**Introduction**

* Linköping Hockey Analytics Conference in Sweden hosts a yearly student competition.
  + <https://www.ida.liu.se/research/sportsanalytics/LINHAC/LINHAC23/studentcompetition.html>
  + For the competition, a dataset is provided to each participating team with event information corresponding to 2020-2021 games played within the Swedish Hockey League
    - As an example for some of the data that may be included, we have insight into the participating teams, the current time elapsed in the game, the player in possession of the puck, the event (shot, pass, carry, dump, save, etc.) that was performed at the listed time, whether the event was successful or failed, the x and y adjusted coordinates denoting puck positioning in the rink among many others.
  + Leveraging this dataset, the competition lists only a general goal to identify patterns related to sequences of events that may lead to some outcome. Whether this outcome includes a scored goal, a high-percentage shot attempt, or a successful penalty kill is left up to the participating teams

**Methods**

* Our current project scope is focused around two major questions and areas to explore.
  + Due to its prevalence as a question in the literature review centered around hockey data analytics, our team will explore success (goal or high-percentage shot on goal) associated with either dumping the puck into the offensive zone or an offensive player carrying the puck in the offensive zone
    - Initial hypothesis, again based on the literature review, suggests increased likelihood of success when the puck is carried in.
    - To investigate this question, there are quite a few options we could explore, but to start:
      * Leverage sequences beginning with carry-in and dump-ins to the offensive zone as input for a supervised model. The end result would be a model that predicts likelihood for success based on the steps taken to get into and within the offensive zone itself
      * Leverage the hockey\_rink library within Python to create and project sequence visualizations on a hockey rink that were and were not successful.
  + The second question we hope to explore is more general but similar to the first, as well, in that we will predict the next event in a sequence based on the previous event in the same sequence.
    - Similarly, supervised learning methods and visualizations will be included in the methods.
      * Specifically, investigate Markov Game models as this was a suggestion in the slack channel but is also quite prevalent in the literature, as well, for this type of problem.

**Results**

* Models that are developed and trained will need to be evaluated.
  + Leverage various evaluation metrics following cross-validation of the model, such as precision, recall, and accuracy as some examples.
* The hockey\_rink library should be a powerful tool to leverage due to the availability of x and y coordinate data, and we therefore expect to effectively plot visualizations that outline our findings.

**Discussion**

* Questions that we should consider within our discussion section
  + Is there a statistically significant difference in offensive zone success when the puck in carried in versus when it is dumped in by the attacking team?
  + Based on the data we have available from the Swedish Hockey League, are we able to effectively predict the next event of a sequence given the prior event?
  + Based on our methods, what areas did we fail to explore or consider and what would be an interesting future area of study?