# Revision and Activity Log History

|  |  |  |
| --- | --- | --- |
| Version | Date | Comment |
| 1 | 10/2/2024 | Fall 2024 Release (Major Overhaul) |
| 2 | 10/2/2024 | Downloaded from Canvas and renamed per standard with is Semester Year – First Last – DSA-1 (e.g., Fall 2024 – Lance Kerwin – DSA-1.docx; or, pdf |
|  |  |  |

## DSA-1 Submission Statement (10/16/2024)

***When you submit your DSA-1, you have an opportunity to make a statement for me and my team of TA’s to consider. You will delete these instructions and insert your statement. The template for your statement is as follows, “I have assessed my overall effort and believe I have performed (below | at | above) the expectations set by Lance. Here are some highlights: (then list 3-5 bullets)***

***I have performed (Answer)***

1. ***Bullet 1***
2. ***Bullet 2***
3. ***Bullet 3-5***

## DSA-2 Submission Statement (11/20/2024)

***When you submit your DSA-2, you have an opportunity to make a statement for me and my team of TA’s to consider. You will delete these instructions and insert your statement. The template for your statement is as follows, “I have assessed my overall effort and believe I have performed (below | at | above) the expectations set by Lance. Here are some highlights: (then list 3-5 bullets)***

***I have performed (Answer)***

1. ***Bullet 1***
2. ***Bullet 2***
3. ***Bullet 3-5***

# Grading Philosophy

DSA-1 is worth 80 points and DSA-2 is worth 200 points. When we grade your work, you are competing against your colleagues. We group assessments by Below Expectations, Met Expectations, Above Expectations. Anyone below expectations starts with a 78%, meeting expectations starts with an 85%, and above expectations starts with a 95%. EXCEPTIONAL work may grade above 95%. Points are then deducted or added from this point.

This document has five sections. You will be able to glean expectations in each of these sections. Many of the expectations will be provided each week and over the life of the project (Weeks 6-12) to encourage weekly engagement and to simulate an iterative experience of building a data story. I will coach you along the way and I have outlined the coaching videos for you in the DSA Resources page.

The biggest risk to running a multi-week project in an online class is that I have little visibility into your weekly commitment, which is why I encourage a time log, and in this framework you have a “Revision History and Log History” which are your opportunity to provide analytics to your work product.

I will honor my NO LATE WORK policy meaning that COMPLETED work includes doing the work, submitting the work on Canvas, and confirming that you uploaded the file that you intend for us to grade. If you attach the wrong file, I’m sorry, there is no remedy for a “wrong file” mistake and the impact is significant to you. You will want to plan for early submission so that if you experience a “Technology Issue” that you can work with the Canvas Help Desk so know their hours. I do not support technology issues.

One other situation that I want to share with you so that I can eliminate this from happening this semester. I will receive emails from students about 24-48 hours before the deadline stating that they are sick. The only way I can even consider working with you in this situation is if you show me significant progress to date. If I give you 21 days to do a task and you come to me in Day 19 stating that you are sick and need an extension, then we actually have a complex problem. You would have to be able to demonstrate that you have completed 19/21 (90%) of the work which is a very tough standard to prove. Please start early and please work in iterations. This methodology will increase your overall quality and position you for an experience that you will be able to leverage in your interviews.

# DSA-1: Setup and Configuration

## 1. Data Visualization Scope and Plan

**Evaluation: grading you on your ability to plan and to apply the concepts/frameworks that we’ve learned to date. I would suggest answering the following questions like an essay, or English paper.**

### What is Business Analytics?

Answer – looking for a thoughtful answer that applies what you are learning…

### Data Gathering Process

*Describe the data gather process.* What data are you interested in researching? What open data platforms are you evaluating?

Answer – journal about you data exploring activities conducted in Week 6 – make sure you listen to all of the material provided in Canvas…

#### Easy Button

*If you are unable to find a dataset of interest based on the video(s) provided, then you have the option to use the data set that I provide in Week 7 copying the steps I put into the coaching video. The highest grade that you can receive with the Easy Button is an 85%. If you use the Easy Button, you need to write an entry into the Activity Log. If you decide to find your own data, then remove Section 1.2.1. Easy Button.*

### Confirming Expectations of the DSA by writing about the Evaluation Process

Describe how Lance and his team will assess the Data Science Activity deliverables (i.e., DSA-1 and DSA-2).

Answer…– *I would expect a summary of the Grading Philosophy and Submission Statement instructions.*

### Business Question

Answer – based on your answer in 1.2. – what do you expect to learn from analyzing the data? What are the potential headlines?.

### Timeline

**Evaluation: grading you on your ability to build a detailed plan and to anticipate what you tasks you think you need to completed in the future.**

The following section will outline your tasks from Weeks 1-12. I would expect you to read through this document and based on your understanding from Week 6 through Week 12, provide a list of activities that you think you need to conduct in order to complete all five sections of the DSA framework. DSA-1 is a milestone assessing how you completed Sections 1 and 2. DSA-2 is the final work product which includes your content for Sections 3-6. I would recommend reading ahead.

#### For the Fall of 2024

I’m asking you to go back and time and capture the activities that should have been completed from Week 1 through Week 5. Take credit for what tasks you completed. Those task that you didn’t complete, you should create tasks starting in Week 6 and beyond to catch up.

**Note: the DSA is a “Living Document” and I will be focused on completing Sections 1 and 2 by insert date. I understand the scope of DSA-2 is Sections 3-6 which is due on insert date.**

The following content is an example outline. I expect you to use this outline and to change the dates that I have in parentheses.

Week 5 Activities (February 6-12):

* Content that relates to YOU

Week 6 Activities (Oct. 3-9):

* Content that relates to YOU

Week 7 Activities (Oct. 10-16):

* Consume lecture materials and identify concepts that I can use in my DSA
* Update Revision History, as necessary
* Continued data gathering each day on Eisenhower Matrix

Week 8 Activities (Oct. 17-xx):

* Add YOUR Activities
* **MILESTONE: DSA-1 due on October 16; but, I plan to submit by Friday October 14 at 5:00pm.**

Week 9 Activities (Month x-x):

* Add YOUR Activities

Week 10 Activities (Month x-x):

* Add YOUR Activities

Week 11 Activities (Month x-x):

* Add YOUR Activities

Week 12 Activities (Month x-x):

* Add YOUR Activities

## Describe your Dataset

**Evaluation: grading you on the level of context and content you are able to provide when evaluating the data types you’ve collected.**

In this section, you will be describing the data that you expect to collect. For example, the options outlined in the DSA Resources page has questions. Those questions have answers. The tables below describe the answers in terms of attributes or columns/fields in a spreadsheet.

**NOTE: Upon your DSA-1 submission, upload the data you wish to use.**

### Data Definition

#### Table 1: Name

Populate the table below per the DSA Launch Video. You may need to add more than one table based on the datasets you find.

|  |  |  |
| --- | --- | --- |
| Attribute / Column Name | Values | Purpose for Collecting / Analyzing Data |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

**During your analysis of the data, you may uncover potential data quality issues or consistencies. You should start keeping a list of data quality issues.**

* Issue
* Issue
* Issue

# DSA-2: Data Preparation and Visualization

The DSA-2 scope includes sections 3-5. Similar to the DSA-1 submission, I provide you with an opportunity to self-assess by providing a ***Submission Statement*** with your work at the beginning of this document. Here is the point breakdown for the DSA-2 sections:

* Section 3: Preparing your Dataset
* Section 4: Storyboard
* Section 5: Storytelling with Data

## Preparing your Dataset

**Evaluation: grading you on your dataset. Was it structured? Scrubbed for quality? Did your explanation make sense?**

Provide a summary of how you modified your dataset with an explanation. You will use this section to document your experience. Questions to consider: how many columns did you use? How many columns did you delete? How many rows did you delete? How was the overall data quality? How many data sources did you have? During your analysis, did you have to gather more data?

You will also be expected to submit your final spreadsheet or dataset AS A SEPARATE FILE with the submission of your paper (i.e., DSA-2). DO NOT EMBED A FILE IN THIS DOCUMENT.

Answer ….

## Storyboard

**Evaluation: grading you on your ability to apply concepts that you’ve learned in the course. You will want to consider leveraging any of the models such as descriptive, prescriptive/decision, or predictive.**

During this course we covered different ways to build an analytics story including several data mining methods (e.g., logical regression, classification). We have also invested time reading from the Knaflic book, Storytelling with Data. My intent on providing a wide range of concepts is to give you a library of methods to consider and choose from. I want to simplify the learning process by letting you explore methods that you find interesting and allow you to select the methods you are comfortable with based on your previous knowledge of math and statistics that you most likely experienced in middle school through high school. Remember, concrete experiences are important to leverage in analytics.

Your data visualization needs to produce four components (e.g., graphs) and needs to tell an overall story. You will insert your data components into a visualization (See Section 5). I would recommend using MS Excel for your tool.

*To complete this section, you need to provide a brief synopsis of how you plan to create your four components INTO ONE VISUAL. Focus on describing your analysis. Maybe you apply some of the concepts from Knaflic and/or other materials in the course. Just specify which ones you used and describe what those methods accomplish. MS Excel also has methods for creating visualizations. You are welcome to explore those features as well.*

**Additionally, you can apply any of the data mining methods such as regression or classification that we have covered. I would also consider the types of visuals/graphs that we learned in Storytelling with Data.**

Answer

### Graph 1: Dog vs. Cat Stress Level

I would like to explore leveraging a vertical bar chart to show the stress levels between those that prefer cats to dogs.

A graph of multiple series

Description automatically generated

### Graph 2: Name

I would like to incorporate a slope graph because…….

A blue lines with dots

Description automatically generated

**[Intentionally left blank]**

### Graph 3: Name

Tell me about …..

**[Intentionally left blank]**

### Graph 4: Name

Tell me about …..

**[Intentionally left blank]**

## Storytelling with Data

**Evaluation: were you able to follow directions? Were you able to use MS Excel to create your four graphs? Did your presentation look compelling and professional? Did you take time to “design” how you presented your data? Did your data tell a story? Did you incorporate your four graphs into a SINGLE visual?**

Provide an overview of your visual below. This is the section where you actually show the insights generated from your analysis. A key to success is making sure that you plan time between Sections 4 and 5 as they are different. You can complete Section 4 without any data, and performing any analysis. Section 4 is a conceptual design component. Section 5 is your actual work product.

A graph of different sizes and colors

Description automatically generated with medium confidence

Provide talking points to support your visualization below:

* Graph 1 tells that Monday and Wednesday….