**Meets Specifications**

**Great job!**

Hi Junwei Tan,

Thanks for fixing all the issues mentioned by the previous reviewers. This is a perfect submission now. **Congratulations!**

You demonstrated your ability to analyze a dataset and create meaningful visualizations from that dataset. This is a tremendously important skill and will prove useful throughout your career in data analytics.

As for further reading, I recommend you to check below links as they show the steps of good explanatory analysis in detail. They will show you how other people approach to data exploration:

* [Comprehensive Data Exploration with Python](https://www.kaggle.com/pmarcelino/comprehensive-data-exploration-with-python)
* [Data Exploration and Analysis Using Python](https://towardsdatascience.com/data-exploration-and-analysis-using-python-e564473d7607)
* [15 Data Exploration Techniques to Go from Data to Insights](https://towardsdatascience.com/15-data-exploration-techniques-to-go-from-data-to-insights-93f66e6805df)

Take pride in the effort you've put into this project and nanodegree.

Kind regards,

Sertac Ozker

**Code Quality**

**All code is functional (i.e. no errors are thrown by the code). Warnings are okay, as long as they are not a result of poor coding practices.**

All your code seems to work well. 👍

**The project uses functions and loops where possible to reduce repetitive code. Comments and docstrings are used as needed to document code functionality.**

You nicely used a function to plot histograms. Perfect! Functions are among the most important concepts and constructs of any language and they generally reduce code duplication. You might check below link to see further examples on function usage.

* <https://www.datacamp.com/community/tutorials/functions-python-tutorial>

However, I feel more explanatory comments are required in between codes. Comments are a good way to tell the underlying idea in a precise manner. There are in some cells but most lack an explanation in that regard. Please check [**this link**](https://www.c-sharpcorner.com/blogs/why-comments-are-important-while-writing-a-code) on why comments are important in coding.

**Exploratory Data Analysis**

**The project appropriately uses univariate, bivariate, and multivariate plots to explore many relationships in the data set. Reasoning is used to justify the flow of the exploration.**

The level of detail in this submission is amazing. You perfectly built a structure going from **univariate** to **bivariate** and **multivariate** plots to explore your dataset. Great. 👍

Since your submission is quite long, you could build up an **introduction** section in your notebook and give hyperlinks to each section there. You can check below link to see how to add hyperlinks to your notebook.

* <https://stackoverflow.com/questions/49535664/how-to-hyperlink-in-a-jupyter-notebook/49717704>

**Questions and observations are placed regularly throughout the report, after each plot or set of related plots.**

In most cases, you nicely made the plot and then wrote your observation. This makes it very clear for a viewer later on. Well done. 👍

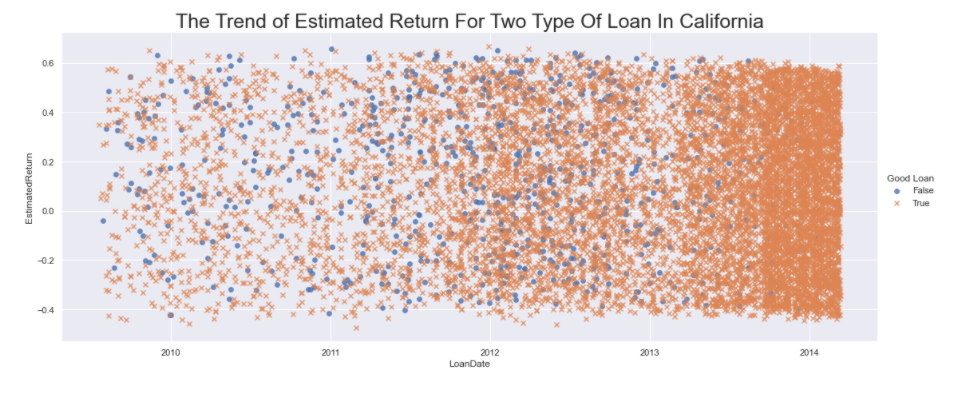
I especially appreciated how accurate your findings were. You did not make over-complex, hard to understand visualizations which further increased your readability.

**Visualizations made in the project depict the data in an appropriate manner that allows plots to be readily interpreted. This includes choice of appropriate plot type, data encodings, transformations, and labels as needed.**

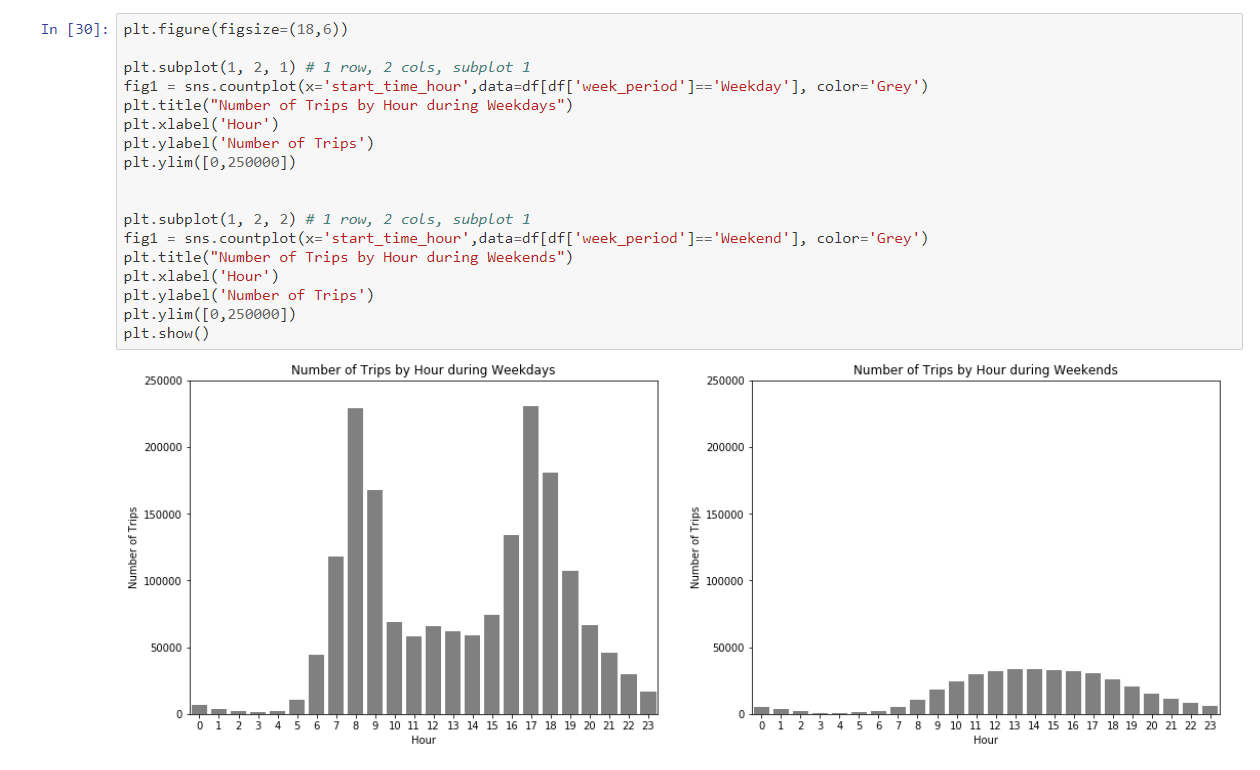
Almost all your plots are appropriately done. 👍

I appreciated how you successfully used different plot types to explore your dataset.

* I wouldn't suggest making scatterplots like this as you may never know how much data is hidden under those points.

[](https://udacity-reviews-uploads.s3.us-west-2.amazonaws.com/_attachments/334640/1613035815/x.png)

Instead, I suggest using subplots and making a separate scatterplot for each category. Here's an example I made:

[](https://udacity-reviews-uploads.s3.us-west-2.amazonaws.com/_attachments/334640/1612466055/subplot_example.png)

If you want to get some ideas or check how to code some particular plots, I suggest you checking this link. There are several cool seaborn plots.

* <https://medium.com/analytics-vidhya/love-the-ocean-love-seaborn-2e8737bef728>

Whenever I'm doing a visualization, I check [seaborn gallery](http://seaborn.pydata.org/examples/index.html) for possible options. Please also check the [controlling figure aesthetics](http://seaborn.pydata.org/tutorial/aesthetics.html) to change the colors or background of the plot to make nicer plots.

**Explanatory Data Analysis**

**A section in the submitted materials includes a summary of main findings that reflects on the steps taken during the data exploration. The section also describes the key insights that are conveyed by the explanatory presentation.**

Your readme file is well-written. 👍

**A slideshow is provided, with at least three visualizations used in the presentation to convey key insights. These key insights match those documented in the summary. Each visualization is associated with comments that accurately depict their purpose.**

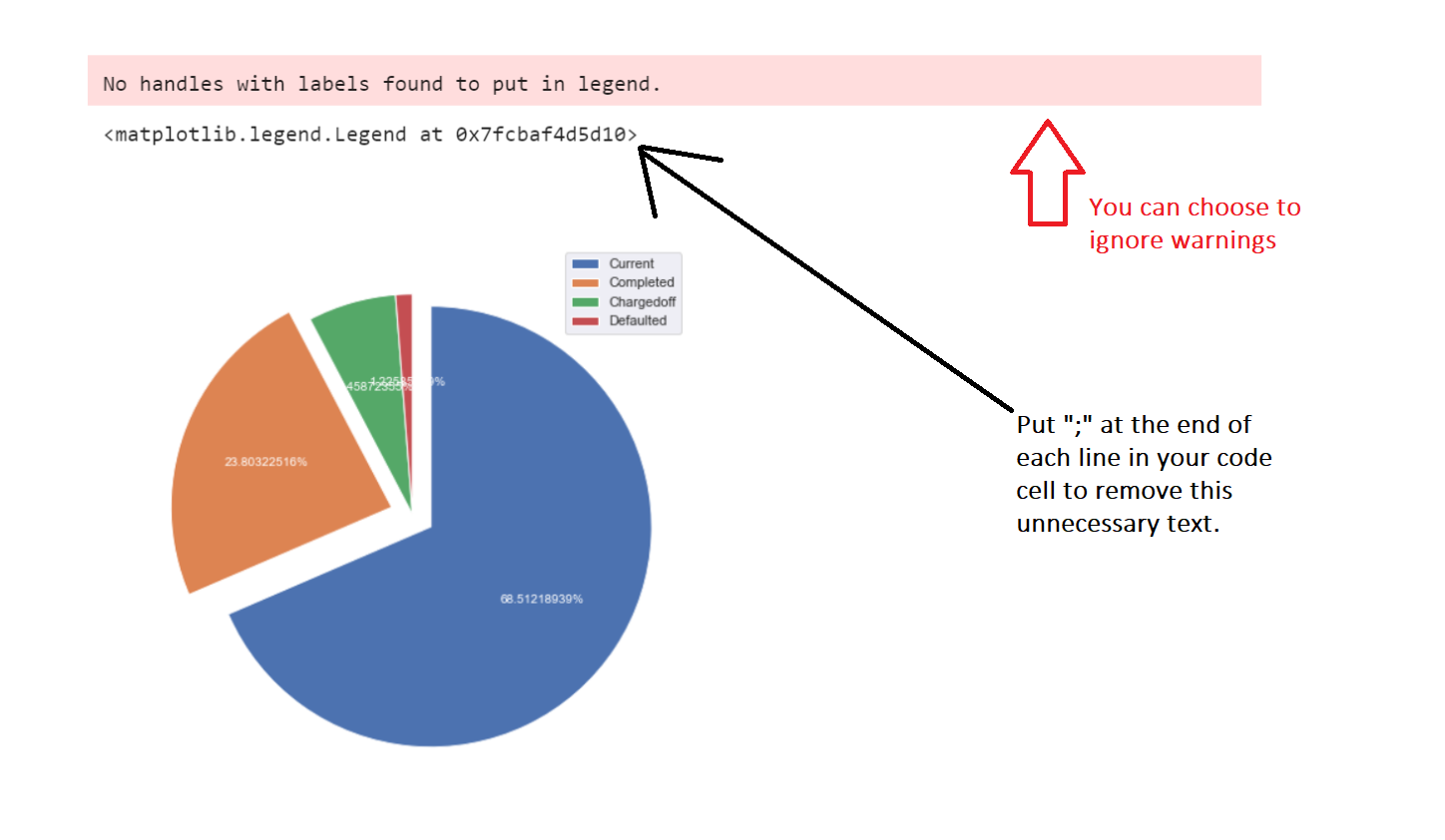
Perfectly done. 👍

I really appreciated how you presented your finding beforehand and showed the visualization so we can quickly come to your point. I also liked the accuracy of your comments.

The **warnings** mentioned in your previous review are still there but I will be lenient on this point. You could simply run below code to get rid of those warnings. You could also add a "**;**" at the end of each code line in your cell to remove the unnecessary text over the plot.

import warnings

warnings.filterwarnings("ignore")

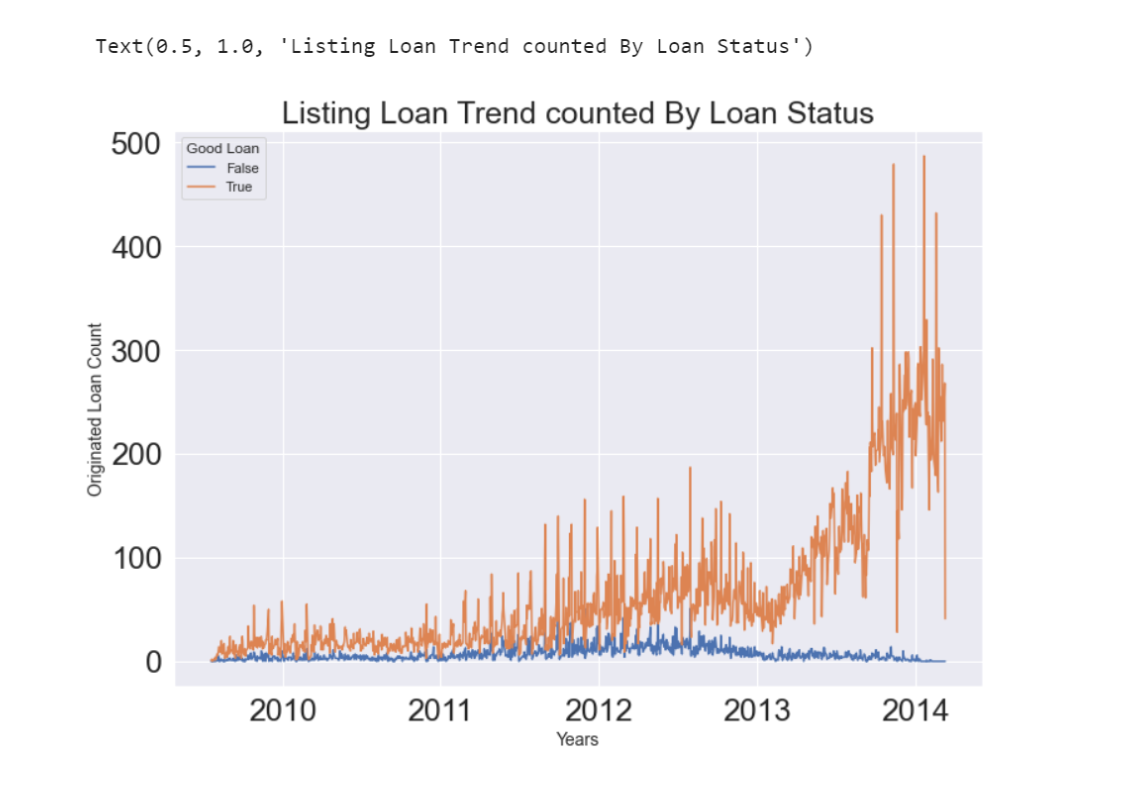
[](https://udacity-reviews-uploads.s3.us-west-2.amazonaws.com/_attachments/334640/1613036137/x.png)

**All plots in the presentation have an appropriate title with labeled axes and legends. Labels include units as needed. Plot type, encodings, and transformations are all appropriate.**

Almost all your plots are perfectly polished with a proper title, labeled axes and a legend when necessary. Great job. 👍

There were some that lacked titles but I believe you've learned the concept so I ignored them.

Just for the below plots, it would be nicer if you build a moving average to smooth your lines. Currently, it's a bit difficult to catch the trend from them.

[](https://udacity-reviews-uploads.s3.us-west-2.amazonaws.com/_attachments/334640/1613036356/x.png)