

Milestone 1 – Dataset Worksheet (Individual)

AI Guidance: Generative AI may not be used for this assignment.

Collaboration Guidance: Collaboration is not required for this assignment. Acknowledgment required for collaboration and reduction in points may occur.

Instructions: Save this assignment to your MA206 folder. Using one of the recommended data resources, identify a dataset you are interested in using for your course project. Answer the below questions about your dataset.

1. What is your research question?

[ANSWER HERE](#)

2. What is the source of your dataset? This doesn't mean "where did you find it?", but rather explains the data collection. An example could be, "this data was collected by Northwestern University researchers from 2017-2022 using patients undergoing treatment for diabetes."



- Website link for your data:

[ANSWER HERE](#)

- Synthetic datasets are not acceptable. You should be able to confidently state the source of your data. See Milestone 1 instructions for more detail.

[ANSWER HERE](#)

3. What does your data look like? Replace the below example screenshot with a screenshot of your own data. You will notice that the example dataset does not have sufficient variables to meet project requirements; ensure that your selected data does.

# Rank	Country	# Visa_Free_Score
The position of a country's passport in the global Henley Passport Index for the year 2025. Countries with the	The name of the country or territory whose passport	The number of destinations the passport holder can travel to without requiring a visa in advance. This includes
 1 99	199 unique values	 25 193
1	Singapore	193
2	Japan	190
2	South Korea	190
3	Denmark	189

4. Your dataset must have a quantitative response variable. What is your response variable? Remember that quantitative variables are numerical variables where arithmetic information is meaningful (average 3-point shot success rate is meaningful, average year of Stanley Cup is not).
- Variable name: **ANSWER HERE**
 - Variable description: **ANSWER HERE**
 - Unit of measure: **ANSWER HERE**
5. In addition, your dataset must have two more quantitative “explanatory variables” that you suspect have some sort of relationship with or influence on the response variable. List them here:
- Variable name: **ANSWER HERE**
 - Variable description: **ANSWER HERE**
 - Unit of measure: **ANSWER HERE**

 - Variable name: **ANSWER HERE**
 - Variable description: **ANSWER HERE**
 - Unit of measure: **ANSWER HERE**
6. Finally, your dataset must have two categorical (non-numerical) explanatory variables that you suspect have some sort of relationship with or influence on the response variable. You may have categorical variables with up to 4 categories.
- Variable name: **ANSWER HERE**
 - Variable description: **ANSWER HERE**
 - Number of categories: **ANSWER HERE**
 - If > 4 categories provided in original file, the categories selected for analysis are: **ANSWER HERE**

 - Variable name: **ANSWER HERE**
 - Variable description: **ANSWER HERE**
 - Number of categories: **ANSWER HERE**
 - If > 4 categories provided in original file, the categories selected for analysis are: **ANSWER HERE**
7. Include the citation for this dataset IAW DAAW here:
ANSWER HERE

Submit this completed worksheet, with appropriate documentation, on the Canvas course site under Assignments → Course Project Milestone 1 → Dataset Worksheet.