Group Members Names:

Eddie Kiernan, Logan Proffitt, Noah Turner

Team Name:

Iron Insight

Problem Space:

- Problem / Inspiration: Lifters can often feel discouraged by their current lifts and might wonder how much they could lift in the long run if they stayed consistent. The presence of performance enhancing drugs (PEDS) has only added to this discouragement, since most lifters have no idea of what is achievable without PEDS. We would like to give lifters a reasonable estimate of how much they might be able to lift with and without PEDS, based on the progression of others (powerlifters competing in events).
- Machine learning would be useful for this problem because there is a ton of data from powerlifting competitions available online. This data contains various features such as bodyweight, age, gender, lifting numbers, success vs failure, tested vs untested, and more. We could leverage machine learning to identify the relationship between various features and predict potential future numbers for others.
    - We could build a model that predicts how long it could take a lifter to increase numbers based on how powerlifters' max number have progressed over time (1 year, 2 years, 10 years, etc.)
        - Look at percentile of lifters that increase number be X_1 in X_2 time
        - How big of increase and how much time it took
    - Could also look at it in terms of tested (natural) lifters and untested (PEDs) lifters
        - Could look at lifters who went from tested to untested to see how their numbers changed
        - We could also look how lifts progressed for tested lifters vs tested lifters
            - Greater degree increases for untested lifters? By how much?

Data / Data Plan:

- Potential data sets
    - https://www.usapowerlifting.com/national-results/
        - All results for USPA past events
    - https://openpowerlifting.gitlab.io/opl-csv/bulk-csv.html
        - Huge dataset of results from various powerlifting competitions
- Using https://openpowerlifting.gitlab.io/opl-csv/bulk-csv.html, we could look at the long-term progression of elite athletes to find how much it is possible to improve, as well as the role of PEDS in this progression

    - We could do this by filtering on athletes with a long history of competing and analyzing their progression over time
    - We could also compare similar athletes of varying ages as well as their history and use this for a general prediction of progression. This could end up being a sort of "How much *could* I lift at X age or after X years" prediction. While the result might be a bit generalized, it could be fun to have an estimate of how much you might be able to lift if you stay consistent. Percentiles could also help to solve this issue.
- Interesting or critical features: The openpowerlifting.org data set contains important features such as date, age, bodyweight, gender, equipment, tested/untested, and various lifts and their success/failure. Features such as location might be excluded.
- Outcome variable(s): The outcome variables for our model would be a predicted weight for a given lift. A range of weights could also potentially be given, with percentiles to indicate the likelihood of various weights.
- Sample size: openpowerlifting.org contains both a 131M and a 54M file, though there may be overlap between these. If this is not enough, we can leverage records from other powerlifting websites as well.