ADHD Study - Milestone 2.2

1. *Any surprises from your domain from these data?*

(This topic got long-winded so I keep the other topics shorter.)

Surprises from research into domain:
- Disability: I learned that ADHD is recognized as a disability under federal legislation; I am not sure how this slipped past me before.

- Parkinson Disease: Commonalities with ADHD include damaged dopamine neurons in the basal ganglia, and cerebellar abnormalities. Taking medicine for ADHD is risky, as it increases the risk of Parkinson's. However simply having ADHD adds risk; we are twice as likely to get Parkinson's than peers if we don't take medication.

- Sensitivity - both to internal feelings of rejection and external stimuli such as noise and light.

  Rejection Sensitivity Dysphoria (RSD):
  "RSD is an intense vulnerability to the perception – not necessarily the reality – of being rejected, or criticized. RSD causes extreme emotional pain that may also be triggered by a sense of failure. Often, people experience RSD as physical pain, like they've been stabbed or struck right in the center of their chest.

  98-99% of adolescents and adults with ADHD acknowledge experiencing RSD. For 30%, RSD is the most impairing aspect of their ADHD." (Dodson, 2020)

  I found those numbers shocking, as it's not what most people think when describing the condition (trouble sitting still, etc).

  Combine RSD with this information: ADHD researchers estimate that by age 12, children with ADHD get 20,000 more negative messages about themselves than other kids their age. I believe this is in part because we have executive dysfunction so we aren't doing what we should, and in part because we fall on the spectrum of Autism (difficulty anticipating the responses that neurotypicals are more likely to give).

Surprises from EDA:
I created a dataframe of the highest correlation variables with target column K2Q31A.
Most of them made a lot of sense to me, such as Autism and Anxiety. There were some new pieces of information though:

- Tourette Syndrome - Research does support this, in fact ADHD is the highest comorbidity with TS: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4737687/#:~:text=Core%20tip%3A%20Tics%20can%20be,exacerbate%20pre%2Dexisting%2C%20tics.

- Cerebral Palsy - Further research supports Cerebral Palsy comorbidity: https://www.smarterparenting.com/specific-diagnosis/when-adhd-co-occurs-with-cerebral-palsy/

2. *The dataset is what you thought it was?*

The dataset is what I thought it was. It did have some complexities that I need to be mindful of. Some of the questions are follow-up questions that are only to be answered if an answer of "Yes" was given for a different question. This means that there would be data leakages if I included the following up questions for K2Q31A (target).

3. *Have you had to adjust your approach or research questions?*.

My research question has gotten more targeted towards individuals without conduct problems, who are likely masking. These are the children who are more likely to be undiagnosed. I have not yet used an unsupervised clustering technique to find similarities in the data points that do have ADHD, however it could be interesting to further isolate subgroups.

4. *Is your method working?*

Yes. My current models have the following statistics:

Highest accuracy models used a subset of features with the highest correlation to ADHD (intrinsic and extrinsic):
0.946 accuracy - does not include K2Q34A "conduct / behavioral problems" subset
0.939 accuracy - includes K2Q34A

ROC - How sure do I need to be that they are diagnosed with ADHD?
(We are looking at a subset that is under diagnosed until adulthood, so some uncertainty in diagnosis is okay):
- 0.808 - using 44 features focused on intrinsic characteristics - does not include K2Q34A (Accuracy using these features = 0.925 and does not include extrinsic identifiers such as parent employer info, do they attend private or public school, etc.)
- 0.916 - 63 features, does not include K2Q34A subset
- 0.935 - 63 features, includes K2Q34A subset
- 0.938 - 301 features

5. *What challenges are you having?*

I am already far off schedule, potentially due to my ADHD. However I am finding success in this project and looking forward to the presentation.

Portfolio: https://holly-e.github.io/

GitHub Project URL: https://github.com/Holly-E/ADHD_Study

References:
Bowman, S. (2020, Jan 15). *"Overdiagnosis or Underdiagnosis of ADHD."*
https://www.rtor.org/2020/01/15/overdiagnosis-or-underdiagnosis-of-adhd/

Dodson, W. (2020, July 31). *"3 Defining Features of ADHD That Everyone Overlooks."*
https://www.additudemag.com/symptoms-of-add-hyperarousal-rejection-sensitivity/