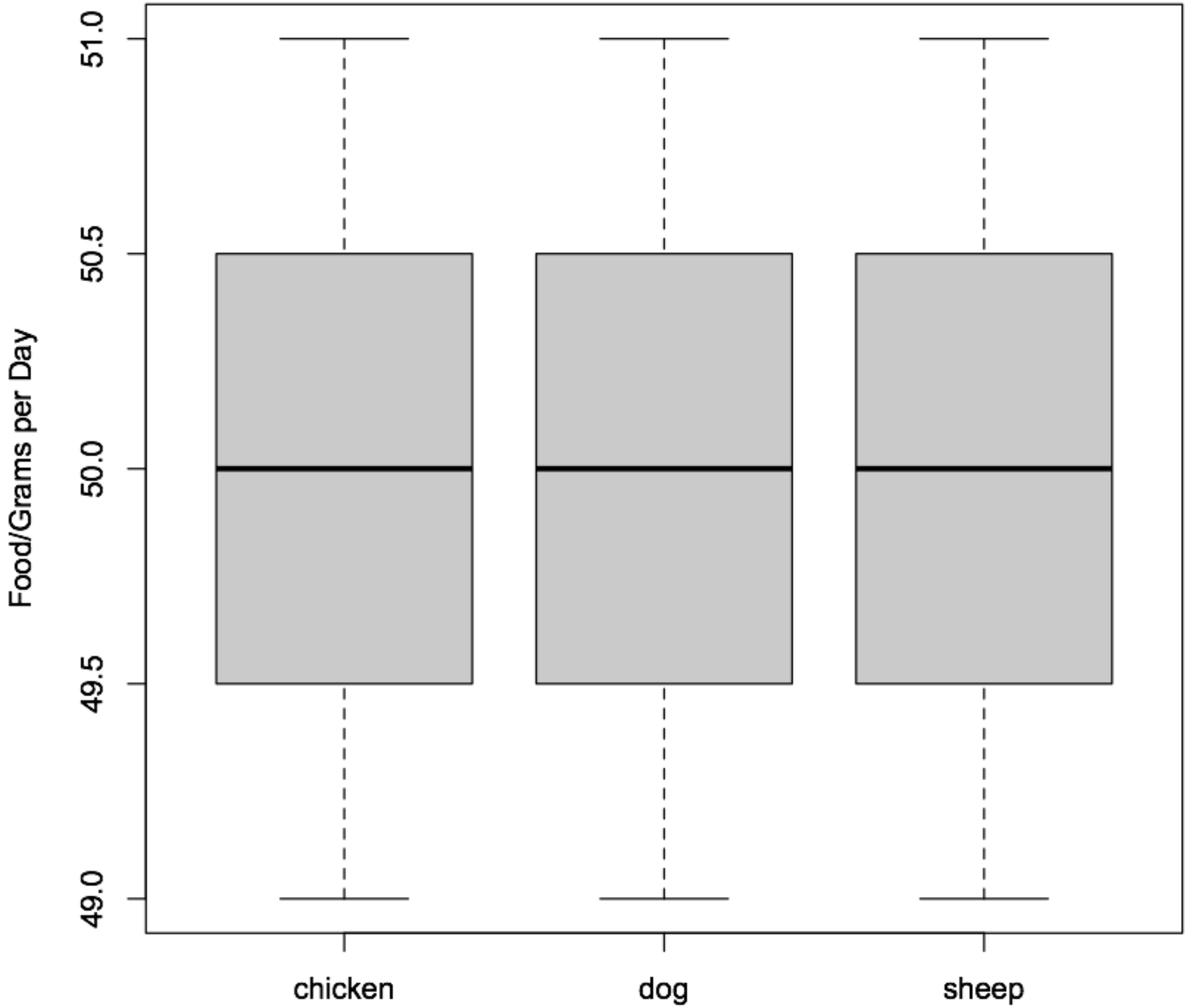
# Yet Another Scenario

| animal | grams/day |
| --- | --- |
| chicken | 49 |
| chicken | 50 |
| chicken | 51 |
| dog | 49 |
| dog | 50 |
| dog | 51 |
| sheep | 49 |
| sheep | 50 |
| sheep | 51 |
| **std** | **0.9** |
| **XBAR** | **50** |

# Yet Another Scenario

| animal | grams/day |
|---|---|
| chicken | 49 |
| chicken | 50 |
| chicken | 51 |
| dog | 49 |
| dog | 50 |
| dog | 51 |
| sheep | 49 |
| sheep | 50 |
| sheep | 51 |
| **std** | **0.9** |
| **XBAR** | **50** |

| | chicken | dog | sheep |
|---|---|---|---|
| | 49 | 49 | 49 |
| | 50 | 50 | 50 |
| | 51 | 51 | 51 |
| **std** | **1** | **1** | **1** |
| **XBAR** | **50** | **50** | **50** |



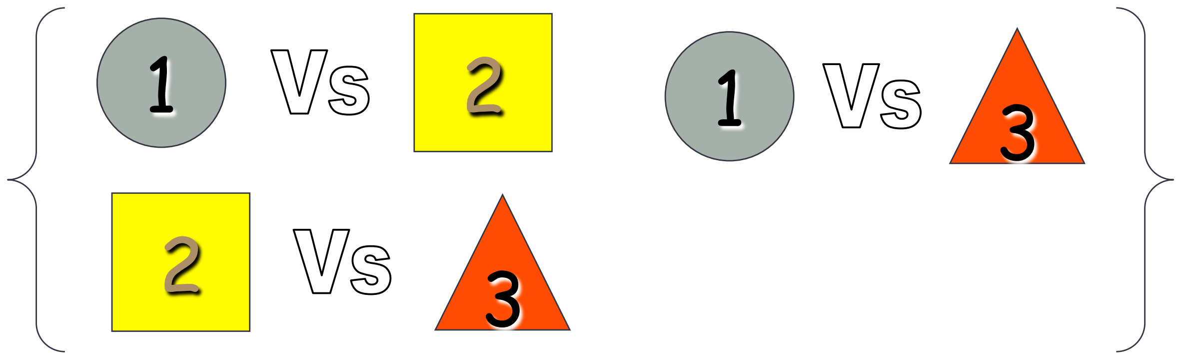
Food/Grams per Day
51.0
50.5
50.0
49.5
49.0
chicken
dog
sheep

# ANOVA: When And Why

- When we want to compare means we can use a *t*-test. This test has limitations:
  - You can compare only 2 means: often we would like to compare means from 3 or more groups.
  - It can be used only with one Predictor/Independent Variable.
- ANOVA
  - Compares several means.

# Why Not Use Lots of *t*-Tests?

- If we want to compare several means why don't we compare pairs of means with *t*-tests?
  - Can't look at several independent variables.
  - Inflates the Type I error rate.


$$\boxed{\text{Familywise Error} = 1 - (0.95)^n}$$

# What Does ANOVA Tell us?

- Null Hyothesis:
  - Like a t-test, ANOVA tests the null hypothesis that the means are the same.
- Research Hypothesis:
  - The means differ.
- ANOVA is an Omnibus test
  - It test for an overall difference between groups.
  - It tells us that the group means are different.
  - It doesn't tell us exactly which means differ.

# Theory of ANOVA

- We calculate how much variability there is between scores
  - Total Sum of squares ($SS_T$).

- We then calculate how much of this variability can be explained by the model we fit to the data
  - How much variability is due to the experimental manipulation, Model Sum of Squares ($SS_B$)...

- … and how much cannot be explained
  - How much variability is due to individual differences in performance, Residual Sum of Squares ($SS_W$).

# Theory of ANOVA

- We compare the amount of variability explained by the Model (experiment), to the error in the model (individual differences)
  - This ratio is called the *F*-ratio.
- If the model explains a lot more variability than it can't explain, then the experimental manipulation has had a significant effect on the outcome (DV).

# Theory of ANOVA

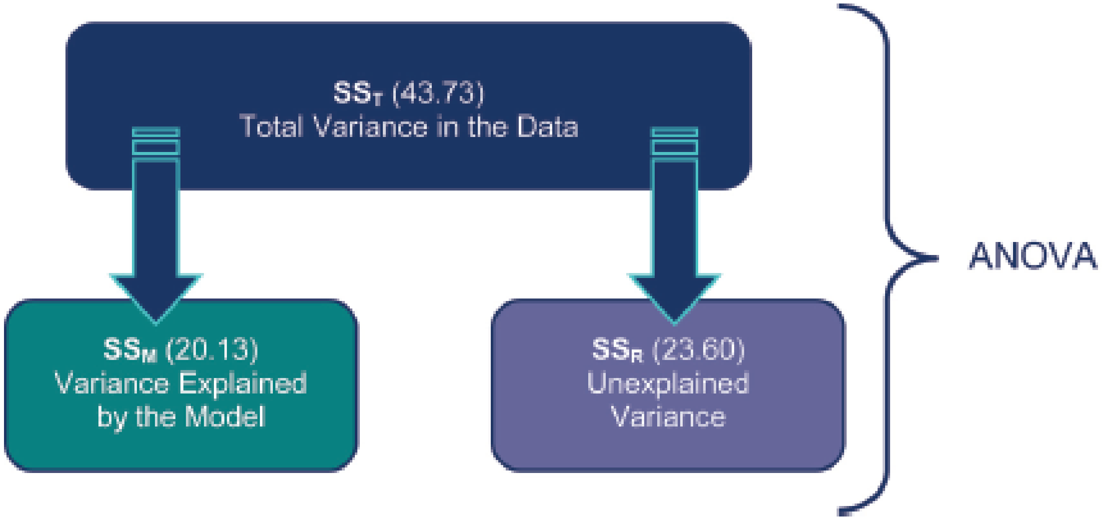



FIGURE 10.3
Partitioning variance for ANOVA

- If the experiment is successful, then the model will explain more variance than it can't
  - $SS_B$ will be greater than $SS_W$

# Test Requirements

- Dependent variable must be Interval-Ratio
- Independent variable must be nominal or ordinal, with 3 or more categories (but usually no more than 5)

# Quiz Yourself

- Could you conduct the test with two interval-ratio variables?  How?

# Follow the 5 Steps

- Step One:
  - Random/EPSEM samples
  - I/R dependent variable, 3-5 categories in independent variable
  - The categories are about the same size (a limitation of ANOVA)

# Follow the 5 Steps

- Step Two, State the Null Hypothesis:

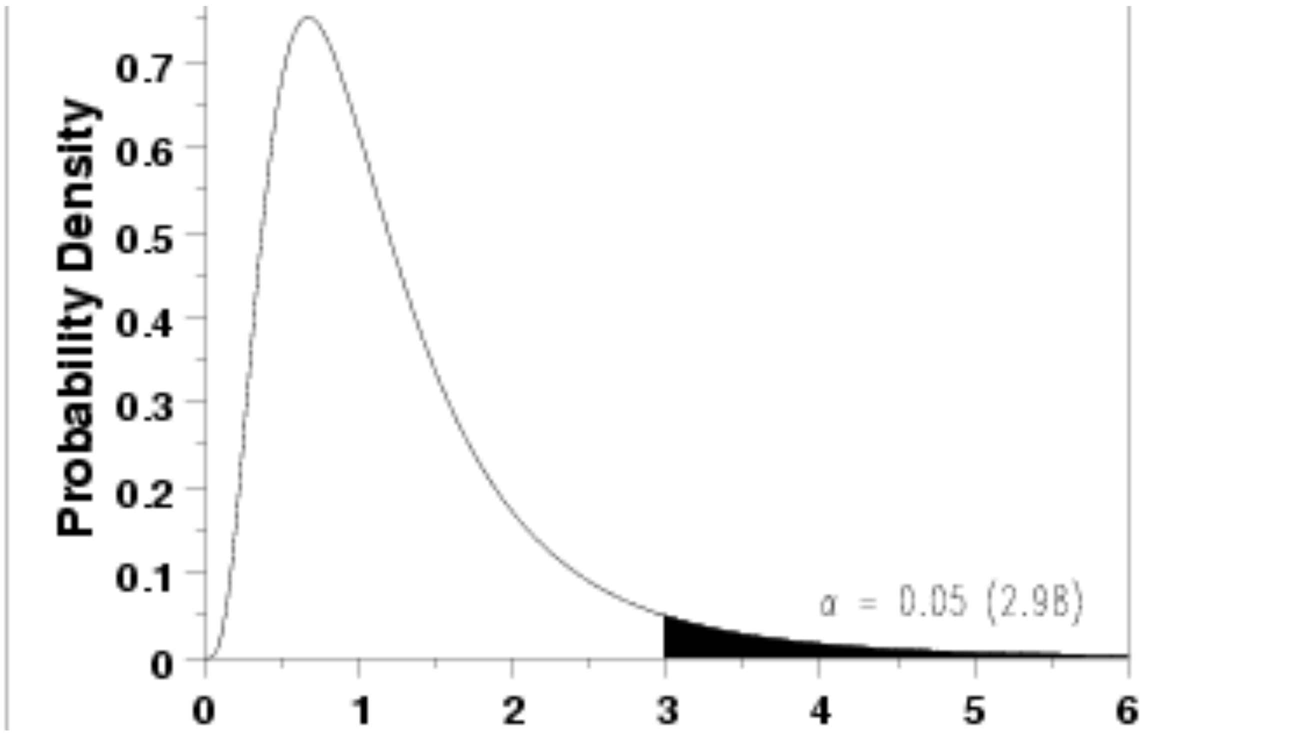
- $H_0$: $\mu_1 = \mu_2 = \mu_n$
- $H_1$: At least one of the means is different.

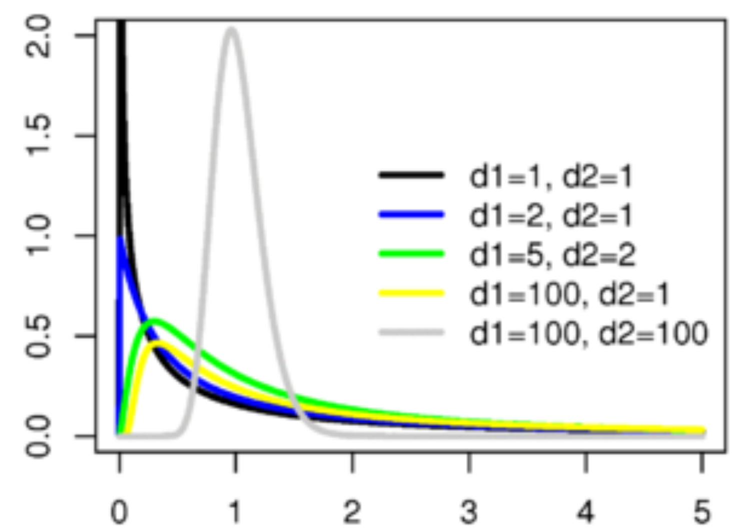
# Follow the 5 Steps

- Step 3:
  - F Distribution (Appendix D)
  - Set Alpha (0.05)
  - Find $F_{critical}$ using dfw & dfb

# Follow the 5 Steps

- Step 4: Compute the Test Statistic
- Step 5: Interpret the Results


Probability Density
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
1
2
3
4
5
6
α = 0.05 (2.98)


2.0
1.5
1.0
0.5
0.0
0
1
2
3
4
5
d1=1, d2=1
d1=2, d2=1
d1=5, d2=2
d1=100, d2=1
d1=100, d2=100

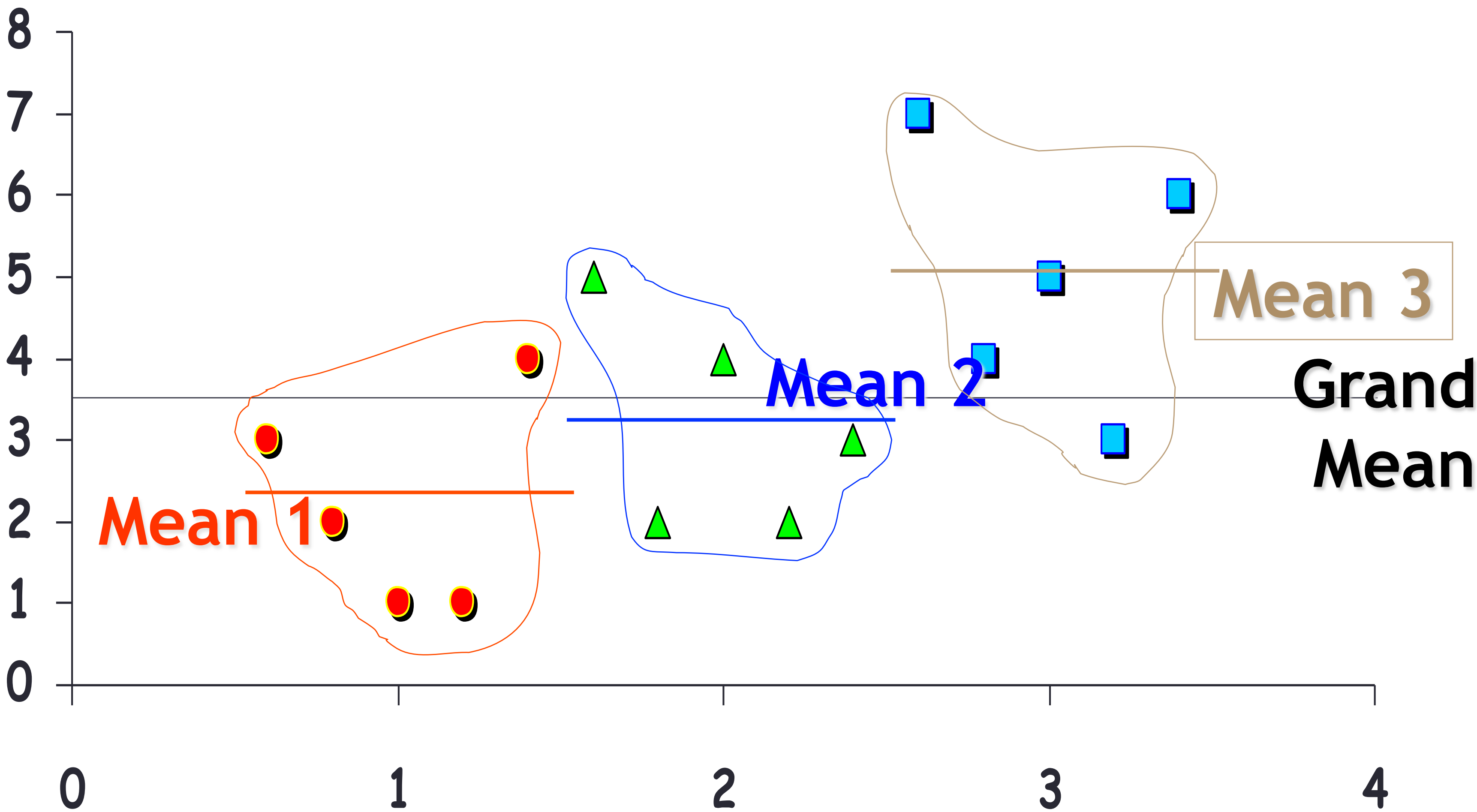
# ANOVA by Hand

- Testing the effects of a medicine using three groups:
  - Placebo (Sugar Pill)
  - Low Dose
  - High Dose
- The Outcome/Dependent Variable (DV) was an objective measure of treatment effect (the higher the score, the more effective the treatment).

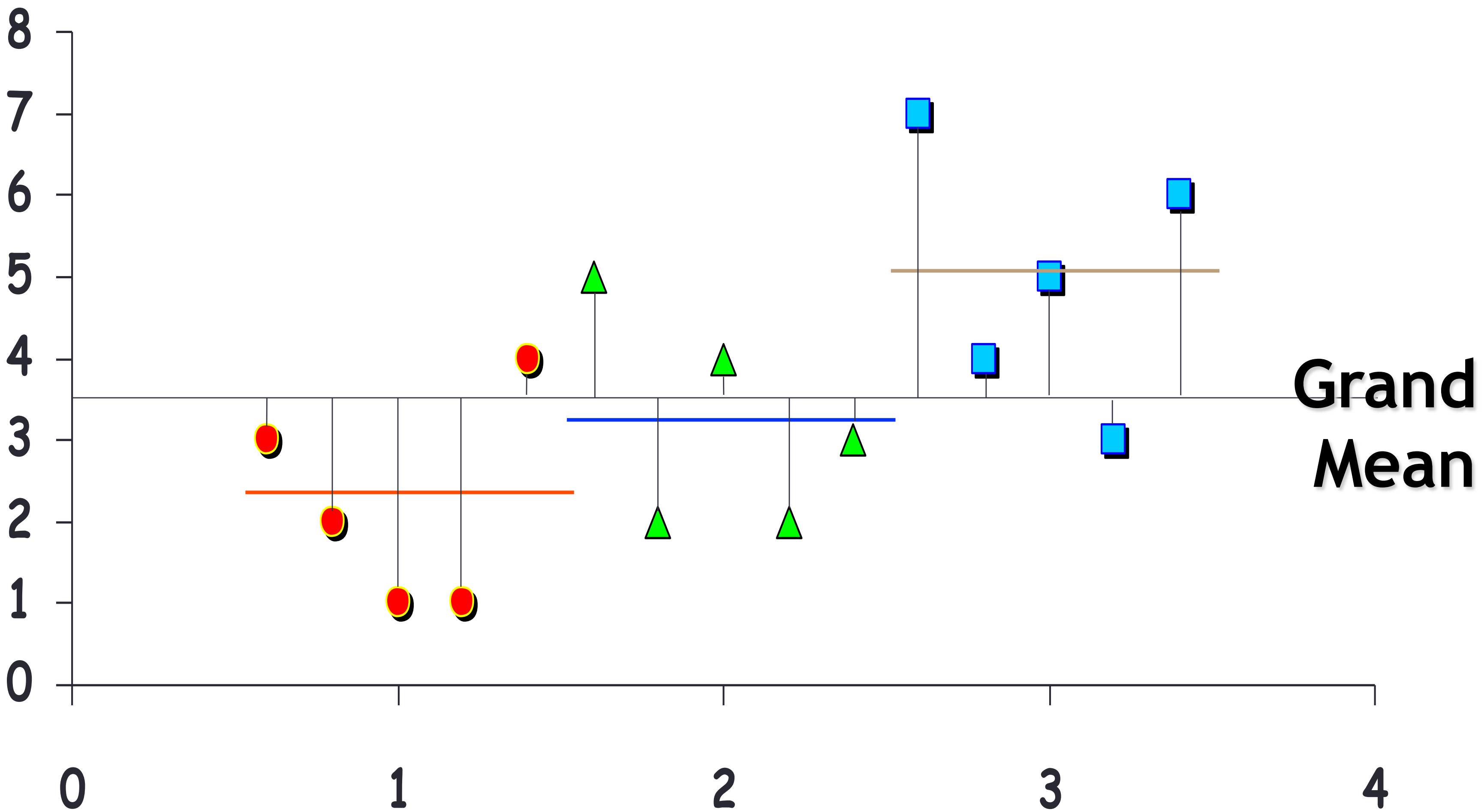
# The Data

| | Placebo | Low Dose | High Dose |
|---|---|---|---|
| | 3 | 5 | 7 |
| | 2 | 2 | 4 |
| | 1 | 4 | 5 |
| | 1 | 2 | 3 |
| | 4 | 3 | 6 |
| $\bar{X}$ | 2.20 | 3.20 | 5.00 |
| $s$ | 1.30 | 1.30 | 1.58 |
| $s^2$ | 1.70 | 1.70 | 2.50 |
| | Grand Mean = 3.467 | Grand SD = 1.767 | Grand Variance = 3.124 |

# The data:



# Total Sum of Squares ($SS_T$):



# Step 1: Calculate $SS_T$

$$SS_T = \sum (x_i - \bar{x}_{grand})^2$$

$$SS_T = \sum x^2 - N\bar{x}^2$$

Computational
formula

# Step 1: Calculate $SS_T$

$$SS_T = \sum x^2 - N\bar{x}^2$$

$$\begin{aligned} SS_T &= 224 - 15(3.467)^2 \\ &= 224 - 15(12.02) \\ &= 224 - 180.3 \\ &= 43.7 \end{aligned}$$

| X | X^2 |
|---|---|
| 3 | 9 |
| 2 | 4 |
| 1 | 1 |
| 1 | 1 |
| 4 | 16 |
| 5 | 25 |
| 2 | 4 |
| 4 | 16 |
| 2 | 4 |
| 3 | 9 |
| 7 | 49 |
| 4 | 16 |
| 5 | 25 |
| 3 | 9 |
| 6 | 36 |
| SUM | 224 |

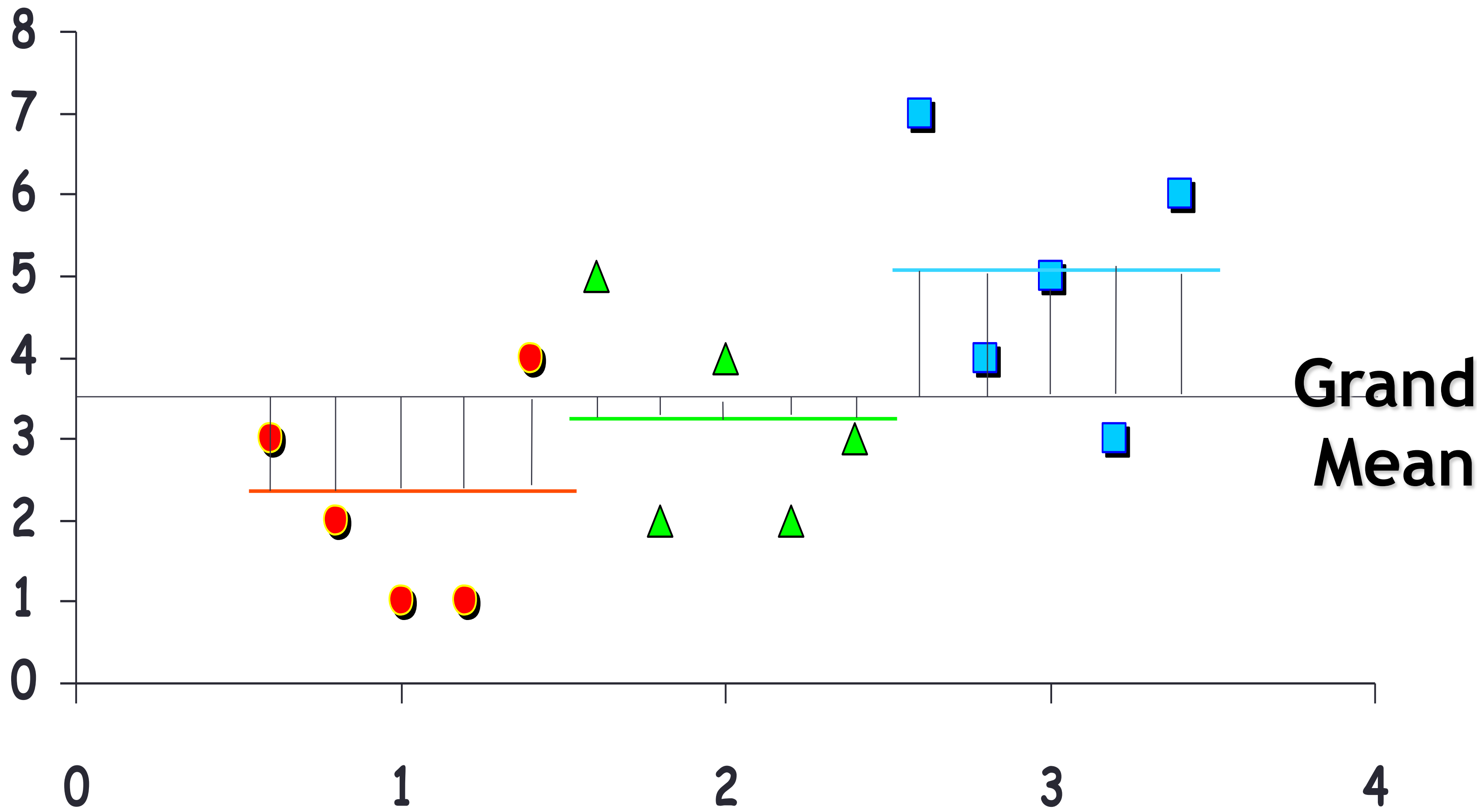
Grand Mean = **3.467** Grand SD = **1.767** Grand Variance = **3.124**

# Degrees of Freedom (df)

- Degrees of Freedom (*df*) are the number of values that are free to vary.
- In general, the *df* are one less than the number of values used to calculate the SS.

$$\boxed{df_T = (N - 1) = 15 - 1 = 14}$$

# Model Sum of Squares ($SS_B$):



# Step 2: Calculate $SS_B$

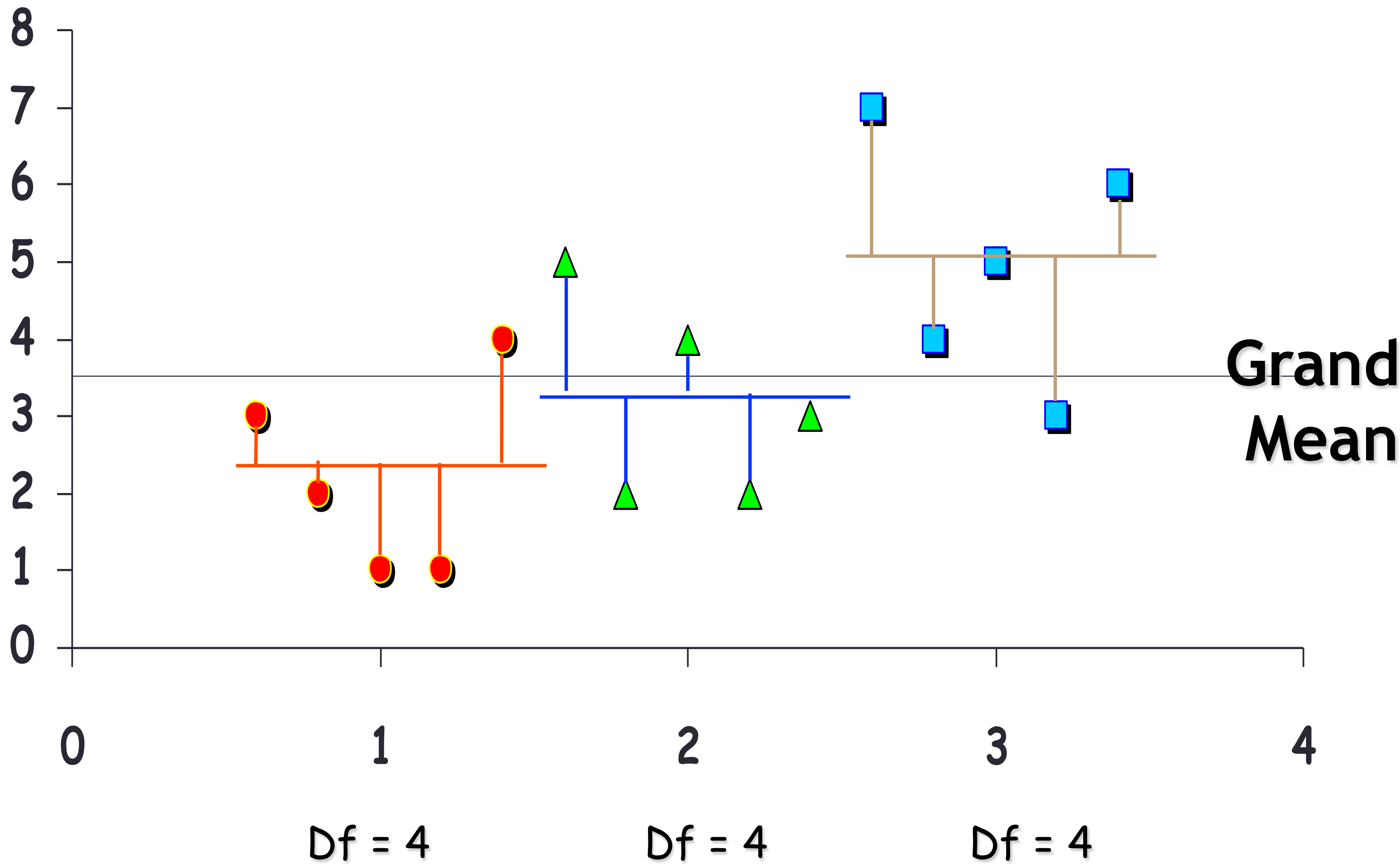
$$SS_B = \sum n_i(\bar{x}_i - \bar{x}_{grand})^2$$

$$\begin{aligned} SS_B &= 5(2.2-3.467)^2 + 5(3.2-3.467)^2 + 5(5.0-3.467)^2 \\ &= 5(-1.267)^2 + 5(-0.267)^2 + 5(1.533)^2 \\ &= 8.025 + 0.355 + 11.755 \\ &= 20.135 \end{aligned}$$

# Model Degrees of Freedom

$$\boxed{df_B = (k-1) = 3 - 1 = 2}$$

# Residual Sum of Squares ($SS_W$):



# Step 3: Calculate $SS_W$

$$\boxed{SS_W = SST - SSB}$$

SSW = 43.73 – 20.14

SSW = 23.59

# Residual Degrees of Freedom

$$df_W = N - k$$

$$= 12$$

# Step 4: Calculate the Mean Squared Error

$$MS_B = \frac{SS_B}{df_B} = \frac{20.135}{2} = 10.067$$

$$MS_W = \frac{SS_W}{df_W} = \frac{23.60}{12} = 1.967$$

# Step 5: Calculate the *F*-Ratio

$$F = \frac{MS_B}{MS_W}$$

$$F = \frac{MS_B}{MS_W} = \frac{10.067}{1.967} = 5.12$$

# Step 6: Construct a Summary Table

| Source | SS | df | MS | *F* |
|---|---|---|---|---|
| Model | 20.14 | 2 | 10.067 | 5.12* |
| Residual | 23.60 | 12 | 1.967 | |
| Total | 43.74 | 14 | | |

# Data in R

```
> > experiment

   treatmentOutcome    dose
1                 3 placebo
2                 2 placebo
3                 1 placebo
4                 1 placebo
5                 4 placebo
6                 5     low
7                 2     low
8                 4     low
9                 2     low
10                3     low
11                7    high
12                4    high
13                5    high
14                3    high
15                6    high
```

# Results in R

```
> aggregate(treatmentOutcome~dose, experiment, mean)

     dose treatmentOutcome
1    high              5.0
2     low              3.2
3 placebo              2.2
```

# Results in R

```
> anova.results <- aov(treatmentOutcome~dose, experiment)
> summary(anova.results)

             Df Sum Sq Mean Sq F value Pr(>F)
dose          2  20.13  10.067   5.119 0.0247 *
Residuals    12  23.60   1.967
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '
' 1
```

# Follow-Up Tests

- The *F*-ratio tells us only that the experiment was successful
  - i.e. group means were different
- It does not tell us specifically which group means differ from which.
- We need additional tests to find out where the group differences lie.

# How?

- Multiple *t*-tests
  - We saw earlier that this is a bad idea
- *Post Hoc* Tests
  - Not Planned (no hypothesis)
  - Compare all pairs of means

# Post Hoc Tests

- Compare each mean against all others.
- In general terms they use a stricter criterion to accept an effect as significant.
  - Hence, control the familywise error rate.

# Post Hoc Test Output

```
> TukeyHSD(anova.results)

  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = treatmentOutcome ~ dose, data
= experiment)

$dose
             diff       lwr        upr     p adj
low-high     -1.8 -4.166241  0.5662412 0.1474576
placebo-high -2.8 -5.166241 -0.4337588 0.0209244
placebo-low  -1.0 -3.366241  1.3662412 0.5162761
```

# Does mood vary by sunshine level?

| Sunshine | Mood |
|---|---|
| High | 8 |
| Low | 3 |
| Medium | 5 |
| Medium | 7 |
| Low | 4 |
| High | 10 |
| High | 10 |
| Medium | 5 |
| Low | 1 |
| Medium | 6 |

# Next:

- Chi-Square
- Chapter 1 & 2 in Tufte book