

STUDENT

Jas Sohi

COURSE

Intro to Data Science v. 2

Congratulations on your project! I have included a few comments for you in the rubric below. Feel free to email nd002-project@udacity.com if you have any questions. I wish you all the best with your Nanodegree studies!

Charlotte and the Udacity team

Communication**Meets Specifications**

- Analysis done using methods learned in the course is explained in a way that would be understandable to a student who has completed the class.
Comment: The project is written in a clear conversational style.

Accuracy and Correctness**Meets Specifications**

- No incorrect conclusions are drawn from the data.

Quality of Visualizations**Meets Specifications**

- Plots depict relationships between two or more variables.
- All plots are of the appropriate type.
Comment: I particularly like your choice of median for the second visualisation.
- All plots are appropriately labeled and titled. Plot is given an appropriate title. X-axis and y-axis are appropriately labeled. Visual cues (colors, size, etc) are easy to distinguish.
Comments: Thank you for your improvements to the plots.

I think the reason for the peak in the line plot is that 15th May was mostly rainy, but there were two data points, corresponding to particularly high ridership, where rain was not recorded. The median (=mean) of only two data points is then taken as the 'non-rainy' point for that day, though it is far from representative of the ridership (or the weather) of the day as a whole. One way to overcome this would be to clean the dataset so that every day was classified as either rainy or non-rainy. A bar chart with two colours, showing the mean numbers of entries would then represent the variation over the month, and allow us to see what happened on rainy days in the month.

Quality of Analysis

Meets Specifications

- Conclusions are correctly justified with data.
Comment: Well done on your improvements to the conclusion.
- The choice of statistical test type, features, and linear regression models are appropriate based on the characteristics of the data.
- Statistical tests or linear regression models are described thoroughly and the reasons for choosing them are clearly articulated.
Comment: Thank you for your improvements to the description of the linear model.
- Some shortcomings of the statistical tests or regression techniques used are appropriately acknowledged.
Comment: To improve this section, you could consider problems with the methods used in your analysis.

PROJECT EVALUATION

Project Meets Specifications