LET’S ARGUE: MAKING PREDICTIONS

Examples of questions to prompt student thinking about creating claims.

PAMELA RAWSON
Baxter Academy for Technology & Science
Portland, Maine
pamela.rawson@gmail.com
@rawsonmath
rawsonmath.com
**Contents**

Ages of Best Actor & Actress Oscar Winners ................................................. 2
Predictions about Movie Data ........................................................................ 3
Class Data - Making Predictions .................................................................. 4
Data about cars: Making Predictions .............................................................. 5
Final CER Paper - Predictions About the Data ............................................ 6
  Data set 1 - Unemployment by Region ...................................................... 6
  Data set 3 – Juvenile detentions ................................................................. 7
  Data set 5 – Unemployment and education .............................................. 8
Ages of Best Actor & Actress Oscar Winners
March 2017

The Oscars have been handing out Best Actor and Best Actress awards since 1929. Sometimes there are ties, though this was more common in the early days of the awards than today. There are 90 winners for each award since 1929 - that's 90 Best Actress awards and 90 Best Actor awards. I want you to focus on the ages of the award winners.

1. What would you expect the distribution of ages of the Best Actress winners to look like?
   a. When thinking about this, consider what the mean age and the range of ages might be.
   b. Would you expect the distribution of ages to be symmetric or skewed? If it’s skewed, in which direction?
   c. Do you think that there might be outliers? If so, would they be on the upper end or the lower end?

2. What would you expect the distribution of ages of the Best Actor winners to look like?
   a. When thinking about this, consider what the mean age and the range of ages might be.
   b. Would you expect the distribution of ages to be symmetric or skewed? If it’s skewed, in which direction?
   c. Do you think that there might be outliers? If so, would they be on the upper end or the lower end?

3. How do you think the distributions of the two groups will compare?
   a. Will the mean of Best Actress ages be less than, greater than, or about the same as the mean of Best Actor ages?
   b. Will the range of Best Actress ages be less than, greater than, or about the same as the range of Best Actor ages?
Predictions about Movie Data
April 2017

I have a bunch of data about movies. All of the movies are part of a movie series. They contain the original movie in a series and the most recent sequel in the series. (At least, the most recent sequel as of December 2016. For example, Jurassic Park (original) and Jurassic World (most recent sequel) are two movies in the data set. There are 40 movies: 20 originals paired with 20 sequels.

The data (variables) for each movie includes:
- the title
- the year it was released
- if it is the original or sequel
- the Rotten Tomatoes Critics Score
- the Rotten Tomatoes Audience Score
- the Opening Weekend box office income (adjusted to 2016 dollars)
- the total Domestic Earnings (adjusted to 2016 dollars)
- the total Worldwide Earnings (adjusted to 2016 dollars)

Take a few minutes to think individually about the data. What predictions do you have? What relationships would you expect to see? Write your ideas below.

Look back over your predictions. Try to write a claim statement from your predictions.

Which variables do you think will be most helpful in proving your claim?
Class Data - Making Predictions

April 2017

Over the last year or so, we’ve gathered lots of data about the students taking this class. We collected the following measures (in cm) about many of the students:

- height, arm span, kneeling height, hand span, forearm length, and wrist circumference

We have consistent data for 54 students.

1. Which pair(s) of variables do you think might show the strongest correlation? (And, what would a strong correlation look like in a scatterplot?)

2. Which pair(s) of variables do you think might show the weakest correlations? (And, what would a weak correlation look like in a scatterplot?)

3. Which variable (from the list above) do you think would be the best predictor of a person’s height (in cm)?

4. Write one claim statement about the class data variables.
Data about cars: Making Predictions

May 2017

I have some data about cars, including

- highway mpg (quantitative)
- curb weight (quantitative)
- engine size (quantitative)
- fuel type (categorical: gas, hybrid, electric)

All of the data is based on 2017 models from various mainstream car companies, like Ford, Chevy, Toyota, etc.

Data for engine size is only available for gas and hybrid models.

Think about how these variables might be related and make some predictions.

1. How might the highway mpg and curb weight be related?

2. How might the highway mpg and engine size be related?

3. How might curb weight and fuel type be related?

4. How might the engine size and fuel type be related?

5. How might highway mpg and fuel type be related?

6. Do you think that curb weight or engine size will have a greater impact on mpg, or will they be about the same?
Data set 1 - Unemployment by Region: Unemployment percentages for selected months between October 2014 and October 2015. Data available: Maine, New England, the Northeast, the South, the Midwest, the West, and the entire US. Source: US Census Bureau.

What is your data about? (I know what it says above, but what is it about?)

What are the variables that you will be analyzing?

Before looking at your data, make some predictions about what you expect to see in the data. Some questions to consider (although you can come up with your own questions, too):

- Which variables will have the strongest relationship?
- Which variables will have the weakest relationship?
- Which region do you think will have the lowest unemployment rate?
- Which region do you think will have the highest unemployment rate?

Look at the predictions that you made. Is there an overall theme? Try to write an overall claim about the data.
Data set 3 – Juvenile detentions: Juvenile detentions (per 100,000 people in the population) in the US for selected years between 1997 and 2013. Data available: Rates of juveniles in residential placement by sex (male, female) and by race (white, black, American Indian, Asian), and offense (violent crime - homicide, violent sexual assault, robbery, aggravated assault; property crime - burglary, theft, auto theft, arson; drug. Source: Department of Justice, Office of Juvenile Justice Programs.

<table>
<thead>
<tr>
<th>What is your data about? (I know what it says above, but what is it about?)</th>
<th>What are the variables that you will be analyzing?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Before looking at your data, make some predictions about what you expect to see in the data. Some questions to consider (although you can come up with your own questions, too):

- Which variables will have the strongest relationship?
- Which variables will have the weakest relationship?
- Will males or females have higher numbers of juvenile detentions?
- Which race/ethnicity will have higher numbers of juvenile detentions for violent crime?
- Which race/ethnicity will have higher numbers of juvenile detentions for property crime?
- Which race/ethnicity will have higher numbers of juvenile detentions for drugs?

Look at the predictions that you made. Is there an overall theme? Try to write an overall claim about the data.
Final CER Paper
March 2017

Data set 5 – Unemployment and education: Unemployment rates for 2014. Data available: rates for each state; rates by level of education (less than high school graduate, high school graduate or equivalent, some college or Associate’s Degree, Bachelor’s Degree or higher). Source: US Census Bureau.

<table>
<thead>
<tr>
<th>What is your data about? (I know what it says above, but what is it about?)</th>
<th>What are the variables that you will be analyzing?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Before looking at your data, make some predictions about what you expect to see in the data. Some questions to consider (although you can come up with your own questions, too):

- Which variables will have the strongest relationship?
- Which variables will have the weakest relationship?
- Which level of education will have the highest unemployment rate?
- Which level of education will have the lowest unemployment rate?

Look at the predictions that you made. Is there an overall theme? Try to write an overall claim about the data.