**Peer Evaluation, due on Thur, Nov 18:**

Each group will be assigned another group with which to do peer evaluations. By class time on Tues, Nov 16, make your latest draft available to your partner group. We will use another Blackboard group for this. Your group will be called Peer Evaluation Group x\_y, with the number of the two groups, x and y. By Thur, Nov 18 at 11:59pm, each member should evaluate the paper of the other group, and upload the evaluation to the same channel, so members of the other group can see them.

The evaluation form is on pages 2 and 3 of this document. For each criterion, please select a score (highlight the box), and enter your comments in the last column. Scores are suggestions; comments are very important; try to give your peer group some good suggestions. Also, offer some overall comments at the bottom. Save the file as *PeerEval\_yourlastname.docx* and upload to the Peer Evaluation Group File Exchange on Blackboard.

**The Four Parts of the Paper:**

1. **Introduction:** At the beginning of your paper, you must describe the data, in a paragraph. Note the following:

* What is the source of the data? Where and when was it created?
* If it is a sample, from what population was it drawn, and how was the sample selected?
* Do you suspect any sampling bias?
* Was it an experiment or an observational study?
* How were measurements taken, or questions asked?
* Do you suspect any bias in the questions or measurements?
* Why is this data of interest to you, and why should the class find it interesting?
* What kind of data cleaning was necessary (R code for this must show…)

1. **Data Analysis:** Write R code to create some relevant graphs, using techniques that we’ve used in class (ggplot, maybe dplyr). About **4 or 5 graphs** should be plenty, depending on complexity. Include some numerical summaries as well. If possible and appropriate, include **a bootstrap confidence interval.**

For each graph and numerical summary, write a paragraph or two summarizing what you see, and suggesting some implications. For example, describe patterns that you observe in a graph, and suggest why they make sense, given what you know about the subject, or if they are unexpected. Do you think there is a cause-effect relationship between any variables? Explain your reasoning.

1. **Conclusions:** Write some overall conclusions – an overall summary of what you learned from your analysis. Summarize in one paragraph.
2. **Limitations / Recommendations:** Write a paragraph describing some of the limitations that are inherent in your study. Also discuss ideas for future research that might build on the work you did in this project. Summarize in one paragraph.

Rubric for Peer Evaluation of Project: Your Name: Eamon Deffner

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Criteria | 0 | 1 | 2 | 3 | Peer Reviewer Comments: |
| Introduction | Data background is not stated. | Data background is not clearly stated and/or is not accurate. | Data background description is reasonably clear. | Data background description is very clear and thorough. In addition, the student gave a compelling reason why the question is interesting. | The background for the data is good. It mentions the dates, the where the data comes from. It also mentions that there is no bias based on time. There is a good reason why this data set is interesting. There is no data setup and cleaning section though. |
| Data Analysis -- Coding | No R code is included in the presentation. | R code for graphs and data summary has errors, or is not completely present. | R code is present and largely correct. Students made use of dplyr, ggplot, techniques learned in class. | R code is correct, uses techniques learned in class, and is well documented with comments in the Rmd script. | There is no R code in this section. It is hidden for now, but in order to see what is done, it would be nice to include it. |
| Data Analysis –  Graphs | No graphs are included in the presentation. | Some graphs are presented, but there is an insufficient number, or some are incorrect or inappropriate for the particular data. | A sufficient number of graphs are presented, appropriate to the data. | A sufficient number of appropriate graphs are presented. Graphs are interesting, attractive, and easy for the audience to interpret. | Most of the graphs make sense, are effective, and show something important. The leaflet plot is especially good, and does a good job of showing call distributions. I would change the number of bins in the hour of the day graph to be 24, one for each hour. On the shift graphs, I would get rid of D and N in the legend, since they aren’t in the graph. |
| Data Analysis –  Summary Statistics | No summary statistics are included in the presentation. | Some summary stats are presented, but there is an insufficient number, or some are incorrect or inappropriate for the particular data. | A sufficient number of summary stats are presented, appropriate to the data. | A sufficient number of appropriate summary stats are presented. Stats are interesting, useful, and easy for the audience to interpret. | Some of these graphs can be a bit confusing because there are no clear summary statistics. I would include some numbers for all of the graphs, except the leaflet plot and maybe the type of call plot. |
| Data Analysis  (Graphs and Numerical Summaries)-- Written Description | There is no written description of the data analysis. | The written description of the data analysis is incorrect or not relevant to answering the research question | The written description of the data analysis is accurate but not complete | The written description of the data analysis is accurate and completely describes the important features of the distribution | The first two analyses are good and include enough information about the plot. In the third graph, it is good to include the type of engine, but there should be more on which types are called most. In the fourth plot, the analysis in incomplete. The last two plots have good analyses, but it should be mentioned the there is overlap in the years that causes more alarms from August to November. |
| Conclusions | There is no written interpretation of the overall project. | There is a written interpretation, but it is incorrect. | Written interpretation is correct, but not clear and/or not in context | Written interpretation is clear and correct and in the context of the research question. | The conclusions are good, and refer to the graphs well. At the end, I’m not sure what was trying to be said, but I think it could just be that this data only affects Burlington. |
| Limitations and Recommendations | There is no discussion of limitations of the project or ideas for future work | There is some discussion of limitations and ideas for future work, but the ideas are unclear and/or don’t make sense | Limitations of the study and ideas for future work are described and are generally sensible but are lackluster | Student sensibly describes limitations of the study and has strong suggestions for future work. | You could mention some ways in which leaflet did not let you get a complete plot of the data points. I like the part about looking for an explanation of the data. One thing that I would mention is how you would use this data if you got it. |
| Quality of Writing | Work is not submitted. | Write up does not use complete sentences and/or uses poor spelling and grammar | Write up uses complete sentences but has quite a few spelling and/or grammatical errors | Write up uses complete sentences and has almost no spelling and/or grammatical errors | Most of the write-up has good grammar and uses complete sentences. There are a few cases of run on sentences, so I would just proofread it once each to make sure you clear these up. |

Overall Comments: This draft has a lot of potential to be something really cool. One thing that needs to happen is to make sure that the R-code is showing in the final draft. Most of the graphs are very good, and the analysis is good for the most part. There are a few things that could be changed, such as mentioning that the data includes the months from August to November twice. The conclusions and limitations sections are good. There are a few awkward wordings in this section but if they are cleared up, they will be good.