**Data Science Program Final Project**

**Executive Summary**

At the end of the Data Science program, students are required to complete a final project of their choice. They are given six weeks to work on the project. Often times, they will be paired up with another fellow-student.

This document is dedicated to Jane and Jessica’s. It will explain the purpose and scope for the project.

**Business Objectives**

To showcase the skills that Jane and Jessica have acquired through the Data Science program. They will be using R, Python, Tableau, SQL, and other programs to wrangle, analyze, and visualize the “Breast Cancer Wisconsin” dataset made available by UCI Machine Learning on Kaggle.

At the end of the project, Jane and Jessica should be able to explain their work in layman’s term, and present their findings to the students, faculty, staff, and potential employers, along with other interested parties via Zoom.

**Background**

As a way to activate and put practical use to what the students have learned, doing a final project is a good way to demonstrate that.

Jane and Jessica have chosen the “Breast Cancer Wisconsin” dataset because they are both interested in healthcare, and preventative care. They hope to glean insight from this document to make actionable suggestions on how to identify breast cancer types (benign and malignant.)

**Scope**

Jane and Jessica will be using the software taught in the program to complete the project. They will be intentional on using tools of their interest or tools that may aid finding a job. They may choose to use additional software/tools, but that is not required.

**Functional requirements**

Data Wrangling: The downloaded dataset should be successfully cleaned up for analyzing. Columns and unusable columns should be removed. As the dataset is fairly large, Jane and Jessica should consider sub-setting the dataset in a proper manner, meaning the subset should be a random selection of the data. The datatypes for each column should also be converted to a usable format for the needed analysis.

Data Analysis: Jane and Jessica will familiarize themselves with the dataset. They should have a good understanding of what each column means, and how the values are measured. They will brainstorm on questions to ask, and what they might gather from the dataset. Then, they will identify the proper functions to create models, predictions, etc.

Data Visualization: Once Jane and Jessica have a comprehensive understanding of and insight gathered from the dataset, they will work on visualizing the findings. They may decide to use Tableau or other graphing programs, and compile the visuals and texts in a Power Point slideshow.

Presentation: Working with school leaders, Jane and Jessica will schedule a time to present their findings via Zoom. They should be able to communicate in a clear and easy-to-understand manner. The presentation should be kept around 20 minutes. They should be dressed professionally for this occasion.

**Personnel requirements**

Jane and Jessica are the two developers. They will need to work closely for this project to succeed. They will touch base once a day via Zoom or Slack to problem-solve or to check in on work progresses. Once a week, they will review the past week workload and plan out the next week. They will take turns being the scrum master, and report their progress to their instructor (Product Owner.)

Once a week, they will meet with their instructor. They should be prepared to ask questions and seek guidance for the next steps.

They may also consult with their coding mentor.

**Delivery schedule**

Week 1: Import dataset into preferred software to begin data wrangling. Any unnecessary columns should be removed. Educate ourselves on breast cancer. Set up Github.

Week 2: Study the dataset and ask questions. What are some possible correlations? Is the data normally distributed? What are some predictive models we can make from it? Visualize the data to see if there is any interesting findings.

Week 3: Modeling/Optimization (Combined Stepwise - Forward and Backward Selection) and Machine Learning (Random Forest.)

Week 4: Review and validate findings from the previous week, and draw insights/conclusions.

Week 5: Compile findings into a Power Point slideshow. Go over it with their instructor and friend/family member to ensure that the presentation is clear and logical. Work on the style and layout of the presentation so it is delightful on the eyes.

Week 6: Make final touches to the Power Point presentation. Jane and Jessica should not attempt to come up with a brand-new analysis. There will not be enough time to verify their findings. They should practice presenting at least a couple times with the two of them, and at least once with their instructor.

**Other requirements**

All programs used should be free of charge. Though Jane and Jessica may decide to use a paid service, such as a more advanced version of Tableau.

**Assumptions**

The software programs and platforms Jane and Jessica use should be available, up-to-date, and not broken.

**Limitations**

# LZ Guarantee

We stand behind our products and services, and we're committed to the highest quality and accuracy. We want to do right by you—if we made a mistake on your order, we’ll correct it at no cost to you. For anything else, you may be eligible for a refund. Learn more below.



### Filed products:

If your order requires us to submit your documents to a government agency, you can [**call or email us**](https://www.legalzoom.com/contact-us) to request a refund within 60 days of purchase so long as your order hasn't been submitted to the agency. Once documents are submitted, we will no longer issue a refund. if multiple products are packaged and purchased together, the entire order is not refundable if any of the products have been submitted to a government agency.

Examples: LLC and corporation formations

**Restrictions that apply to our filed products:**

* **Attorney-led trademark registration:** Attorney-led trademark registration packages are refundable within 7 days of purchase if no action has been taken on the order. After 7 days or if you’ve attended a consultation with an intake agent, a minimum $150 cancellation fee applies. Or, if you cancel after you’ve attended a consultation with an intake agent and LZ Legal Services has started a trademark search for your application, then a $200 cancellation fee applies. No refunds are available after 30 days of purchase or if you complete your consultation with an LZ Legal Services attorney.
* **Other intellectual property products:** Attorney-assisted patents are not covered by this guarantee. Refer to your legal services agreement for more information.

### Non-filed products:

If your order doesn’t require us to submit your documents to a government entity, you can [**call or email us**](https://www.legalzoom.com/contact-us) to request a refund within 60 days of purchase so long as your order hasn’t been completed. If your order has been completed, meaning it is available for download in your account or received in the mail, you have 7 days from completion to request a refund.

Examples: operating agreements and estate plan bundles

**Restrictions that apply to our non-filed products:**

* Estate plan bundles are refundable up to 7 days after the first estate plan product is completed.

### Subscription services:

You can [**call or email us**](https://www.legalzoom.com/contact-us) to request a refund within 30 days of purchase unless one of the conditions listed below applies. after 30 days, we will no longer issue a refund for a subscription service.

**Restrictions that apply to refunds for subscription services during the 30-day refund period:**

* **LZ Virtual Mail:** Your monthly subscription fee is refundable, but you will still be responsible for all usage fees incurred while your subscription is active, such as storage fees, shipping fees, and check deposit services.
* **Legal plans:** Your subscription fees are refundable unless you have completed an attorney consultation.
* **LZ Tax®:** if your plan includes tax advisory services only, your subscription fees are refundable unless you have completed a consultation with an LZ Tax expert. If your plan also includes tax prep, your subscription fees are refundable unless you receive a completed tax return and give consent for LZ Tax to file it on your behalf.
* **Legal forms:** Your subscription fees are refundable unless you have downloaded a form.

**Please note:**  
Refunds do not include third-party processing fees, such as payments to our service partners. And, we can't refund filing fees once we submit your paperwork to the government. Fees paid directly to attorneys are not covered.

**Risks**

The risks that may arise are such like natural disasters, power outages, family emergencies or broken software/hardware. Jane and Jessica are eager to complete the program so there should be no motivation issues. The instructor and mentor are phenomenal so there is no concern of no help from them. The risk of this project being incomplete is minimal. They will be successful in completing this project!