Midcourse Project Requirements

The midcourse project is an effort for the students to combine what they have learned into an individually designed, executed, and presented project. The students will apply their skills to a project that has not already been defined. The goal is for the students to take ownership and stretch themselves. This could include use of multiple technologies or datasets, or taking deep dives into one. The product will be something that will be a keystone of their portfolio. The instructor's involvement is to support progress and arrival at an MVP defined by the student. This includes ensuring students meet deadlines, supporting and troubleshooting student efforts toward data acquisition, analysis, visualization, or presentation, conducting regular check-ins and providing feedback to students.

Requirements:

- The student will choose data that is representative of real world data which will showcase their skills in data wrangling and effectively answering data questions.
    - The chosen dataset will allow the student to demonstrate cleaning, merging, and data preparation skills.
    - The primary data set cannot come from Kaggle.
    - The primary data set cannot be class project data set.
- The student will ask and answer appropriately scoped but sufficiently complex question(s).
    - The data question(s) will be maintained and clearly defined throughout the process of working.
    - The data question(s) will be sufficiently different from previous class projects.
- The student will have all code in GitHub.
    - demonstration of analysis
    - at least one push per class
    - include informative and representative README
    - ensure no visible PII
    - Exception: use of proprietary data. In this case, the student must have instructor review code.
- Communication of findings using visualizations that are relevant to the storytelling
    - Interactive R Shiny app that allows for exploration and discovery by the user
- Midcourse project work demonstrated and approved.