Critique Analysis 1a

**Overall Score: 10**

Rubric

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Criteria | Ratings | | | |
| Questions & Hypothesis | 3pts  Excellent | 2pts  Fair Quality | 1pts  Poor Quality | 0 pts  Missing Elements |
| Analysis | 3pts  Excellent | 2pts  Fair Quality | 1pts  Poor Quality | 0 pts  Missing Elements |
| Graphical Summary | 3pts  Excellent | 2pts  Fair Quality | 1pts  Poor Quality | 0 pts  Missing Elements |
| Interpretation | 3pts  Excellent | 2pts  Fair Quality | 1pts  Poor Quality | 0 pts  Missing Elements |
| Presentation | 3pts  Excellent | 2pts  Fair Quality | 1pts  Poor Quality | 0 pts  Missing Elements |

## Question and Hypothesis:

Does the analysis contain a clearly stated question that is properly motivated? If statistical hypothesis tests are being performed in the analysis (i.e., p-values are used) then are the appropriate null and alternative hypotheses stated in mathematical form with a corresponding significance level?

The question for the analysis is not clearly stated, instead it is inferred by the test and the object names. This could be remedied by stating, “The question for this analysis is this, ‘can the high air temperature on January 12th be used to explain the high air temperature on January 14th in Rexburg, Idaho?’” The hypotheses are not stated, mathematically in the document.

## Analysis:

Does the analysis properly perform, demonstrate, and describe all numerical and graphical statistical analysis being performed? This is the hardest part of the critique as you must check the "correctness" of the work performed.

According to my understanding, the statistical analysis was performed correctly. A linear model was created on the formula “MondayHigh ~ Saturday,” which predicts the high air temperature on January 14th, 2019 will be 26.3 F. I think the names of the columns, “MondayHigh” and “Saturday” are a bit misleading, because January 14th doesn’t always fall on a Monday.

## Graphical Summary:

Does the graphic meet the stand-alone principal? In other words, if it were removed from the analysis and viewed all by itself, does it properly communicate all the necessary information and conclusions? Often, better labeling of the graphic is needed to achieve this principal.

The regression line is not labeled on the graphic, and the predicted temperature on January 14, 2019 is not on the graph. The tic mark labels on the x-axis of the first scatterplot found at the top of the document are impossible to read. I suggest fixing this by turning the text on an angle, using the following code in r: plot + opts(axis.text.x = theme\_text(angle = 45))

## Interpretation:

Does the analysis bring the reader to a definitive conclusion that is well supported by either outside sources or proper statistical analysis or both? Is the original question well answered?

I think the analysis is lacking a little in the interpretation. I think there should have been more explanation of what the coefficient values mean in context of the question. The alpha and the prediction are stated, but there is not much more than that.

## Presentation:

Was the analysis well presented? In other words, easy (even enjoyable) to read and understand? A few spelling errors isn't a big deal, but any errors that lead to confusion or difficulty or extra time on your part in trying to understand the analysis is cause for a lower score. The analysis should read from question to answer very easily.

The first scatterplot doesn’t stand alone, but I understand that it was meant to be used for exploring the data and not explaining it. I still think more time should have been spent on cleaning it up a little since it was included in the document.