

REVIEW

The goal of this unit was looking for relationships of two data sets (bivariate data). We can give surveys to collect data and interpret the results using frequency tables. We can create scatterplots from two data sets to look for both linear and non-linear relationships. We can create models to represent data and evaluate the strength of said model using the correlation coefficient (r -value) and check residual plots to ensure a linear model is appropriate.

Use the survey results to make a two table. Include marginal frequency.

1. Bob asked 146 7th and 8th graders “Who is your favorite superhero?”

- Of the 60 7th graders surveyed, 20 chose Hulk
- 12 7th and 22 8th graders chose Wonder Women
- 42 total students chose Hulk
- 12 8th graders chose Thor
- 10 7th graders chose Batman

		Favorite Superhero				
		Hulk	Wonder Women	Thor	Batman	Total
Grade	7 th Grade					
	8 th Grade					
	Total					

Convert the two way frequency tables from above into relative frequency tables. Round to nearest hundredth.

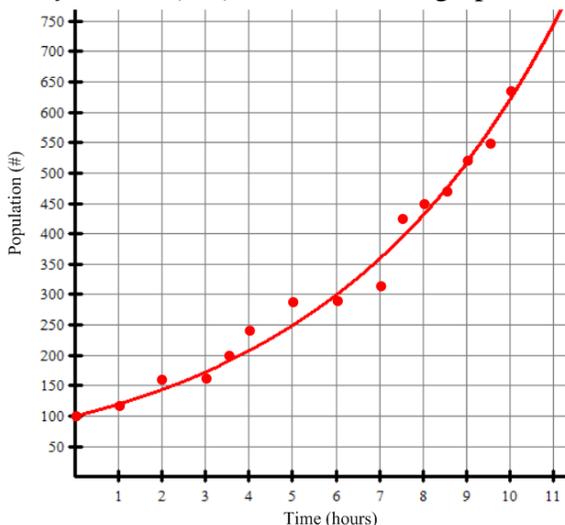
2.

		Favorite Superhero				
		Hulk	Wonder Women	Thor	Batman	Total
Grade	7 th Grade					
	8 th Grade					
	Total					

- What percent prefer Batman?
- What percent of those surveyed are 7th graders that chose Thor?
- Who is the least favorite superhero for 8th graders?

Use the scatterplot and equation for the best fit line/curve to answer the following.

3. The scatterplot shows the population of a certain bacteria over time. The equation of the best fit curve is $y = 100(1.2)^x$ and is shown graphed below.



- Use the equation of best fit to predict the population of the bacteria at 6.5 hours. (round to hundredths)
- The actual bacteria population at 5 hours was 290. How far off is the model’s prediction at this time? (round to hundredths)
- Explain why predicting the bacteria population at 11 hours is extrapolation of the data.

Construct a scatterplot on a graphing calculator of the data below. State the window that you used to view the graph. Make a rough sketch and answer the questions.

4. The price of a ticket given how far it is from the football game.

Rows Back (#)	Ticket Price (\$)
2	195
8	175
10	170
15	145
20	125
25	100
30	65
32	77
40	51
48	22

```

WINDOW
Xmin=
Xmax=
Xscl=
Ymin=
Ymax=
Yscl=
↓Xres=1
    
```

SKETCH:



a. Find and graph a linear regression equation that models the data. (Round to nearest hundredth)

EQUATION: _____

CORRELATION COEFFICIENT: _____

b. Use a sentence to explain what the slope means.

c. Use a sentence to explain what the y-intercept means.

d. Explain the meaning of the *r*-value.

e. Use the linear regression model to predict the cost of a seat 18 rows back.

f. Use the linear regression model to predict the row of a ticket that costs \$42.

g. Find the residual for the price of ticket in 20 rows back.

h. Analyze the residual plot to determine if a linear model is appropriate for this data set. Explain why or why not.

