Statistical Reasoning Name \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

WS 4.5: Residuals Date \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Day \_\_\_\_\_\_\_\_\_

With the growth of different internet service providers, a researcher decided to examine whether there is a correlation between the cost of internet service per month (rounded to nearest dollar) and the degree of customer satisfaction (on a scale of 1-5 with 5 being the highest). The researcher only included programs with comparable types of services. A sample of the data is provided below.



20

103.5

3.6

3.5

|  |  |  |
| --- | --- | --- |
| **Provider** | **Cost per month** | **Satisfaction Rating** |
| AT&T | $15 | 3.92 |
| CenturyLink | $23 | 3.75 |
| CableOne | $56 | 3.59 |
| Cox | $10 | 3.76 |
| Frontier | $20 | 3.70 |
| MediaCom | $25 | 3.70 |
| Optimum | $45 | 3.89 |
| RCN | $10 | 4.04 |
| Spectrum | $30 | 3.84 |
| SuddenLink | $35 | 3.90 |
| Windstream | $40 | 3.75 |
| Xfinity | $20 | 3.87 |

1. Construct a scatterplot
2. Describe the overall pattern (direction, form, strength)
3. Find the regression equation, r, and r2
4. What does r and r2 tell us about the data.
5. Find the residual value for each cost amount. Does the equation overpredict or underestimate our data.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Provider** | **Cost per month** | **Observed****Satisfaction Rating** | **Predicted Satisfaction****Rating (use equation)** | **Residual = Observed - Predicted** |
| AT&T | $15 | 3.92 |  |  |
| CenturyLink | $23 | 3.75 |  |  |
| CableOne | $56 | 3.59 |  |  |
| Cox | $10 | 3.76 |  |  |
| Frontier | $20 | 3.70 |  |  |
| MediaCom | $25 | 3.70 |  |  |
| Optimum | $45 | 3.89 |  |  |
| RCN | $10 | 4.04 |  |  |
| Spectrum | $30 | 3.84 |  |  |
| SuddenLink | $35 | 3.90 |  |  |
| Windstream | $40 | 3.75 |  |  |
| Xfinity | $20 | 3.87 |  |  |
|  |  |  |  | TOTAL: |

The table below shows the weight (in grams) of a bar of soap used to shower. Before showering, the bar of soap is weighed each day. The weight goes down as the soap is used. Notice that some days the soap was not weighed.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Day | Weight | Day | Weight | Day | Weight |
| 1 | 124 | 8 | 84 | 16 | 27 |
| 2 | 121 | 9 | 78 | 18 | 16 |
| 5 | 103 | 10 | 71 | 19 | 12 |
| 6 | 96 | 12 | 58 | 20 | 8 |
| 7 | 90 | 13 | 50 | 21 | 6 |

1. Find the regression equation, r, and r2
2. What does the slope mean?
3. Predict what the weight of the soap was on Day 4?
4. What is the weight of the soap on Day 25?
5. Find the residual value for each day. Does the equation overpredict or underestimate our data.

|  |  |  |  |
| --- | --- | --- | --- |
| **Day** | **Observed****Weight** | **Predicted Weight (use equation)** | **Residual = Observed - Predicted** |
| 1 | 124 |  |  |
| 2 | 121 |  |  |
| 5 | 103 |  |  |
| 6 | 96 |  |  |
| 7 | 90 |  |  |
| 8 | 84 |  |  |
| 9 | 78 |  |  |
| 10 | 71 |  |  |
| 12 | 58 |  |  |
| 13 | 50 |  |  |
| 16 | 27 |  |  |
| 18 | 16 |  |  |
| 19 | 12 |  |  |
| 20 | 8 |  |  |
| 21 | 6 |  |  |
|  |  |  | TOTAL: |