

Data Science Online Data Science Bootcamp

Module 4 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 zillow prices dataset, constructed the business case around it well.	Business case not clearly articulated. Answered an obvious business question.	<input type="checkbox"/>	Business case constructed clearly. And answered an obvious question such as, 'What zip code increased the most in price?'	<input type="checkbox"/>	Created original and meaningful work - Created a unique business case for the zillow prices dataset. For example, analyzed the dataset from the point of view of sellers, buyers, a startup that would use this business model, etc.	<input type="checkbox"/>	
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.	<input type="checkbox"/>	Explored different methods.	<input type="checkbox"/>	Handled especially tricky issues. Explored different methods with benchmarking.	<input type="checkbox"/>	
Describe data	Use data exploration, visualizations (at least 3 different kinds), distributions, and discussed the distributions as pertaining to the given dataset. Calculated few statistics, models, and analyses.	Inadequate visualizations (less than 3 different ones), fit inappropriate distributions (or no discussion of distributions), inadequate descriptive statistics	<input type="checkbox"/>	Visualizations (at least 3) enough to understand data. Explored enough to understand data. Fit the most appropriate distributions.	<input type="checkbox"/>	Created novel distributions. Compared multiple distributions. Fit distributions outside of Module./class. Found insights in data.	<input type="checkbox"/>	
Fit models/Hypothesis testing	Fit at least one model. Summarize model impact and meaning.	Attempted basic model fitting (or forgot to model fit). Incorrect application. Misinterpreted results.	<input type="checkbox"/>	Correctly fit a single model. Correctly interpreted model results. Summarized model meaning & impact.	<input type="checkbox"/>	Compared multiple models. Fit models outside of class materials. Detailed numerical and visual analysis of models.	<input type="checkbox"/>	
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.	<input type="checkbox"/>	Engaging talk with insights & lessons. Explained code examples.	<input type="checkbox"/>	Live demo! Ran code and changed parameter values.	<input type="checkbox"/>	

Write quality code	Code does what the analysis says it does. It is clear, concise, easy to read and understand.	Code is incomplete. Code NOT in GitHub. Code does NOT work. Code is hard to read. Code does not have README. Commit messages are not helpful.	<input type="checkbox"/>	Repeated some analyses covered in sections/class. Showed some creativity.	<input type="checkbox"/>	Code has comments and tests. Professional level/ pep 8. GitHub repo is public (if appropriate).	<input type="checkbox"/>	
Conclusion	Notebook contains a conclusion with business recommendations that are driven by analysis.	No conclusion present.	<input type="checkbox"/>	Conclusion present but only states findings and contains 1 or 2 relevant business recommendations.	<input type="checkbox"/>	Conclusion is present and contains at least 3 recommendations that are business relevant.	<input type="checkbox"/>	
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.	<input type="checkbox"/>	Showed creativity.	<input type="checkbox"/>	Ground breaking.	<input type="checkbox"/>	

Non-Technical Presentation

Project Specifications	Metric for success	Developing		Accomplished		Exemplary (X-Factor)		Notes
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.	<input type="checkbox"/>	Presentation uses technical terms without succinct explanations once or twice.	<input type="checkbox"/>	Presentation does not use technical terms or provides succinct explanations when using them.	<input type="checkbox"/>	
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.	<input type="checkbox"/>	Slides are organized and tell a story, but contain too much text at times, especially when a visualization will suffice.	<input type="checkbox"/>	Slides are organized, contain visualizations that relay information and slides tell a story.	<input type="checkbox"/>	
Duration	Your presentation should be between 5 and 8 minutes.	Presentation is over 10 minutes or under 3 minutes.	<input type="checkbox"/>	Presentation is over 8 minutes or under 5 minutes.	<input type="checkbox"/>	Presentation is between 5 and 8 minutes.	<input type="checkbox"/>	
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.	<input type="checkbox"/>	Test results are shown and made clear to business case.	<input type="checkbox"/>	Test results are shown, made relevant to business case and also highlight deeper insights into the business.	<input type="checkbox"/>	
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.	<input type="checkbox"/>	Slides contain visualizations that are relevant to the story but hard to interpret.	<input type="checkbox"/>	Slides contain visualizations that are relevant and easy to understand.	<input type="checkbox"/>	
Recommendations	A great presentation contains business recommendations and steps moving forward.	No recommendations are made	<input type="checkbox"/>	At least 3 recommendations are made, but are not driven by data analysis or model.	<input type="checkbox"/>	At least 3 recommendations are made and are driven by analysis and model.	<input type="checkbox"/>	

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.	<input type="checkbox"/>	Future work slide content not well defined and/or articulated.	<input type="checkbox"/>	Future work clearly articulated, explored, and its potential business impact (s) described.	<input type="checkbox"/>	
Thank You Slide	Thank your audience for their time, it's a great practice.	Thank You Slide is not present.	<input type="checkbox"/>	Thank You Slide is present.	<input type="checkbox"/>	Thank You Slide is present. Appendix includes additional work.	<input type="checkbox"/>	

Qualitative Assessment

- 1. Problem Statement - how well was it defined for this project
- 2. Things you did well:
- 3. Things to work on/ consider :
- 4. Action items: