ROI – CBSC Trends Over Time

My final project in the class, CBSC Trends Over Time, allowed me to further my data science skills in several ways. The project was centered around a daily survey that the class took before the start of every class period (2 times a week). From the survey, I was to pick a topic of analysis that interested me, form a hypothesis, analyze the data, and then draw conclusions based on what I had found. For this project, the data was obtained through responses from the daily survey. I was interested in studying the weather, how much time someone spends outside, and mood.

For the analysis, I used R to put my data in the form that I needed it to be in – I selected only the variables I needed from the dataset and then got rid of any missing information. Then, after cleaning the data, I made visualizations to get a sense of the responses I was looking at. I made histograms and boxplots for this as they show the frequency and skewness of the data. Finally, I used R to run descriptive statistics on the data, looking at the mean, min, and max of each variable in my dataset.

When performing my analysis, I was looking at two things: how the weather impacted how much time people spent outside and how much time someone spent outside affected their mood. For the first question, I hypothesized that the nicer the weather, the more time someone will spend outside. For this question, I ran a mixed model ANOVA test in R. For the second question that I asked, I hypothesized that the more time someone spends outside, the better their mood will be. For the analysis of this question, I decided to run a time series analysis in R.

Finally, I had to present my findings in a poster. I learned how to communicate the most important information concisely with this because there is not much room on a poster. I also

learned to pick the visualization that looked the best and provided the most information for the viewer when deciding what visualization to put on the poster.

Overall, with this project, I furthered my data science skills by obtaining my class survey data and then being able to transform it into a dataset that I could use for analysis. With this project, I practiced how to deal with missing information in datasets and how to get a large dataset into the form needed with just a few select variables. Furthermore, I practiced using my analysis skills as I ran a mixed model ANOVA and time series analysis in R to answer my research questions. Finally, because the outcome of the project was a poster, I learned how to concisely communicate the work I did. Overall, this project allowed me to showcase my data science skills, and I learned a lot while doing it!