Explorations in Data Science - Algorithm Bias Group - Project Plan

- Paper Topic
  Our paper will discuss the major challenges of fairly using algorithms that affect individual people. We'll briefly discuss the various problems, then provide some concrete examples of the problems, and then wrap up with some possible solutions to said problems.

- Research Planned
  These are the specific topics we researched in preparation for our presentation and for the final paper
    - General topics:
        - data collection bias
        - algorithmic bias
        - An overview of Amazon Rekognition, its training and availability
    - Gender and NLP
        - overarching theme has to do with how gender bias ends up infiltrating NLP systems
        - real world examples of this occurring
        - why it's a problem
        - how researchers are attempting to fix gender bias in these systems
        - Specific NLP areas:
            - Word embeddings
            - Machine translation
            - Sentiment analysis systems
    - Facial Recognition Issues in LGBTQ and Race recognition
        - accuracy of facial recognition systems at determining genders of the non-binary population
        - why is this a problem?
        - is it even necessary to perform the task the programs are designed to do?
        - how do we improve?
    - Racial Bias in Criminal Justice
        - Facial Recognition disproportionately misidentifies minorities
        - Algorithmically recommended sentencing guidelines based on likelihood of recidivism
        - Not all bad
        - We can do better, there are current techniques, and it's an ongoing process
    - Current and Potential AI Implications in Medical Field

        - Current advancements of AI in Healthcare
        - Potential biases of AI in Machine Learning Algorithms using Health Record Data
        - Missing Data and Patients not identified by algorithms.
        - Sample size and Underestimation
        - Misclassification and measurement error
        - Best practices and policies to mitigate the above bias

- <u>Team Structure</u>
    - Hannah Galbraith - Gender Bias in Natural Language Processing
    - Will Mass - Intro, overview, definitions of biases (algorithmic and data collection, with generic examples), and research into Rekognition (how it's trained, and its general accessibility)
    - Annie Merlin - Algorithm Bias in Medical Field
    - Trina Rutz - Facial Recognition Software & Automatic Gender Recognition
    - Chad Tolleson - Racial Bias in Criminal Justice

- <u>Milestones</u>
    1. July 2nd @ 4pm - meeting w/ Kristin - TREC Conference Room - it is FAB 170-05 - the room is about halfway down the hallway between the FAB Fourth Ave entrance and the Computer Science Dept on the left
    2. July 2nd to 7th - team members research their subtopics and create dialogue, questions, and related slides for in-class presentation
    3. July 7th to 9th - team consolidates subtopic materials and create a comprehensive presentation, timed trial run of presentation
    4. July 9th - presentation
    5. July 19th - provide project plan / update to Prof. Tufte
    6. July 19th to August 13th - additional research as necessary, write and edit the final research paper
    7. August 13th to 15th - final paper and presentation due