The purpose of this assignment is to brainstorm thesis ideas and to think about the potential ethics/privacy/security issues of your topic. First, you will think about 3 potential topics for your thesis and explore those options. Then, you will pick your top choice and think of ethical issues. If you already decided on your thesis topic, you may go with one topic with more details.

1. Think about thesis ideas that interest you. Write down those three ideas in the following table. Why are they of interest to you? Any drawbacks or difficulties to study the topic as your thesis? (25 point)

|  |  |  |
| --- | --- | --- |
| Topic | Interesting aspects | Any drawbacks/difficulty |
| 1.  Solving for Bayesian Equlibrium in Predicting Pit Stops in Formula1 Race | I think the use of predicting a decision to make a pit stop or not is perfect for Bayesian statistics as each lap a pit stop is not made increases the likelihood of a pitstop.  Bayesian Equilibrium comes into play as this is a strategic decision.  Predicting a decision is the data science aspect. | The data had to be gathered in an arduous manner, with each race from each year being pulled individually in several different ways to get all the data.  The data is also only consistent dating back 6 years so there is some slight restriction on timing data (but with 20-ish races/yr and 20-ish drivers/yr, there is a large enough sample size). |
| 2. |  |  |
| 3. |  |  |

1. For those topics you listed above, explore your own 2-3 research questions and find potential data sources. Spend an hour or two looking for data sources and report what you found. (25 point)

|  |  |  |
| --- | --- | --- |
| Topic | Research questions | Where to find the Dataset(s) |
| 1. | Solving for Bayesian Equlibrium in Predicting Pit Stops in Formula1 Race | https://theoehrly.github.io/Fast-F1/ |
| 2. |  |  |
| 3. |  |  |

1. Reflect on the three topics you compared in your table in A and B. Pick your favorite thesis topic. Find three to five articles about the topic at least one of which should mention ethics/privacy/security issues related to the topic and reflect..

a. What is your favorite thesis topic? Why? (10 point)

I ended up going with the only one listed above because that was initiated several months ago. I have plenty of work-related ideas including using time-series weather data to predict labor needs or classification to better hiring decisions. Outside of work, I have an idea to create a jigsaw puzzle solving algorithm or to do some geospatial analysis to find the best neighborhood for me based upon proximity to grocery stores, swimming pools, etc. Another idea is to create a highly personalized book recommendation system for my father who has 1000+ books he has read (or listened to) all on an excel list with ratings and a synopsis of his thoughts.

b. What are three to five articles? Provide the link or article information (close to what you can use in Bibliography). (20 point)

1. Bruce Bueno de Mesquita - The Predictioneer’s Game - <https://www.amazon.com/Predictioneers-Game-Brazen-Self-Interest-Future/dp/081297977X>
2. Bruce Bueno de Mesquita - A New Model for Predicting Policy Choices - <https://www.jstor.org/stable/26275398>
3. Norman Schofield - Instability of Simple Dynamic Games - <https://www.jstor.org/stable/2297259>
4. Avinash Dixit and Barry Nalebuff - Thinking Strategically, the Competitive Edge in Business, Politics, and Everyday Life - <https://terpconnect.umd.edu/~pswistak/GVPT%20100/Dixit%20and%20Nalebuff.pdf>
5. Erik Jan van Kesteren and Tom Bergkamp - Bayesian Analysis of Formula One Race Results: Disentangling Driver Skill and Constructor Advantage - <https://arxiv.org/pdf/2203.08489.pdf>
6. Analysis of tire changes, yellow flags, etc -  
   <https://users.cs.duke.edu/~cynthia/docs/TulabandhulaRu14Racing.pdf>

c. Pick at least one article ethics/privacy/security. Write a reflection in 2-3 paragraphs. Note that these wordings can be used in your thesis. (20 point)

I’m not really sure if there is a lot of ethical concerns. This is a sport and we are using prior action in the same situation to predict future decisions. The worst thing I can think of will occur if this is massively successful that it gets picked up on and gets extrapaloted and used for more serious situations than a sporting event.