MATH 530-630: Final Project

Momma Bear, Papa Bear, Baby Bear

12/4/2017

Paper:

Przybylski, A. K.. & Weinstein, N. "A Large-Scale Test of the Goldilocks Hypothesis." *Psychological Science* 28, 204-215 (2017)

Paper Summary:

In this paper, Pryzbylski and Weinstein explore the potential relationship between digital-screen usage and mental well-being scores, with the goal of being able to empirically quantify what would be considered just the right amount of digital-screen usage - hence, the goldilocks hypothesis.

Their data set consisted of 120,115 English adolescents who self-reported the criterion variable, mental well-being score, from the Warwick-Edinburgh Mental Well-Being Scale. They further divided their explanatory variable, digital-screen time, into four subtypes: time spent using computers, watching tv and movies, playing videogames, and using smart phones. They additionally separated their data into whether the engagement took place on the weekend or the weekday, and examined the potential effects of this difference in time of engagment during their analysis steps.

They chose their control and confounding variables to be gender, ethnicity, and economic factors, as these have all been linked to both mental well-being scores and digital-screen time engagement by past research. For ethnicity, they assessed whether their ethnic background corresponded with a minority status. For economic factors, they included whether or not the participant lived in what was considered to be a deprived local-authority district.

Their results contained evidence that there is a quadractic relationship between mental well-being scores and digital-screen time and that this relationship further varies if the digital-screen time is spent on the weekend or weekday. They concluded by remarking that digital-screen time can be harmless and even beneficial, but only when engaged with in moderation.

Exploratory Data Analysis Report

Discussion of issues unconvered in data quality review:

During out data quality review we noted multiple issues with the given data set:

1. **Values are missing**: On our initial look through of the data, we noticed the presence of a lot of "NA" values. We were not able to find any explicit explanation of these missing values in the materials associated with the paper, but an example of the survey given to the participants to elicit the data points allowed skipping of a few questions. This could be what caused the "NA" values, but nothing is officialy noted. There was also a lack of explanation about what they did with these missing values when analyzing the data. After trying several things, our replicated analyses that excluded these values appeared to match those from the paper.
2. **Units are not specified**: Further exploration of the data suggested that the units used were hours.
3. **Field names are ambiguous**: We were able to figure out what the field names meant with further after a closer look at the data.

4. **Text has been converted to numbers**: The variable in which this occured was not used during our analysis process, but this could be an issue if further analysis of the data was done.

5. **Data were entered by humans**: We were concerened about the units not being specified and wondered about inconsistencies during the data entry process. We were unable to explore this concern further with the information that we were given.

6. **Aggregations were computed on missing values**: Our omission of the missing values during our data exploration and analysis would resolve this issue.

7. **Results have been p-hacked**: The choices of analyses seemed to be reasonable and not indicative of p-hacking.

Descriptive statistics and plots presented in paper:

Importing, cleaning, and tidying the data:

```
source("01-initialization.R") #import and tidy goldilocks data
```

Figure 1:

This figure shows the average mental well being score as a function of digital-screen time spent watching TV and movies, playing video games, using computers, and using smartphones with the separation of time spent during the weekend or the weekday. We were only able to replicate their graphs by first omitting the "NA" values from our tidy dataset, and then working with the result data set. Using the dplyr package in R, we grouped the tidy dataset without "NA"s by time (2 levels), and pred(4 levels), with pred denoting the type of activity. We used the summarise function to find the mean of the engagement hours and calculated the error values for a 95% confidence interval. We also chose to create a smaller summary data frame for this plot, to avoid repeatedly plotting the same points during graphing.

Using ggplot from the ggplot2 package we plotted engagement time vs. mental well-being score grouping the data by time, and used the facet wrap command to create subplots for each activity type. The error bars are also included. From this plot, it is clearly visible that there is no negative monotonic relationship between the mental well-being score and the engagment time.

```
fig_1_summary <- gold_no_na %>%
  group_by(pred, engagement, time) %>%
  summarise(avg_mwbi = mean(mwbi),
            error = qt(0.975, df = n()-1)*sd(mwbi)/sqrt(n()))
levels(fig_1_summary$time) <- c("Weekday", "Weekend")
levels(fig_1_summary$pred) <- c("Using Computers", "Gaming", "Using Smartphones", "Watching")
fig_1_summary$pred <- factor(fig_1_summary$pred,
                        levels = c("Watching", "Gaming", "Using Computers", "Using Smartphones"))
ggplot(fig_1_summary, aes(x = engagement,
                          y = avg_mwbi,
                          group = time,
                          color = time)) +
  geom_point() +
  geom_line() +
  ylim(40, 50) +
  xlab("Daily Digital-Screen Engagement (hr)") +
  ylab("Average Mental Well-Being") +
  facet_wrap(~pred) +
  geom_errorbar(aes(ymin = avg_mwbi-error,
                    ymax = avg_mwbi+error),
                    width = .25)
```

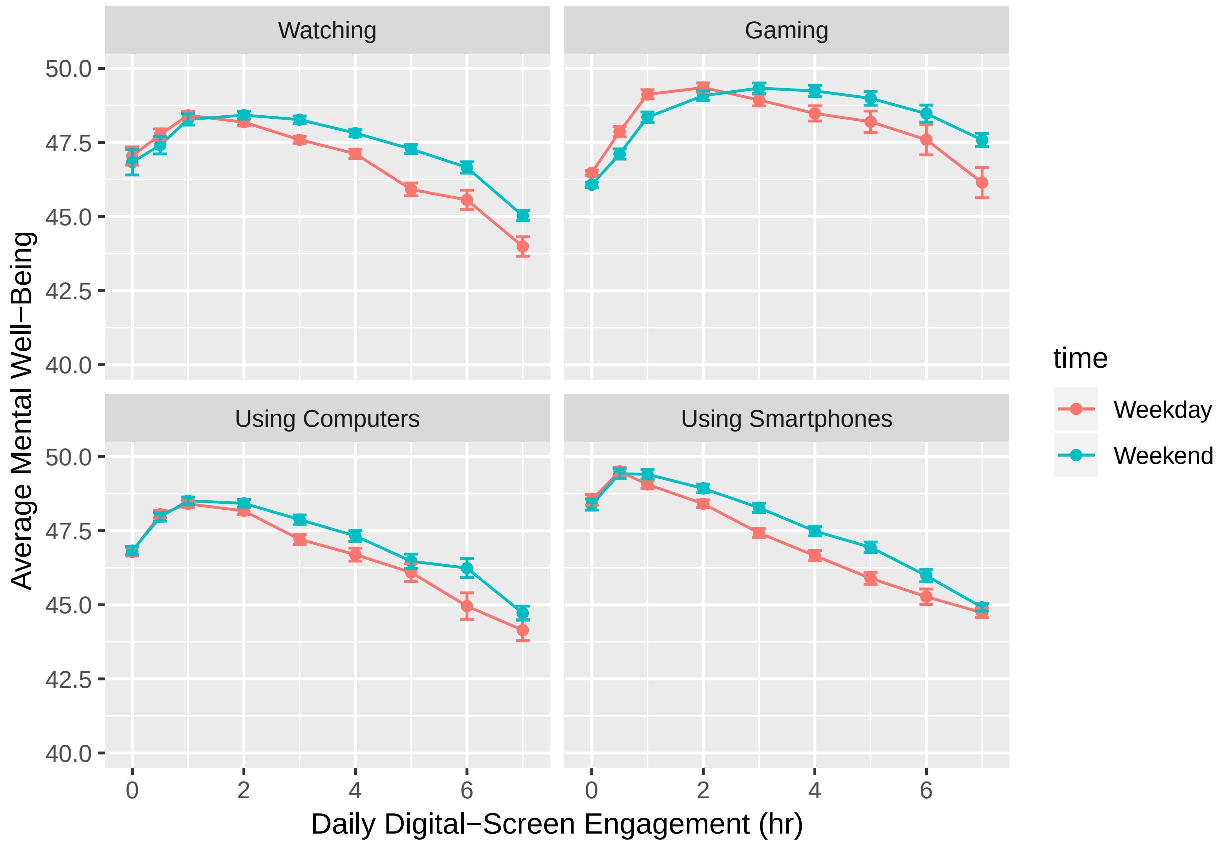


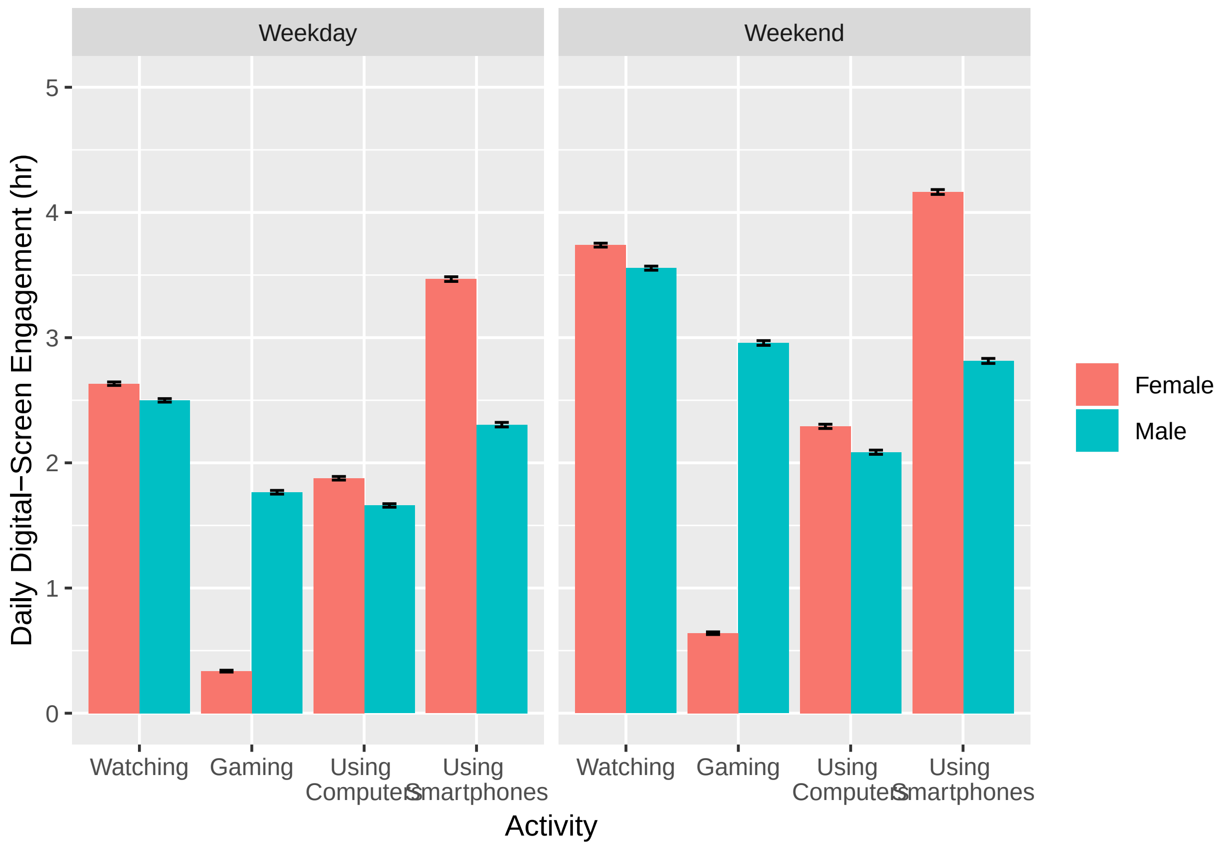
Figure 2:

This figure illustrates the relationship between the average digital-screen time and different types of digital activities included in the data and shows the difference in engagement levels for females versus males. Two separate histograms are created for time spent during the weekend or weekday. As with Figure 1, we had to omit the "NA" values in order to replicate the figure from the paper, and again created a smaller summary dataframe to prevent repeated ploted. Using the dplyr package, we grouped by gender, time (weekend or weekday), and pred(type of activity), and calculated the average engagement amount and error bars for a 95% confidence interval.

From these figures, we can see that female adolescents devoted more time to watching TV and movies, using smartphones, and using computers, during both the weekend and weekday. We can also see that boys adolscents spent more time playing video games than female adolescents reported.

```
fig_2_summary <- gold_no_na %>%
  group_by(male, pred, time) %>%
  summarise(avg_engagement = mean(engagement),
            error = qt(0.975, df = n()-1)*sd(engagement)/sqrt(n()))
fig_2_summary$male <- as.factor(fig_2_summary$male)
levels(fig_2_summary$male) <- c("Female", "Male")
levels(fig_2_summary$time) <- c("Weekday", "Weekend")
levels(fig_2_summary$pred) <- c("Using \nComputers", "Gaming", "Using \nSmartphones", "Watching")
fig_2_summary$pred <- factor(fig_2_summary$pred,
                             levels = c("Watching", "Gaming", "Using \nComputers", "Using \nSmartphones"))
ggplot(fig_2_summary, aes(x = pred,
                          y = avg_engagement,
                          group = male,
                          fill = as.factor(male))) +
```

```
  geom_bar(stat = "identity", position = "dodge") +
  facet_wrap(~time) +
  ylab("Daily Digital-Screen Engagement (hr)") +
  xlab("Activity") +
  ylim(0, 5) +
  guides(fill=guide_legend(title=NULL)) +
  geom_errorbar(aes(ymin = avg_engagement-error,
                    ymax = avg_engagement+error),
                    position = position_dodge(0.9),
                    width = .25)
```



Below is our t-test analysis replication for all four activities - watching, playing, using computers, using smartphones - on both weekday and weekend comparing male and female.

From the values from the two sample t-test, we can clearly see girls devoted spending more time using smart phones, watching TV, and using computers. Boys spent more time playing console/video games during both weekdays and weekends (for all, p-value < 2.2e-16)

```
#Watching TV programs during weekdays for two groups, male and female
males_watch_we<-filter(gold_no_na, male==1 & pred=="watch" & time=="we")
females_watch_we<-filter(gold_no_na, male==0 & pred=="watch" & time=="we")
t.test(males_watch_we$engagement,females_watch_we$engagement, var.equal= TRUE)
```

```
##
##  Two Sample t-test
##
## data:  males_watch_we$engagement and females_watch_we$engagement
```

```
## t = -16.028, df = 115770, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.2065345 -0.1615273
## sample estimates:
## mean of x mean of y
##  3.555051  3.739082
```

```
#Watching TV programs during weekends for two groups, male and female
males_watch_wd<-filter(gold_no_na, male==1 & pred=="watch" & time=="wd")
females_watch_wd<-filter(gold_no_na, male==0 & pred=="watch" & time=="wd")
t.test(males_watch_wd$engagement,females_watch_wd$engagement, var.equal= TRUE)
```

```
##
##  Two Sample t-test
##
## data:  males_watch_wd$engagement and females_watch_wd$engagement
## t = -13.382, df = 116280, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.1527602 -0.1137294
## sample estimates:
## mean of x mean of y
##  2.499090  2.632335
```

```
#Playing console games during weekdays for two groups, male and female
males_play_we<-filter(gold_no_na, male==1 & pred=="play"&time=="we")
females_play_we<-filter(gold_no_na, male==0 & pred=="play"& time=="we")
t.test(males_play_we$engagement,females_play_we$engagement, var.equal= TRUE)
```

```
##
##  Two Sample t-test
##
## data:  males_play_we$engagement and females_play_we$engagement
## t = 215.02, df = 116280, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  2.297216 2.339480
## sample estimates:
## mean of x mean of y
## 2.9576086 0.6392608
```

```
#Playing console games during weekends for two groups, male and female
males_play_wd<-filter(gold_no_na, male==1 & pred=="play"&time=="wd")
females_play_wd<-filter(gold_no_na, male==0 & pred=="play"& time=="wd")
t.test(males_play_wd$engagement,females_play_wd$engagement, var.equal= TRUE)
```

```
##
##  Two Sample t-test
##
## data:  males_play_wd$engagement and females_play_wd$engagement
## t = 175.8, df = 116600, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.412867 1.444726
## sample estimates:
```

```
## mean of x mean of y
##  1.765018  0.336221
```

```r
#Using Computers during weekdays for two groups, male and female
males_comp_we<-filter(gold_no_na, male==1 & pred=="comp"&time=="we")
females_comp_we<-filter(gold_no_na, male==0 & pred=="comp"& time=="we")
t.test(males_comp_we$engagement,females_comp_we$engagement, var.equal= TRUE)
```

```
##
##  Two Sample t-test
##
## data:  males_comp_we$engagement and females_comp_we$engagement
## t = -16.903, df = 116040, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.2302177 -0.1823756
## sample estimates:
## mean of x mean of y
##  2.085074  2.291371
```

```r
#Using Computers during weekends for two groups male and female

males_comp_wd<-filter(gold_no_na, male==1 & pred=="comp"&time=="wd")
females_comp_wd<-filter(gold_no_na, male==0 & pred=="comp"& time=="wd")
t.test(males_comp_wd$engagement,females_comp_wd$engagement, var.equal= TRUE)
```

```
##
##  Two Sample t-test
##
## data:  males_comp_wd$engagement and females_comp_wd$engagement
## t = -21.364, df = 116450, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.2371201 -0.1972677
## sample estimates:
## mean of x mean of y
##  1.659605  1.876799
```

```r
#Using Smartphones during weekdays for two groups male and female
males_sp_we<-filter(gold_no_na, male==1 & pred=="sp" & time=="we")
females_sp_we<-filter(gold_no_na, male==0 & pred=="sp"& time=="we")
t.test(males_sp_we$engagement,females_sp_we$engagement, var.equal= TRUE)
```

```
##
##  Two Sample t-test
##
## data:  males_sp_we$engagement and females_sp_we$engagement
## t = -95.699, df = 116450, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -1.377119 -1.321842
## sample estimates:
## mean of x mean of y
##  2.814039  4.163520
```

```r
#Using Smartphones during weekends for two groups male and female
males_sp_wd<-filter(gold_no_na, male==1 & pred=="sp" & time=="wd")
```

```
females_sp_wd<-filter(gold_no_na, male==0 & pred=="sp"& time=="wd")
t.test(males_sp_wd$engagement,females_sp_wd$engagement, var.equal= TRUE)
```

```
##
##  Two Sample t-test
##
## data:  males_sp_wd$engagement and females_sp_wd$engagement
## t = -87.56, df = 116630, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -1.188876 -1.136817
## sample estimates:
## mean of x mean of y
##  2.304908  3.467754
```

Additional comments:

- They reported the number of total particpants in the final data set to be 120,115, which matches the size of our tidy data set. Of the 120,115 participants, 62,962 of them are female while 57,153 are male.

```
raw_gold %>%
  tabyl(male)%>%
  adorn_totals("row") %>%
  kable()
```

male	n	percent
0	62962	0.524181
1	57153	0.475819
Total	120115	1.000000

- Though we were able to replicated the descriptive statistics, and plots in the paper by excluded all "NA" values, we had concern that this choice could impact our later analyses replication steps. The lack of documentation about how the data set was constructed and then processed is also concerning and could present challenges later on.
- Though they reported their sample size to be 120,115, since we ended up removing all "NA" values for the analysis section, the amount of data we were working with changed and it also became uneven when divided by type of activity and time of week. The table below shows the size of the data set for each pairing of variables. This mismatch in length could create problems later on, as it already has with our attempt at replicating their pair t-test, and could limit the types of analysis that can be performed.

```
gold_no_na %>%
  group_by(pred, time) %>%
  summarise(n = n()) %>%
  kable()
```

pred	time	n
comp	wd	116452
comp	we	116046
play	wd	116599
play	we	116279
sp	wd	116630
sp	we	116454
watch	wd	116286
watch	we	115772

Replication Report:

Table 1:

"Results of Models Linking Mental Well-Being to Daily Digital Screen Engagement Without Adjustments for the Control Variables"
In this table, linear and quadratic regression models were run to further examine the relationship between mental well-being scores and weekday and weekend digital-screen usage without the inclusion of the control variables. For each model, they have reported the Estimated Coeffecient as "b" , Standard Error of the Coefficient Estimate as"SE", 95% confidence interval as "95% CI", Variable p-valuea as "p", and Cohen's d as "d".

To replicate this section, we used "dplyr" package to filter the data based on Daily digital screen engagement activities and the time they spent during weekday or weekend. After filtering we checked for the "mean" of each subset and compared it with values represented in the Graph 2 and they were matched. Then we fitted both linear and quadratic models with 95% confidence intervals. The values for:(b, SE, P) are directly reported from the regression models. The "95% CI" is obtained by using "confit" command. For Cohen's d, because they have not mention what variables are used to calculate the effect size, we explored it based on different variables. Since the effect size is different for each activity at diffrent time of week, and based on the results we got by exploring diffrenet variables, our assumption is that the values are calculated for male and female Daily digital screen engagement time during weekday and weekend separately. We initially calculated cohen's d (M1-M1/ Pooled SE) but later after talking with the instructor we used "tes" function from "compute.es" package to convert t-test values to effect size to replicate this section.

Observations:
The Standard Error values are much lower for quadratic models and it suggests the non linear relationship between the Mental well-being and Daily digital screen engagement time.
While our values for our quadratic regression model matched those in the paper relatively closely, we were unable to get the values for our linear regression model to match even after attempting it from multiple different approaches. When taking a closer look at their results, we realized that all of the slope values that they calculated for the linear model were positive, but from figure 1 all four graphs appear to have a negative slope. We think that perhaps they excluded the data from certain participants before performing the linear regression model, but nothing about this is mentioned in the paper. Since we removed the "NA" values from our data set, we chose to include the degrees of freedom for each model.

```
list_pred = c("watch", "play", "comp", "sp")
list_time = c("wd", "we")
list_pred_name = c("Watching films, TV programs, etc.",
                   "Playing games", "Using computers for Internet, e-mail,etc.",
                   "Using smartphones for socialnetworking, chatting, etc.")
i = 1
for(pred_var in list_pred){
  tbl_1 = matrix(c("", "", "", "", "", "", ""), ncol = 7, byrow = TRUE)
  colnames(tbl_1) = c("", "b", "SE", "95% CI", "p", "|d|", "df")
  for (time_var in list_time){
    #Linear Regression
    gold_subset <- filter(gold_no_na, time == time_var & pred == pred_var)
    lm_model <- lm(mwbi ~ engagement, data = gold_subset)
    tidy_lm <- tidy(lm_model)
    slope_lm <- round(tidy_lm$estimate[2], digits = 3)
    std_err_lm <- round(tidy_lm$std.error[2], digits = 3)
    conf_low_lm <- round(confint(lm_model, level = 0.95)[2], digits = 3)
    conf_up_lm <- round(confint(lm_model, level = 0.95)[4], digits = 3)
    conf_int_lm <- paste0("[", conf_low_lm, ", ", conf_up_lm, "]")
    p_val_lm <- round(tidy_lm$p.value[2], digits = 3)
```

```r
    df_lm <- df.residual(lm_model)
    male_lm<- filter(gold_subset, male==1)
    female_lm<- filter(gold_subset, male==0)
    tval<-t.test(male_lm$engagement,female_lm$engagement)
    tess <- tes(tval$statistic,
                as.double(n_distinct(male_lm)),
                as.double(n_distinct(female_lm)),
                level = 95, cer = 0.2, dig = 2,
                verbose = FALSE, id=NULL, data=NULL)
    cohens_lm = tess[4]
    cohens_lm = abs(cohens_lm$d)
    cohens_lm = round(cohens_lm, digits = 3)
    linear_values <- c("Linear", slope_lm, std_err_lm,
                       conf_int_lm, p_val_lm, cohens_lm, df_lm)
    tbl_1 <- rbind(tbl_1, linear_values)


    #Quadratic Regression
    gold_subset <- gold_subset %>%
      mutate(engagement2 = engagement^2)
    qd_model <- lm(mwbi ~ engagement + engagement2, data = gold_subset)
    tidy_qd <- tidy(qd_model)
    slope_qd <- round(tidy_qd$estimate[3], digits = 3)
    std_err_qd <- round(tidy_qd$std.error[3], digits = 3)
    conf_low_qd <- round(confint(qd_model, level = 0.95)[3], digits = 3)
    conf_up_qd <- round(confint(qd_model, level = 0.95)[6], digits = 3)
    conf_int_qd <- paste0("[", conf_low_qd, ", ", conf_up_qd, "]")
    p_val_qd <- round(tidy_qd$p.value[3], digits = 3)
    df_qd <- df.residual(qd_model)
    male_lm <- gold_subset %>%
      filter(male ==1) %>%
      mutate(engagement_2 = engagement^2)
    female_lm <- gold_subset %>%
      filter(male == 0) %>%
      mutate(engagement_2 = engagement^2)
    tval<-t.test(male_lm$engagement_2,female_lm$engagement_2)
    tess<-tes(tval$statistic, as.double(n_distinct(male_lm)),
              as.double(n_distinct(female_lm)),
              level = 95, cer = 0.2, dig = 2,
              verbose = FALSE, id=NULL, data=NULL)
    cohens_qd = tess[4]
    cohens_qd= abs(cohens_qd$d)
    quadratic_values <- c("Quadratic", slope_qd, std_err_qd,
                          conf_int_qd, p_val_qd, cohens_qd, df_qd)
    tbl_1 <- rbind(tbl_1, quadratic_values, c("", "", "", "", "", "", ""))
  }
  tbl_1 <- tbl_1[-7,]
  rownames(tbl_1) = c("Weekday", "", "", "Weekend", "", "")
  print(kable(tbl_1, caption = list_pred_name[i]))
  i = i + 1
}
```

```
##
##
```

```
## Table: Watching films, TV programs, etc.
##
##                     b        SE      95% CI              p     |d|     df
## --------  ----------  -------  ------  -----------------  ---  -----  -------
## Weekday
##           Linear      -0.497   0.016   [-0.529, -0.465]   0    0.08   116284
##           Quadratic   -0.138   0.008   [-0.153, -0.123]   0    0.08   116283
## Weekend
##           Linear      -0.42    0.014   [-0.448, -0.392]   0    0.09   115770
##           Quadratic   -0.159   0.007   [-0.173, -0.146]   0    0.09   115769
##
##
## Table: Playing games
##
##                     b        SE      95% CI              p     |d|     df
## --------  ----------  -------  ------  -----------------  ---  -----  -------
## Weekday
##           Linear      0.403    0.018   [0.368, 0.438]     0    1      116597
##           Quadratic   -0.295   0.009   [-0.312, -0.277]   0    0.66   116596
## Weekend
##           Linear      0.405    0.013   [0.38, 0.43]       0    1.23   116277
##           Quadratic   -0.232   0.007   [-0.245, -0.219]   0    0.93   116276
##
##
## Table: Using computers for Internet, e-mail,etc.
##
##                     b        SE      95% CI              p     |d|     df
## --------  ----------  -------  ------  -----------------  ---  -----  -------
## Weekday
##           Linear      -0.378   0.016   [-0.409, -0.347]   0    0.13   116450
##           Quadratic   -0.161   0.007   [-0.176, -0.146]   0    0.11   116449
## Weekend
##           Linear      -0.304   0.013   [-0.33, -0.278]    0    0.1    116044
##           Quadratic   -0.167   0.006   [-0.18, -0.155]    0    0.1    116043
##
##
## Table: Using smartphones for socialnetworking, chatting, etc.
##
##                     b        SE      95% CI              p        |d|     df
## --------  ----------  -------  ------  -----------------  ------  -----  -------
## Weekday
##           Linear      -0.683   0.012   [-0.706, -0.66]    0       0.52   116628
##           Quadratic   -0.021   0.006   [-0.032, -0.009]   0.001   0.45   116627
## Weekend
##           Linear      -0.633   0.011   [-0.655, -0.611]   0       0.56   116452
##           Quadratic   -0.092   0.006   [-0.104, -0.081]   0       0.52   116451
```

Here is our t-test analysis replication for engagement time versus time of the week (weekend or weekday). From the t-test analysis between the digital screen time and tiem (weekday and weekend), we can see that for all four activities, screen time was longer on weekends than during the weekday.

```r
#watching TV programs during weekdays and weekends
weekday_watch<-filter(gold_no_na, pred=="watch" & time=="wd")
weekend_watch<-filter(gold_no_na, pred=="watch" & time=="we")
t.test(weekday_watch$engagement,weekend_watch$engagement, var.equal= TRUE)
```

```
##
##  Two Sample t-test
##
## data:  weekday_watch$engagement and weekend_watch$engagement
## t = -142.54, df = 232060, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -1.097552 -1.067779
## sample estimates:
## mean of x mean of y
##  2.568766  3.651431
```

```
#playing console games during weekdays and weekends
weekday_play<-filter(gold_no_na,  pred=="play"&time=="wd")
weekend_play<-filter(gold_no_na,  pred=="play"& time=="we")
t.test(weekday_play$engagement,weekend_play$engagement, var.equal= TRUE)
```

```
##
##  Two Sample t-test
##
## data:  weekday_play$engagement and weekend_play$engagement
## t = -92.874, df = 232880, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.7426538 -0.7119564
## sample estimates:
## mean of x mean of y
##  1.018165  1.745470
```

```
#Using Computers during weekdays for two groups male and female
weekday_comp<-filter(gold_no_na,  pred=="comp"&time=="wd")
weekend_comp<-filter(gold_no_na,  pred=="comp"& time=="we")
t.test(weekday_comp$engagement,weekend_comp$engagement, var.equal= TRUE)
```

```
##
##  Two Sample t-test
##
## data:  weekday_comp$engagement and weekend_comp$engagement
## t = -52.858, df = 232500, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.4354664 -0.4043271
## sample estimates:
## mean of x mean of y
##  1.773001  2.192898
```

```
#Using Smartphones during weekdays for two groups male and female
weekday_sp<-filter(gold_no_na,   pred=="sp" & time=="wd")
weekend_sp<-filter(gold_no_na, pred=="sp"& time=="we")
t.test(weekday_sp$engagement,weekend_sp$engagement, var.equal= TRUE)
```

```
##
##  Two Sample t-test
##
## data:  weekday_sp$engagement and weekend_sp$engagement
## t = -60.634, df = 233080, p-value < 2.2e-16
```

```
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.6271750 -0.5878983
## sample estimates:
## mean of x mean of y
##  2.912334  3.519871
```

Table 2:

"Results of Models Linking Mental Well-Being to Daily Digital-Screen Engagement With Adjustments for the Control Variables"

This table also examines the relationship between mental well-being scores and digital-screen usage on both the weekend and weekday, but the control variables - gender (male), economic factors (deprivity), and technology access (minority) - were included when calculating both linear and quadratic regression models.

```r
i=1

for (pred_var in list_pred){
  tbl_2 = matrix(c("", "", "", "", "", "", ""), ncol = 7, byrow = TRUE)
  colnames(tbl_2) = c("", "b", "SE", "95% CI", "p", "|d|", "df")
  for (time_var in list_time){
    #linear regression
    gold_subset <- filter(gold_no_na, time == time_var & pred == pred_var)
    lm_model <- lm(mwbi ~ engagement+male+minority+deprived, data = gold_subset)
    tidy_lm <- tidy(lm_model)
    slope_lm <- round(tidy_lm$estimate[2], digits = 3)
    std_err_lm <- round(tidy_lm$std.error[2], digits = 3)
    conf_low_lm <- round(confint(lm_model, level = 0.95)[2], digits = 3)
    conf_up_lm <- round(confint(lm_model, level = 0.95)[4], digits = 3)
    conf_int_lm <- paste0("[", conf_low_lm, ", ", conf_up_lm, "]")
    p_val_lm <- round(tidy_lm$p.value[2], digits = 3)
    df_lm <- df.residual(lm_model)
    df_lm <- df.residual(lm_model)
    male_lm<- filter(gold_subset, male==1)
    female_lm<- filter(gold_subset, male==0)
    tval<-t.test(male_lm$engagement,female_lm$engagement)
    tess <- tes(tval$statistic, as.double(n_distinct(male_lm)),
                as.double(n_distinct(female_lm)),
                level = 95, cer = 0.2, dig = 2,
                verbose = FALSE, id=NULL, data=NULL)
    cohens_lm = tess[4]
    cohens_lm = abs(cohens_lm$d)
    cohens_lm = round(cohens_lm, digits = 3)
    linear_values <- c("Linear", slope_lm, std_err_lm,
                       conf_int_lm, p_val_lm, cohens_lm, df_lm)
    tbl_2 <- rbind(tbl_2, linear_values)

    #Quadratic Regression
    gold_subset <- gold_subset %>%
      mutate(engagement2 = engagement^2,
             male2 = male^2,
             minority2 = minority^2,
             deprived2 = deprived^2)
    qd_model <- lm(mwbi ~ engagement+engagement2 +
                     male+male2 + minority+minority2 +
```

```r
                      deprived+deprived2, data = gold_subset)
    tidy_qd <- tidy(qd_model)
    slope_qd <- round(tidy_qd$estimate[3], digits = 3)
    std_err_qd <- round(tidy_qd$std.error[3], digits = 3)
    conf_low_qd <- round(confint(qd_model, level = 0.95)[3], digits = 3)
    conf_up_qd <- round(confint(qd_model, level = 0.95)[6], digits = 3)
    conf_int_qd <- paste0("[", conf_low_qd, ", ", conf_up_qd, "]")
    p_val_qd <- round(tidy_qd$p.value[3], digits = 3)
    df_qd <- df.residual(qd_model)
    male_lm <- gold_subset %>%
      filter(male ==1) %>%
      mutate(engagement_2 = engagement^2)
    female_lm <- gold_subset %>%
      filter(male == 0) %>%
      mutate(engagement_2 = engagement^2)
    tval<-t.test(male_lm$engagement_2,female_lm$engagement_2)
    tess<-tes(tval$statistic, as.double(n_distinct(male_lm)),
              as.double(n_distinct(female_lm)),
              level = 95, cer = 0.2, dig = 2,
              verbose = FALSE, id=NULL, data=NULL)
    cohens_qd = tess[4]
    cohens_qd= abs(cohens_qd$d)
    quadratic_values <- c("Quadratic", slope_qd, std_err_qd,
                          conf_int_qd, p_val_qd, cohens_qd, df_qd)
    tbl_2 <- rbind(tbl_2, quadratic_values, c("", "", "", "", "", "", ""))
  }
  tbl_2 <- tbl_2[-7,]
  rownames(tbl_2) = c("Weekday", "", "", "Weekend", "", "")
  print(kable(tbl_2, caption = list_pred_name[i]))
  i = i + 1
}
```

```
##
##
## Table: Watching films, TV programs, etc.
##
##                          b         SE       95% CI             p     |d|     df
## --------  ----------  -------  ------  ----------------  ---  -----  -------
## Weekday
##            Linear      -0.425   0.016   [-0.456, 0.232]   0     0.08    116281
##            Quadratic   -0.127   0.008   [-0.142, 0.252]   0     0.08    116280
## Weekend
##            Linear      -0.352   0.014   [-0.379, 0.316]   0     0.09    115767
##            Quadratic   -0.163   0.007   [-0.176, 0.298]   0     0.09    115766
##
##
## Table: Playing games
##
##                          b         SE       95% CI             p     |d|     df
## --------  ----------  -------  ------  ----------------  ---  -----  -------
## Weekday
##            Linear      -0.411   0.02    [-0.449, 0.102]   0     1       116594
##            Quadratic   -0.055   0.009   [-0.073, 0.124]   0     0.66    116593
## Weekend
```

```
##           Linear      -0.288   0.015   [-0.317, 0.166]   0    1.23   116274
##           Quadratic   -0.069   0.007   [-0.082, 0.159]   0    0.93   116273
##
##
## Table: Using computers for Internet, e-mail,etc.
##
##                       b        SE      95% CI            p     |d|     df
## --------  ----------  -------  ------  ----------------  ---  -----  -------
## Weekday
##           Linear      -0.287   0.015   [-0.317, 0.333]   0    0.13   116447
##           Quadratic   -0.158   0.007   [-0.173, 0.254]   0    0.11   116446
## Weekend
##           Linear      -0.245   0.013   [-0.27, 0.32]     0    0.1    116041
##           Quadratic   -0.16    0.006   [-0.172, 0.215]   0    0.1    116040
##
##
## Table: Using smartphones for socialnetworking, chatting, etc.
##
##                       b        SE      95% CI            p     |d|     df
## --------  ----------  -------  ------  ----------------  ---  -----  -------
## Weekday
##           Linear      -0.432   0.012   [-0.455, 0.177]   0    0.52   116625
##           Quadratic   -0.058   0.006   [-0.069, 0.173]   0    0.45   116624
## Weekend
##           Linear      -0.379   0.011   [-0.401, 0.204]   0    0.56   116449
##           Quadratic   -0.109   0.006   [-0.12, 0.175]    0    0.52   116448
```

Two way ANOVA for Table 2:
Factors: Pred (4 levels) and time (2 levels)

First let's see how the mean response changes based on the two main effects:

```
aggregate(engagement ~ time, data = gold_no_na, mean)
```

```
##   time engagement
## 1   wd   2.067879
## 2   we   2.777036
```

```
aggregate(engagement ~ pred, data = gold_no_na, mean)
```

```
##    pred engagement
## 1  comp   1.982583
## 2  play   1.381318
## 3    sp   3.215873
## 4 watch   3.108899
```

```
with(gold_no_na, tapply(engagement, list(time, pred), mean))
```

```
##        comp     play       sp    watch
## wd 1.773001 1.018165 2.912334 2.568766
## we 2.192898 1.745470 3.519871 3.651431
```

By looking at the means of the individual factors, we can see that neither of them has any effect on the response variable. But by looking at the interaction, we can see that weekend screen times were higher compared to the weekday screen times spent by the adolescents.

For visualizing interactions we can use interaction.plot. It basically plots the means we just examined and connects them with lines. The first argument, is the variable we want on the x-axis. The second variable, is how we want to group the lines it draws. The third argument, response, is our response variable. From the resulting plot, we can tell the slopes are not parallel at the end.

A common method for analyzing the effect of categorical variables on a continuous response variable is the Analysis of Variance, or ANOVA. We can fit a linear model to these data using the lm function. By looking at the table of coefficients, we can tell everything is significant. This just means the coefficients are significantly different from 0.

Running an ANOVA on these data reveals a significant interaction, as we expected, but we also notice the main effects are significant as well. This just means the effects of time and pred individually explain a fair amount of variability in the data.

From the ANOVA results, we can conclude the following, based on the p-values and a significance level of 0.05:

- The p-value of pred is < 2e-16 (significant), which indicates that the levels of pred are associated with significant different engagement time.
- The p-value of time is < 2e-16 (significant), which indicates that the levels of time are associated with significant different engagement time.
- The p-value for the interaction between pred*time is < 2e-16 (significant), which indicates that the relationships between pred and engagement time depends on the time method.

In ANOVA test, a significant p-value indicates that some of the group means are different, but we don't know which pairs of groups are different.

We performed multiple pairwise-comparison, to determine if the mean difference between specific pairs of group are statistically significant.

Since we have time (2 levels), we can just do it with the pred (4 levels) variable. We used the general linear hypothesis test to compare the means between the groups. From the result, we can clearly see the difference between the pred - time pairs were all significant (p < 0.001).

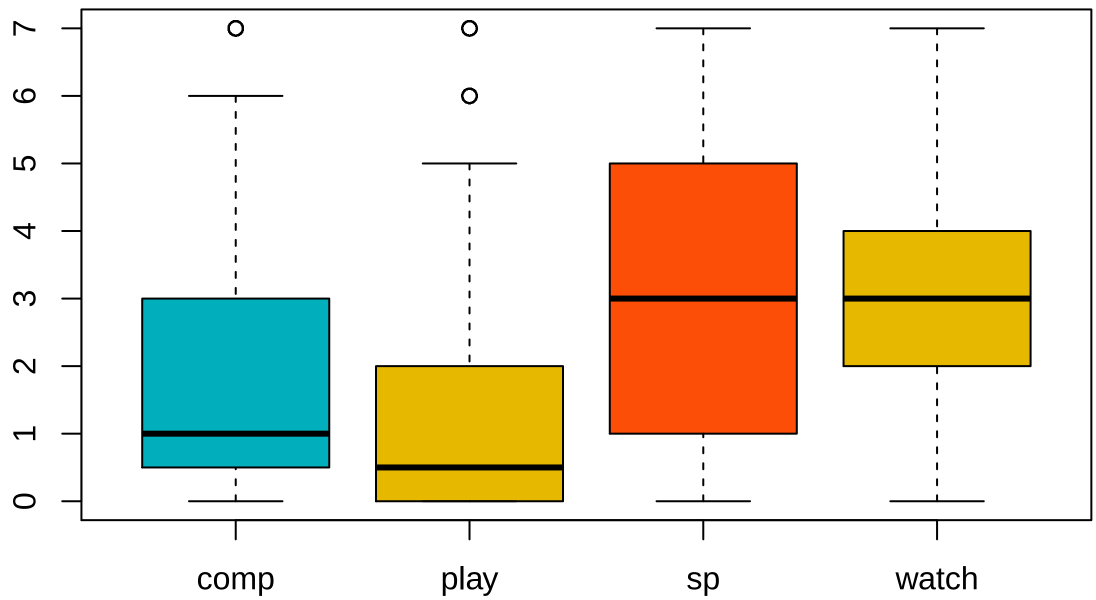
```
group_by(gold_no_na, pred) %>%
  summarise(count = n(),
            mean = mean(engagement, na.rm = TRUE),
            sd = sd(engagement, na.rm = TRUE))
```

```
## # A tibble: 4 x 4
##   pred   count  mean    sd
##   <fct>  <int> <dbl> <dbl>
## 1 comp  232498  1.98  1.93
## 2 play  232878  1.38  1.92
## 3 sp    233084  3.22  2.44
## 4 watch 232058  3.11  1.91
```

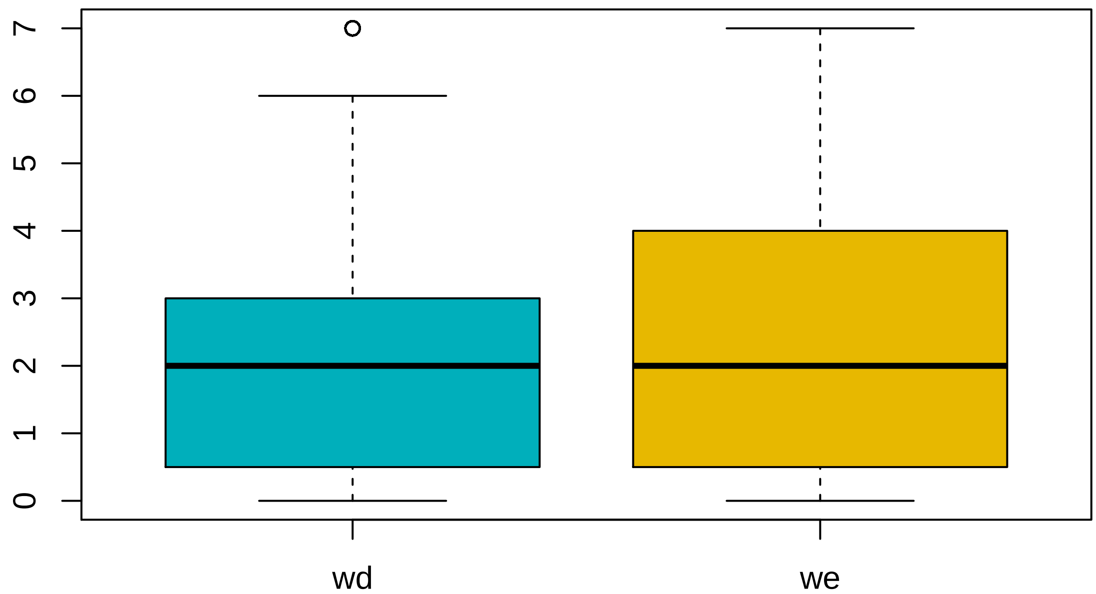
```
group_by(gold_no_na, time) %>%
  summarise(count = n(),
            mean = mean(engagement, na.rm = TRUE),
            sd = sd(engagement, na.rm = TRUE))
```

```
## # A tibble: 2 x 4
##   time   count  mean    sd
##   <fct>  <int> <dbl> <dbl>
## 1 wd    465967  2.07  2.00
## 2 we    464551  2.78  2.34
```

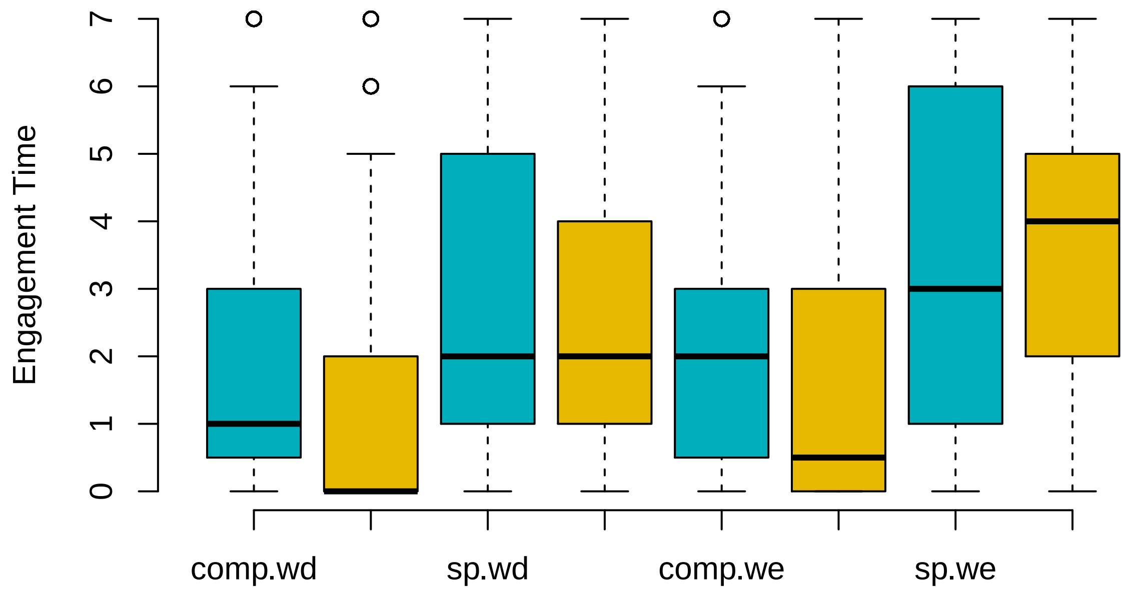
```
boxplot(engagement ~ pred,data=gold_no_na,
        col=c("#00AFBB", "#E7B800", "#FC4E07","#E7B800"))
```



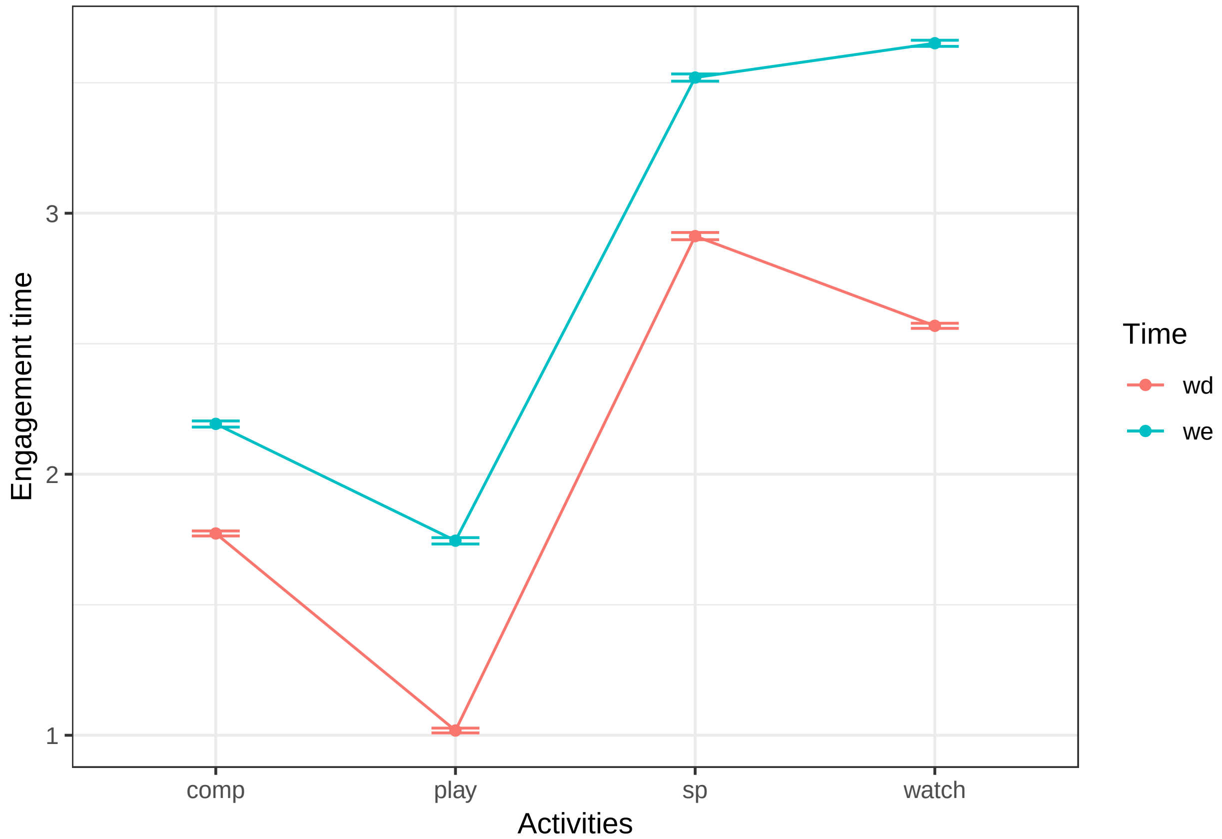

```r
boxplot(engagement ~ time,data=gold_no_na,
        col=c("#00AFBB", "#E7B800"))
```




```r
boxplot(engagement ~ pred * time, data=gold_no_na, frame = FALSE,
        col = c("#00AFBB", "#E7B800"), ylab="Engagement Time")
```




```
ggplot(gold_no_na, aes(x = pred, y = engagement, colour = time)) +
  stat_summary(fun.y = mean, geom = "point") +
  stat_summary(fun.y = mean, geom = "line", aes(group = time))+
  stat_summary(fun.data = mean_cl_boot, geom = "errorbar", width=0.2) +
  labs(x = "Activities", y = "Engagement time", colour = "Time") +
  theme_bw()
```




```
anov_mod <- lm(engagement ~ pred*time, data = gold_no_na)
summary(anov_mod)
```

```
##
## Call:
## lm(formula = engagement ~ pred * time, data = gold_no_na)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.6514 -1.5688 -0.5688  1.2545  5.9818
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.773001   0.005941  298.44   <2e-16 ***
## predplay           -0.754836   0.008399  -89.87   <2e-16 ***
## predsp              1.139333   0.008398  135.66   <2e-16 ***
## predwatch           0.795765   0.008405   94.68   <2e-16 ***
## timewe              0.419897   0.008409   49.93   <2e-16 ***
## predplay:timewe     0.307408   0.011887   25.86   <2e-16 ***
## predsp:timewe       0.187640   0.011885   15.79   <2e-16 ***
## predwatch:timewe    0.662769   0.011898   55.70   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.027 on 930510 degrees of freedom
## Multiple R-squared:  0.1517, Adjusted R-squared:  0.1517
```

```
## F-statistic: 2.378e+04 on 7 and 930510 DF,  p-value: < 2.2e-16
```

```r
Anova(anov_mod,type=c("II"))
```

```
## Anova Table (Type II tests)
##
## Response: engagement
##            Sum Sq     Df F value    Pr(>F)
## pred       553480      3 44888.7 < 2.2e-16 ***
## time       116977      1 28461.4 < 2.2e-16 ***
## pred:time   13578      3  1101.2 < 2.2e-16 ***
## Residuals 3824417 930510
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
pairwise.t.test(gold_no_na$engagement,gold_no_na$pred, p.adjust.method="BH", pool.sd=F)
```

```
##
##  Pairwise comparisons using t tests with non-pooled SD
##
## data:  gold_no_na$engagement and gold_no_na$pred
##
##       comp   play   sp
## play  <2e-16 -      -
## sp    <2e-16 <2e-16 -
## watch <2e-16 <2e-16 <2e-16
##
## P value adjustment method: BH
```

```r
pairwise.t.test(gold_no_na$engagement,gold_no_na$time, p.adjust.method="BH", pool.sd=F)
```

```
##
##  Pairwise comparisons using t tests with non-pooled SD
##
## data:  gold_no_na$engagement and gold_no_na$time
##
##    wd
## we <2e-16
##
## P value adjustment method: BH
```

```r
gold_no_na$tw <- with(gold_no_na, interaction(pred, time))
cell <- lm(engagement ~ tw - 1, data = gold_no_na)
val.glht <- (glht(cell, linfct = mcp(tw = "Tukey")))
summary(val.glht)
```

```
##
##   Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: Tukey Contrasts
##
##
## Fit: lm(formula = engagement ~ tw - 1, data = gold_no_na)
##
## Linear Hypotheses:
##                             Estimate Std. Error  t value Pr(>|t|)
## play.wd - comp.wd == 0   -0.754836   0.008399  -89.872   <0.001 ***
```

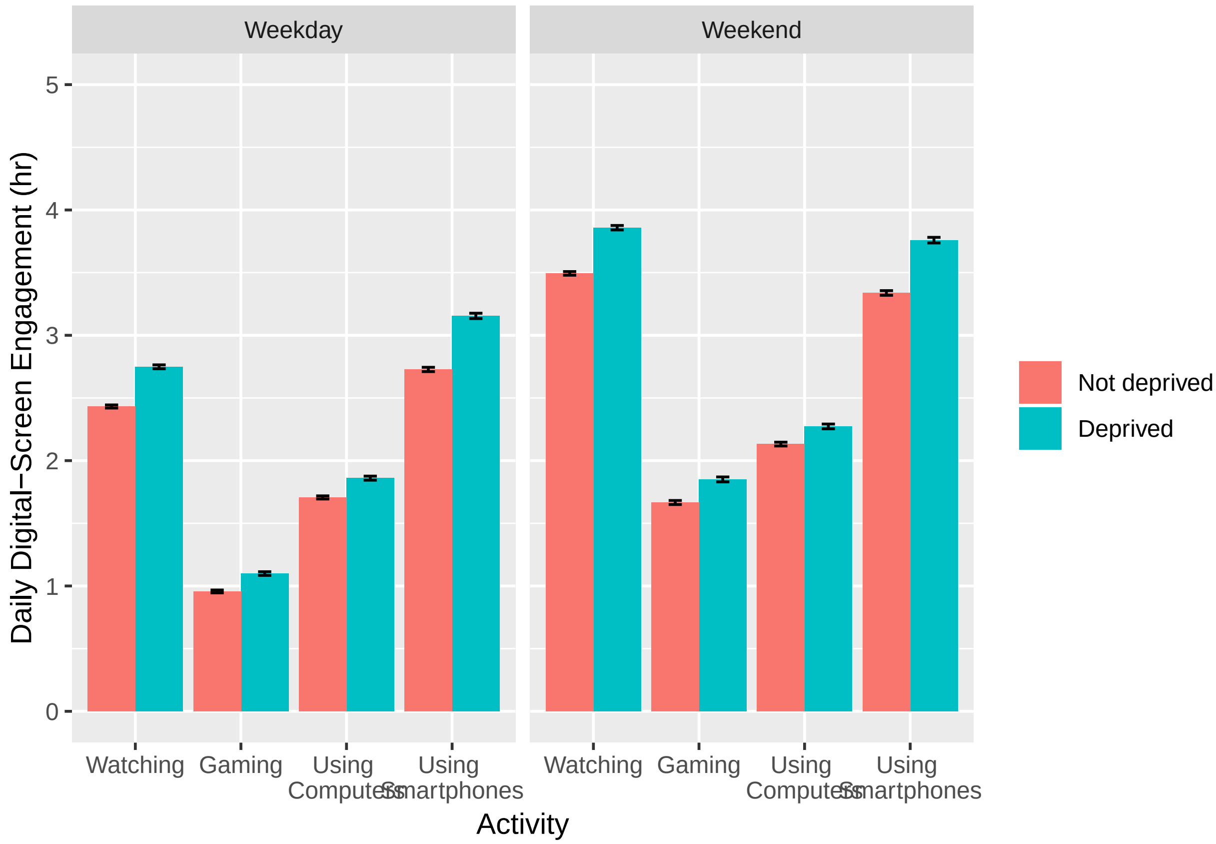
```
## sp.wd - comp.wd == 0      1.139333   0.008398  135.660   <0.001 ***
## watch.wd - comp.wd == 0   0.795765   0.008405   94.682   <0.001 ***
## comp.we - comp.wd == 0    0.419897   0.008409   49.934   <0.001 ***
## play.we - comp.wd == 0   -0.027531   0.008405   -3.276   0.0232 *
## sp.we - comp.wd == 0      1.746870   0.008402  207.921   <0.001 ***
## watch.we - comp.wd == 0   1.878430   0.008414  223.252   <0.001 ***
## sp.wd - play.wd == 0      1.894169   0.008396  225.610   <0.001 ***
## watch.wd - play.wd == 0   1.550601   0.008402  184.552   <0.001 ***
## comp.we - play.wd == 0    1.174733   0.008406  139.744   <0.001 ***
## play.we - play.wd == 0    0.727305   0.008402   86.562   <0.001 ***
## sp.we - play.wd == 0      2.501706   0.008399  297.859   <0.001 ***
## watch.we - play.wd == 0   2.633266   0.008411  313.062   <0.001 ***
## watch.wd - sp.wd == 0    -0.343568   0.008401  -40.894   <0.001 ***
## comp.we - sp.wd == 0     -0.719436   0.008406  -85.588   <0.001 ***
## play.we - sp.wd == 0     -1.166864   0.008402 -138.887   <0.001 ***
## sp.we - sp.wd == 0        0.607537   0.008398   72.340   <0.001 ***
## watch.we - sp.wd == 0     0.739097   0.008411   87.875   <0.001 ***
## comp.we - watch.wd == 0  -0.375868   0.008412  -44.683   <0.001 ***
## play.we - watch.wd == 0  -0.823296   0.008408  -97.921   <0.001 ***
## sp.we - watch.wd == 0     0.951105   0.008405  113.165   <0.001 ***
## watch.we - watch.wd == 0  1.082665   0.008417  128.629   <0.001 ***
## play.we - comp.we == 0   -0.447428   0.008412  -53.189   <0.001 ***
## sp.we - comp.we == 0      1.326973   0.008409  157.805   <0.001 ***
## watch.we - comp.we == 0   1.458534   0.008421  173.196   <0.001 ***
## sp.we - play.we == 0      1.774401   0.008405  211.120   <0.001 ***
## watch.we - play.we == 0   1.905961   0.008417  226.440   <0.001 ***
## watch.we - sp.we == 0     0.131561   0.008414   15.636   <0.001 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- single-step method)
```

Extension Report:

Our question:
How does engagement time for each activity, on both the weekend and the weekday, vary between those who lived in the deprived local area and the rest of the area? Postal code data was used to identify whether the partcipants lived in the deprived area or not.

Our exploratory plot:
This figure plots the digital screen time for all the four activities - playing games, watching tv, using computers, using smartphones - grouped by the deprived factor. We used the dplyr::summarise function to find the mean of the engagement hours and calculated the error values for a 95% confidence interval.

Using ggplot from the ggplot2 package, we plotted engagement time vs. pred (watching TV; playing games; using computers; using smartphones) grouping the data by deprived factor, and used the facet wrap command to create subplots for each activity type. The error bars are also included.



Below is our t-test analysis for all four activities - watching, playing, using computers, using smartphones - on both weekday and weekend comparing those who were coded as deprived or not deprived.

```
#Watching TV programs during weekdays for two groups, male and female
deprived_watch_we<-filter(gold_no_na, deprived==1 & pred=="watch" & time=="we")
ndeprived_watch_we<-filter(gold_no_na, deprived==0 & pred=="watch" & time=="we")
t.test(deprived_watch_we$engagement,ndeprived_watch_we$engagement, var.equal= TRUE)
```

```
##
##  Two Sample t-test
##
## data:  deprived_watch_we$engagement and ndeprived_watch_we$engagement
## t = 31.588, df = 115770, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.3418795 0.3871130
## sample estimates:
## mean of x mean of y
##  3.858451  3.493955
```

```
#Watching TV programs during weekends for two groups, male and female
deprived_watch_wd<-filter(gold_no_na, deprived==1 & pred=="watch" & time=="wd")
ndeprived_watch_wd<-filter(gold_no_na, deprived==0 & pred=="watch" & time=="wd")
t.test(deprived_watch_wd$engagement,ndeprived_watch_wd$engagement, var.equal= TRUE)
```

```
##
##  Two Sample t-test
##
```

```
## data:  deprived_watch_wd$engagement and ndeprived_watch_wd$engagement
## t = 31.638, df = 116280, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.2969097 0.3361261
## sample estimates:
## mean of x mean of y
##  2.748522  2.432004
```

```
#Playing console games during weekdays for two groups, male and female
deprived_play_we<-filter(gold_no_na, deprived==1 & pred=="play"&time=="we")
ndeprived_play_we<-filter(gold_no_na, deprived==0 & pred=="play"& time=="we")
t.test(deprived_play_we$engagement,ndeprived_play_we$engagement, var.equal= TRUE)
```

```
##
##  Two Sample t-test
##
## data:  deprived_play_we$engagement and ndeprived_play_we$engagement
## t = 14.331, df = 116280, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.1588680 0.2092074
## sample estimates:
## mean of x mean of y
##  1.850036  1.665998
```

```
#Playing console games during weekends for two groups, male and female
deprived_play_wd<-filter(gold_no_na, deprived==1 & pred=="play"&time=="wd")
ndeprived_play_wd<-filter(gold_no_na, deprived==0 & pred=="play"& time=="wd")
t.test(deprived_play_wd$engagement,ndeprived_play_wd$engagement, var.equal= TRUE)
```

```
##
##  Two Sample t-test
##
## data:  deprived_play_wd$engagement and ndeprived_play_wd$engagement
## t = 15.482, df = 116600, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.1245065 0.1606013
## sample estimates:
## mean of x mean of y
## 1.0991364 0.9565825
```

```
#Using Computers during weekdays for two groups, male and female
deprived_comp_we<-filter(gold_no_na, deprived==1 & pred=="comp"&time=="we")
ndeprived_comp_we<-filter(gold_no_na, deprived==0 & pred=="comp"& time=="we")
t.test(deprived_comp_we$engagement,ndeprived_comp_we$engagement, var.equal= TRUE)
```

```
##
##  Two Sample t-test
##
## data:  deprived_comp_we$engagement and ndeprived_comp_we$engagement
## t = 11.422, df = 116040, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.1165290 0.1648044
```

```
## sample estimates:
## mean of x mean of y
##  2.272818  2.132151
```

```r
#Using Computers during weekends for two groups, male and female
deprived_comp_wd<-filter(gold_no_na, deprived==1 & pred=="comp"&time=="wd")
ndeprived_comp_wd<-filter(gold_no_na, deprived==0 & pred=="comp"& time=="wd")
t.test(deprived_comp_wd$engagement,ndeprived_comp_wd$engagement, var.equal= TRUE)
```

```
##
##  Two Sample t-test
##
## data:  deprived_comp_wd$engagement and ndeprived_comp_wd$engagement
## t = 15.021, df = 116450, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.1340201 0.1742436
## sample estimates:
## mean of x mean of y
##  1.860506  1.706374
```

```r
#Using Smartphones during weekdays for two groups, male and female
deprived_sp_we<-filter(gold_no_na, deprived==1 & pred=="sp" & time=="we")
ndeprived_sp_we<-filter(gold_no_na, deprived==0 & pred=="sp"& time=="we")
t.test(deprived_sp_we$engagement,ndeprived_sp_we$engagement, var.equal= TRUE)
```

```
##
##  Two Sample t-test
##
## data:  deprived_sp_we$engagement and ndeprived_sp_we$engagement
## t = 28.631, df = 116450, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.3924621 0.4501442
## sample estimates:
## mean of x mean of y
##  3.759142  3.337839
```

```r
#Using Smartphones during weekends for two groups, male and female
deprived_sp_wd<-filter(gold_no_na, deprived==1 & pred=="sp" & time=="wd")
ndeprived_sp_wd<-filter(gold_no_na, deprived==0 & pred=="sp"& time=="wd")
t.test(deprived_sp_wd$engagement,ndeprived_sp_wd$engagement, var.equal= TRUE)
```

```
##
##  Two Sample t-test
##
## data:  deprived_sp_wd$engagement and ndeprived_sp_wd$engagement
## t = 31.034, df = 116630, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.4002503 0.4542140
## sample estimates:
## mean of x mean of y
##  3.154801  2.727569
```

Additional Analysis:
We wanted to see if the deprived factor influenced the engagement time for all the four activities. By looking

at only the means of the individual factors, neither of them appear to have any effect on the response variable.

A common method for analyzing the effect of categorical variables on a continuous response variable is the Analysis of Variance, or ANOVA. We can fit a linear model to these data using the lm function. By looking at the table of coefficients, we can tell everything is significant. This just means the coefficients are significantly different from 0.

Running an ANOVA on these data reveal a significant interaction, and we also notice the main effects are significant as well. This just means the effects of deprived and pred individually explain a fair amount of variability in the data.

From the ANOVA results, we can conclude the following, based on the p-values and a significance level of 0.05: * The p-value of pred is p$<$ 2e-16 (significant), which indicates that the levels of pred are associated with significant different engagement time. * The p-value of deprived is p $<$ 2e-16 (significant), which indicates that the levels of deprived are associated with significant different engagement time. * The p-value for the interaction between pred*deprived is p $<$ 2e-16 (significant), which indicates that the relationships between pred and engagement time depends on the deprived factor.

In the ANOVA test, a significant p-value indicates that some of the group means are different, but it does not tell us which pairs of groups are different. We performed multiple pairwise-comparison to determine if the mean differences between specific pairs of group are statistically significant.

Since we are looking at deprived (2 levels), we can just do it with the pred (4 levels) variable. We used the general linear hypothesis test to compare the means between the groups. From the result, we can clearly see that the difference between the pred - deprived pairs were all significant ($p < 0.001$).

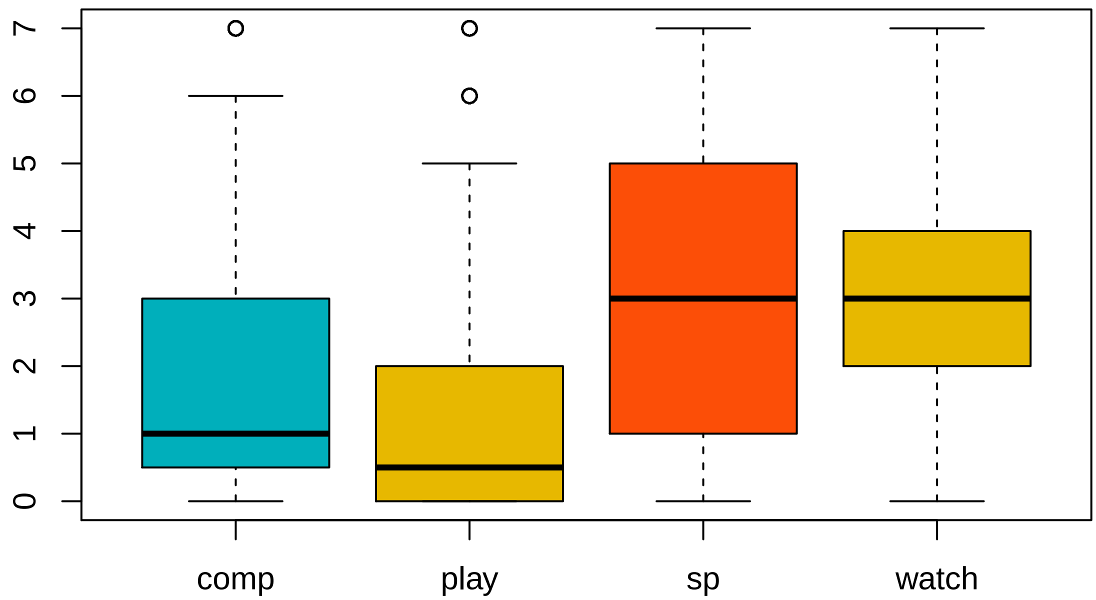
```
group_by(gold_no_na, pred) %>%
  summarise(count = n(),
            mean = mean(engagement, na.rm = TRUE),
            sd = sd(engagement, na.rm = TRUE))
```

```
## # A tibble: 4 x 4
##   pred   count  mean    sd
##   <fct>  <int> <dbl> <dbl>
## 1 comp  232498  1.98  1.93
## 2 play  232878  1.38  1.92
## 3 sp    233084  3.22  2.44
## 4 watch 232058  3.11  1.91
```

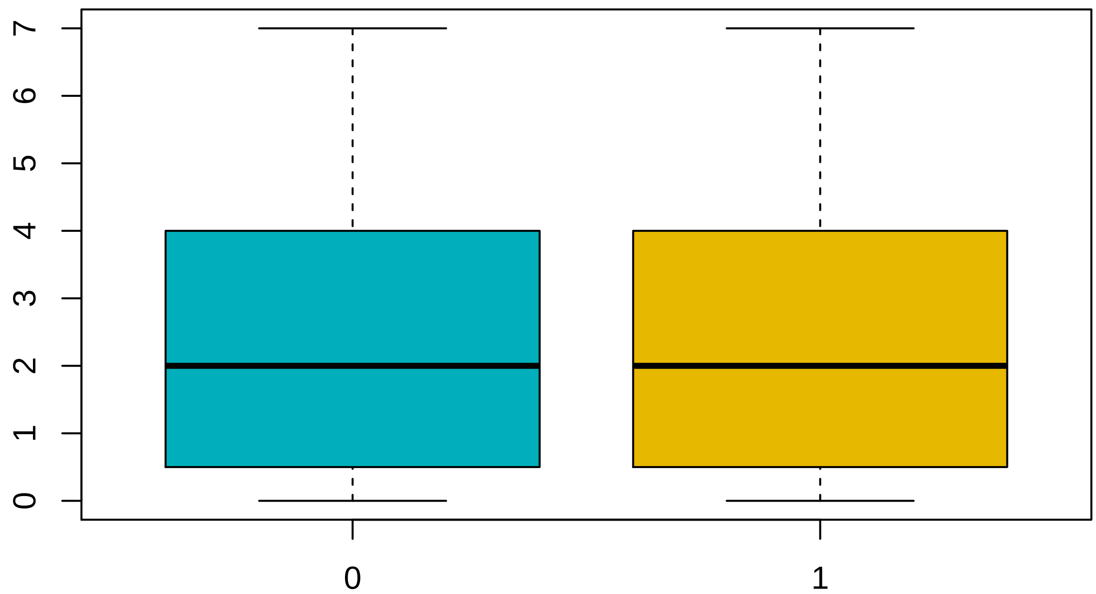
```
group_by(gold_no_na, deprived) %>%
  summarise(count = n(),
            mean = mean(engagement, na.rm = TRUE),
            sd = sd(engagement, na.rm = TRUE))
```

```
## # A tibble: 2 x 4
##   deprived  count  mean    sd
##      <int>  <int> <dbl> <dbl>
## 1        0 528465  2.31  2.11
## 2        1 402053  2.57  2.31
```

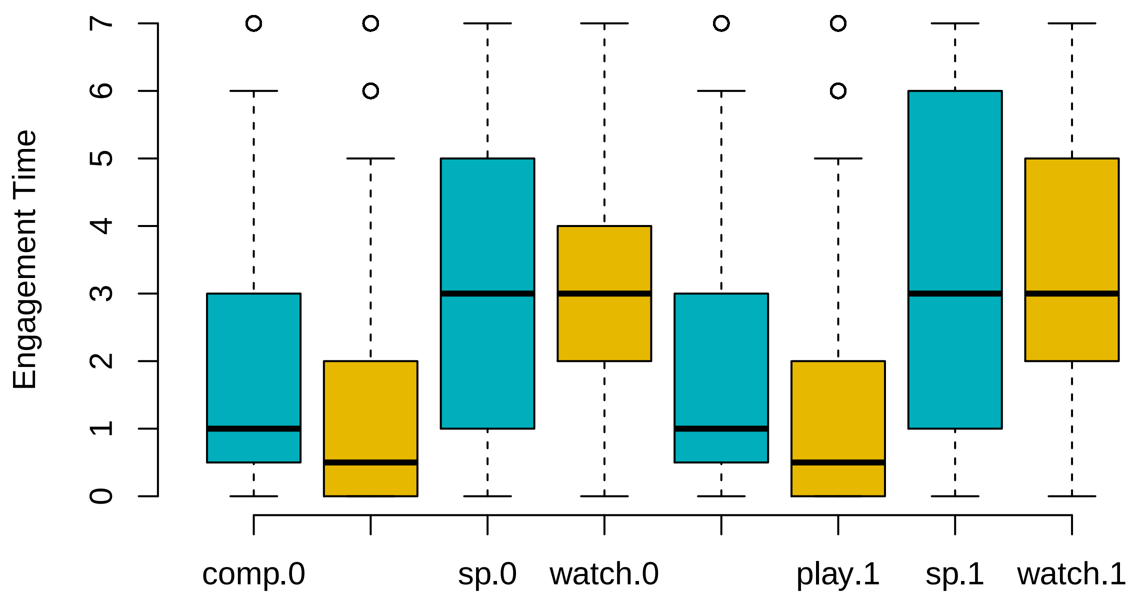
```
boxplot(engagement ~ pred,data=gold_no_na,
        col=c("#00AFBB", "#E7B800", "#FC4E07","#E7B800"))
```



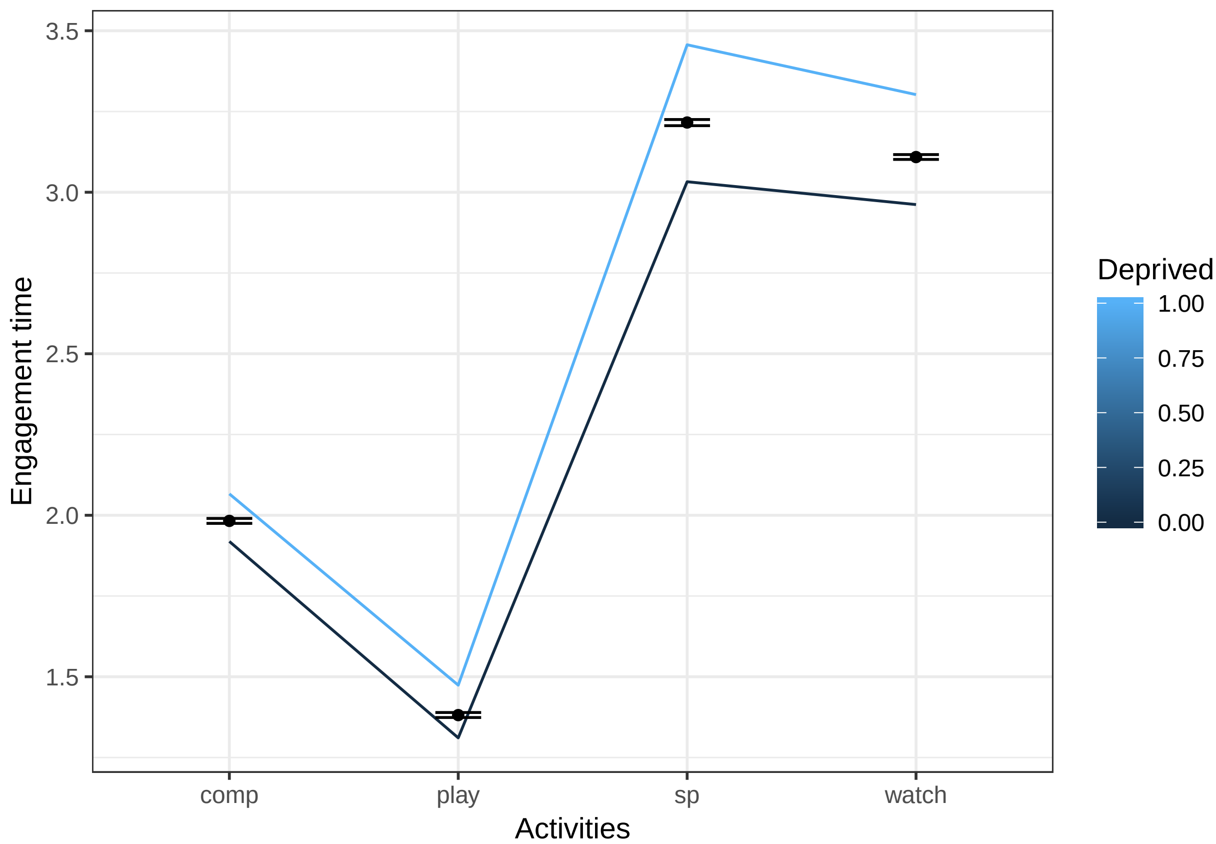

```r
boxplot(engagement ~ deprived,data=gold_no_na,
        col=c("#00AFBB", "#E7B800"))
```




```
boxplot(engagement ~ pred * deprived, data=gold_no_na, frame = FALSE,
        col = c("#00AFBB", "#E7B800"), ylab="Engagement Time")
```




```
ggplot(gold_no_na, aes(x = pred, y = engagement, colour = deprived)) +
  stat_summary(fun.y = mean, geom = "point") +
  stat_summary(fun.y = mean, geom = "line", aes(group = deprived)) +
  stat_summary(fun.data = mean_cl_boot, geom = "errorbar", width = 0.2) +
  labs(x = "Activities", y = "Engagement time", colour = "Deprived") +
  theme_bw()
```




```
anov_mod <- lm(engagement ~ pred*deprived, data = gold_no_na)
summary(anov_mod)
```

```
##
## Call:
## lm(formula = engagement ~ pred * deprived, data = gold_no_na)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -3.457 -1.457 -0.474  1.081  5.689
##
## Coefficients:
##                        Estimate Std. Error t value Pr(>|t|)
## (Intercept)            1.918971   0.005659 339.078   <2e-16 ***
## predplay              -0.608115   0.008000 -76.016   <2e-16 ***
## predsp                 1.113611   0.007999 139.214   <2e-16 ***
## predwatch              1.042852   0.008007 130.237   <2e-16 ***
## deprived               0.147229   0.008610  17.100   <2e-16 ***
## predplay:deprived      0.015911   0.012172   1.307    0.191
## predsp:deprived        0.276792   0.012168  22.747   <2e-16 ***
## predwatch:deprived     0.193178   0.012182  15.858   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.057 on 930510 degrees of freedom
## Multiple R-squared:  0.1271, Adjusted R-squared:  0.1271
```

```
## F-statistic: 1.936e+04 on 7 and 930510 DF,  p-value: < 2.2e-16
```

```r
Anova(anov_mod,type=c("II"))
```

```
## Anova Table (Type II tests)
##
## Response: engagement
##                 Sum Sq     Df  F value    Pr(>F)
## pred            553453      3 43621.36 < 2.2e-16 ***
## deprived         16489      1  3898.85 < 2.2e-16 ***
## pred:deprived     3151      3   248.37 < 2.2e-16 ***
## Residuals      3935332 930510
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
gold_no_na$tw <- with(gold_no_na, interaction(pred, deprived))
cell <- lm(engagement ~ tw - 1, data = gold_no_na)
val.glht <- (glht(cell, linfct = mcp(tw = "Tukey")))
summary(val.glht)
```

```
##
##   Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: Tukey Contrasts
##
##
## Fit: lm(formula = engagement ~ tw - 1, data = gold_no_na)
##
## Linear Hypotheses:
##                           Estimate Std. Error  t value Pr(>|t|)
## play.0 - comp.0 == 0     -0.608115   0.008000  -76.016   <2e-16 ***
## sp.0 - comp.0 == 0        1.113611   0.007999  139.214   <2e-16 ***
## watch.0 - comp.0 == 0     1.042852   0.008007  130.237   <2e-16 ***
## comp.1 - comp.0 == 0      0.147229   0.008610   17.100   <2e-16 ***
## play.1 - comp.0 == 0     -0.444974   0.008607  -51.701   <2e-16 ***
## sp.1 - comp.0 == 0        1.537632   0.008603  178.741   <2e-16 ***
## watch.1 - comp.0 == 0     1.383259   0.008615  160.573   <2e-16 ***
## sp.0 - play.0 == 0        1.721726   0.007995  215.337   <2e-16 ***
## watch.0 - play.0 == 0     1.650967   0.008004  206.279   <2e-16 ***
## comp.1 - play.0 == 0      0.755344   0.008606   87.766   <2e-16 ***
## play.1 - play.0 == 0      0.163140   0.008603   18.963   <2e-16 ***
## sp.1 - play.0 == 0        2.145747   0.008599  249.534   <2e-16 ***
## watch.1 - play.0 == 0     1.991374   0.008611  231.260   <2e-16 ***
## watch.0 - sp.0 == 0      -0.070758   0.008003   -8.841   <2e-16 ***
## comp.1 - sp.0 == 0       -0.966382   0.008606 -112.293   <2e-16 ***
## play.1 - sp.0 == 0       -1.558585   0.008603 -181.173   <2e-16 ***
## sp.1 - sp.0 == 0          0.424021   0.008599   49.313   <2e-16 ***
## watch.1 - sp.0 == 0       0.269648   0.008611   31.316   <2e-16 ***
## comp.1 - watch.0 == 0    -0.895623   0.008613 -103.980   <2e-16 ***
## play.1 - watch.0 == 0    -1.487827   0.008610 -172.797   <2e-16 ***
## sp.1 - watch.0 == 0       0.494779   0.008606   57.492   <2e-16 ***
## watch.1 - watch.0 == 0    0.340407   0.008618   39.499   <2e-16 ***
## play.1 - comp.1 == 0     -0.592204   0.009173  -64.558   <2e-16 ***
## sp.1 - comp.1 == 0        1.390403   0.009169  151.636   <2e-16 ***
## watch.1 - comp.1 == 0     1.236030   0.009181  134.636   <2e-16 ***
```

```
## sp.1 - play.1 == 0      1.982606   0.009166  216.291   <2e-16 ***
## watch.1 - play.1 == 0   1.828234   0.009178  199.206   <2e-16 ***
## watch.1 - sp.1 == 0    -0.154373   0.009174  -16.828   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- single-step method)
```