Team Name: Student's t-eam

Team Member Names: Noah Cline, Carissa French, William Laubscher, Brendan Pinkerton

We started by thinking about the Study's purpose: How did the game affect overall understanding of sexual health and substance abuse? The gameplay is focused on a cyclical loop of choosing the wrong choice, analyzing the environment, and using those clues to find the eventual correct choice. Sprinkled throughout are minigames that teach the same lessons about substance abuse and sexual health as the game's main scenarios. As such, we found that there is not much to conclude about the order of events completed or the minigames completed, as 100% completion of the game simply requires everything to be done in a certain way. We decided that in order to see the game's impact on understanding of the topics, we would need to use the S5 assessment scores. Our goal was to see if they were related. If so, how?

We removed players without S5 assessment score data, outliers in total playtime (as determined via IQR criteria), players with anomalous event_time_dbl data, entries for event_id and variables specific to minigame data and panning the scene. This left us with 32 potential variables of interest, 178,999 observations, and data for 43 students.

Five out of seven coefficient estimates for Time_Spent are negative, which could indicate that higher playtime doesn't mean higher S5 assessment score.

Initial residual plots from a model fitted with first-order predictor terms demonstrated significant non-normality and non-constant variance. A LOOCV indicated the best remedy to be a seven-degree model of the predictors, and a Box-Cox transformation indicated a $\lambda = 2$ transformation of the response would help as well.

Our final conclusion is that the number of events a player interacts with has explanatory power when it comes to predicting S5 assessment score, but only minimally so.