**Model Comparison - Crime Data**

Open the “Crime Data CA” in the StatCrunch group.



We wish to analyze the relationship between the number of Murders in CA and the number of Motor Vehicle Theft. The analyst wants to show that the more thefts there are in a town, the more murders there are.

1. What should be your explanatory and response variables? Why did you choose this way?
2. Use StatCrunch to make a scatterplot of the ordered pairs, post your graph below.
3. Which models do you think would be a good fit and why?

|  |  |  |
| --- | --- | --- |
| **Model** | **Appears to be a good fit? Yes/No?** | **Why?** |
| Linear |  |  |
| Quadratic |  |  |
| Exponential |  |  |
| Logarithmic |  |  |

Find the following graph, put the graphs in the following with the regression line: (Write next to which model you think is the best and why.)

|  |  |
| --- | --- |
| **Model** | **Graph – with regression Line** |
| Linear |  |
| Quadratic |  |
| Exponential |  |
| Logarithmic |  |

Now find the residual plots for each: (Describe each residual plot and identify which model is the best based on the residual plots.)

|  |  |
| --- | --- |
| **Model & Residual Plot Description** | **Residual Plots** |
| Linear |  |
| Quadratic |  |
| Exponential |  |
| Logarithmic |  |

Then find the histograms of the residuals: (Describe each histogram and identify which model is the best based on the histograms.)

|  |  |
| --- | --- |
| **Model & Histogram Description** | **Graph of the histogram of the residuals** |
| Linear |  |
| Quadratic |  |
| Exponential |  |
| Logarithmic |  |

Fill in the following table

|  |  |  |
| --- | --- | --- |
| **Model** | **R-Squared** | **Se (Including units.)** |
| Linear |  |  |
| Quadratic |  |  |
| Exponential |  |  |
| Logarithmic |  |  |

1. Which do you think is the best model? Explain why using the information obtained above.