**Descriptive Statistics Investigation 1 - Turag**

For the first descriptive statistics investigation, we aimed to compare media consumption between two groups of users based on their deprivation status. Dataset 1 includes a column labelled ‘deprived’ to indicate whether a user lived in a high-deprivation area—characterized by high unemployment rates, crime, poor public services, and housing barriers. Users from these areas are marked with a 1, while those from non-deprived areas are marked with a 0.

To test this hypothesis, we began by loading the necessary libraries and datasets. We created two separate lists to track users from both deprived and non-deprived areas, using the shared ‘ID’ column from both datasets to identify these users. Next, we utilised Dataset 2, which records the amount of time users spent on various electronic devices, focusing specifically on the video games and TV columns.

We then used conditional statements to filter users by their shared IDs, separating them into variables for deprived and non-deprived users. We calculated the average time spent on video games and TV for both groups during the weekdays, weekend and the overall average of both. These averages allowed us to compare media consumption patterns between the two groups.

To present these comparisons visually, we created scatterplots that illustrate the differences in screen time between deprived and non-deprived users. Our findings suggest that screen time might influence behaviour and psychological symptoms. Specifically, users from non-deprived areas tended to have lower screen time, whereas those from deprived areas had higher screen time. Thus, we can assume that media consumption could play a role in relation to individuals’ poor behavioural choices and patterns. Things such as high unemployment, high crime rates etc. could potentially be helped via a lessening of overall screentime.

**Descriptive Statistics Investigation 2 – Manoj**

For the second descriptive statistics investigation, our goal was to explore the relationship between media consumption and psychological or behavioural attributes. Specifically, we focused on identifying outliers in media consumption and examining how these individuals felt. We utilised two out of the three datasets: Dataset 2, which tracks overall screen time across various electronic devices, and Dataset 3, which includes users' responses to statements rated on a scale from 1 to 5, where 1 indicates the lowest level of positive sentiment and 5 the highest.

Dataset 3 contained statements with positive connotations, which allowed us to use these responses to assess whether users displayed overall ‘negative feelings.’ To begin, we loaded the necessary libraries and datasets. We then developed a function to calculate the upper bound for outliers using the Interquartile Range (IQR). This function was applied to all relevant columns in Dataset 2, except for the ID column. Outliers in media consumption were identified and their unique IDs were recorded in a variable named `meeting\_ids`.

Next, we used these IDs to filter Dataset 3 and identify users who reported feeling scores of 3 or below, indicating negative feelings. We then analysed these filtered results to determine the frequency and median scores for the users in question.

To present our findings, we created scatterplots to visualize the distribution of scores for the behavioural attributes of the identified outliers. This visualization helps in understanding how extreme media consumption might be associated with negative feelings, offering insights into the potential psychological impact of high media use.

**Inferential Statistics Investigation 1 - Sasmit**

For the first inferential statistics investigation, the focus is on examining the relationship between the number of hours spent watching TV and the self-reported feeling of usefulness among the sample. This analysis utilises two datasets: Dataset 2, which tracks overall screen time across various electronic devices, and Dataset 3, which contains responses to statements rated on a scale from 1 to 5. Specifically, we concentrate on the feeling of usefulness, represented by the `Usef` column in Dataset 3.

The investigation begins by loading the necessary libraries and datasets. A new column, `T\_hours`, is then created in Dataset 2. This column combines the maximum hours spent watching TV from two columns `T\_we` (weekday TV hours) and `T\_wk` (weekend TV hours) to provide a consolidated measure of total TV viewing time. The dataset is filtered to include only those IDs with the highest TV watching times. These IDs are then merged with Dataset 3, which contains the corresponding `Usef` scores, using the shared ID column.

With the merged dataset, the correlation between TV watching hours and the feeling of usefulness is calculated. The Pearson correlation coefficient and the p-value of the correlation test are computed to determine the strength and significance of the relationship. The p-value helps assess whether the observed correlation is statistically significant, with a common significance level of 0.05 being used for this purpose.

In this analysis, the null hypothesis is tested. The results indicate a significant correlation, leading to the rejection of the null hypothesis. This suggests that there is a meaningful relationship between the amount of time spent watching TV and the feeling of usefulness.

**Inferential Statistics Investigation 2 - Adrian**

For the second inferential statistics investigation, we conducted a regression analysis to explore the relationship between TV viewing hours and various psychological attributes.

We began by importing the necessary libraries and datasets: dataset 2, which tracks overall screen time across various electronic devices, and dataset 3, which includes responses to statements rated on a scale of 1 to 5.

As in our previous analysis, we combined the weekday and weekend TV viewing hours from dataset 2 into a single column, `T\_hours`. We then merged this dataset with dataset 3 using the common ID column. Our focus was on individuals who reported scores of 3 or lower in three psychological attributes: `Usef` (usefulness), `Thcklr` (thinking clearly), and `Dealpr` (dealing with problems).

In the regression analysis, `T\_hours` was used as the dependent variable, while the three psychological attributes served as independent variables. The aim was to assess how well these attributes predicted total hours spent watching TV.

The results indicated that the model's R-squared and Adjusted R-squared values were extremely low, suggesting that the psychological attributes included have minimal explanatory power for the variation in TV viewing hours.

Specifically, `Usef` and `Thcklr` had negative coefficients with very low p-values, indicating a significant negative relationship with TV viewing hours. This suggests that as scores for feeling useful and thinking clearly decrease, TV viewing hours increase.

On the other hand, `Dealpr` had a coefficient very close to zero and a higher p-value in comparison, indicating that it does not significantly influence TV viewing hours. This result implies that the ability to deal with problems is not meaningfully related to the amount of time spent watching TV.

Overall, while the psychological attributes of usefulness and clarity of thought show some relationship with TV viewing, their practical significance is limited by the low predictive power of the model.