AI Study Data Analysis Outline

1. Download Qualtrics survey data and chatplat chat log data.
   1. Qualtrics data should come separated by close/small and predictor/predictee
   2. Chatplat data should come separated by close/small

On Qualtrics data:

* Consolidate data – combine all spreadsheets into one large spreadsheet
  + Make sure the columns match
  + Add columns indicating close/small, AI/human, and predictor/predictee conditions
* Add column that sums the number of movies picks that they were correctly able to guess.
  + Create column that shows difference between their guess on how many they would get correctly and how many they actually got correct.
* Combine the two columns that ask for their confidence in whether the person they were talking to was an AI or human into one column.
  + On the scale, human should be 100 and AI should be -100
  + Create column that shows difference between predictor and predictee answer to above.
* Using chatplat data, match predictors and predictees by group using their unique ID (R\_####)
  + Combine data by row – each row should be one group, with the responses of both participants in one row.
  + Add the two columns containing the predictor’s and predictee’s ID.
* At some point, filter out those who failed attention check

On Chatplat data:

* Add two columns containing the predictor’s and predictee’s ID.
* Filter out conversations that did not complete most of the questions.
* Some chats may have weird symbols, be out of order, have no spaces, etc. – try to clean chats using code, otherwise clean the rest manually.
* Format data into rows such that each row is the conversation of one group
  + Separate each response into a column. Format should be (Q1-A1a-A1b-Q2-A2a-a2b…)
  + Try using code to do this, check the rest manually.

1. Combine Qualtrics and Chatplat data into one spreadsheet.
   1. Each row should have the chat and responses of both participants in each group.