Agenda

Homework (AP)
Pg.158-159 #1-5
DATA COLLECTION, use link

- Warm Up — 5 min
- Explanatory vs. response — 15 min
- Scatterplots — 10 min
- Direction, Form, Strength (DFS) — 10 min
- Car data — 15 min
  - *Copies*
  - *Skip?*
- Hand out tests/project — 15 min
  - ~~*Update*~~
  - ~~*Project handout → check, copies*~~
  - *Find project examples*
- Exit Pass — 5 min

Warm Up

1. (Test #10) A soft-drink machine can be regulated so that it discharges an average number of ounces per cup. If the ounces of fill are normally distributed with a standard deviation of 0.4 ounces, what should the average be set at so that 6-ounce cups will overflow only 2% of the time?
   A. 5.18
   B. 5.60
   C. 6.00
   D. 6.01
   E. 6.82

tinyurl.com/602housesurvey

2. Describe two variables that might vary directly.

3. Describe two variables that might vary inversely (indirectly).

4. Stretch out the fingers of your left hand as far as you can. Use a ruler to measure the distance (inches) from the tip of your thumb to the tip of your smallest finger.

5. Convert your height into inches.

- Due tomorrow

tinyurl.com/602housesurvey

Chapter 3

- Chapter 1 and 2 → "Univariate" (SINGLE) data
  - 1 → Distribution of univariate data
  - 2 → Individuals within univariate data

- Chapter 3 → "Bivariate" (PAIRED) data
  - Two variables and how they might co-vary or predict each other

Chapter 3 summary

Test
Friday 2/7

- Commenting (DFS) on scatterplots
- Correlation coefficient (r)
- Coefficient of determination (r^2)
- Regression lines
- Residuals and residual plots

Notes

Explanatory vs. Response

- Response variable measures outcome
- Explanatory variable helps explain (influence, predict) response variable
- Also called “independent” and “dependent” ☹
- *Example: Parents’ income and the years of education their eldest child completes.*

Which is which? (explanatory or response)

1. The amount of time a student spends studying for a statistics exam. Their grade on the exam.
2. Inches of rain in the growing season. Yield of corn in bushels per acre.
3. The weight of a person. The height of a person.
4. A student's grade in AP Statistics. A student's grade in AP Biology.
5. The number of wars a country has been involved with in the last 100 years. A country's average price of a Big Mac from McDonald's.
6. The life expectancy of a person. The number of shoes they own.

ANSWERS, explanatory underlined

1. The amount of time a student spends studying for a statistics exam. Their grade on the exam.
2. Inches of rain in the growing season. Yield of corn in bushels per acre.
3. The weight of a person. The height of a person.
4. A student's grade in AP Statistics. A student's grade in AP Biology. (could be argued)
5. The number of wars a country has been involved with in the last 100 years. A country's average price of a Big Mac from McDonald's. (*totally* could be argued)
6. The life expectancy of a person. The number of shoes they own.

Notes

2 of 3

Scatterplots

- Used to display relationship between two quantitative variables
  - Explanatory on x-axis, response on y-axis
- 1st of 6 "graph types" on calc, looks like dots. Enter data into L_1 and L_2.
- Geogebra (show)

Explanatory	Response
1	5
2	6
3	3
3	5
7	1

Sleep	GPA	Sleep	GPA
8	3.32	5	2.9
6	4.21	5	3.5
6	3.75	7	3
5	3.9	4	3.4
4	4.13	6	3.8
8	3.8	7.5	3.6
0	1.7	6	3.5
6	3.25	6	3.25
9	4.15	6	3.5
8	3.25	7	3.7
8	3.75	9	3.79
5	3.8	7	3
4	3.8	7	3.8
8.5	3.07	6	4.1
7	3.36	5	3.33
5	2.75	3	4.41
7	3.8	2	4.03
7	4.13	4	3.5
8	4.25	6	4.44
5	3.74	6	4.03
7	3.1	4	4.38
5.5	3.9	4	3.5
8	3.8	8	4.38
7	2.75	6	3.95
9	3.75	8	3.5
7	3.5	8	4
7	4.13	0	4.23
4	4	6	3.75
6	3.5	4	4.16
8	3.75	6	3.56
5	4	5.5	4.25
10	4.15	6	3.75
2	3.6	8	3.2
9	4.2	6	4
8	3.4	7	3
5	3.9	4	3.7
3	3.5	6	4.07
5	4.23	6	3.5
8	4.57	7	3.5
6	3.3	8	3.75
1	4.13	9	3.75
7	4.06	8	3.76
8	4.25	8	3.5
7	4	9.5	3.8
8	2.47	8	3.83
7	2.5	8	3.8
7	3.4	7	4
5	3	8	3.34
8	2.5	6.5	3.1
5	3.5	7	4.06
8	3.5	4	4.22
6	3.02	4	3.75
2	3.5	4	3.43
8.5	3	4	3.75
5	3.25	7	2.5
7	3.6	7	3.2

You try!

- Beers and BAC

n=16

Beers	BAC
5	0.1
2	0.03
9	0.19
8	0.12
3	0.04
7	0.095
3	0.07
5	0.06
3	0.02
5	0.05
4	0.07
6	0.1
5	0.085
7	0.09
1	0.01
4	0.05

Notes

Interpreting Scatterplots = **DFS**

- **D**irection
  - positive vs. negative
  - direct vs. inverse
- **F**orm
  - Shape (linear vs. nonlinear)
  - Clusters, outliers
- **S**trength
  - Use ***r*** if you can. If not, use phrases like “fairly weak,” “fairly strong”, or “very strong.”

Worksheet: Practice with DFS

- I will give you and your partner a shared worksheet.
- Fill in the table using the letters of the scatterplots. Rank them by their direction and strength. Argue, disagree, etc.
- Pick 2 scatterplots each (4 total). Take turns verbally commenting on your scatterplots.
  - EXAMPLE: *"A" is vaguely negative, not very linear, and quite weak.*

	Negative				**None**	**Positive**			
	Strongest			Weakest	None	Weakest			Strongest
Letter of scatter plot									

Worksheet: Practice with DFS

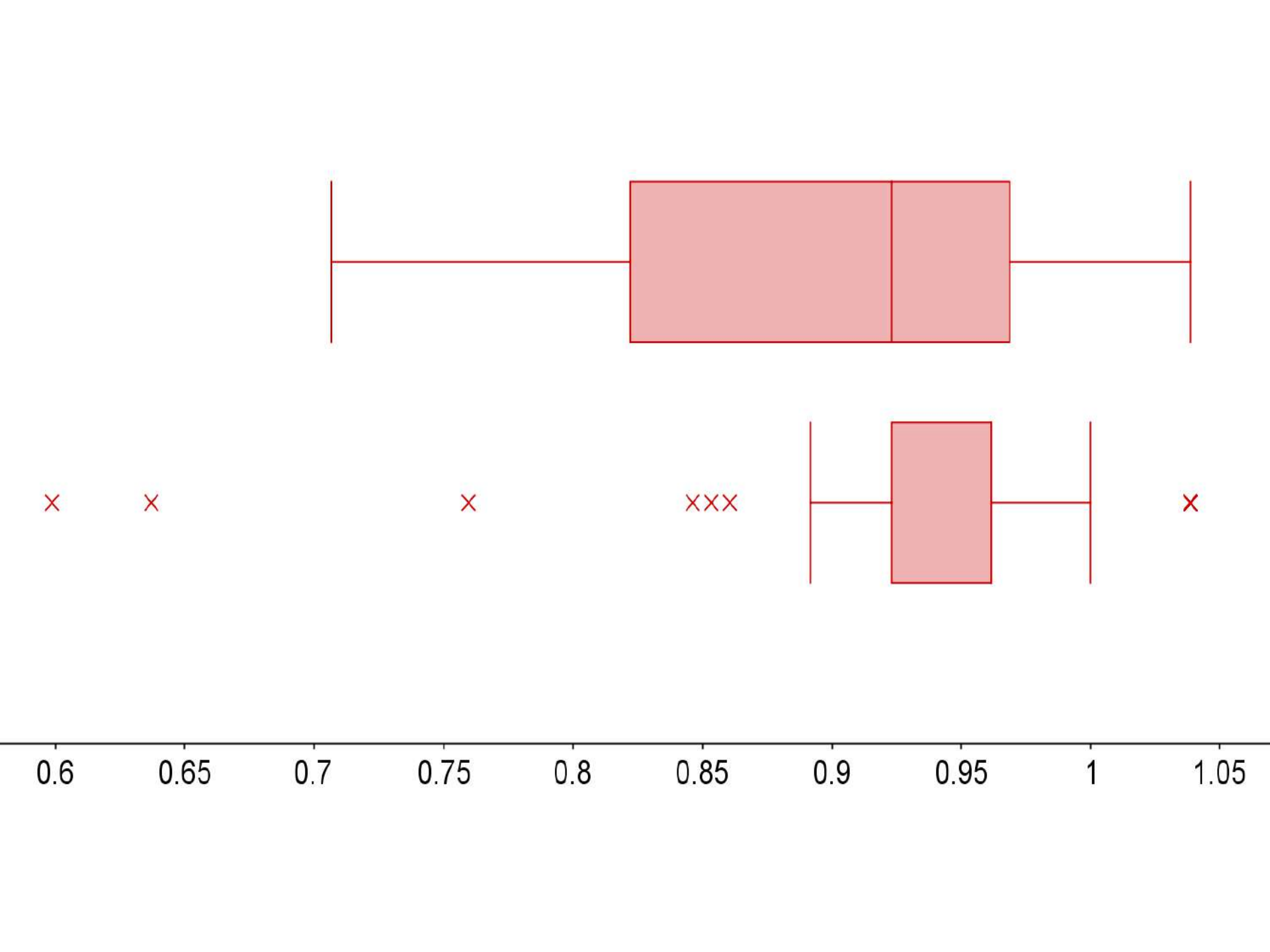
½ piece of paper.

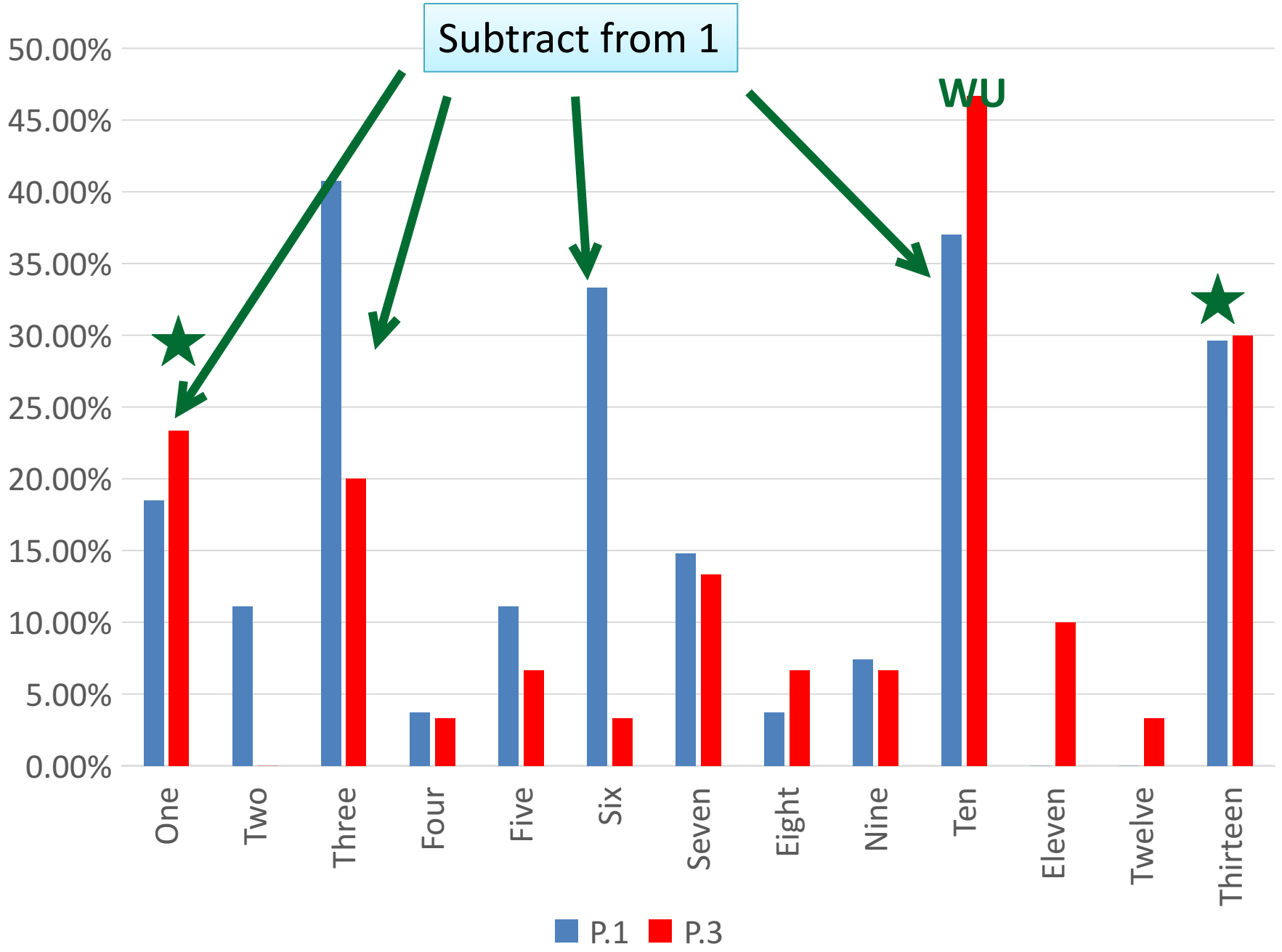
1. Copy the table below
2. Fill it in using the letters of the scatterplots, ranking them by their direction and strength.

	Negative				**None**	**Positive**			
	Strongest			Weakest	None	Weakest			Strongest
Letter of scatter plot									

3. Using your separate piece of paper, comment briefly (one sentence) on each scatterplot. I have done Scatterplot A for you below as an example.
   - *EXAMPLE. "A" is vaguely negative, not very linear, and quite weak.*

Unit 2 Test


0.6
0.65
0.7
0.75
0.8
0.85
0.9
0.95
1
1.05


Subtract from 1
WU
50.00%
45.00%
40.00%
35.00%
30.00%
25.00%
20.00%
15.00%
10.00%
5.00%
0.00%
One
Two
Three
Four
Five
Six
Seven
Eight
Nine
Ten
Eleven
Twelve
Thirteen
P.1
P.3

Unit 2 test: Common Mistakes, 1 of 2

1. Items produced by a manufacturing process are supposed to weigh 90 grams. However, there is variability in the manufacturing process, and they do not all weigh exactly 90 grams. The distribution of weights can be approximated by a normal distribution with a mean of 90 grams and a standard deviation of 1 gram. What percentage of the items will either weigh less than 87 grams or more than 93 grams?
   A. 1%
   B. 3%
   C. 6%
   D. 94%
   E. 99%

13. A stemplot of a set of data is roughly symmetric, but the data do not appear to be linear on a normal quantile plot. We conclude that the data are
   F. Normal, but not standard normal
   G. Standard normal
   H. Not normal
   I. Normal
   J. Skewed in both directions.

Unit 2 test: Common Mistakes, 2 of 2

The best male long jumpers for State College since 1973 have jumped an average of 263.0 inches with a standard deviation of 14.0 inches. The best female long jumpers have averaged 201.2 inches with a standard deviation of 7.7 inches. This year Joey jumped 275 inches and his sister, Carla, jumped 207 inches.

14a. Which athlete had the more impressive performance? Explain briefly.

JOEY

$$\frac{275 - 263}{14} = 0.857$$

CARLA

$$\frac{207 - 201.2}{7.7} = 0.753$$

Joey did better than 85.7% of males, while Carla did better than 75.3% of females. Joey had a more impressive performance.

Project #2 → Predicting the Uncertain

- *Due Friday 2/7. Individual, no groups. 32 points.*
- Choose 2 quantitative variables that might be correlated.
- Collect a sample of 25+ pairs of data.
- Create numerical summaries for each of your two sets of data.
- Create four separate graphical displays.
  - An appropriate graphical display for your explanatory variable
  - An appropriate graphical display for your response variable
  - A scatterplot for both variables simultaneously
  - A residual plot. *For the residual plot, show your calculations for all twenty-five residual values.*
- If you collected your own data, describe your data collection process, and how you tried to make the collection of data fair and unbiased.
- Write an analysis of your data. Comment on your scatterplot. Identify/ interpret r, r^2, and the least-squares equation. Interpret your residual plot. Comment broadly on your data.

Tests + Project #2

- Test explanations
  - Due next Friday 2/7. Separate paper.
  - Why is the right answer right?
  - Why is your wrong answer wrong?

- Look at Project #2 handout

Homework (AP)
Pg.158-159 #1-5
DATA COLLECTION, use link

tinyurl.com/602housesurvey

Exit Pass

Homework (AP)
Pg.158-159 #1-5
DATA COLLECTION, use link

tinyurl.com/602housesurvey

Let's say you're analyzing IQ scores and GPA's.

1. Which is more likely to be the explanatory variable, and why?

2. What do you think the direction of the relationship will be? Why?