The main focus of the project proposal (update 1) is (1) defining a research question and (2) doing exploratory analysis.

Proposals should be prepared in slide form through RMarkdown, with up to 10 slides. You must submit both your Rmd file and your slides. Your proposal should include the following components:

1. Introduction (1-2 slides): Describe the dataset, the data collection, the scientific context, and the *key* variables that were collected. Describe the disorders being studied. **State your SPECIFIC research question.** Your research question should seek to answer the main question of interest to the collaborator (the relationship between motor skills and social deficits in children with ADHD and Autism), but you can come at this from many different angles.

You can read the paper titled, “Evidence for Specificity of Motor Impairments in Catching and Balance in Children with Autism” by Ament, Mejia, et al. (available in the Data Analysis Project folder on Canvas) to understand more about the scientific context and to see an example of a research paper introduction. (Make sure not to copy anything directly from the paper – this would be considered plagiarism and a violation of the IU code of ethics, and would result in a grade of zero on the proposal.)

2. Exploratory data analysis (7-8 slides): Describe any data cleaning or organization steps you performed. Summarize the number of observations in each disease group and sub-group, and the main variables of interest. Check that the main disease groups are balanced on each variable. This information should be summarized in a table (this is typically "table 1" in a scientific paper).

Begin to examine and visualize the data to explore your research question. Here are some ideas: boxplots of motor performance scores by disease group, plots of age versus motor performance, use a smoother for each disease type/subtypes. Plot before and after performance for those subjects with repeated visits. Use color and grouping/smoothers to add a third variable to a plot. Make MANY plots to explore different relationships between variables. (You will only keep the important ones for your final report.) Each plot should explore a DIFFERENT question or relationship – don’t build up or repeat plots. After EACH plot, state your observations. What relationships do you observe? **You should not fit any models at this point.** Your exploratory data analysis will help you formulate the models for your final report.

3. Conclusions (1 slide): Based on your exploratory analysis, what do you BELIEVE about the nature of the relationship between motor skills and social skills in children with autism and ADHD compared with typically developing children? We will not have learned a lot of the statistical lingo yet, so you just need to describe your beliefs in layman terms.

Tips:

- Divide and conquer. Have different group members investigate different things. Recombine to discuss your results and iterate.

- Although the divide-and-conquer approach is encouraged, your report should be one cohesive document, not three or four disconnected analyses. Make sure you take the time to meet together and SYNTHESIZE the analysis each team member has undertaken.

- Remember to focus on **the main relationship of interest (motor skills (X) and social deficits (Y) in each population).** You should explore this relationship in detail for the different diagnosis groups. While there are many other variables (Z) that may influence this relationship, your end goal should always be to understand that primary relationship. There are many ways to do this: look at the residuals of Y vs. X and see how they relate to Z; use color, faceting or grouping to see how the relationship between X and Y changes with Z. For each plot, make sure to relate your findings to the overall goal of studying the relationship between motor skills and social deficits in autism and ADHD.

- Regarding the length of the report: 10 slides is a general guideline. However, I'd rather see a slightly shorter presentation that is well-organized and cohesive, than a longer presentation that lacks a clear structure and narrative. Don't focus too much on the length – focus on quality.