Share your accomplishment

Explore and Summarize Data

Student Notes | Code Review | Project Review | Project History

Does Not Meet Specifications

Code Functionality

SPECIFICATION
All code is functional (e.g. No Error is produced and RMD document is not prevented from bring knit.)

MEETS SPECIFICATION

Reviewer Comments
Well done producing a functional RMD file.

SPECIFICATION
The project almost never uses repetitive code where a function would be more appropriate. The code references variables by name instead of using constants or column numbers.

MEETS SPECIFICATION

Reviewer Comments
Well done consistently referencing columns by their variable names.

OPTIONAL: Although this is not covered much in the course, it is possible to apply a function to execute repeated code on ggplot and cor tests. This would make your ggplot sections less repetitive, because the function could be called with various parameters using aes_string instead of aes inside ggplot. For examples of how to do this, see the top solution in the following post:

> http://stackoverflow.com/questions/6488748/passing-parameters-to-ggplot

Requirements to exceed the specification ›

Project Readability

SPECIFICATION

All complex code is adequately explained with comments. It is always clear what the code is doing.

DOES NOT MEET SPECIFICATION

Reviewer Comments

There was not much commenting in your code. Even though certain sections may seem self-explanatory, comments can help any reader follow along. Try to include comments explaining the purpose of more of your R chunks. Here are the Google guidelines related to commenting:

> http://google-styleguide.googlecode.com/svn/trunk/Rguide.xml#comments

Requirements to exceed the specification ›

SPECIFICATION

The code uses formatting techniques to improve code readability. All lines are shorter than 80 characters.

DOES NOT MEET SPECIFICATION

Reviewer Comments

Please ensure lines of R code are less than 80 characters! This guideline is important to ensure your code displays well in all types of consoles. Assuming you are using RStudio, you can select Tools-->Options-->Code Editing to set a line at the 80 character mark. For more information on programming best practices, please refer to Google guidelines for R programming:

> http://google-styleguide.googlecode.com/svn/trunk/Rguide.xml

Requirements to exceed the specification ›

SPECIFICATION

Markdown syntax is used in the code to improve readability of the knitted file.

There are no large sections of the knitted HTML file with bad readability.

DOES NOT MEET SPECIFICATION

Reviewer Comments

Please avoid displaying messages and warnings in your HTML document. The final HTML document should include results from the analysis, figures and discussion. You can easily do it by setting the parameter "message=FALSE" as in this example:

```
{r echo=FALSE, message=FALSE, warning=FALSE, packages}
```

Optional: The HTML document is readable and clear yet you might want to consider using underlined between paragraphs, or between charts that depict different ideas or features. You might find this link useful:

> http://rmarkdown.rstudio.com/authoring_pandoc_markdown.html

Requirements to exceed the specification ›

Quality of Analysis

SPECIFICATION

The project appropriately uses univariate, bivariate, and multivariate plots to explore most of the expected relationships in the data set.

DOES NOT MEET SPECIFICATION

Reviewer Comments

Good job in your Univariate section.

In order to meet specifications here, please expand on your Bivariate analysis by evaluating additional relationships. In your Multivariate section, try to include some plots that utilize color to depict a third variable. For example, in a ggplot statement, create a scatterplot comparing two variables (e.g. density and residual sugar) and then set color equal to a certain variable of interest (such as quality).

Requirements to exceed the specification ›

SPECIFICATION

Questions and findings are placed between blocks of R code regularly so it is clear what the student was thinking throughout the analysis.

DOES NOT MEET SPECIFICATION

Reviewer Comments

You should endeavor to include questions and findings on a regular basis throughout your analysis. It is important to include the reasoning behind your decisions so that the reader can follow along and stay aligned to your thought process and key takeaways.

SPECIFICATION

Reasoning is provided for the plots made throughout the analysis. Plots made follow a logical flow. Comments following plots accurately reflect the plots' contents.

DOES NOT MEET SPECIFICATION

Reviewer Comments

As mentioned in the last criterion, an important aspect of EDA is including reasoning, questions, and findings throughout your analysis. Please reference the second and third items under "Quality of Analysis" in the Project Rubric located here:

> https://docs.google.com/document/d/1L2Wwofs6D8Crd0QLZ1-RxBHlVoBZ3mec2xWgxrmUs5I/pub

Put yourself into the shoes of the reader and consistently answer the question "Why are you taking

that next step in the analysis?"

SPECIFICATION

The project contains at least 20 visualizations. The visualizations are varied and show multiple comparisons and trends. Relevant statistics such as means, medians, quartiles, or confidence intervals are computed throughout the analysis when an inference is made about the data.

DOES NOT MEET SPECIFICATION

Reviewer Comments

It appears you are very close to the minimum number of visualizations but not quite there. I recommend exploring additional relationships in your Bivariate section to increase the count. Also, EDA is strongest when your plots are backed up by statistics. I see that you calculate the correlations so try to reference those more often throughout your analysis.

Requirements to exceed the specification ›

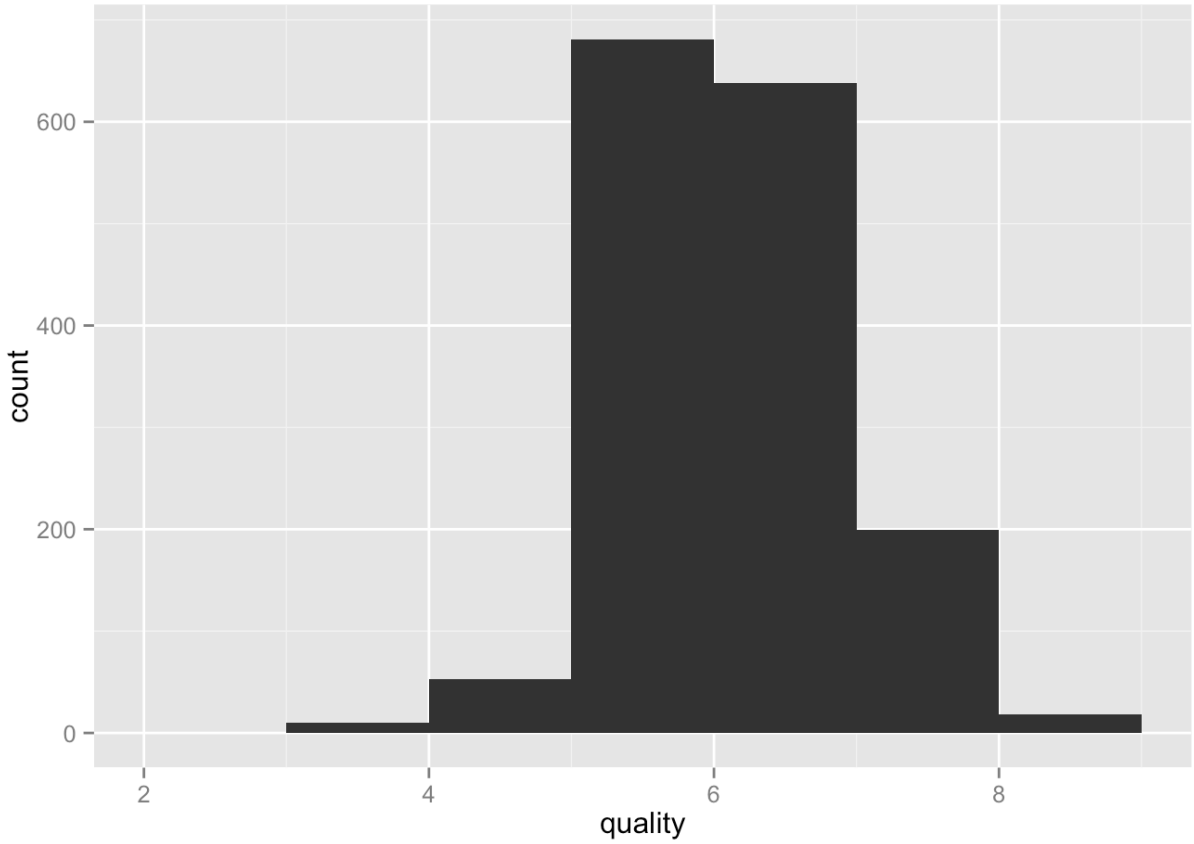
SPECIFICATION

Visualizations made in the project depict the data in an appropriate manner that allows plots to be readily interpreted. Choices of plot type and parameters aid interpretability.

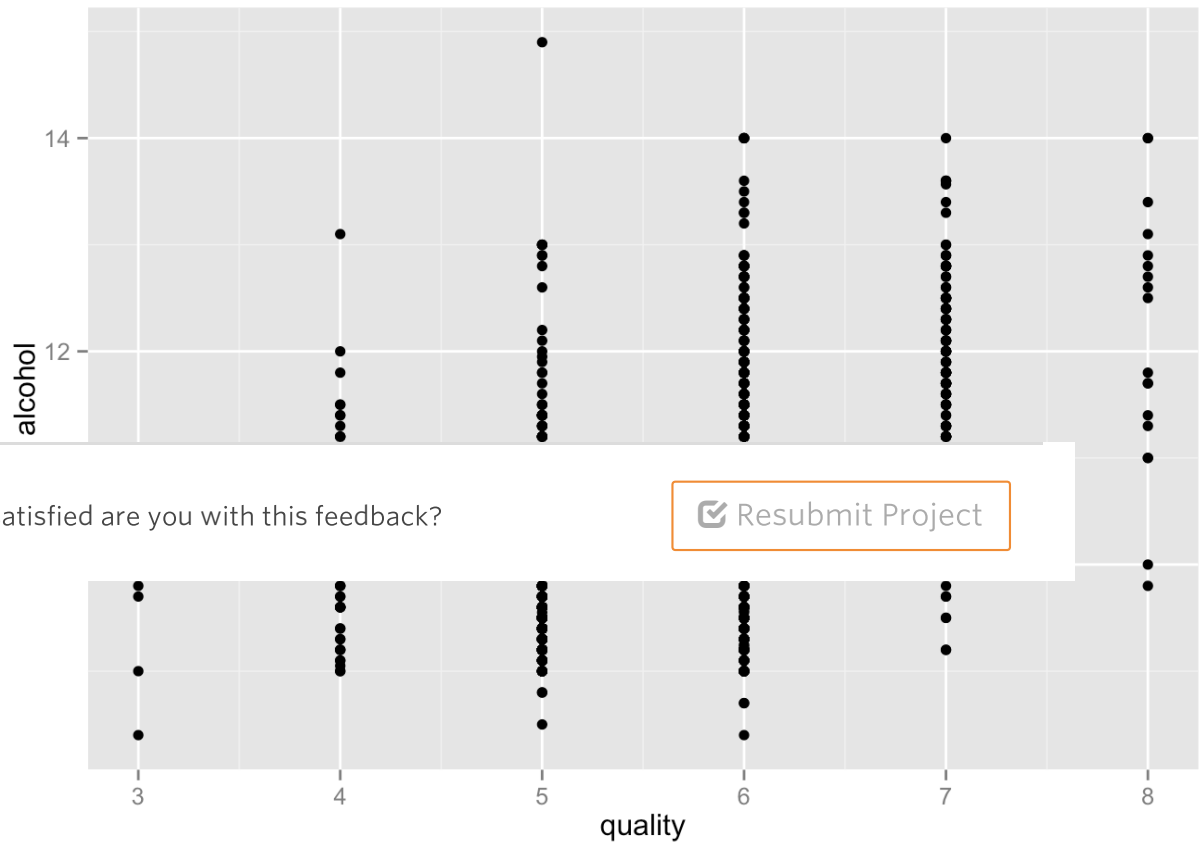
DOES NOT MEET SPECIFICATION

Reviewer Comments

I have a couple suggestions for improving your visualizations. First, in your Univariate histogram for Quality, it is difficult to determine the distribution because the x-axis numbers are not aligned appropriately to the bars. For example, is the most common quality 5 or 6? You can easily remedy this by turning Quality into a categorical variable using the factor() function, and then reapplying the histogram.



Secondly, the scatterplots with Quality on the x-axis are not very insightful with the current layout. Since Quality is more appropriate as a categorical variable, I recommend using a boxplot instead of a scatterplot. As an alternative option, you can apply geom_jitter() so that you can more accurately see the distribution of the values.



How satisfied are you with this feedback?

Resubmit Project

Requirements to exceed the specification ›

Final Plots and Summary

SPECIFICATION

The project includes a Final Plots and Summary section containing three plots and commentary. All plots in this section reflect what has been explored in the main body of the analysis.

DOES NOT MEET SPECIFICATION

Reviewer Comments

I did not see your Final Plot 2 or 3 in the body of your analysis. The Final Plots is meant as a summary of your key findings, so please add them to the flow of your exploration. It is important for the reader to get a better understanding of how you decided to create them and why they are important.

SPECIFICATION

The plots are well chosen and the plots fulfill at least 2 of the criteria. The plots are varied and reveal interesting trends and relationships.
Given criteria:

- Draw comparisons.
- Identify trends.
- Engage a wide audience.
- Explain a complicated finding.
- Clarify a gap between perception and reality.
- Enable the reader to digest large amounts of information.

MEETS SPECIFICATION

Reviewer Comments

The plots are varied and reveal interesting trends. In order to Exceed Specifications, the plots would need to communicate their findings more clearly.

Requirements to exceed the specification ›

SPECIFICATION

All plots have appropriately selected variables and are plotted in a way that accurately conveys the data/information (i.e findings in Final Plot 1 do not depend on the findings of Final Plot 2).

DOES NOT MEET SPECIFICATION

Reviewer Comments

Please make the correction to Final Plot 1 as noted earlier in this analysis.

Regarding your second plot, it appears that the distribution of pH is relatively wide for high quality, so I do not believe the chart demonstrates your finding. Can you think of a different way to display the data or decide on a different finding?

SPECIFICATION

All plots are labeled appropriately (axis labels, plot titles, axis units) and can be read and interpreted easily. Plots are scaled appropriately.

DOES NOT MEET SPECIFICATION

Reviewer Comments

Please include units of measurement in your axis labels!

It is a little unusual to have plots go above the top of the y-axis as seen in your Final Plot 2. I recommend ensuring all of the relevant data is included in the final view.

SPECIFICATION

The reasoning and findings from each plot are explained and the text about each plot is descriptive enough to stand alone. Comments reflect the contents of the plots that they are associated with.

DOES NOT MEET SPECIFICATION

Reviewer Comments

Please be more descriptive when discussing the final plots. Why did you choose each plot and how are the important? Are there outliers? What are the correlations between variables? Is there enough data for all of the qualities to make good conclusions? Keep in mind that the final plot descriptions should 'stand alone'. Try to imagine that you want to describe the chart to someone that does not have it in front of their eyes.

Requirements to exceed the specification ›

Reflection

SPECIFICATION

The project includes a Reflection section discussing the analysis performed.

MEETS SPECIFICATION

SPECIFICATION

The section reflects on how the analysis was conducted and reports on the struggles and successes throughout the analysis. The section provides at least one idea or question for future work.

DOES NOT MEET SPECIFICATION

Reviewer Comments

Please expand on your Reflection. Try to turn this into a broader reflection on your project. In order to pass this metric, you need to include areas that you struggled and also ideas for future work that build on your analysis. A good guideline is 2-3 paragraphs.

Requirements to exceed the specification ›

Download project

Best practices for your project resubmission

Ben shares 5 helpful tips to get you through revising and resubmitting your project.

Watch Video (3:01)

Have a question about your review? Email us at review-support@udacity.com and include the link to this review.

NANODEGREE PROGRAMS

Front-End Web Developer
Full Stack Web Developer
Data Analyst
iOS Developer
Android Developer
Intro to Programming
Tech Entrepreneur

STUDENT RESOURCES

Blog
Help & FAQ
Catalog
Veteran Programs

PARTNERS & EMPLOYERS

Georgia Tech Program
Udacity for Business
Hire Nanodegree Graduates