**Health Benefits of Fruit and Vegetables**

Author: Jose Manuel Alvarez Gonzalez

Email Id: jmanuel.ag89@gmail.com

Date: 14/3/2022

**Introduction**

This experiment focuses on the health benefits of consumption of fruit and vegetables. The goal of the research is to measure the health benefits between groups that consume a higher amount of fruits and vegetable, versus a control group with a low fruit and vegetable intake, this is done during a period of 14 days. This information will be used to help developing an app with the purpose of encouraging adults to eat more fruits and vegetables.

**Methods**

**Experiment Design**

The experiment design was carried during 14 clinical days. The experimental units were 171 students from the University of Central Otago between the age of 18 to 25, the experimental units were self-identified as low fruit and vegetable consumers.

The treatment factor of the experiment is the type of intervention that the units were subject during the 14 days, with three different levels:

* **Control** condition, usual diet.
* **EMI**, were given a challenge to increase their FV consumption to at least 5 serves a day, they got vouchers for food and two targeted text a day by the app.
* **FVI**, participants were asked to consume one additional serve of fruit and one additional serve of vegetables every day, they also received a supply of fruits and vegetables to consume.

**Groups homogeneity**

The experimental units were randomly assigned to each treatment level, three features can be identified to describe the homogeneity of each group:

* **Gender**, 56 men and 115 women
* **BMI**, a continuous variable from 15 to 40.
* **Ancestry,** 64% European descent and 36% non-European.

**Response variables**

The experiment will focus on different scores to measure the effectiveness of treatment, at the same time it will measure the total fruit consumption of each unit during the 14 days. The scores we will work:

* Happiness
* Energy
* Fatigue
* Vitamin C, measured pre and post experiment.
* Carotenoids, measured pre and post experiment.

**Analysis Methods**

Descriptive statistics 🡪 Understand dispersion and centrality of variables.  
Shappiro Test 🡪 Measure normality of response variables.  
Bartlett Test 🡪 Measure variance homogeneity between groups.  
T-Test 🡪 Used as a paired test to measure difference significance on an A/B test.  
Wilcox Test 🡪 Non parametric Test to measure difference significance on an A/B test.  
Power Analysis 🡪 Measure power of test and Cohen’s D difference on results.   
Tukey’s Test 🡪 Post-hoc analysis of between group differences.   
Linear Model 🡪 Post-hoc analysis of between and within group differences, also used to identify interaction with continuous variable.  
ANOVA 🡪 Measure differences between intervention groups and interaction with other factor.  
ANOVA, Type = III 🡪 Paired test to measure difference within intervention group.  
ANCOVA 🡪 Measure interaction with a continuous variable.

**Results and Discussion**

**Data Introduction**

For this analysis, we worked with two data sets, one looks at the response variables of each unit. The other one looks at the individuals daily fruit consumption during the experiment. For analysis, we had to remove units that failed to include some of the response variables. For the daily fruit consumption, we also had to remove units that failed to include daily consumption across the 14 days. We finished with the following designs:

|  |  |  |
| --- | --- | --- |
| **Intervention Design** | | |
|  |
| **Treatment** | **Block** | **Subjects** |  |
| Control | 0 | 20 |  |
| Control | 1 | 38 |  |
| EMI | 0 | 17 |  |
| EMI | 1 | 37 |  |
| FVI | 0 | 18 |  |
| FVI | 1 | 37 |  |

|  |  |
| --- | --- |
| **Diary Design** | |
|  |
| **Treatment** | **Subjects** |  |
| Control | 16 |  |
| EMI | 9 |  |
| FVI | 18 |  |

*Table 1, design for Intervention and Diary data sets*

For the intervention design, we have the following descriptive statistics across the variables:



Table 2, Descriptive Statistics of response variables

In Figure 1, we can illustrate the density distribution across each response. Here I’m including the differences for Carotenoids and Vitamin C. For the histogram we have average total consumption per days recorded, this is before removing non-compliance of 14 days.

Chart, histogram

Description automatically generatedChart, shape, radar chart

Description automatically generated

Figure , Density distribution for main response variables and histogram of average FV consumption per individual (before removing 14 days non-compliance)

From the last visual and the table, we can identify that most of the data is not normally distributed. In Figure 2, we have a first look on how each response variables differs from treatment level.

Graphical user interface

Description automatically generated

Figure , Boxplot for response variables difference by intervention

In Table 3, we can see that for most cases, most of the assumptions are broken for parametric Test. To solve this, we will create a complete block design when required, otherwise we will do a non-parametric.

Table 3, Assumption results for both data sets we will be working with

**BMI**

We want to identify if is a BMI mean difference across the three intervention levels, from figure 1, we can identify BMI is not normally distributed and has some severe outliers on the Control group. As we will only analyse with visuals, I will remove the extreme outliers to visualise. In Figure 2, top right visual, seems to be a narrow gap between FVI and Control, closer on the equal means result. If we also add gender as a block, we can confident say is no difference of means on BMI levels across all intervention levels. For the next part of the report, we will consider that is not a statistical difference. This will allow us to don’t consider BMI as a covariate for every time we want to measure a difference between responses.

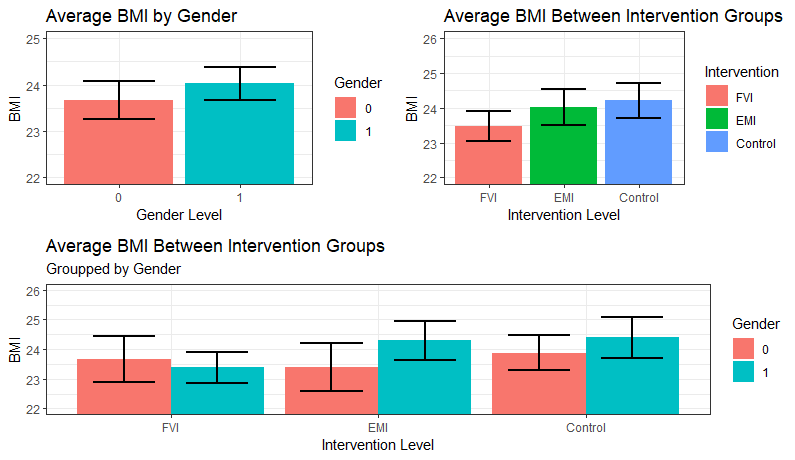


Figure , BMI across interaction and gender, extreme outliers were removed of visuals.

Next, we want to measure if is any interaction between the levels of Energy and BMI across the different treatment levels, I’m using the full data for this analysis. To identify an interaction, we use an ANCOVA model, in the test we failed to identify any interaction between them, with a p-value = 0.273. I tested the residuals of the ANCOVA test for normality and we obtained p-value = 0.703, meaning that normality is not broken. I also used a Linear Model and failed to find any statistically significant value for interaction across all intervention levels.

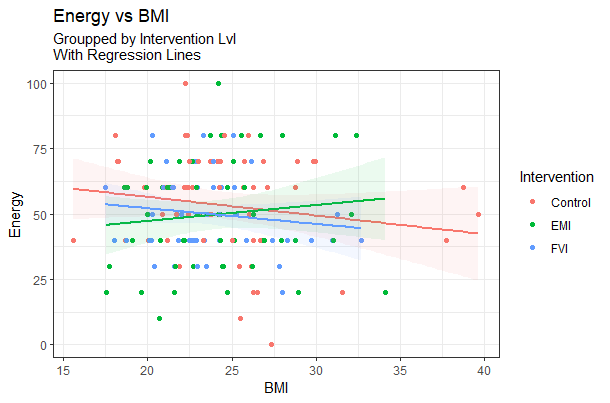
 

Figure , left, Point plot showing energy scores according to BMI, grouped by Intervention level and linear model with standard errors. Right, the ANCOVA table of the test.

In figure 4, we can find that is a tendency from EMI energy to increase and for the other groups to reduce, but the standard errors don’t allow this to make it statistically significant enough.

**Complete Block Design and Replication**

From the assumptions check, normality and homogeneity of variance is broken in most cases. To fix this, we will create a complete block design with gender as a block, the subjects will be randomly selected, the units level will be set at 17 as the treatment level EMI for male units. At the same time, we will replicate this design and the ANOVA tests to ensure enough power on the result of our tests. For each test replication we will do a new random CBD and measure the amount of times p-values were under 0.05 for the sum of the run, with runs from 1, 10, 25, 50, 100 and 500 replicates.

First, we will look at any interaction between happiness and gender for each intervention level, first we run an ANOVA model before the CBD, here we can find a statistically significant interaction with   
p-value = 0.019, but normality of residuals is broken. When I do the CBD and replicate the ANOVA experiment, this result is only statistically significant 51% of the time on 500 replicates. We can find the results on Table 4.

Table 4, ANOVA table with the original design, following with a table of results for the CBD and the different times the experiment was replicated.

Figure 6, identifies the different clusters of Happiness level across each group (particularly control group, gender 1) and why only one CBD might not be enough to draw a real conclusion.

Chart

Description automatically generated

Figure 5, Dot plot for each intervention level and gender, measuring happiness levels. The groups also show the density distribution and quantile lines.

Next, we want to measure if is any difference between treatment levels for Energy and Fatigue with gender as a block. We will follow the previous procedure with a random CBD selection and different replications on the ANOVA test to measure power. We failed to find any statistical difference between groups in both response variables.



Table 5, ANOVA results for the CBD, measuring difference between intervention level with gender as a block for Fatigue and Energy

**A/B Paired Test**

Next, we want to measure any difference between Vitamin C and Carotenoids group means before and after the experiment for FVI and EMI interventions. This is a paired design, and we will do a T-Test and a Wilcox Test for the response that don’t comply with the assumptions check.



Table 6, Results for the different A/B paired tests, with type of test and power Analysis

From table 6, we can identify that for treatment EMI the difference on mean pre and post result on Vitamin C is statistical significant with a p-value = 0.034, this shows a Cohen’s D = 0.277 which is a small difference. The difference is an increase of 5.16 units, if we look at the post Vitamin C range (2.7 – 121.5), this is about a 4.34% increase. At the same time FVI difference on Carotenoids is close to be statistically significant with a p-value = 0.052 and Cohen’s D = 0.24, which is an increase of 0.256 units, if we look at the post Carotenoids range, this is a 4.38% increase.

**Daily Fruit Consumption**

Now we are going to change data set and work with amount of FV consumed per day of study, focusing on days 1, 7 and 14. We will work with the reduced design previously detailed on Table 1. First, we will do a Paired ANOVA test to identify any difference of mean in the three days for each intervention level.



Table 7, ANOVA results to identify a difference for FV consumption across the three different days, this test is run independently for each treatment level.

In Table 7, we can find that we couldn’t identify any statistical difference of FV consumption on days 1, 7 and 14, for each intervention group. From the test we can find that Sphericity was broken in all cases and Greenhouse–Geisser correction was applied on the p-values.

At the same time, I applied a Linear Model on each intervention to identify differences across the three days, I couldn’t find any statistically significant result, but I did find that for the linear model on the FVI intervention, FV consumption increases after day 1 and for EMI FV consumption decreases after day 1. We can find the coefficients on Figure 6.



Figure 6, Plot and Linear Model for each intervention across the 1, 7 and 14 days of study. With coefficients of linear model

For the last Analysis of the report, we will work with the previous data set of daily FV consumption, but in this case, we will work across the 14 days. For this part we will do an average consumption of each experimental unit across the 14 days. We will measurer if is any difference across the three different treatment levels on the total consumption of FV.

This is the only analysis that complies with assumptions and allows a full ANOVA test.



Table 8, ANOVA result for difference of Average FV consumption between intervention levels

The test identifies that is no statistical significance with a p-value = 0.08, at the same time I want to further study the difference between groups and I applied a Tukey’s test, showing that the biggest difference is between Control and FVI group with a p-value = 0.072 with a difference of 0.911 units. This result says that FVI consumed on average 0.911 more FV a day than the control group. If we look at the total average consumption of the control group of 2.74 FV with se = 0.363, this is actually a 32% more FV consumption per day, might not be statistical significant according to the Tukey’s test result but is a good practical increase in 14 days.



Figure 7, Tukey’s analysis, and table of results

In Figure 8, I decided to include a bar plot with standard errors for each intervention, here it looks like is a difference but this is not accounting Tukey’s adjustment.

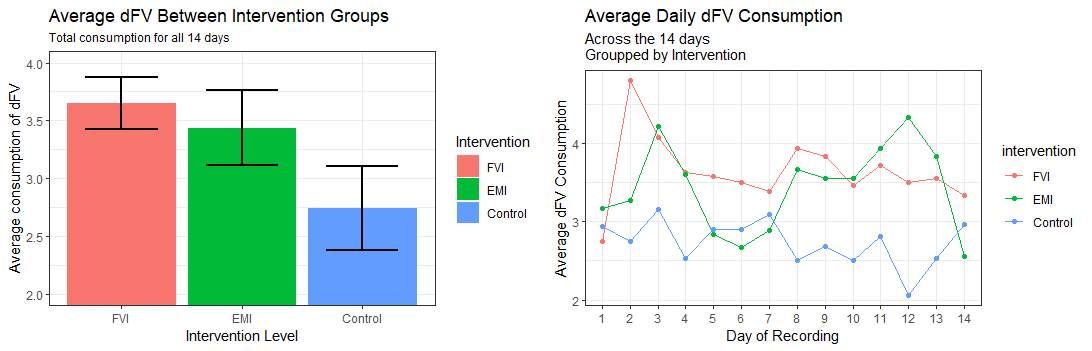


Figure 8, left Bar plot with average unit consumption during the 14 days, right is a line plot with average daily consumption per unit on each day across the three intervention levels.

**Recommendations & Conclusions**

In the previous experiment, we failed to prove a major health benefit by increasing the consumption of fruit and vegetables during 14 days. We did find a statistically significant increase of Vitamin C in the EMI group. These increases might be statistically significant but on reality is only a 4.34% increase on the relative level.

For the daily fruit and vegetable consumption, we failed to identify any statistically significant difference on average consumption between each intervention group. We did find a difference of daily consumption of 0.911 more FV by the FVI group compared to the Control group. This difference accounts for a 32% more FV. For the EMI vs the control group, the difference is 0.697 units, compared to the control group is a 25% more FV consumed on a day.

The linear models we applied during the average daily food consumption, showed a tendency for the FVI group to consume more FV as the days progress and a reduction from Day 1 on FV consumption for the EMI group. This results are not statistical significant and more days of study will be required to make the previous models statistical significant.

The experiment demonstrates that 14 days is not enough time to measure any major health benefit on fruit and vegetable consumption, it will be beneficial to increase the length of study for future experiments. At the same time the sample sizes need to increase to ensure enough power of results, from the one positive results we had on the experiment, Vitamin C increase on EMI group, it only has a power of 52%.  
  
Another key benefit on the experiment will be a more rigorous process in the way the samples record the information as we lost a high amount of units by failing to report the fruit and vegetables daily consumption. At the same time, it will be beneficial to try to reduce any blocking factors by checking homogeneity of groups after the random selection of subjects.

What we can take from the previous experiment is that we can incentivize adults to increase their fruit and vegetables consumption and maintain it for 14 days, we just need to find the right strategy to keep them excited about it. Maybe new cooking recipes with the targeted messages can help to increase and maintain consumption.

**Appendix – R Code**

library(tidyverse)

library(rstatix)

library(moments)

library(effsize)

library(pwr)

library(gridExtra)

# ------------------------ Read the Data --------------------------------------

# Read the data and convert to NA the null values

diary <- read\_csv("./Fruit-Veg-diary.csv", na = "#NULL!")

intervention <- read\_csv("./Fruit-veg-intervention.csv", na = "#NULL!")

# ------------------------ Custom Functions ------------------------------------

# Creating a Shapiro test function that can check on each treatment level for

# a desired factor

Normality\_test <- function(factor, levels){

  result <- list()

  for (level in unique(levels)){

    to\_test <- factor[levels == level]

    result[[level]] <- shapiro.test(to\_test)

  }

  return(result)

}

# Creating a function that we will use for the measures of centrality and spread

summary.fun = function(x){

  round(c(min = min(x), max = max(x),

          median = median(x), mean = mean(x), SD = sd(x), kurtosis = kurtosis(x),

          SK = skewness(x)),4)

}

# ------------------------ Cleaning & Wrangling --------------------------------

# Check if NA are present

diary %>% summarise(across(everything(), ~ sum(is.na(.))))

intervention %>% summarise(across(everything(), ~ sum(is.na(.))))

# look at the rows with NA

diary[!complete.cases(diary), ]

intervention[!complete.cases(intervention), ]

# For Diary I will drop NA as we can see that weekly recordings are inconsistent

# between different ID

diary <- na.omit(diary)

# For intervention I will drop the 4 ID, I will keep record of dropped ID's

ID\_dropped <- intervention$ID[!complete.cases(intervention)]

intervention <- na.omit(intervention)

# Create a difference field for Vitamin C and Cartenoids

intervention$VitC\_diff <- intervention$VitC\_post - intervention$VitC\_pre

intervention$Carot\_diff <- intervention$Carot\_post - intervention$Carot\_pre

# Convert to Factors

diary %>% convert\_as\_factor(ID, Diary) -> diary

intervention$intervention <-factor(intervention$intervention,

                                   labels = c("Control", "EMI", "FVI"))

intervention %>% convert\_as\_factor(ID, gender, ancestry) -> intervention

# Create a long table for intervention

intervention[, c(-9,-10,-11,-12)] %>%

  gather(key = "Factor", value = "Level", BMI:Carot\_diff) -> intervention\_L

# ------------------------ Descriptive Summary of Data ------------------------

# Summary Statistics of diary

summary(diary)

# Find how many weeks each ID recorded fruit intake

table(diary$ID)

diary %>% group\_by(ID) %>% tally() %>% summarise(range(n))

# People recorded from 8 to 14 weeks of fruit intake

# Summary Statistics of intervention

summary(intervention)

# Stats of Factors

lapply(intervention[,5:14], summary.fun)

# ------------------------ Initial Visualizations -------------------------------

# General view on factors according to intervention

intervention\_L %>% ggplot(aes(x=intervention, y=Level, col=intervention)) +

  geom\_boxplot() +

  facet\_wrap(vars(Factor), scales = "free") +

  xlab("Intervention") +

  labs(title = "Resonse Variables Results",

       subtitle = "Groupped by each Intervention Lvl",

       col = "Intervention") +

  theme\_bw()

# General view on factors distribution

intervention\_L %>% ggplot(aes(x=Level, fill=intervention)) +

  geom\_density(alpha = .5) +

  facet\_wrap(vars(Factor), scales = "free") +

  xlab("Intervention") +

  labs(title = "Distribution Density for Factors Results",

       subtitle = "Groupped by each Intervention Lvl",

       col = "Intervention") +

  theme\_bw()

# Most factors look not normally distributed

# General view on factors with gender as a block

intervention\_L %>% ggplot(aes(x=intervention, y=Level, col=gender)) +

  geom\_boxplot() +

  facet\_wrap(vars(Factor), scales = "free") +

  xlab("Intervention") +

  labs(title = "Main Factors With Gender as Block ",

       subtitle = "Groupped by each Intervention Lvl",

       col = "Intervention") +

  theme\_bw()

# We can see that gender makes a difference but lets see if is statistical significant

# dFV histogram distribution

diary %>% ggplot(aes(x=dFV)) +

  geom\_histogram(bins=10, color = "blue", fill = "red") +

  ylab("Frequency") +

  xlab("dFV Consumption") +

  labs(title = "Histogram of dFV Consumption per Day",

       subtitle = "For all records") +

  theme\_bw()

# dFV histogram on total average consumption per member

diary %>% group\_by(ID) %>% summarise("dFV" = mean(dFV)) %>%

  ggplot(aes(x=dFV)) +

  geom\_histogram(bins=10, color = "blue", fill = "red") +

  ylab("Frequency") +

  xlab("dFV Consumption") +

  labs(title = "Histogram of Total dFV Consumption",

       subtitle = "Average for all records") +

  theme\_bw()

# dFV Average consumption across the 14 days

diary %>% group\_by(Diary) %>% summarise(Total = mean(dFV)) %>%

  ggplot(aes(x=Diary, y=Total, group = 1)) + geom\_line(col="red") +

  geom\_point(col="blue") +

  ylab("Average dFV Consumption") +

  xlab("Day of Recording") +

  labs(title = "Average Daily dFV Consumption",

       subtitle = "Across the 14 days") +

  theme\_bw()

# We can see that most of the average consumption decreasead from the first

# couple of days, we will figure it out if is the same for al treatments.

# ------------------------ BMI Visualization -----------------------------------

# Plot densities of the BMI according to intervention Lvl

intervention %>% ggplot(aes(x=BMI, fill=intervention)) +

  geom\_histogram(bins = 25) +

  ylab("Frequency") +

  labs(title = "Histogram of BMI",

       subtitle = "Groupped by Intervention Lvl",

       fill = "Intervention") +

  theme\_bw()

# Identify outliers by treatment

intervention %>%

  group\_by(intervention) %>%

  identify\_outliers(BMI) -> BMI\_outliers

# Remove Extreme outliers

int\_BMI\_Nout <- intervention[!(intervention$ID %in%

                                 BMI\_outliers$ID[BMI\_outliers$is.extreme == T]),]

# We have two extreme outliers and will remove ONLY for the visuals

# Creating a Gender table for BMI

int\_BMI\_Nout %>% group\_by(gender) %>%

  summarise(mean = mean(BMI),

            se = sd(BMI)/sqrt(length(BMI))) -> gender\_BMI\_mse

# Creating an Intervention table for BMI

int\_BMI\_Nout %>% group\_by(intervention) %>%

  summarise(mean = mean(BMI),

            se = sd(BMI)/sqrt(length(BMI))) -> intervention\_BMI\_mse

# Creating a Intervention and Gender table for BMI

int\_BMI\_Nout %>% group\_by(intervention, gender) %>%

  summarise(mean = mean(BMI),

            se = sd(BMI)/sqrt(length(BMI))) -> gender\_int\_BMI\_mse

# First visualization with gender on the total experimental units

dodge <- position\_dodge(width=0.9)

ggplot(gender\_BMI\_mse, aes(x = gender, y = mean, fill = gender)) +

  geom\_bar(stat = "identity", position = dodge) +

  geom\_errorbar(aes(ymin = mean - se,

                    ymax = mean + se),

                position = dodge,

                width  = .6,

                size = .8) +

  coord\_cartesian(ylim = c(22, 25)) +

  ylab("BMI") +

  labs(title = "Average BMI by Gender",

       x = "Gender Level",

       fill= "Gender") +

  theme\_bw() -> bmi\_1

# BMI between intervention groups

ggplot(intervention\_BMI\_mse,

              aes(x = intervention, y = mean, fill = intervention)) +

  geom\_bar(stat = "identity", position = dodge) +

  geom\_errorbar(aes(ymin = mean - se,

                    ymax = mean + se),

                position = dodge,

                width  = .6,

                size = .8) +

  coord\_cartesian(ylim = c(22, 26)) +

  ylab("BMI") +

  xlab("Intervention Level") +

  labs(title = "Average BMI Between Intervention Groups",

       fill = "Intervention") +

  theme\_bw()  ->bmi\_2

# BMI between intervention groups and gender

ggplot(gender\_int\_BMI\_mse,

              aes(x = intervention, y = mean, fill = gender)) +

  geom\_bar(stat = "identity", position = dodge) +

  geom\_errorbar(aes(ymin = mean - se,

                    ymax = mean + se),

                position = dodge,

                width  = .6,

                size = .8) +

  coord\_cartesian(ylim = c(22, 26)) +

  ylab("BMI") +

  xlab("Intervention Level") +

  labs(title = "Average BMI Between Intervention Groups",

       subtitle = "Groupped by Gender",

      fill = "Gender") +

  theme\_bw() -> bmi\_3

# Place them together

grid.arrange(arrangeGrob(bmi\_1, bmi\_2, ncol=2), bmi\_3,nrow = 2)

# We can confidently say is the same by gender and with gender as a block

# I can also say that in no difference between treatments but only a test will

# statistically determine, visually it looks the same

# ------------------------ Assumption check for all factors --------------------

# Creating a vector with factors we want to test

factors <- c("BMI","happy","Energy","Fatigue","VitC\_diff","Carot\_diff")

factors\_normality <- list()

# Normality test on all factors by treatment

for (factor in factors){

  factors\_normality[[factor]] <- Normality\_test(intervention[[factor]],

                                                intervention$intervention)

}

# Check result

factors\_normality

# Looks like we are going need to fix this, I will look at normality of each test

# CBD will be my fix, I don't like to transform, I find like forcing something to happen.

# Test homogeneity for all factors

var\_homogenity\_intervention <- list()

for (factor in factors){

  var\_homogenity\_intervention[[factor]] <- bartlett.test(intervention[[factor]] ~

                                              intervention$intervention)

}

# Test homogeneity for gender

var\_homogenity\_gender <- list()

for (factor in factors){

  var\_homogenity\_gender[[factor]] <- bartlett.test(intervention[[factor]] ~

                                              intervention$gender)

}

# check result

var\_homogenity\_intervention

var\_homogenity\_gender

# Not as bad but not perfect.

# ------------------------ Creating a CBD by gender ----------------------------

# Find the distribution of gender for each Treatment

intervention %>% group\_by(intervention, gender) %>%

  summarise(unique = n()) -> gender\_design

# Check the minimum value for the gender block design

min(gender\_design$unique)

# Create a complete block design for gender, couldn't functionalise it, it will

# sorry for the copy paste in future Analysis.

new\_inds <- c()

for (i in 1:6) {

  w <- which(intervention$intervention == gender\_design$intervention[i] &

               intervention$gender == gender\_design$gender[i])

  s <- sample(c(1:length(w)), size = 17, replace = F)

  new\_inds <- c(new\_inds, w[s])

}

cbd\_intervention <- intervention[new\_inds, ]

# Check if worked

cbd\_intervention %>% group\_by(intervention, gender) %>% summarise(n())

# I will also work with replication, as this still not normally distributed on

# each level

# ------------------------ Analysis for Happiness and Gender Interaction -------

# Analysis

aov\_happ <- aov(happy ~ intervention \* gender, data = intervention)

summary(aov\_happ)

# It looks to be an interaction using the full data, we know that this is not

# totally correct by the broken assumptions.

# Checking Normality of Residuals from original ANOVA model

qqnorm(aov\_happ$residuals)

qqline(aov\_happ$residuals)

# Normality test

shapiro.test(aov\_happ$residuals)

# normality is broken and we can't trust the test fully.

# Test multiple times with a CBD

p\_values <- length(500)

for (test in 1:500){

  # Creating a new CBD

  new\_inds <- c()

  for (i in 1:6) {

    w <- which(intervention$intervention == gender\_design$intervention[i] &

                 intervention$gender == gender\_design$gender[i])

    s <- sample(c(1:length(w)), size = 17, replace = F)

    new\_inds <- c(new\_inds, w[s])

  }

  cbd\_intervention <- intervention[new\_inds, ]

  # Testing

  aov\_results <- summary(aov(happy ~ intervention \* gender,

                             data = cbd\_intervention))

  p\_values[test] <- aov\_results[[1]]$`Pr(>F)`[3]

}

# Result of replicates

sum(p\_values < .05)/length(p\_values)

# This is interesting and only about 50% of the time is a difference

# this is due to the different pockets of results of happines across genders

# I will try to make it visually to understand

# Interaction plot

interaction.plot(intervention$intervention, intervention$gender,

                 intervention$happy,

                 xlab = "Intervention Lvl",

                 trace.label = "Gender", ylab = "Happines")

# We can see how it looks to be an interaction, sadly the data is not reliable enough

# To make it a statistical fact.

# Dot plot graph with a Box plot

intervention %>% ggplot(aes(x=intervention, y=happy, fill=gender)) +

  geom\_dotplot(binaxis='y', stackdir='center', position=position\_dodge(.9)) +

  geom\_boxplot(position=position\_dodge(0.8), alpha=.1, show.legend = F)  +

  theme\_bw()

# Dot plot graph with densities and quantiles

intervention %>% ggplot(aes(x=intervention, y=happy, fill=gender)) +

  geom\_violin(alpha=.1, draw\_quantiles = c(0.25, 0.75), linetype = "dashed",

              colour = "red", size = 1, show.legend = F) +

  geom\_violin(alpha=.05, draw\_quantiles = 0.5, size = 1, show.legend = F) +

  geom\_dotplot(binaxis='y', stackdir='center', position=dodge,

               alpha = .7) +

  ylab("Hippnes Lvl") +

  xlab("Intervention") +

  labs(title = "Distribution Result of Happines",

       subtitle = "Groupped by Intervention\nQuantile divisions",

       fill= "Gender") +

  theme\_bw()

# In the dot plot we can see how the Control for gender 1 is really bad distributed

# and cluster on results above the median and below

# ------------------------ Analysis for Energy and BMI interaction -------------

# Create an Ancova test to identify any interaction

ancova\_model <- aov(Energy ~ intervention \* BMI, data = intervention)

summary(ancova\_model)

Anova(ancova\_model, type=3)

# No interaction, I'm going to check normality of residuals to understand the

# reliability of model

# Normality of residuals

shapiro.test(ancova\_model$residuals)

# Residuals seem to be normal distributed

# Test a linear model with each intervention level as a baseline

# I'm doing this to find any statistical significance interaction

mylm <- list()

for (x in levels(intervention$intervention)){

  intervention$intervention <- relevel(intervention$intervention, ref = x)

  mylm[[x]] <- summary(lm(Energy ~ intervention \* BMI, data = intervention))

}

# Check result

mylm

# Seems to be a bit of an interaction with the control compared to EMI, but

# not statistical significant.

# Checking Normality of Residuals from Linear model

qqnorm(mylm[[1]]$residuals, col="blue")

qqline(mylm[[1]]$residuals, col="blue")

# Residuals look normal distributed

hist(mylm[[1]]$residuals, xlab = "Residuals", ylab="Frequency",

     main = "Histogram for Resiguals", col = "blue")

# Test normality of residuals

shapiro.test(mylm[[1]]$residuals)

# Graph for the interaction with Intervention and BMI fro Energy

intervention %>% ggplot(aes(x=BMI, y=Energy, col=intervention)) +

  geom\_point() + geom\_smooth(method='lm', show.legend = F,

                             aes(fill=intervention), alpha = .08) +

  labs(title = "Energy vs BMI",

       subtitle = "Groupped by Intervention Lvl\nWith Regression Lines",

       col="Intervention") +

  theme\_bw()

# Here we can see how visually looks to be an increase on Enegry by BMI on the

# EMI, making it interact different to the other groups. At the same time

# we can see that the error difference make this not a statistical difference.

# ------------------------ Analysis for Fatigue with Gender as Block  ----------

# Analysis

aov\_fatigue <- aov(Fatigue ~ intervention + gender, data = intervention)

summary(aov\_fatigue)

# No really a difference

# Checking Normality of Residuals from original ANOVA model

qqnorm(aov\_fatigue$residuals)

qqline(aov\_fatigue$residuals)

# Normality test

shapiro.test(aov\_happ$residuals)

# The full data is not the best model

# Test multiple times with a CBD

p\_values <- length(500)

for (test in 1:500){

  # Creating a new CBD

  new\_inds <- c()

  for (i in 1:6) {

    w <- which(intervention$intervention == gender\_design$intervention[i] &

                 intervention$gender == gender\_design$gender[i])

    s <- sample(c(1:length(w)), size = 17, replace = F)

    new\_inds <- c(new\_inds, w[s])

  }

  cbd\_intervention <- intervention[new\_inds, ]

  # Testing

  aov\_results <- summary(aov(Fatigue ~ intervention + gender,

                             data = cbd\_intervention))

  p\_values[test] <- aov\_results[[1]]$`Pr(>F)`[1]

}

# Result of replicates

sum(p\_values < .05)/length(p\_values)

# Only 10% of the time is a difference

# ------------------------ Analysis for Energy with Gender as Block  -----------

# Analysis

aov\_energy <- aov(Energy ~ intervention + gender, data = intervention)

summary(aov\_energy)

# Not at all

# Checking Normality of Residuals from original ANOVA model

qqnorm(aov\_energy$residuals)

qqline(aov\_energy$residuals)

# Normality test and broken on full ANOVA test

shapiro.test(aov\_happ$residuals)

# Test multiple times with a CBD

p\_values <- length(500)

for (test in 1:500){

  # Creating a new CBD

  new\_inds <- c()

  for (i in 1:6) {

    w <- which(intervention$intervention == gender\_design$intervention[i] &

                 intervention$gender == gender\_design$gender[i])

    s <- sample(c(1:length(w)), size = 17, replace = F)

    new\_inds <- c(new\_inds, w[s])

  }

  cbd\_intervention <- intervention[new\_inds, ]

  # Testing

  aov\_results <- summary(aov(Energy ~ intervention + gender,

                             data = cbd\_intervention))

  p\_values[test] <- aov\_results[[1]]$`Pr(>F)`[1]

}

# Result of replicates

sum(p\_values < .05)/length(p\_values)

# only a difference 4% of the time if lucky

# ------------------------ Analysis for Vitamin C for FVI group ----------------

# Simplify

FVI\_VitC\_pre <- intervention$VitC\_pre[intervention$intervention == "FVI"]

FVI\_VitC\_post <- intervention$VitC\_post[intervention$intervention == "FVI"]

# Simple paired T-Test

t.test(FVI\_VitC\_pre, FVI\_VitC\_post, paired = TRUE)

# No real difference

# Calculate Effect size

cd <- cohen.d(FVI\_VitC\_pre, FVI\_VitC\_post, paired = T)

# Calculate Power of Analysis

pwr.t.test(n= length(FVI\_VitC\_pre), d= cd$estimate, sig.level=0.05,

           type="paired")

# Really low power

# ------------------------ Analysis for Vitamin C for EMI group ----------------

# Normality check for the group

EMI\_VitC\_pre <- intervention$VitC\_pre[intervention$intervention == "EMI"]

EMI\_VitC\_post <- intervention$VitC\_post[intervention$intervention == "EMI"]

# Simple paired T-Test

t.test(EMI\_VitC\_pre, EMI\_VitC\_post, paired = TRUE)

# Yay, finally we got something, a real statistical difference

# Calculate Effect size

cd <- cohen.d(EMI\_VitC\_pre, EMI\_VitC\_post, paired = T)

# Calculate Power of Analysis

pwr.t.test(n= length(EMI\_VitC\_pre), d= cd$estimate, sig.level=0.05,

           type="paired")

# Might be a difference but is low and only 5% practical difference. Also really

# low power

# ------------------------ Analysis for Cartenoids for FVI group ----------------

# Normality check for the group

FVI\_Cart\_pre <- intervention$Carot\_pre[intervention$intervention == "FVI"]

FVI\_Cart\_post <- intervention$Carot\_post[intervention$intervention == "FVI"]

# Simple paired T-Test

t.test(FVI\_Cart\_pre, FVI\_Cart\_post, paired = TRUE)

# No difference at alpha 5%, close to be

# Calculate Effect size

cd <- cohen.d(FVI\_Cart\_pre, FVI\_Cart\_post, paired = T)

# Calculate Power of Analysis

pwr.t.test(n= length(FVI\_Cart\_pre), d= cd$estimate, sig.level=0.05,

           type="paired")

# The difference still low and power is only 41%, not perfect

# ------------------------ Analysis for Cartenoids for EMI group ----------------

# Normality check for the group

EMI\_Cart\_pre <- intervention$Carot\_pre[intervention$intervention == "EMI"]

EMI\_Cart\_post <- intervention$Carot\_post[intervention$intervention == "EMI"]

# Wilcox Test

wilcox.test(EMI\_Cart\_pre, EMI\_Cart\_post, paired = TRUE)

# No difference by the non parametric test

# Calculate Effect size

cd <- cohen.d(EMI\_Cart\_pre, EMI\_Cart\_post, paired = T)

# Calculate Power of Analysis

pwr.t.test(n= length(EMI\_Cart\_pre), d= cd$estimate, sig.level=0.05,

           type="paired")

# ------------------------ Data prep for dFV across day 1, 7 and 14 ------------

# Sorry I'm more of a pandas thinker, Identify which ID had 14 days of diary

diary %>% group\_by(ID) %>% summarise(total = n()) -> diary\_weeks

# Only 44 out of the 171, that's a really big loss 75% !!!!

# Select only the ID with 14 days

diary\_weeks <- diary\_weeks[diary\_weeks$total == 14,]

# Now we collect the ID values

diary\_14 <- diary[diary$ID %in% diary\_weeks$ID, ]

# Now comes the fun, identify their Intervention level

diary\_14 <- left\_join(diary\_14, intervention[,c("ID","intervention")])

diary\_14 <- na.omit(diary\_14)

# Lets hope is balanced enough

diary\_14 %>% group\_by(ID, intervention) %>%

  summarise(total = n()) %>%

  group\_by(intervention) %>%

  summarise(total = n()) -> diary\_design

# Now we keep only the days we are interested 1, 7, 14

days <- c(1,7,14)

# First I save a new variable for days 1 to 14, I will use at the end

diary\_1\_to\_14 <- diary\_14

# Now only keep days 1, 7, 14

diary\_14 <- diary\_14[diary\_14$Diary %in% days,]

# Vuala clean data

diary\_14

# ------------------------ Assumptions check for dFV across day 1, 7 and 14 ----

# Assumptions check, yay!

diary\_normality <- list()

# Normality test on all factors by treatment

for (level in unique(diary\_14$intervention)){

  diary\_normality[[level]] <- Normality\_test(diary\_14$dFV[diary\_14$intervention == level],

                                             diary\_14$Diary)

}

# Check results

diary\_normality

# Not as bad, we are working with paired data so the Anova test will check shpericity

# Difference of variance

diary\_homogenity <- list()

for (level in unique(diary\_14$intervention)){

  diary\_homogenity[[level]] <- bartlett.test(diary\_14$dFV[diary\_14$intervention == level],

                                               diary\_14$Diary[diary\_14$intervention == level])

}

# Check Results

diary\_homogenity

# Check for outlier

diary\_14 %>% group\_by(intervention, Diary) %>%

  identify\_outliers(dFV)

# Last steps is just to understand the data a bit more but probably not fully needed

# ------------------------ Analysis for dFV across day 1, 7 and 14 -------------

# Boxplot

diary\_14 %>% ggplot(aes(x = Diary, y = dFV, col = intervention)) +

  geom\_boxplot()

# Is a paired data so we can't fully asume by this visuals, but gives an idea

# Point Plot and Linear models

diary\_14 %>% ggplot(aes(x = Diary, y = dFV, fill = intervention)) +

  geom\_dotplot(binaxis='y', stackdir='center', position=position\_dodge(.9),

               alpha = .7) + geom\_boxplot(alpha = .1, width = .2,

                                          position=position\_dodge(.9),

                                          show.legend = F) +

  facet\_wrap(vars(intervention)) +

  xlab("Day of Recording") +

  labs(title = "dFV Consumption per Unit",

       subtitle = "Across the [1, 7, 14] days",

       fill= "Intervention") +

  theme\_bw()

# Paired ANOVA

aov\_diary <- list()

for (level in unique(diary\_14$intervention)){

  diary\_14[diary\_14$intervention == level,] %>%

    anova\_test(dv = dFV, wid = ID, within = c(Diary)) -> aov\_diary[[level]]

}

# Check results

aov\_diary

# Nothing statistical, Control stayed almos the same, FVI had a difference but not

# statistically, if we had more subjects maybe it could be. Sphericity broken in all

# cases but not that badly

# Get tables

for (level in aov\_diary){

  print(get\_anova\_table(level))

}

# Differences by pairwise T-Test

diary\_14 %>% group\_by(intervention) %>%

    pairwise\_t\_test(dFV ~ Diary, paired = TRUE,

                    p.adjust.method = "bonferroni")

# just to have an idea between changes, FVI change increase between day 1 and 14

# we can also see a reduction on consumption on the EMI group. But sadly again

# not statistically, more subjects next time or rigurosity as we lost so many subjects.

# Differences per day on intervention by Linear Model

mylm <- list()

for (level in unique(diary\_14$intervention)){

  summary(lm(data = diary\_14[diary\_14$intervention == level,],

             dFV ~ Diary)) -> mylm[[level]]

}

# Check results

mylm

# No differences statistically significant, but we can see the tendencies of before

# in the coefficients, FVI consumption increases and EMI decreases

# Visual for 1, 7 and 14 days

diary\_14$Day <- as.numeric(diary\_14$Diary)

diary\_14 %>%

  ggplot(aes(x=Day, y=dFV, col=intervention)) +

  geom\_point() + geom\_smooth(method='lm', show.legend = F,

                             aes(fill=intervention), alpha = .08) +

  labs(title = "dFV vs Days [1 - 7 - 14]",

       subtitle = "Groupped by Intervention Lvl\nWith Regression Lines",

       col="Intervention") +

  scale\_x\_continuous(breaks=c(1,7,14)) +

  theme\_bw()

# Here the standard errors show why we can't see any statistical difference

# -------------------- Analysis of Average consumption of dFV per treatment ----

# Create an average total consumption per ID

diary\_1\_to\_14 %>%

  group\_by(ID) %>%

  summarise(dFV\_avg = mean(dFV),

            intervention = unique(intervention)) -> diary\_avg

# Check normality

Normality\_test(diary\_avg$dFV\_avg, diary\_avg$intervention)

#Yay finally something normal

# Check Variance homogenity

bartlett.test(diary\_avg$dFV\_avg, diary\_avg$intervention)

# Great news, all assumptions pass

# Anova

diary\_avg\_aov <- aov(data = diary\_avg, dFV\_avg ~ intervention)

#Almost

summary(diary\_avg\_aov)

# thats close to be a difference with a p-value of 0.08

# Lm for differences

lm\_avg\_dFV <- lm(data = diary\_avg, dFV\_avg ~ intervention)

summary(lm\_avg\_dFV) # Interesting, but as ANOVA was used I will report Tukey

# This is interesting, here we have a statistical significant difference with the

# FVI and Control with a p-value of 0.288

# Normality Model

shapiro.test(lm\_avg\_dFV$residuals)

# Residuals are normal distributed

# Tukey test

tukey\_ad<- TukeyHSD(diary\_avg\_aov)

# Now we don't have a difference, most liklely this is with the adjusment on the

# Tukey test, I will report on this finding as we follwed an ANOVA test with no

# assumptions broken

# lets visualise

par(mar = c(1, 6, 2, 1)+ 2)

plot(tukey\_ad, las = 1)

# Create a table with Average and Margin of Error

diary\_avg %>% group\_by(intervention) %>%

  summarise(mean = mean(dFV\_avg),

            se = sd(dFV\_avg)/sqrt(length(dFV\_avg))) -> diary\_dFV\_mse

# Create a bar plot with error bars

ggplot(diary\_dFV\_mse,

       aes(x = intervention, y = mean, fill = intervention)) +

  geom\_bar(stat = "identity", position = dodge) +

  geom\_errorbar(aes(ymin = mean - se,

                    ymax = mean + se),

                position = dodge,

                width  = .6,

                size = .8) +

  coord\_cartesian(ylim = c(2, 4)) +

  ylab("Average consumption of dFV") +

  xlab("Intervention Level") +

  labs(title = "Average dFV Between Intervention Groups",

       subtitle = "Total consumption for all 14 days",

       fill = "Intervention") +

  theme\_bw() +

  theme(plot.subtitle=element\_text(size=9)) -> dFV\_1

# Line plot for all 14 days by Intervention

diary\_1\_to\_14 %>% group\_by(intervention, Diary) %>% summarise(Total = mean(dFV)) %>%

  ggplot(aes(x=Diary, y=Total, col=intervention, group=intervention)) + geom\_line() +

  geom\_point() +

  ylab("Average dFV Consumption") +

  xlab("Day of Recording") +

  labs(title = "Average Daily dFV Consumption",

       subtitle = "Across the 14 days\nGroupped by Intervention") +

  theme\_bw() -> dFV\_2

# Place them together

grid.arrange( dFV\_1, dFV\_2,nrow = 1)

# Here seems to be a difference as is no adjusments, is also interesting to see

# the daily averages, more like a time series analysis.

# --------------------- Analysis of Happiness with Ancestry and Interactions----

## I won't include this on the Report but I leave it here

# Find the distribution of gender for each Treatment

intervention %>% group\_by(intervention, ancestry) %>%

  summarise(unique = n()) -> ancestry\_design

# We will have to create a new complet block desing by ancestry, with a level of

# 15 units

# Analysis

aov\_happ\_anc <- aov(happy ~ intervention \* ancestry, data = intervention)

summary(aov\_happ\_anc)

# No difference on full data

# Checking Normality of Residuals from original ANOVA model

qqnorm(aov\_happ\_anc$residuals)

qqline(aov\_happ\_anc$residuals)

# Normality is broken adain

shapiro.test(aov\_happ\_anc$residuals)

# Test multiple times with a CBD

p\_values <- length(500)

for (test in 1:500){

  # Creating a new CBD

  new\_inds <- c()

  for (i in 1:6) {

    w <- which(intervention$intervention == ancestry\_design$intervention[i] &

                 intervention$ancestry == ancestry\_design$ancestry[i])

    s <- sample(c(1:length(w)), size = 15, replace = F)

    new\_inds <- c(new\_inds, w[s])

  }

  cbd\_intervention <- intervention[new\_inds, ]

  # Testing

  aov\_results <- summary(aov(happy ~ intervention + ancestry,

                             data = cbd\_intervention))

  p\_values[test] <- aov\_results[[1]]$`Pr(>F)`[1]

}

# Result of replicates

sum(p\_values < .05)/length(p\_values)

# No difference at all, maybe only one time in the 500 tests

# Interaction plot

interaction.plot(intervention$intervention, intervention$ancestry,

                 intervention$happy,

                 xlab = "Intervention Lvl",

                 trace.label = "Gender", ylab = "Happines")

# not much really, more drastic the lower on happiness by the gender 0 on control group