FLATIRON SCHOOL

Student Name: Akshay Ghalsasi
Instructor: Jeff Herman
Date: 1/24/2020
Attempt Number: 2

Data Science Online Data Science Bootcamp
Module 3 Final Project Review

Technical Notebook

Project Specifications	Metric for success	Developing		Accomplished		Exemplary (X-Factor)		Notes
Pick a novel interesting problem at the appropriate challenge level.	For the Northwinds Database dataset, constructed the business case around it well.	Business case not clearly articulated. Answered an obvious business question.	☐	Business case constructed clearly. And answered an obvious question, like clearly articulated the business stakeholder requirements that the project aims to accomplish. Combine questions in unique ways like how do discounts affect product sales?	☐	Created original and meaningful work - Created a unique business case for the Northwind Database dataset. For example, analyzed the dataset from the point of view of sales, business stakeholders, a startup that would use this business model, want to expand into new geographical regions, specialize in different product lines, etc.	☑	Nice intro to your problem
Preprocess data	Import the data and preprocess the data that includes cleaning, scrubbing, handling missing values, etc.	Data not fully ready for later analysis. 100% correctly structured data. Handled missing values.	☐	Explored different methods.	☐	Handled especially tricky issues. Explored different methods with benchmarking.	☑	Nice SQL querries. Only pulled relevant datal - great work!
Describe data	Use EDA to create meaningful visualizations that describe your data. Plotting words to show cosine similarity, showing plots for class imbalance, etc.	No visualizations are present in the notebook	☐	1 or 2 visualizations are present in the notebook and visualizations are relevant to the project in a technical or business sense.	☐	3 or more visualizations are present in the notebook and visualizations are relevant to the project in a technical or business sense.	☑	Impressed with the plot that you made showing the p value vs the alpha level
Fit models/Hypothesis testing	Hypothesis Tests are ran correctly taking into account standard deviations and normality of distributions.	Hypothesis tests are ran incorrectly. The null and alternative hypotheses tests are not defined.	☐	Correct hypotheses tests are ran but not supported using analysis. In other words the correct test is run on data, but no proof of the data's normality and standard deviation.	☐	Correct hypotheses tests are ran and are supported using analysis. The subsequent business decisions following the hypotheses are correctly identified, and articulated. Future work is explored.	☑	Good use of Monte Carlo Simulation with non-normal data
Present to technical audience	Present work done to a technical audience with code, insights, summary, future work, and even a live demo (for extra credit).	Unintelligible, hard to follow. Unclear. Incomplete.	☐	Engaging talk with insights & lessons. Explained code examples.	☐	Live demo! Ran code and changed parameter values.	☑	Were comfortable walking through your code

Write quality code	Code is non-repetitive and uses OOP when necessary to avoid repetition. Custom methods/classes contain docstrings to help the reader understand what is happening. Variables have names that are relevant to what they represent.	Code is unorganized, lacks docstrings, variables are not named intentionally, and code repeats itself.	☐	Code lacks docstrings, but does not repeat itself and uses custom methods to do repetitive tasks. Code follows pep-8 standards.	☐	Code follows pep-8 standards, contains docstrings/comments, does not repeat itself and uses custom classes methods for tasks.	☑	Add docstrings
Conclusion	Notebook contains a conclusion with business recommendations that are driven by analysis.	No conclusion present.	☐	Conclusion present but only states findings and contains 1 or 2 relevant business recommendations.	☑	Conclusion is present and contains at least 3 recommendations that are business relevant.	☐	Include a conclusion cell walking through the results of each of your hypothesis tests
X - factor: Did something out of the box	Went above and beyond to research some additional topic, concept, Python package(s).	Routine project. Repeated analysis covered in class/sections of the module.	☐	Showed creativity.	☑	Ground breaking.	☐	

Non-Technical Presentation

Project Specifications	Metric for success	Developing		Accomplished		Exemplary (X-Factor)		Notes
Present to non-technical audience	Present work done to a non-technical (business focused) audience with problem statement, business value, methodology explained simply, business recommendations, summary, and future work.	Unintelligible, hard to follow. Unclear. Incomplete. Slides are too verbose, slide notes non existent.	☐	Engaging talk with insights & lessons. Explained methodology. Slides have images, less text, slide notes present on slide that mirror the script of the presenter. One slide for each of the following - Problem statement, business value, methodology, business recommendations (each recommendation on a separate slide), future work/next steps.	☐	Additional slides like findings, or use of engaging images, graphics, material showing expertise in communicating to business stakeholders.	☑	Good intro to the project and explaining why each of the questions you came up with are important
Slide Quality	Slides are light on text, engaging and tell a story.	Slides are very text heavy or highly unorganized and all over the place.	☐	Slides are organized and tell a story, but contain too much text at times, especially when a visualization will suffice.	☐	Slides are organized, contain visualizations that relay information and slides tell a story.	☑	Slides follow a clear story; Opportunity - incldue a conclusion with key insights from all 4 tests
Duration	Your presentation should be between 5 and 8 minutes.	Presentation is over 10 minutes or under 3 minutes.	☑	Presentation is over 8 minutes or under 5 minutes.	☐	Presentation is between 5 and 8 minutes.	☐	13 minutes 3 seconds
Non Technical	Presentation contains great data science that is delivered using non technical language.	Presentation uses technical terms without succinct explanations more than 3 times.	☐	Presentation uses technical terms without succinct explanations once or twice.	☐	Presentation does not use technical terms or provides succinct explanations when using them.	☑	Kept it at an appropriate level
Test Results	Hypothesis test results are shown and made relevant to the business, driving the recommendations from the project.	No tests are shown or tests shown do not relate to business.	☐	Test results are shown and made clear to business case.	☐	Test results are shown, made relevant to business case and also highlight deeper insights into the business.	☑	Results are clearly stated for each of your questions

Visualizations	Slides contain visualizations that take the place of text and give the viewer insight.	Slides do not contain visualizations or the visualizations present are not relevant to the story.	☐	Slides contain visualizations that are relevant to the story but hard to interpret.	☑	Slides contain visualizations that are relevant and easy to understand.	☐	Make the fontsize larger for your graphs tick labels and axis labels
Recommendations	A great presentation contains business recommendations and steps moving forward.	No recommendations are made	☐	At least 3 recommendations are made, but are not driven by data analysis or model.	☑	At least 3 recommendations are made and are driven by analysis and model.	☐	Include a find conclusion; what should we do now that we know the answer to all of these questions
Future Work	A data scientist will never have enough time to explore all aspects of dataset. If you had more time, what other aspects of the dataset would you explore?	No slide on Future work.	☑	Future work slide content not well defined and/or articulated.	☐	Future work clearly articulated, explored, and its potential business impact (s) described.	☐	What are some things you wish you would have tested, but did not either due to running out of time or resources
Thank You Slide	Thank your audience for their time, it's a great practice.	Thank You Slide is not present.	☐		☐	Thank You Slide is present. Appendix includes additional work.	☑	Verbal thank you

Qualitative Assessment

1. Problem Statement - how well was it defined for this project

Came up with interesting questions to test and did a great job relating them to the business

2. Things you did well:

Visualizing the results of your test; verifying the multiple-comparison problem is not going to be an issue

3. Things to work on/ consider :

Think about what takeaways you want the audience to have and add a conclusion to notebook and slide deck

4. Action items:

None - Congrats on Passing!