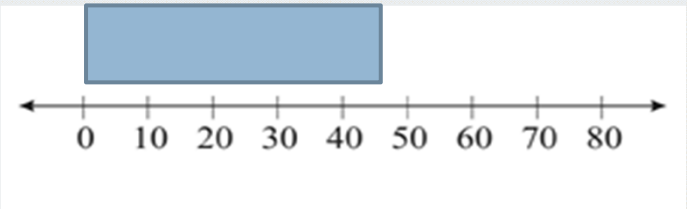
Organizing and Analyzing Data

Measurement, Analysis, Display & Significance

Written by and used with permission from Dana Gonzalez, Greater New Orleans STEM Initiative

Measurement

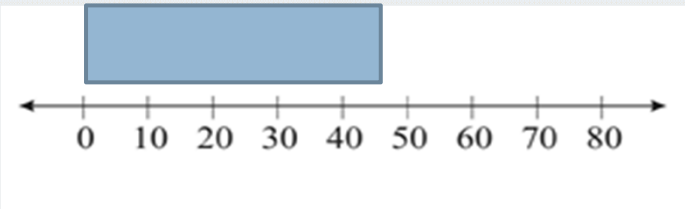
- Accuracy
  - 45.324905832 from this instrument?



  - No → most precise estimate could be to the tenths place because instrument's markings are only specific enough to allow estimation of one digit

Measurement

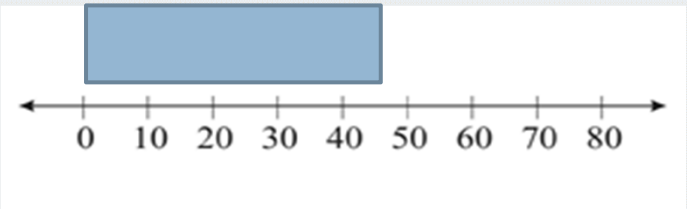
- The rule:
  - Report a measurement to the precision of the instrument plus one estimated digit



  - What would the reading for this measurement be?

Measurement

- The rule:
  - Report a measurement to the precision of the instrument plus one estimated digit



  - What would the reading for this measurement be?

  - Answer: 45.3

Data Tables

- Tables need to be titled
- Data tables must include units of measurement

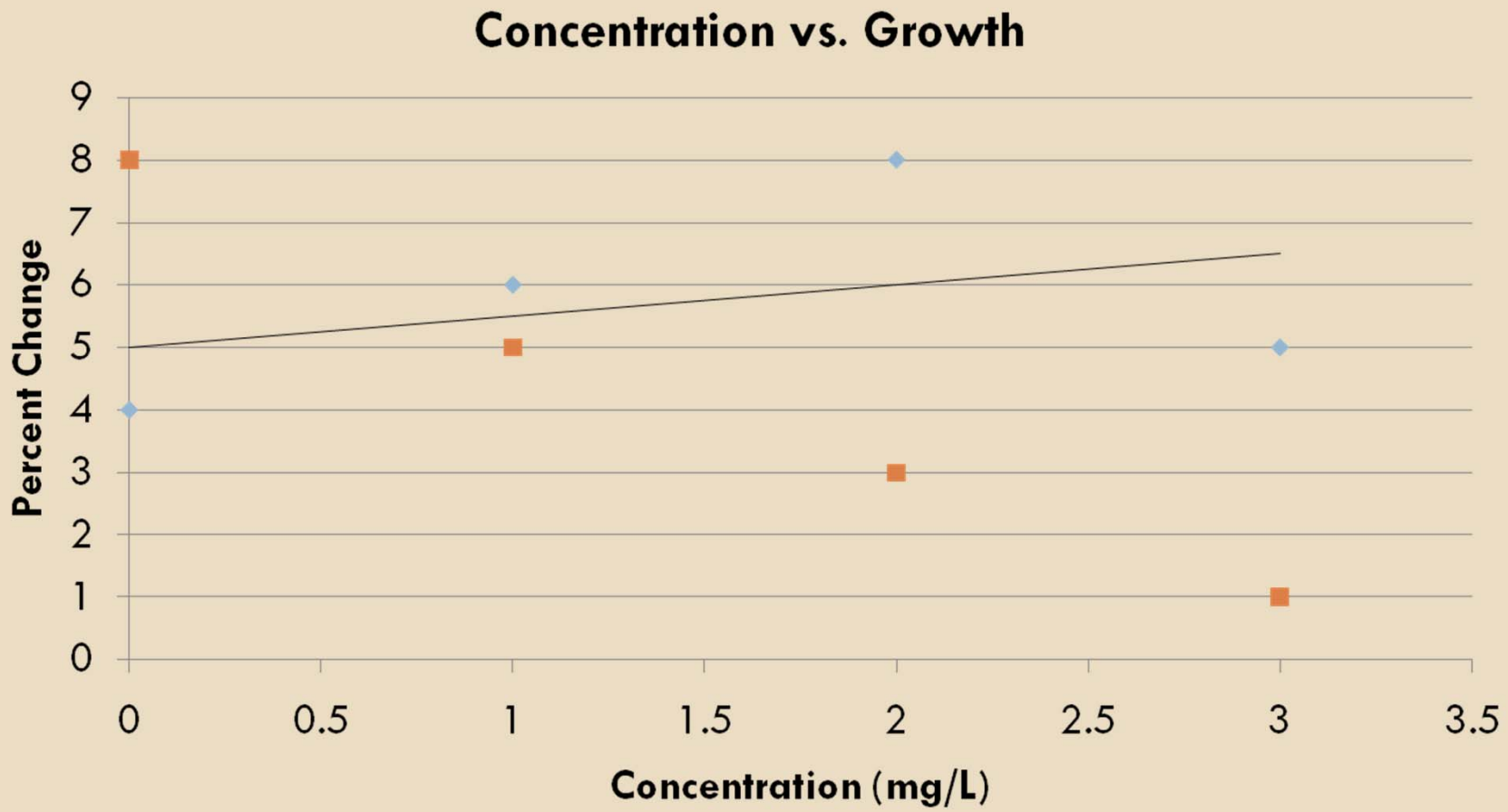
Trial #	Treatment #1 Mass (g)	Treatment #2 Mass (g)
1	9	13
2	10	14
3	8	12

Graphing Data

- Categorical data: does not fit a mathematical formula
  - Use a bar graph
  - Use a histogram for categorical range data (i.e. grades)

- Continuous data: can be described by a line with a slope or a smooth curve
  - Scatter plot
  - Use best-fit line, don't connect the dots

Graphing: Best-Fit Line



Graphing: Variables

- What type of data are these?
- Which values are plotted on the x-axis?
- Which values are plotted on the y-axis?

Concentration	Percent Change
0%	4
1%	6
2%	8
3%	5

Memory device: "IX" and "DY"

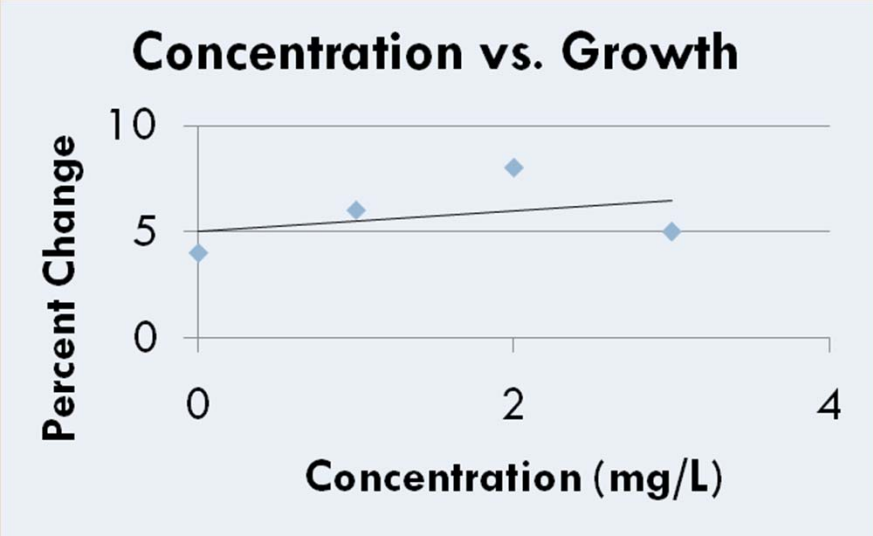
Graphing: Variables

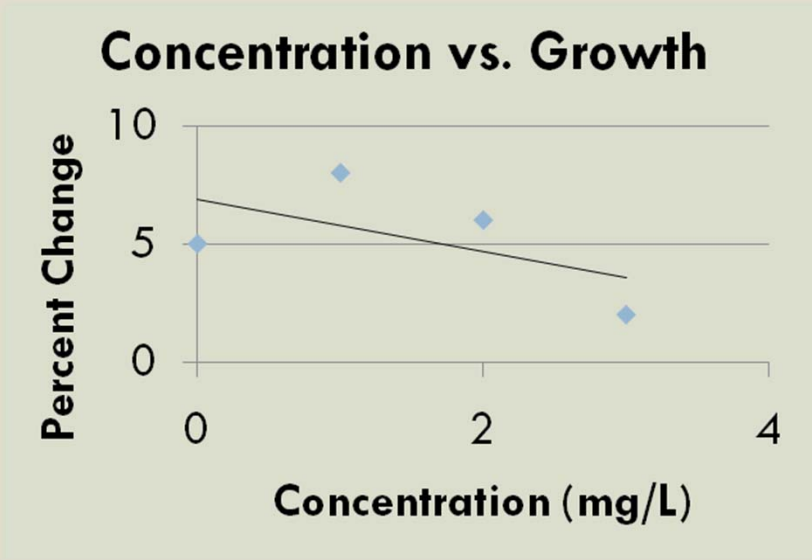
- What type of data are these? Continuous data
- Which values are plotted on the x-axis?
  - Concentration, the independent variable
- Which values are plotted on the y-axis?
  - Percent change, the dependent variable

Concentration	Percent Change
0%	4
1%	6
2%	8
3%	5

Memory device: "IX" and "DY"

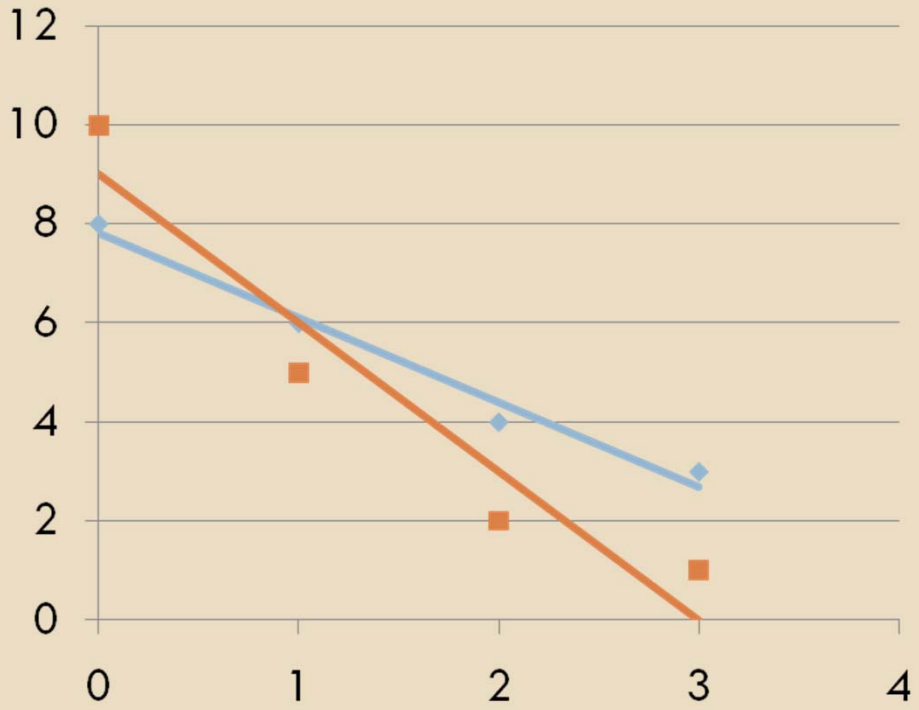
Graph: Relationship of *x* to *y*





- As x-values increase, what is happening to the y-values?
  - Which graph shows an inverse proportion?
  - Which graph shows a direct proportion?

Graphical to Algebraic Representation



Rate (Slope) = Rise / Run

m (Slope) = Δy / Δx

y = mx + b

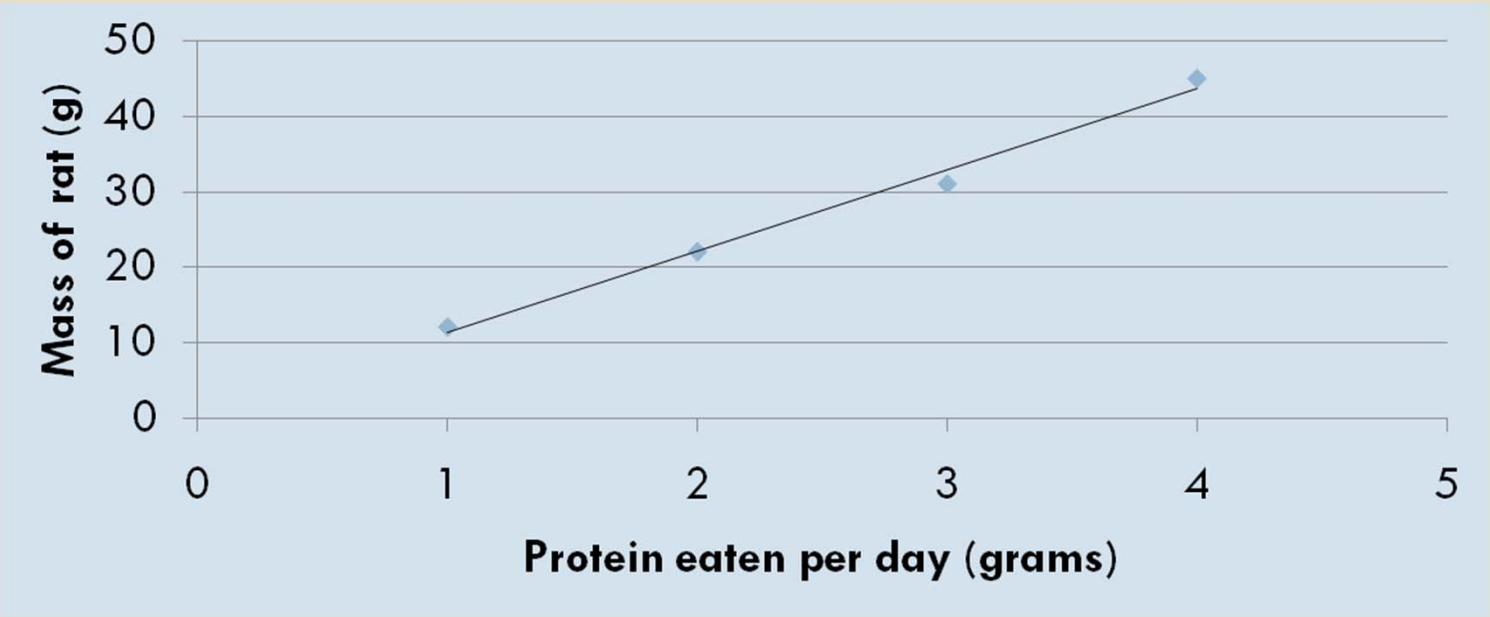
- Which line shows the fastest rate?
- *"Ski slope" trick: steeper graph = fastest rate*

Graphing: Interpolation

- Interpolation: using a graphical relationship to predict a measurement you *haven't actually measured*
  - **Interpolation is between data points!**
  - Use slope to interpolate data
  - Extrapolation not allowed for experimental data

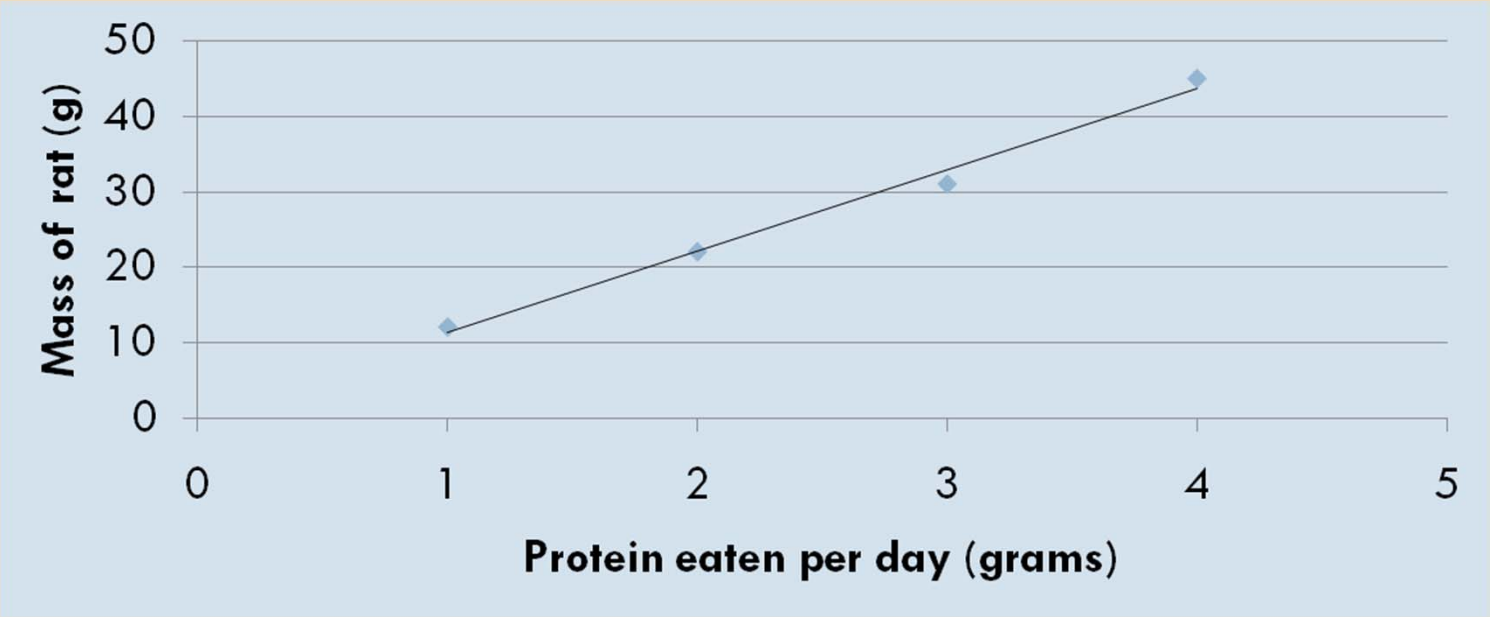
- You are **not** allowed to use the "projected value" or values outside the range of data

Graphing: Interpolation



- A rat of mass 25 g ate how much protein per day?
- Eating 3.6 g of protein per day would result in a rat of what mass?

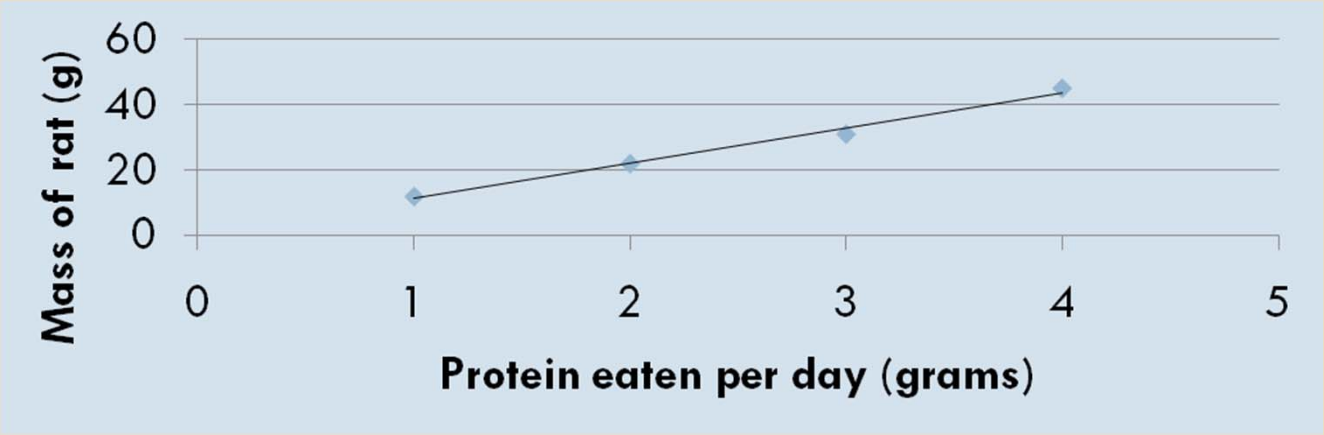
Graphing: Interpolation



- A rat of mass 25 g ate how much protein per day?
  - About 2.3 g (estimated from graph)
- Eating 3.6 g of protein per day would result in a rat of what mass?
  - A rat with a mass of about 40 grams (estimated from graph)

Graphing: Extrapolation

- *Extrapolation* involves making predictions outside of the data range
- Not always valid because the relationship is not necessarily linear and predictable



- A rate having a mass of 50 grams would eat how many grams of protein a day?

Extrapolation

- In order to extrapolate, you have to re-measure the data in order to include the range you want to measure
- ***Reasons not to extrapolate:***
  - *Data changes at high and low extremes*
  - *High and low values often approach zero or plateau in value*

Statistics: Evaluating a Hypothesis

- Hypothesis should express a cause and effect relationship:
  - The dependent variable will occur as a result of an independent variable applied to a system as measured by a method.
- You will need to write a null hypothesis and an alternate hypothesis
- A null hypothesis is always in the form of:
  - The dependent variable will NOT occur as a result of an independent variable applied to a system as measured by a method.

  OR
  - The dependent variable will have no effect of the independent variable in the system as measured by a method.

Statistics: Evaluating a Hypothesis

- **Example:**

  - Carbohydrate consumption will <u>cause</u> an increase in the mass of rats.

  - *What is the independent variable?*
  - *What is the dependent variable?*
  - *Is this a null hypothesis or an alternate hypothesis?*

Statistics: Evaluating a Hypothesis

- **Example:**

  - Carbohydrate consumption will <u>cause</u> an increase in the mass of rats.

  - *What is the independent variable?*
  - *What is the dependent variable?*
  - *Is this a null hypothesis or an alternate hypothesis?*

Statistics: Evaluating a Hypothesis

- **Example:**

  - Carbohydrate consumption will cause an increase in the mass of rats.

  - *What is the independent variable?*
    - *Carbohydrate consumption*
  - *What is the dependent variable?*
    - *The mass of the rats*
  - *Is this a null hypothesis or an alternate hypothesis?*
    - *This is an alternate hypothesis, since it implies a connection between the IV and DV.*

Statistics: Evaluating a Hypothesis

- **Example of a null hypothesis:**

  - Carbohydrate consumption will *not* cause an increase in the mass in rats.

    OR

  - There is no correlation between carbohydrate consumption and mass change in rats.

Statistics: Evaluating a Hypothesis

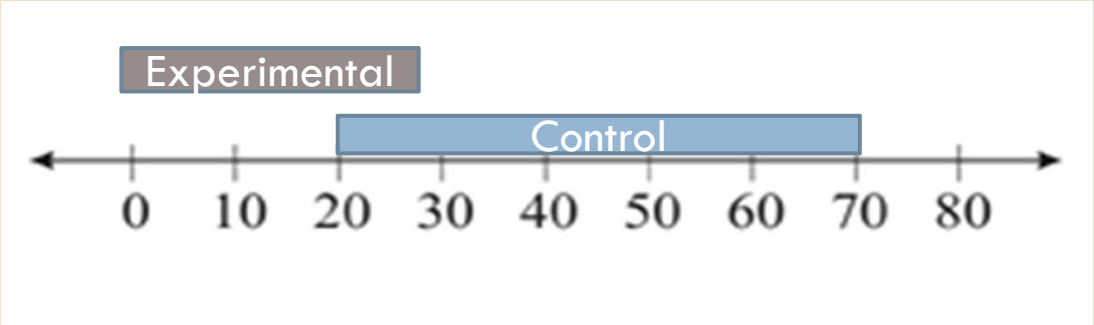
- Are the two means significantly different from each other?

Group	Mean Mass
Control- 10 mice (no carbohydrate diet)	60 g
Experimental- 10 mice (carbohydrates diet)	45 g
10% of control average	(0.1) x (60) = 6

- ***The 10% Rule:*** if the experimental mean (values) are more than 10% different from the control mean (values), the data can be considered *significant.*
  - *Not that this is NOT the same as "statistically significant."*

Statistics: Evaluating a Hypothesis

- Hypothesis: Carbohydrate consumption will NOT cause an increase in the mass of rats.



*Note that you evaluate the "null" hypothesis first

If the difference between control and experimental data is greater than 10% of the control, there is a significant difference between them.

Choosing a Statistical Test

- Define the nature of the independent and dependent variables
  - Categorical
    - Can't put in mathematical formula
    - Cannot interpolate
    - Bar graphs/histograms
  - Continuous
    - Goes in mathematical formula
    - C*an* interpolate

Choosing a Statistical Test

- Categorical variable values cannot be sequentially ordered or differentiated from each other using a mathematical method
  - *Interpolation makes no sense!*
- Categorical data examples include:
  - Location of samples taken
  - Treatments-
    - Amounts of chemicals added to different test groups
  - Hurricane damage by Category storm

Choosing a Statistical Test

- Continuous variables are numeric values that can be ordered sequentially and that do not naturally fall into discrete ranges.
  - Data can be any value within reason
  - *Can interpolate!*
- Continuous data examples include:
  - Weight or mass
  - Number of seconds it takes to perform a task
  - Number of words on a page

Choosing a Statistical Test

Independent Variable	Dependent Variable	
	Categorical	Continuous
Categorical	Chi-square	t-test, ANOVA
Continuous	Linear discriminate analysis Quadratic discriminate analysis	Regression

- For your IRP data, you will probably be using either a chi-square test, a t-test or an ANOVA
- If you are not sure