Equal Salaries for Equal Work?

Mathematical Goals

- Represent data on a scatter plot
- Describe how two variables are related
- Informally assess the fit of a function by plotting and analyzing residuals
- Fit a linear function for a scatter plot that suggests a linear association

Common Core GPS

MCC9-12.S.ID.6 Represent data on two quantitative variables on a scatter plot, and describe how the variables are related.

MCC9-12.S.ID.6a Fit a function to the data; use functions fitted to data to solve problems in the context of the data. Use given functions or choose a function suggested by the context. Emphasize linear and exponential models.

MCC9-12.S.ID.6b Informally assess the fit of a function by plotting and analyzing residuals.

MCC9-12.S.ID.6c Fit a linear function for a scatter plot that suggests a linear association.

Interpret linear models

MCC9-12.S.ID.7 Interpret the slope (rate of change) and the intercept (constant term) of a linear model in the context of the data.

MCC9-12.S.ID.8 Compute (using technology) and interpret the correlation coefficient of a linear fit.

Common Core State Standards for Mathematical Practice

1. Make sense of problems and persevere in solving them.
2. Reason abstractly and quantitatively.
3. Construct viable arguments and critique the reasoning of others.
4. Model with mathematics.
5. Use appropriate tools strategically.
6. Attend to precision.
7. Look for and make use of structure.
8. Look for and express regularity in repeated reasoning.

Introduction

This task asks students to compare additive and multiplicative growth (represented by linear and exponential models) to make predictions and solve problems within the context of gender-based salary differences. In doing this task, students analyze data sets, create scatter plots, determine the most appropriate mathematical model, and justify their model selection.
This task provides a good example of how data points can appear to be linear over a relatively small domain, but how a different type of mathematical model might be more appropriate over a larger domain. This is an opportunity for students to discuss strengths and limitations of using mathematical functions to model real data. One discussion might arise as to whether other types of mathematical functions might sometimes be used for different types of data, perhaps leading students to look for patterns in data they might gather from sources like newspapers or books of world records.

Note that students will need to make a decision about the initial value representing the year. For example, it would be reasonable to assign the year 1984 (the first year in the table) as Year 0. The sample solutions below are based on this assumption.

**Prerequisites**
Students must have knowledge of using the graphing calculator to create linear and exponential models and to analyze residuals. It is important that students understand how to assess the fit of a function to data and choose a function suggested by context.

**Learning Targets**
When making statistical models, technology is valuable for varying assumptions, exploring consequences and comparing predictions with data. Students will interpret the correlation coefficient and show understanding of strengths and limitations of using mathematical functions to model real data.

**Time Required**
1 class period

**Materials**
Pencil and (graphing) paper; graphing calculator or statistical software package.
The data table shows the annual median earnings for female and male workers in the United States from 1984 to 2004. Use the data table to complete the task. Answer all questions in depth to show your understanding of the standards.

<table>
<thead>
<tr>
<th>Year</th>
<th>Women’s median earnings (in dollars)</th>
<th>Men’s median earnings (in dollars)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1984</td>
<td>8675</td>
<td>17026</td>
</tr>
<tr>
<td>1985</td>
<td>9328</td>
<td>17779</td>
</tr>
<tr>
<td>1986</td>
<td>10016</td>
<td>18782</td>
</tr>
<tr>
<td>1987</td>
<td>10619</td>
<td>19818</td>
</tr>
<tr>
<td>1988</td>
<td>11096</td>
<td>20612</td>
</tr>
<tr>
<td>1989</td>
<td>11736</td>
<td>21376</td>
</tr>
<tr>
<td>1990</td>
<td>12250</td>
<td>21522</td>
</tr>
<tr>
<td>1991</td>
<td>12884</td>
<td>21857</td>
</tr>
<tr>
<td>1992</td>
<td>13527</td>
<td>21903</td>
</tr>
<tr>
<td>1993</td>
<td>13896</td>
<td>22443</td>
</tr>
<tr>
<td>1994</td>
<td>14323</td>
<td>23656</td>
</tr>
<tr>
<td>1995</td>
<td>15322</td>
<td>25018</td>
</tr>
<tr>
<td>1996</td>
<td>16028</td>
<td>25785</td>
</tr>
<tr>
<td>1997</td>
<td>16716</td>
<td>26843</td>
</tr>
<tr>
<td>1998</td>
<td>17716</td>
<td>28755</td>
</tr>
<tr>
<td>1999</td>
<td>18440</td>
<td>30079</td>
</tr>
<tr>
<td>2000</td>
<td>20267</td>
<td>30951</td>
</tr>
<tr>
<td>2001</td>
<td>20851</td>
<td>31364</td>
</tr>
<tr>
<td>2002</td>
<td>21429</td>
<td>31647</td>
</tr>
<tr>
<td>2003</td>
<td>22004</td>
<td>32048</td>
</tr>
<tr>
<td>2004</td>
<td>22256</td>
<td>32483</td>
</tr>
</tbody>
</table>

Data provided by U.S. Census Bureau
1. Create two scatter plots, one for women’s median earnings over time and one for men’s median earnings over time. Describe two things you notice about the scatter plots.

Each scatter plot below is graphed with the following window:

**Window:**
- **Xmin**: (-)2
- **Xmax**: 22
- **Xscl**: 5
- **Ymin**: 4500
- **Ymax**: 37000
- **Yscl**: 2500
- **Xres**: 1

![Women's Data](image1)

![Men's Data](image2)

![Both Data Sets](image3)

*Answers may vary.*

*Possible answers: From 1984 to 2004, median earnings for both men and women increased. In each of these years, men’s median earnings were greater than women’s median earnings.*

2. Terry and Tomas are trying to decide what type of model will most accurately represent the data. Terry thinks that a linear model might be most appropriate for each scatter plot. Help Terry find reasonable linear function rules for each scatter plot. Explain how you found these.

*Answers may vary*

*One solution method:*

*To find a linear model of women’s median earnings, use the starting earnings figure for women, $8675, and the average rate of change of $680 per year. (To find the average rate of change, find successive differences and then find the average of the successive differences.) The linear model is \( m(x) = 680x + 8675 \), where \( x \) represents years and \( m(x) \) gives the median earnings. To find a linear model of men’s median earnings, use the starting earnings figure for men, $17,026, and the average rate of change of $773 per year. The linear model is \( m(x) = 773x + 17026 \), where \( x \) represents years and \( m(x) \) gives the median earnings.*
Another solution method:
Using a graphing calculator to determine a regression line, women’s median earnings
could be represented by the function \( y = 703x + 8181 \).

Using a graphing calculator to determine a regression line, men’s median earnings could
be represented by the function \( y = 814x + 16709 \).

3. Using the linear models, will women’s annual median earnings ever equal those of men?
Why or why not?

Using the linear models created from the data provided, women’s annual median
earnings will never equal men’s annual median earnings. The men’s linear model has a
larger y-intercept and a larger slope, meaning the men start out earning more money
and also experience a faster rate of increase in earnings.

4. Tomas thinks that an exponential model might be most appropriate for each scatter plot.
Help Tomas find reasonable exponential function rules for each scatter plot. Explain how
you found these.

Answers may vary.

One solution method:
To find an exponential model of women’s median earnings, use the starting income for
women, $8675, and the average quotient, 1.048. (To find the average quotient, find
successive quotients then find the average of the successive quotients.) The exponential
model is \( m(x) = 8675(1.048)^x \), where \( x \) represents years and \( m(x) \) gives the median salary. To
find an exponential model of men’s median earnings, use the starting earnings figure for
men, $17,026, and the average quotient, 1.033. The exponential model is \( m(x) = 17026(1.033)^x \), where \( x \) represents years and \( m(x) \) gives the median earnings.

Another solution method:
Calculating an exponential regression function on a graphing calculator, women’s median
earnings could be represented by the function \( y = 9087(1.049)^x \).

Calculating an exponential regression function on a graphing calculator, men’s median
earnings could be represented by the function \( y = 17479(1.034)^x \).

5. Using the exponential models, will women’s annual median earnings ever equal those of
men? Why or why not?

Using the exponential models, women’s annual median earnings will eventually equal
those of men. The exponential model of men’s earnings has a base of 1.034, and the
exponential model of women’s earnings has a base of 1.049. Since the women’s model has
a higher base, their earnings are increasing at a faster rate and will eventually surpass
men’s earnings. These functions can also be graphed to determine their intersection (45.7, 79533.88), demonstrating that at some point during the year 2029, women’s annual median earnings will overtake men’s annual median earnings.

6. If you answered yes to either question 3 or question 5, use that model to determine the first year women will have higher median earnings than men. Explain how you found your answer.

Using the exponential models, women’s annual median earnings will eventually equal those of men. The exponential model of men’s earnings has a base of 1.034, and the exponential model of women’s earnings has a base of 1.049. Since the women’s model has a higher base, their earnings are increasing at a faster rate and will eventually surpass men’s earnings. These functions can also be graphed to determine their intersection (45.7, 79533.88), demonstrating that at some point during the year 2029, women’s annual median earnings will overtake men’s annual median earnings.

7. For each year listed in the table, find the ratio of women’s to men’s annual median earnings expressed as a percentage. Use the data to create a scatter plot of percentage versus year. Based on this graph, do you think women’s annual median earnings will ever equal those of men? Why or why not?

The scatter plot has a positive correlation. This means that women’s annual earnings are approaching those of men and (if the trend continues) will eventually catch up to men’s annual median earnings.

8. Considering the results of the scatter plot in question 7 above, do you think the linear model or exponential model makes more sense? Why?

Answers will vary. Generally speaking, the exponential model makes more sense because the gap between men’s earnings and women’s earnings is decreasing, as shown in the percentage-versus-time scatter plot. This more closely represents the real situation. The linear model shows the gap widening — an inaccurate representation of what is actually happening.

Data on earnings by gender provided by:
Performance Task: Equal Salaries for Equal Work!

Common Core State Standards for Mathematical Practice
1. Make sense of problems and persevere in solving them.
2. Reason abstractly and quantitatively.
3. Construct viable arguments and critique the reasoning of others.
4. Model with mathematics.
5. Use appropriate tools strategically.
6. Attend to precision.
7. Look for and make use of structure.
8. Look for and express regularity in repeated reasoning.

Common Core GPS
MCC9-12.S.ID.6 Represent data on two quantitative variables on a scatter plot, and describe how the variables are related.

MCC9-12.S.ID.6a Fit a function to the data; use functions fitted to data to solve problems in the context of the data. Use given functions or choose a function suggested by the context. Emphasize linear and exponential models.

MCC9-12.S.ID.6b Informally assess the fit of a function by plotting and analyzing residuals.

MCC9-12.S.ID.6c Fit a linear function for a scatter plot that suggests a linear association.

Interpret linear models

MCC9-12.S.ID.7 Interpret the slope (rate of change) and the intercept (constant term) of a linear model in the context of the data.

MCC9-12.S.ID.8 Compute (using technology) and interpret the correlation coefficient of a linear fit.

MCC9-12.S.ID.9 Distinguish between correlation and causation.