AMS 572 Group 2 project

Matt, Priyansh, Junhao, Saurabh, Amy

Introduction

Mental Health is an important aspect of our lives that all of us should take care of. It influences how we think, feel and act as we cope with life. It also helps determine how we handle stress, relate to others, and take decisions. Therefore, mental health is significant at every stage of life, from childhood and adolescence through adulthood and aging. Mental Health is closely connected to Physical Health. It plays a major role in people's ability to maintain good physical health. Metal illnesses affects people's ability to participate in health-promoting behaviors. We have collected a dataset for two different types of Therapy, Cognitive Behavioral Therapy and Psychodynamic Psychotherapy. Cognitive Behavioral Therapy is a psychotherapy in which negative patterns of thought about self and world are challenged in order to alter unwanted behavior patterns or treat mood disorders such as depression. CBT ultimately aims to teach patients to be their own therapist, by helping them to understand their current ways of thinking and behaving, and by equipping them with the tools to change their maladaptive cognitive and behavioral patterns. Psychodynamic psychotherapy involves the interpretation of mental and emotional processes rather than focusing on behavior. Psychodynamic therapists attempt to help patients find patterns in their emotions, thoughts and beliefs in order to gain insight into their current self.

Our dataset (https://teachpsychscience.org/index.php/design-data-which-therapy-is-best-for-treating-eating-disorders-mixed-design-and-mixed-anova/432/research-methods/) includes 63 participants who went through either one of the therapies. 32 participants who went through the Cognitive Behavioral Therapy and 31 participants for Psychodynamic Psychotherapy. From our dataset we wanted to know how do the treatment types and different emotional triggers such a Relationship, Work or Future impacts disordered eating symptoms. After the participants had been treated by respective therapies, they were exposed to the three different emotional triggers (Relationship, Work, Future) while having a bowl of grapes placed in front of them. For getting inference on the effectiveness of the therapies, the data on number of grapes consumed by each was gathered, also the distance at which the participants moved the bowl of grapes away from them. The data was collected for participants of 5 different Ethnicities, which are Asian or Eastern Ancestry, Black or African Ancestry, Hispanic or Spanish Ancestry, European Ancestry or Other. Considering the gender of the participants, our data only includes three male participants as to 60 female participants.

The age of the participants ranges from 18 years to 40 years with the majority of participants falling in the range of 18 to 30 years of age. The final column of the data

represents the mean of overall grapes consumed by a single participant under each emotional trigger.

Aim

In this dataset we plan to perform the following hypothesis:
1. Hypothesis 1: Eating habits of both the therapies are same
2. Hypothesis 2: Mixed Design Analysis of Variance
i. Effect of therapy had same eating patterns on all three stress situations
ii.There is an interaction between therapy and distance of the bowl on all three stress situations

Along with these hypotheses, we also generate missing values in the data at random (MCAR) and non-random manner (MNAR), and discuss the effect of two types of missing values on the data analysis.

Methodology

The dataset contains 63 rows of participants data and 12 feature columns - participant identification (*Participant*), type of therapy (*Therapy*), number of grapes eaten after being exposed to relationship stress (*Rel_grapes*), number of grapes eaten after being exposed to relationship stress (*Work_grapes*), number of grapes eaten after being exposed to relationship stress (*Future_grapes*), distance of the bowl after being exposed to relationship stress (*Rel_distance*), distance of the bowl after being exposed to relationship stress (*Work_distance*), distance of the bowl after being exposed to relationship stress (*Future_distance*), Gender, Age, Ethnicity, mean of the grapes eaten after being exposed to all three stresses (*x_OverallGrapes*).

Loading the data and required libraries

The data was in .sav format and the *foreign* package was utilized to load the data. We also utilized *mice* and *missMethods* packages for missing values imputation and generation, *tidyverse*,*dplyr*,*rstatix* for easy handling and tidying of data, *ggpubr* for plotting, and *WRS2* for model analysis.

Hypothesis 1 : Eating habits of both the therapies are same

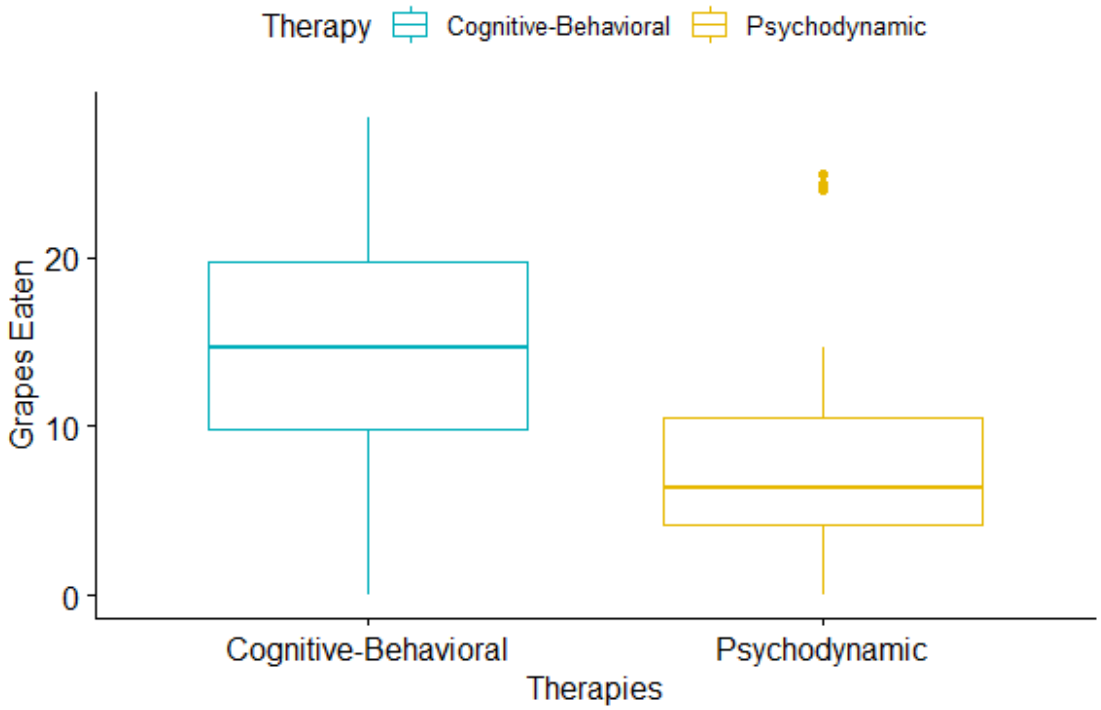
The participants completing Cognitive-Behavioral therapy will eat the same amount of grapes as those completing Psychodynamic therapy. This can be tested using the two-sample t-test on the overall grapes eaten on both the therapies. But to use this test, we need to check whether the samples are normally distributed and check whether the variance of both the samples are equal or not. These tests and first hypothesis will be performed on the dataset and then on two kinds of missing dataset.

Case 1: Using complete data

Summary Statistics of the Data

```
## # A tibble: 2 x 4
##   Therapy              count  mean    sd
##   <fct>                <int> <dbl> <dbl>
## 1 Cognitive-Behavioral    32 14.4   7.46
## 2 Psychodynamic           31  8.27  6.54
```

Visualize the Data



Normality Assumption

Null Hypothesis: The data are normally distributed.
Alternative Hypothesis: The data are not normally distributed.
Significance Level: 0.05

```
##
##  Shapiro-Wilk normality test
##
## data:  x_OverallGrapes[Therapy == "Cognitive-Behavioral"]
## W = 0.97525, p-value = 0.6546

##
##  Shapiro-Wilk normality test
##
## data:  x_OverallGrapes[Therapy == "Psychodynamic"]
## W = 0.8531, p-value = 0.0005903
```

With respect to Cognitive-Behavioral therapy (1), since the p-value = 0.6546 > 0.05 = alpha, we fail to reject the null hypothesis and conclude that the data are normally distributed according to the Shapiro-Wilk normality test.

With respect to Psychodynamic therapy (2), since the p-value = 0.0005903 < 0.05 = alpha, we reject the null hypothesis and conclude that the data are not normally distributed according to the Shapiro-Wilk normality test.

However, since the sample sizes of Cognitive-Behavioral therapy and Psychodynamic therapy are 32 and 31 respectively, both of these values exceed 30 (n > 30), therefore the Central Limit Theorem can be applied to these samples. Due to this, it can be assumed that both therapy samples from the data approach the normal distribution and are therefore normally distributed.

Homogeneity in Variance Assumption

Now that we have verified that both of the therapy samples are normally distributed, we must check for homogeneity in their variances using the F-test.

Null Hypothesis: The two therapy populations have same variances.
Alternative Hypothesis: The two therapy populations have different variances.
Significance Level: 0.05

```
##
##  F test to compare two variances
##
## data:  x_OverallGrapes by Therapy
## F = 1.3022, num df = 31, denom df = 30, p-value = 0.4718
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.6303919 2.6784453
## sample estimates:
## ratio of variances
##           1.302223
```

Since the p-value = 0.4718 > 0.05 = alpha, we fail to reject the null hypothesis and conclude that there is no significant difference between the variances of the two therapy populations.

Unpaired Two-Samples t-test

Null Hypothesis: Those completing Cognitive-Behavioral therapy will eat the same amount of grapes as those completing Psychodynamic therapy (equal means).
Alternative Hypothesis: Those completing Cognitive-Behavioral therapy will not eat the same amount of grapes as those completing Psychodynamic therapy (unequal means).
Significance Level: 0.05

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

```
## data:  x_OverallGrapes by Therapy
## t = 3.4788, df = 61, p-value = 0.0009354
## alternative hypothesis: true difference in means between group Cognitive-
Behavioral and group Psychodynamic is not equal to 0
## 95 percent confidence interval:
##  2.618493 9.698040
## sample estimates:
## mean in group Cognitive-Behavioral        mean in group Psychodynamic
##                           14.427083                           8.268817
```

Since the p-value = 0.0009354 < 0.05 = alpha, we reject the null hypothesis and conclude that the means of grapes eaten by the two therapy populations are different.

Case 2: Using MCAR data

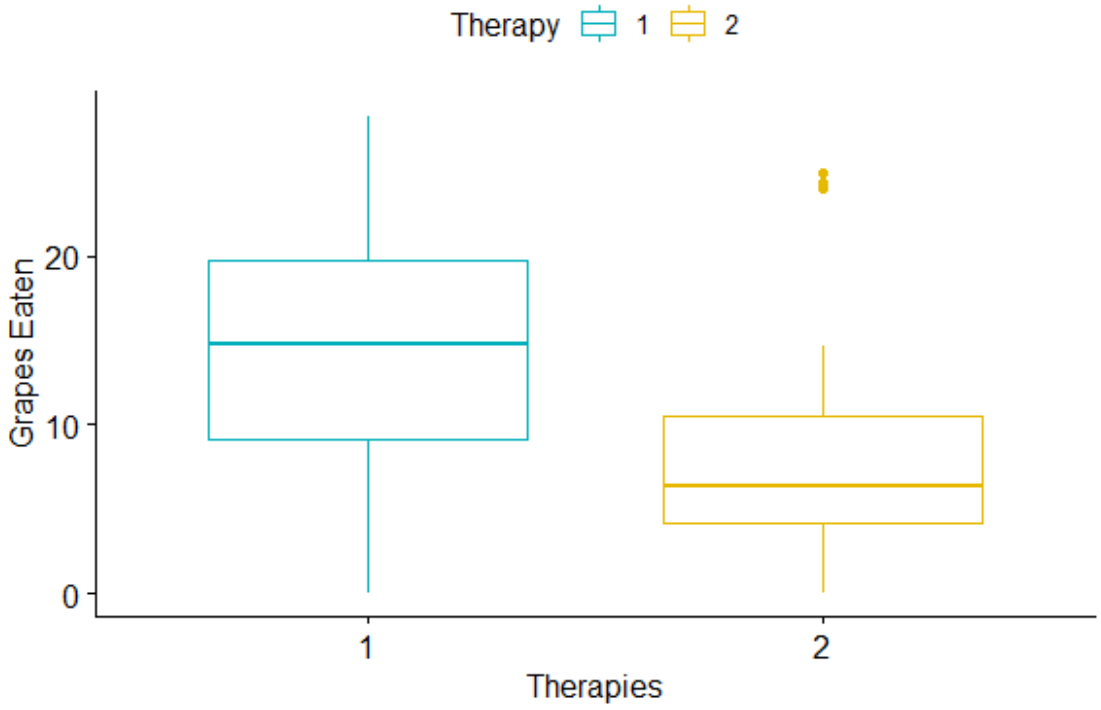
Missing values generated Completely at Random (MCAR)

Since we had a dataset with no missing values we generate missing values completely at random and use the Predictive mean matching (pmm) method to impute the misssing values from the *MICE* package.

Summary Statistics of the Data

```
## # A tibble: 2 x 4
##   Therapy count  mean    sd
##     <dbl> <int> <dbl> <dbl>
## 1       1    32 14.4   7.54
## 2       2    31  8.26  6.55
```

Visualize the Data



Normality Assumption

Null Hypothesis: The data are normally distributed.
Alternative Hypothesis: The data are not normally distributed.
Significance Level: 0.05

```
##
##  Shapiro-Wilk normality test
##
## data:  x_OverallGrapes[Therapy == 1]
## W = 0.97414, p-value = 0.6205

##
##  Shapiro-Wilk normality test
##
## data:  x_OverallGrapes[Therapy == 2]
## W = 0.8526, p-value = 0.0005758
```

With respect to Cognitive-Behavioral therapy (1), since the p-value = 0.6205 > 0.05 = alpha, we fail to reject the null hypothesis and conclude that the data are normally distributed according to the Shapiro-Wilk normality test.

With respect to Psychodynamic therapy (2), since the p-value = 0.0005758 < 0.05 = alpha, we reject the null hypothesis and conclude that the data are not normally distributed according to the Shapiro-Wilk normality test.

However, since the sample sizes of Cognitive-Behavioral therapy and Psychodynamic therapy are 32 and 31 respectively, both of these values exceed 30 (n > 30), therefore the Central Limit Theorem can be applied to these samples. Due to this, it can be assumed that both therapy samples from the data approach the normal distribution and are therefore normally distributed.

Homogeneity in Variance Assumption

Now that we have verified that both of the therapy samples are normally distributed, we must check for homogeneity in their variances using the F-test.

Null Hypothesis: The two therapy populations have same variances.
Alternative Hypothesis: The two therapy populations have different variances.
Significance Level: 0.05

```
##
##  F test to compare two variances
##
## data:  x_OverallGrapes by Therapy
## F = 1.3273, num df = 31, denom df = 30, p-value = 0.4404
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.642533 2.730031
## sample estimates:
## ratio of variances
##           1.327303
```

Since the p-value = 0.4404 > 0.05 = alpha, we fail to reject the null hypothesis and conclude that there is no significant difference between the variances of the two therapy populations.

Unpaired Two-Samples t-test

Null Hypothesis: Those completing Cognitive-Behavioral therapy will eat the same amount of grapes as those completing Psychodynamic therapy (equal means).
Alternative Hypothesis: Those completing Cognitive-Behavioral therapy will not eat the same amount of grapes as those completing Psychodynamic therapy (unequal means).
Significance Level: 0.05

```
##
##  Two Sample t-test
##
## data:  x_OverallGrapes by Therapy
## t = 3.4569, df = 61, p-value = 0.001001
## alternative hypothesis: true difference in means between group 1 and group
2 is not equal to 0
## 95 percent confidence interval:
##  2.596157 9.721047
## sample estimates:
```

```
## mean in group 1 mean in group 2
##       14.416667        8.258065
```

Since the p-value = 0.001001 < 0.05 = alpha, we reject the null hypothesis and conclude that the means of grapes eaten by the two therapy populations are different.

Case 3: Using MNAR data

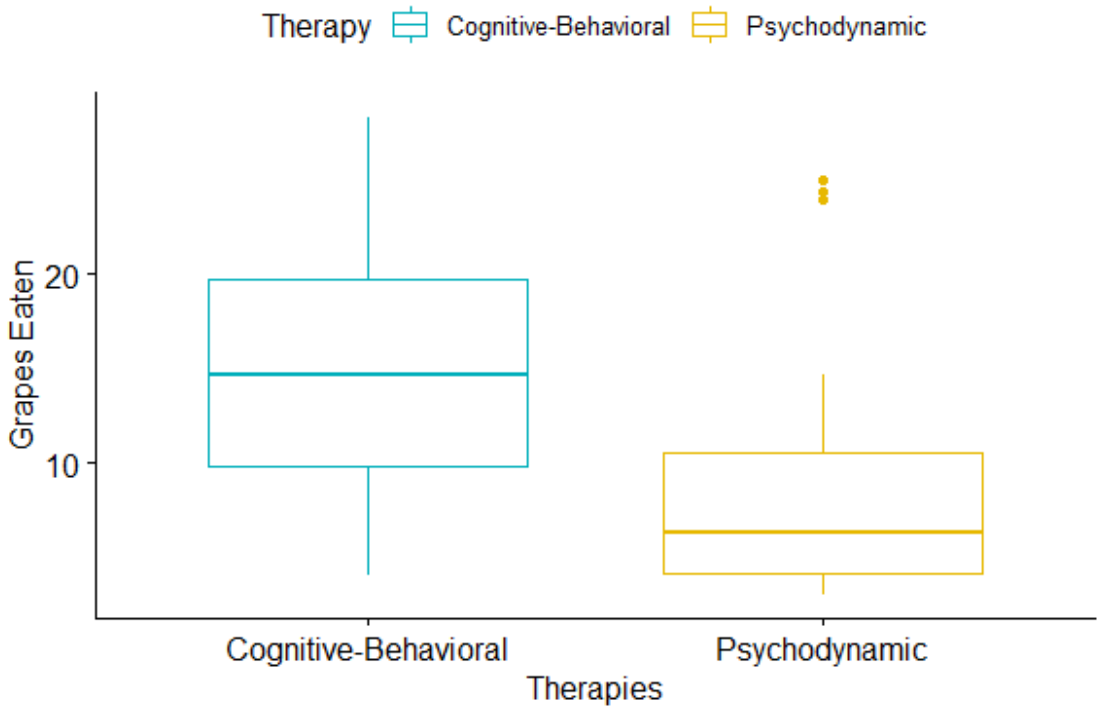
Missing values generated not at Random (MNAR)

Here we generate missing values not at random, that is these missing values will be non-ignorable ones. We use *delete_MNAR_censoring* from *missMethods* to generate these missing values and use the Predictive mean matching (pmm) method to impute the missing values from the *MICE* package.

Summary Statistics of the Data

```
## # A tibble: 2 x 4
##   Therapy              count  mean    sd
##   <fct>                <int> <dbl> <dbl>
## 1 Cognitive-Behavioral    32 14.7   7.01
## 2 Psychodynamic           31  8.60  6.18
```

Visualize the Data



Normality Assumption

Null Hypothesis: The data are normally distributed.
Alternative Hypothesis: The data are not normally distributed.
Significance Level: 0.05

```
##
##  Shapiro-Wilk normality test
##
## data:  x_OverallGrapes[Therapy == "Cognitive-Behavioral"]
## W = 0.95747, p-value = 0.2339

##
##  Shapiro-Wilk normality test
##
## data:  x_OverallGrapes[Therapy == "Psychodynamic"]
## W = 0.78208, p-value = 2.447e-05
```

With respect to Cognitive-Behavioral therapy (1), since the p-value = 0.2339 > 0.05 = alpha, we fail to reject the null hypothesis and conclude that the data are normally distributed according to the Shapiro-Wilk normality test.

With respect to Psychodynamic therapy (2), since the p-value = 2.447e-05 < 0.05 = alpha, we reject the null hypothesis and conclude that the data are not normally distributed according to the Shapiro-Wilk normality test.

However, since the sample sizes of Cognitive-Behavioral therapy and Psychodynamic therapy are 32 and 31 respectively, both of these values exceed 30 (n > 30), therefore the Central Limit Theorem can be applied to these samples. Due to this, it can be assumed that both therapy samples from the data approach the normal distribution and are therefore normally distributed.

Homogeneity in Variance Assumption

Now that we have verified that both of the therapy samples are normally distributed, we must check for homogeneity in their variances using the F-test.

Null Hypothesis: The two therapy populations have same variances.
Alternative Hypothesis: The two therapy populations have different variances.
Significance Level: 0.05

```
##
##  F test to compare two variances
##
## data:  x_OverallGrapes by Therapy
## F = 1.2892, num df = 31, denom df = 30, p-value = 0.4888
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.6240896 2.6516677
## sample estimates:
```

```
## ratio of variances
##           1.289204
```

Since the p-value = 0.4888 > 0.05 = alpha, we fail to reject the null hypothesis and conclude that there is no significant difference between the variances of the two therapy populations.

Unpaired Two-Samples t-test

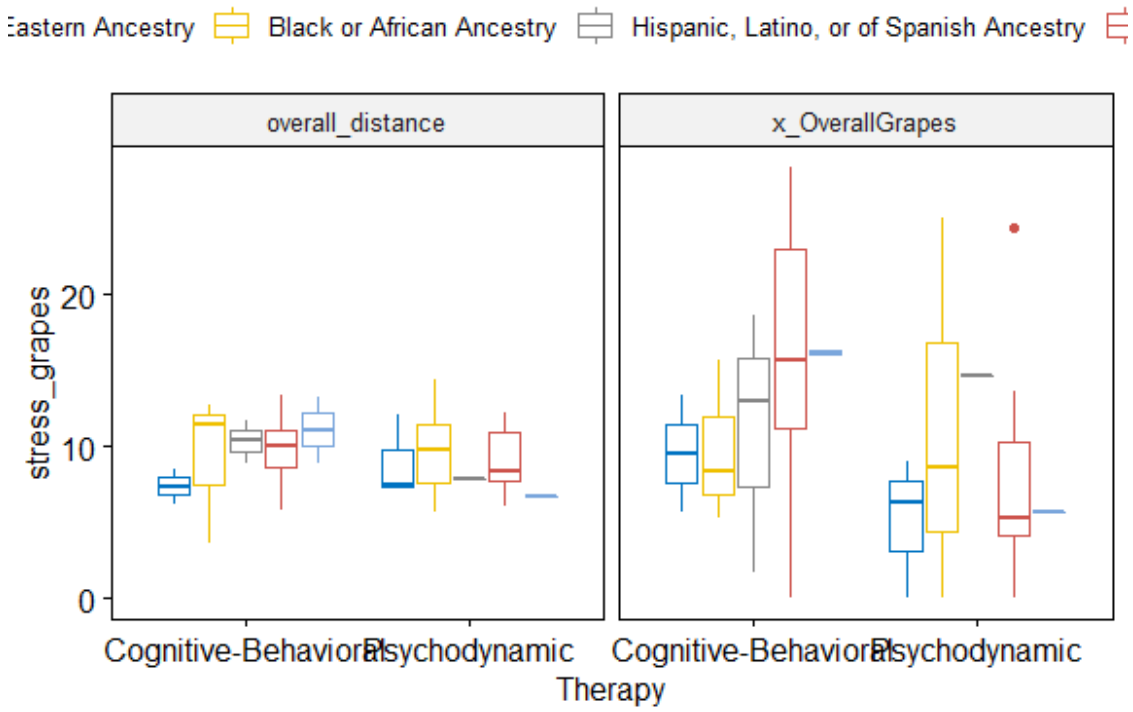
Null Hypothesis: Those completing Cognitive-Behavioral therapy will eat the same amount of grapes as those completing Psychodynamic therapy (equal means).
Alternative Hypothesis: Those completing Cognitive-Behavioral therapy will not eat the same amount of grapes as those completing Psychodynamic therapy (unequal means).
Significance Level: 0.05

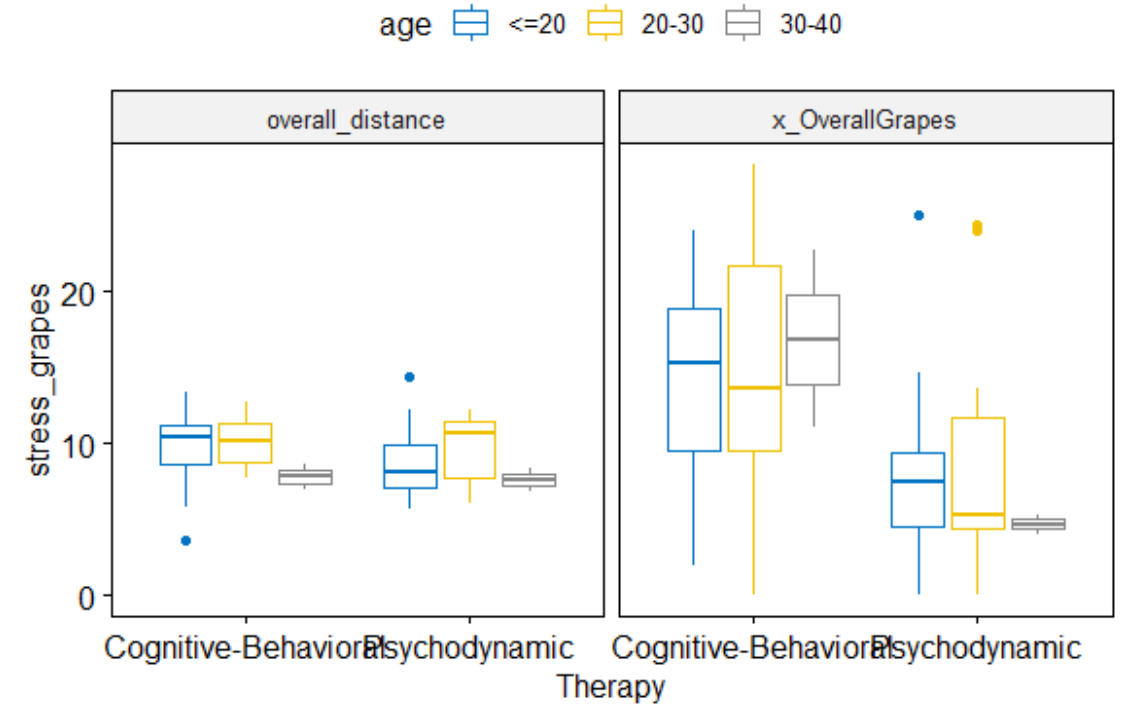
```
##
##  Two Sample t-test
##
## data:  x_OverallGrapes by Therapy
## t = 3.6497, df = 61, p-value = 0.0005461
## alternative hypothesis: true difference in means between group Cognitive-
Behavioral and group Psychodynamic is not equal to 0
## 95 percent confidence interval:
##  2.751255 9.419444
## sample estimates:
## mean in group Cognitive-Behavioral        mean in group Psychodynamic
##                           14.687500                           8.602151
```

Since the p-value = 0.0005461 < 0.05 = alpha, we reject the null hypothesis and conclude that the means of grapes eaten by the two therapy populations are different.

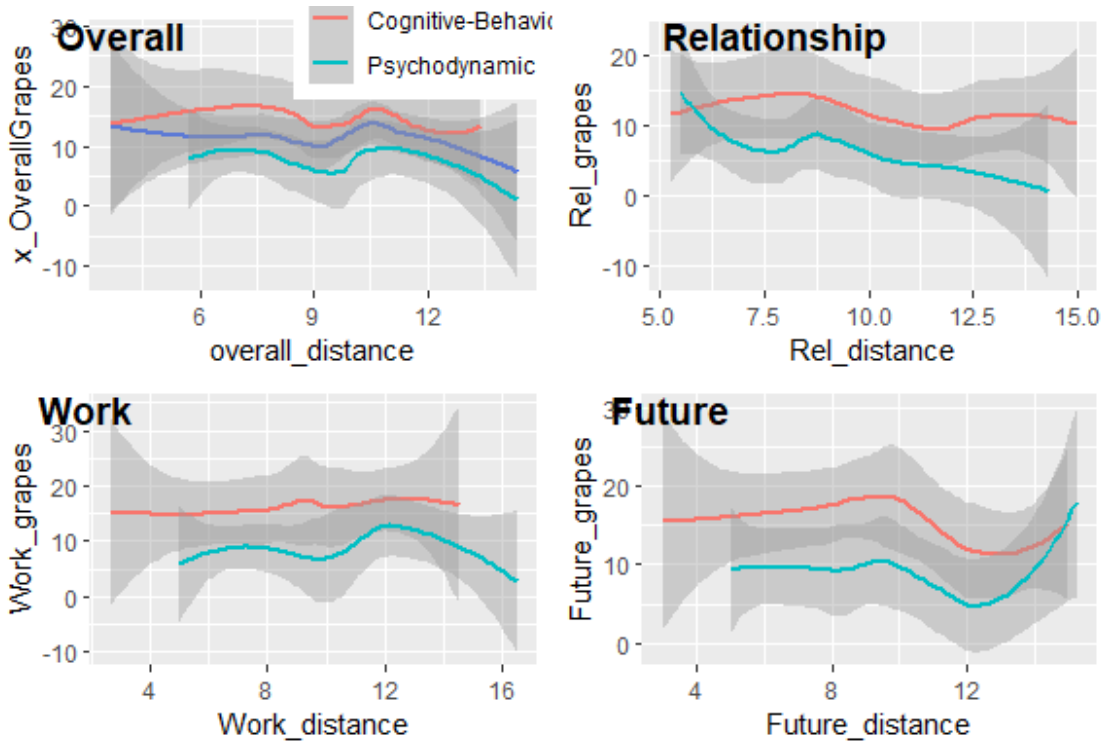
Visualizing the data

In the following plot we notice that there is no pattern for age and ethnicity is the number of grapes eaten or the ditance of the bowl.





We plan to do a mixed design model where we check if there is a relation between the kind of stress and therapy undergone based on the number of grapes eaten and the distance of the bowl. We now check if there is any relationship or dependence between number of grapes eaten and the distance of the bowl.



From the plots, we notice that there is no noticeable relationship between number of grapes eaten and the distance of the bowl for participants exposed to each stresses and the overall stress. Therefore, in our study we will analyze effects of number of grapes eaten and distance of the bowl separately.

Hypothesis 2 : Mixed Design Analysis of Variance Model

A mixed-design analysis of variance model, also known as a split-plot ANOVA, is used to test for differences between two or more independent groups whilst subjecting participants to repeated measures. Thus, in a mixed-design ANOVA model, one factor (a fixed effects factor) is a between-subjects variable, which have independent categories and the other (a random effects factor) is a within-subjects variable which have related categories also known as repeated measures. Thus, overall, the model is a type of mixed-effects model.

We use two-way anova to analyze th effect of therapy on all the different stress situations. This model has the following assumptions:
1. **No significant outliers** in any cell of the design
2. **Normality:** the outcome (or dependent) variable should be approximately normally distributed in each cell of the design
3. **Homogneity of variance assumption**
4. **Assumption of sphericity:** the variance of the differences between within-subjects groups should be equal

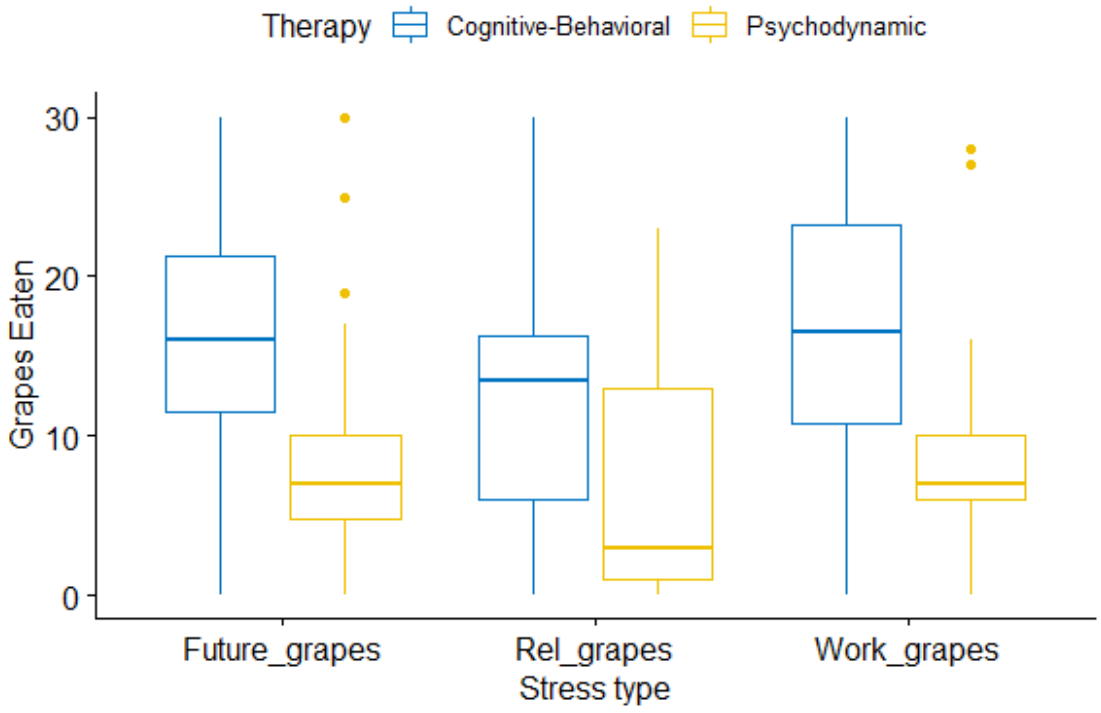
Case 1: Using complete data

Grapes eaten due to stress

Here we have the two therapies as between groups effects and the three stresses as within group effects. Using this method we check whether there was any noticeable difference in the number of grapes eaten to the two therapies participants after being exposed to each of the stresses. The following graph shows that the participants receiving Psychodynamic treatment have a smaller median and inter-quartile range compared to its counterpart and two of them(future and work) have several outliers which can be verified using identify_outliers() command.

Assumptions

Checking for outliers

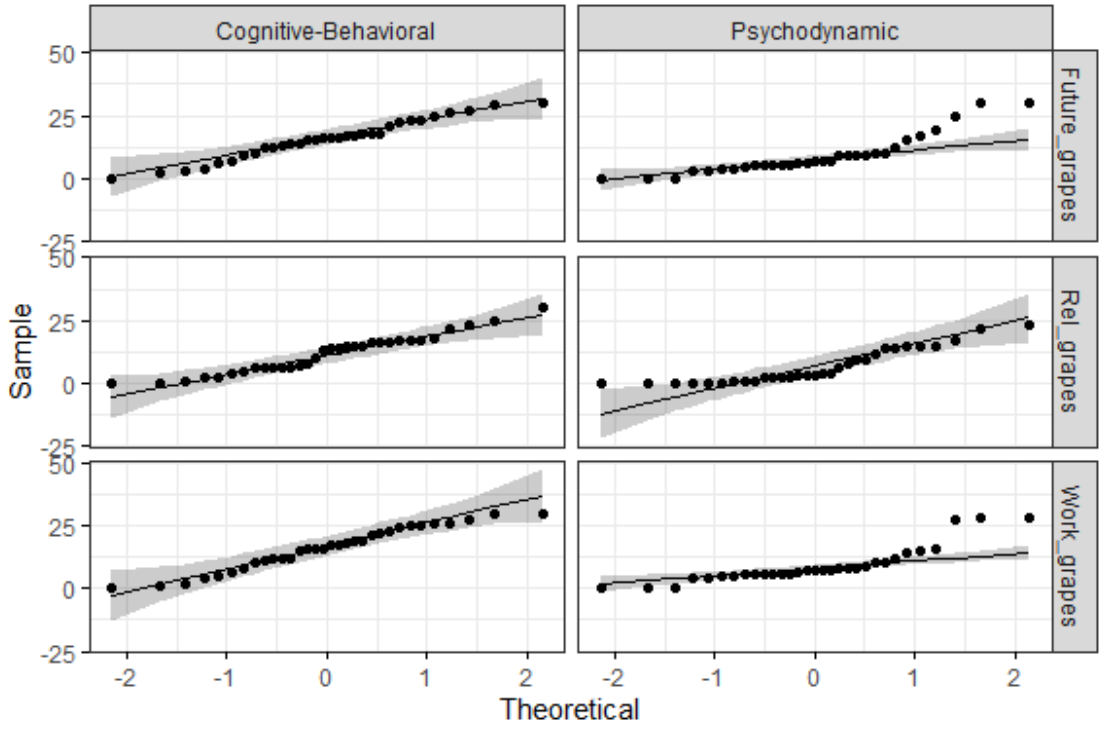



```
## # A tibble: 7 x 6
##   Therapy       stress_type   Participant stress_grapes is.outlier
is.extreme
##   <fct>         <fct>         <fct>               <dbl> <lgl>      <lgl>
## 1 Psychodynamic Future_grapes 58                     19 TRUE       FALSE
## 2 Psychodynamic Future_grapes 44                     30 TRUE       TRUE
## 3 Psychodynamic Future_grapes 60                     30 TRUE       TRUE
## 4 Psychodynamic Future_grapes 9                      25 TRUE       FALSE
## 5 Psychodynamic Work_grapes   44                     27 TRUE       TRUE
## 6 Psychodynamic Work_grapes   60                     28 TRUE       TRUE
## 7 Psychodynamic Work_grapes   9                      28 TRUE       TRUE
```

We notice that there are a few ouliers using the identify_outliers()but we still move forward with the analysis with the outliers.

Normality

We see that all three stresses receiving Cognitive-Behavioral treatment and the relationship stress for Psychodynamic treatment fall approximately along the reference line, but work and future stress are not normal for Psychodynamic treatment. Therefore, we log transform the data. But the shapiro_test() inspired from the Shapiro-Wilk test on the log transformed data shows that p-values are less than 0.05. Therefore, the normality assumption is still not satisfied. Therefore, we do robust ANOVA on the original data from the WRS2 package.




```
## # A tibble: 6 x 5
##   Therapy              stress_type   variable      statistic         p
##   <fct>                <fct>         <chr>             <dbl>     <dbl>
## 1 Cognitive-Behavioral Future_grapes stress_grapes     0.818 0.0000902
## 2 Psychodynamic        Future_grapes stress_grapes     0.906 0.0100
## 3 Cognitive-Behavioral Rel_grapes    stress_grapes     0.876 0.00163
## 4 Psychodynamic        Rel_grapes    stress_grapes     0.915 0.0173
## 5 Cognitive-Behavioral Work_grapes   stress_grapes     0.816 0.0000829
## 6 Psychodynamic        Work_grapes   stress_grapes     0.843 0.000354
```

Homogeneity of variance

The homogeneity of variances is verified by Levene's test ($p > 0.05$).

```
## # A tibble: 3 x 5
##   stress_type     df1   df2 statistic      p
##   <fct>         <int> <int>     <dbl>  <dbl>
## 1 Future_grapes     1    61     0.359 0.551
## 2 Rel_grapes        1    61     0.632 0.430
## 3 Work_grapes       1    61     3.19  0.0792
```

Homogeneity of covariance

There was homogeneity of covariances, as assessed by Box's test of equality of covariance matrices ($p > 0.001$).

```
## # A tibble: 1 x 4
##   statistic p.value parameter method
##       <dbl>   <dbl>     <dbl> <chr>
## 1      1.12   0.289         1 Box's M-test for Homogeneity of Covariance
Matric~
```

Assumption of sphericity

The assumption of sphericity will be automatically checked during the computation of the ANOVA test using the R function anova_test(). The Mauchly's test is internally used to assess the sphericity assumption. By using the function get_anova_table() to extract the ANOVA table, the Greenhouse-Geisser sphericity correction is automatically applied to factors violating the sphericity assumption

Two-way mixed ANOVA test

A significant two-way interaction indicates that the impact that one factor has on the outcome variable depends on the level of the other factor (and vice versa). We check the interactions using the t2way, robust ANOVA function from the WRS2 package. Post hoc tests can be applied using the mcp2atm function

Robust ANOVA

```
## Call:
## t2way(formula = stress_grapes ~ stress_type * Therapy, data =
d_Stress_type)
##
##                       value p.value
## stress_type          7.0622   0.037
## Therapy             48.5656   0.001
## stress_type:Therapy  1.2481   0.543
```

The number of grapes eaten between therapies and stress type are pretty much same. We provide this conclusion because the p-value we received from ANOVA is greater than 0.05. If we have a significant result we would use mcp2atm to check the comparisons as to which groups wold have different grapes eaten.

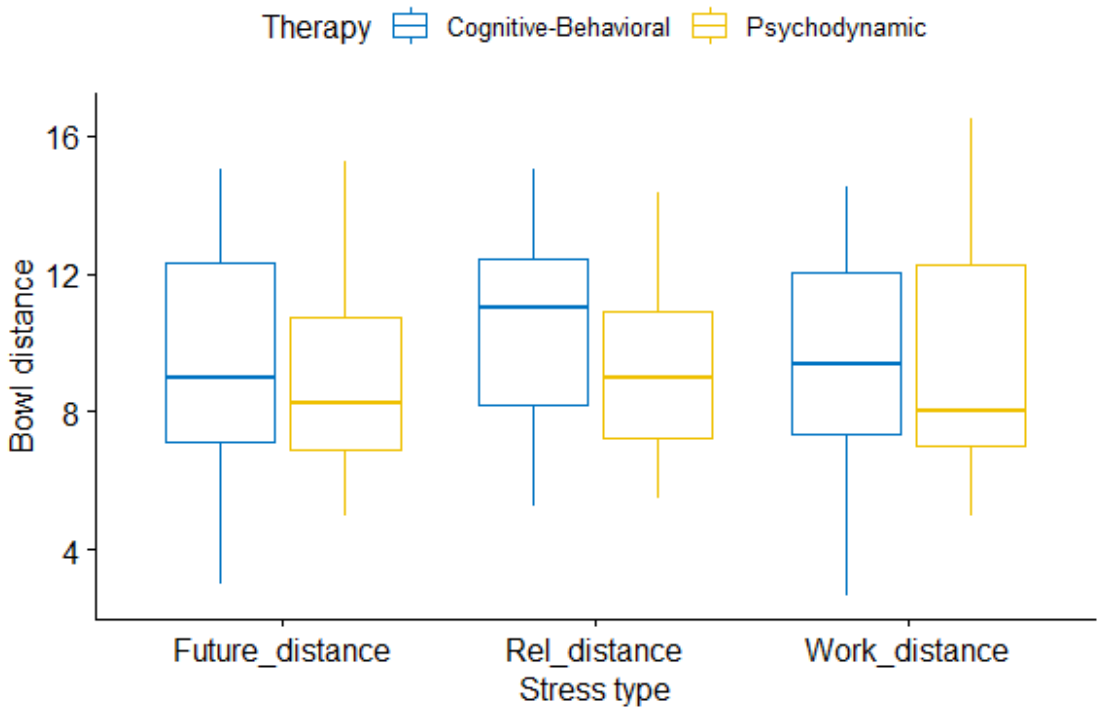
```
## Call:
## mcp2atm(formula = stress_grapes ~ Therapy * stress_type, data =
```

```
d_Stress_type)
##
##                            psihat  ci.lower ci.upper p-value
## Therapy1                 24.11316  17.23954 30.98677 0.00000
## stress_type1              6.22456  -0.63565 13.08478 0.03017
## stress_type2             -1.14825  -7.73032  5.43383 0.66966
## stress_type3             -7.37281 -14.66875 -0.07686 0.01615
## Therapy1:stress_type1     2.57544  -4.28478  9.43565 0.36268
## Therapy1:stress_type2    -0.55175  -7.13383  6.03032 0.83747
## Therapy1:stress_type3    -3.12719 -10.42314  4.16875 0.29820
```

Let us focus on the interaction first by starting at the bottom. The last effect tells us that the difference number of grapes eaten for relationship stress vs. work stress does not differ significantly in Cognitive-Behavioral and Psychodynamic Therapy. Similarly, the second to last effect tells us that there is no significant difference in number of grapes eaten for relationship vs future stress, and future vs work stress considering both therapies. However, both the therapies do differ differently in the overall manner.

Distance of the bowl after exposed to stress

By using the same method for stress and therapy interaction using number of grapes eaten, we check whether there was any noticeable difference in the distance of the bowl moved to the two therapies participants after being exposed to each of the stresses.



The graphs shows that there is no noticeable difference in the distance of the bowl moved after being exposed to all three kinds of stresses in both the treatments.

Assumptions

All the assumptions are satisfied.

Two-way mixed ANOVA test

Since the assumption are satisfied, we use the normal ANOVA model for the analysis

```
##                       Df Sum Sq Mean Sq F value Pr(>F)
## Therapy                1   15.9  15.888   1.838  0.177
## stress_type            2   14.4   7.216   0.835  0.436
## Therapy:stress_type    2   17.5   8.740   1.011  0.366
## Residuals            183 1581.9   8.644
```

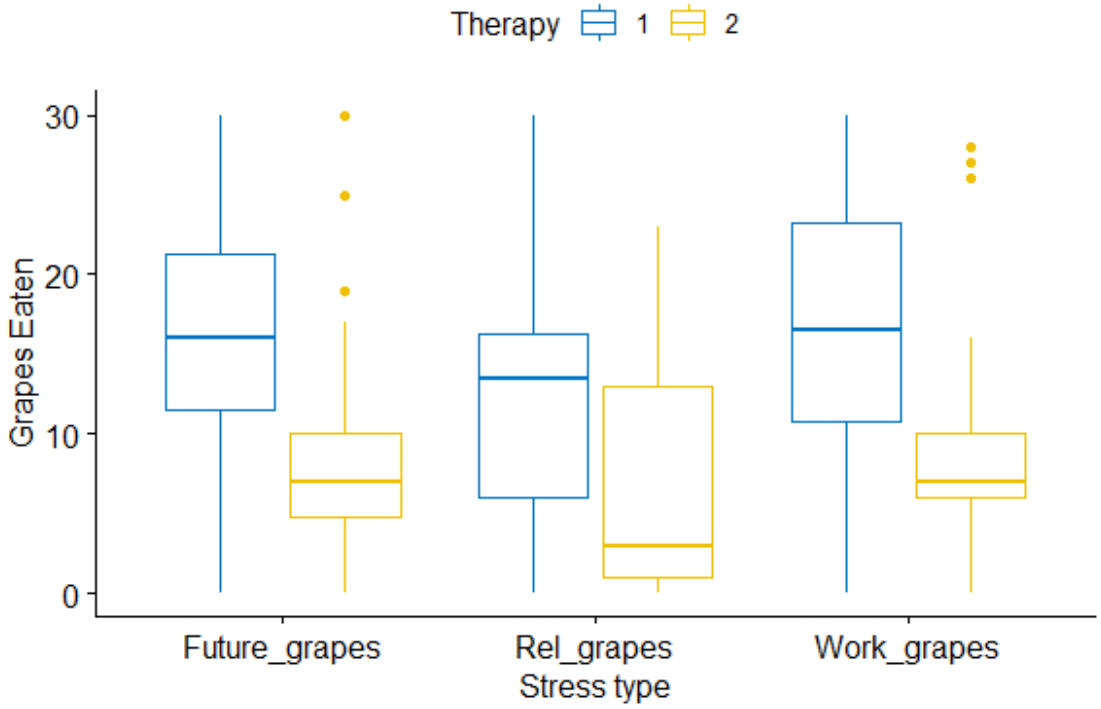
From the ANOVA table we can conclude that both therapy and stress type are not statistically significant. These results would lead us to believe that changing therapies or tress type, will not impact significantly the mean distance of the bowl. If the result was statistically different, we would have done the bonferroni correction test to decide which groups have statistical significance.

Case 2: Using MCAR data

We now use the data set where the missing values were generated completely at random and imputed missing values using the MICE package. We proceed with the same analysis as was done with the complete dataset.

Grapes eaten due to stress

From the plot we notice that number of grapes eaten after being exposed to future and work stress is approximately same.



Assumptions

All the assumptions are satisfied except that there are a few outliers and normality. The p-value under shapiro test for the log transformed data is still less than 0.05.Therefore, similar to the complete case, we moved ahead with Robust ANOVA.

Two-way mixed ANOVA test

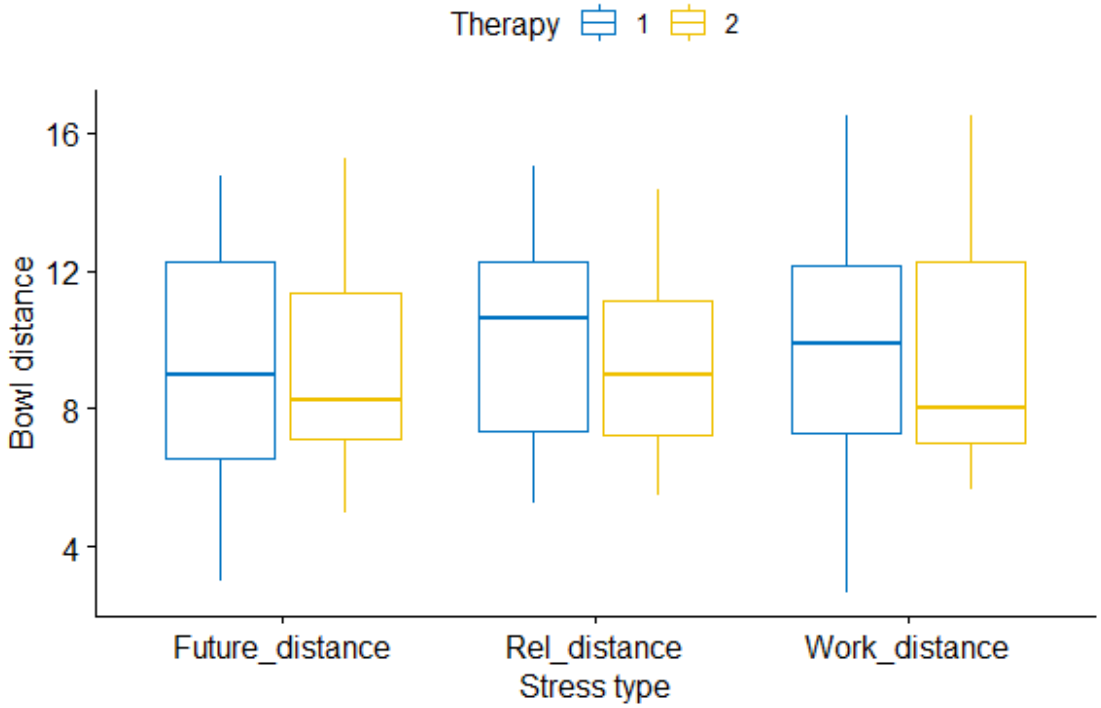
Robust ANOVA

```
## Call:
## t2way(formula = stress_grapes ~ stress_type * Therapy, data =
d_Stress_type_mcar)
##
##                        value p.value
## stress_type           7.4926   0.031
## Therapy              50.5643   0.001
## stress_type:Therapy   1.3329   0.521
```

The number of grapes eaten between therapies and stress type are pretty much same. We provide this conclusion because the p-value we received from ANOVA is greater than 0.05. This is the same conclusion as complete/original data.

Distance of the bowl after exposed to stress

By using the same method as before we check whether there was any noticeable difference in the distance of the bowl moved to the two therapies participants after being exposed to each of the stresses.



The graphs shows that there is no noticeable difference in the distance of the bowl moved after being exposed to all three kinds of stresses in both the treatments.

Assumptions

All the assumptions are satisfied.

Two-way mixed ANOVA test

Since the assumption are satisfied, we use the normal ANOVA model for the analysis

```
##                       Df Sum Sq Mean Sq F value Pr(>F)
## Therapy                1    3.6   3.552   0.402  0.527
## stress_type            2    6.1   3.032   0.343  0.710
## Therapy:stress_type    2    7.5   3.753   0.425  0.654
## Residuals            183 1615.8   8.830
```

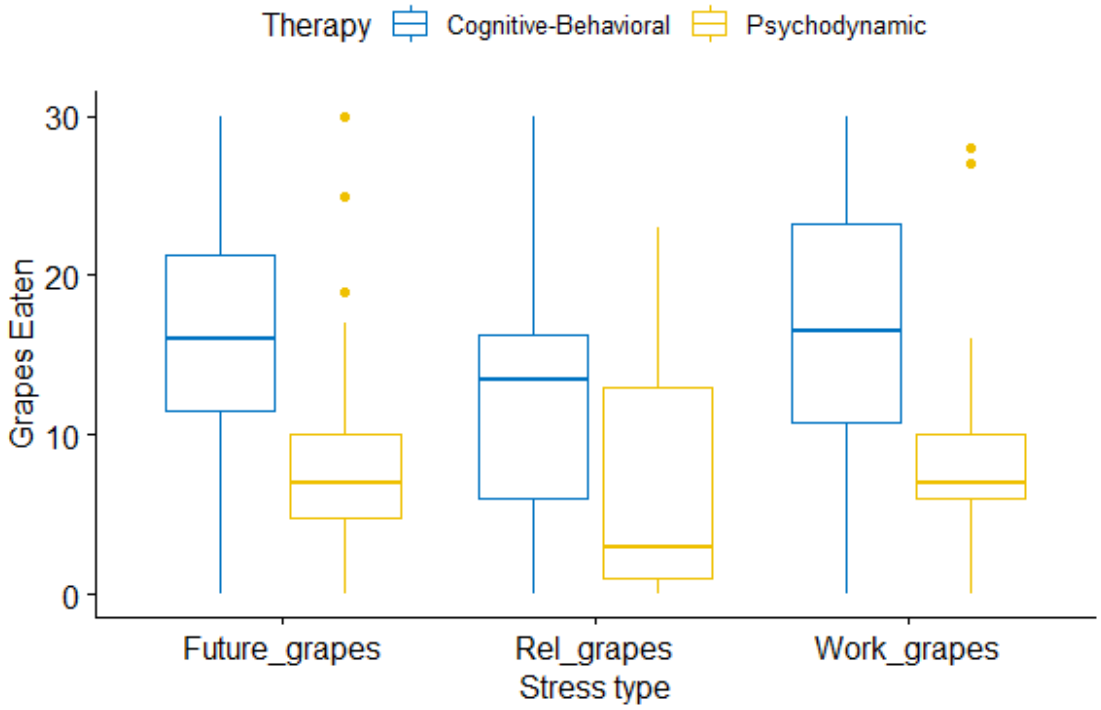
From the ANOVA table we can conclude that both therapy and stress type are not statistically significant. These results would lead us to believe that changing therapies or tress type, will not impact significantly the mean distance of the bowl. This is the same conclusion as complete/original data.

Case 3: Using MNAR data

We now use the data set where the missing values were generated not completely at random and imputed missing values using the MICE package. We proceed with the same analysis as was done with the complete dataset.

Grapes eaten due to stress

From the plot we notice that number of grapes eaten after being exposed to future and work stress is approximately same.



Assumptions

All the assumptions are satisfied except that there are a few outliers and the p-value under shapiro test of number of grapes eaten after being exposed to future and work stress after Psychodynamic therapy is less than 0.05. This was the same situation that occurred in the complete data. But we move ahead with the analysis.

Two-way mixed ANOVA test

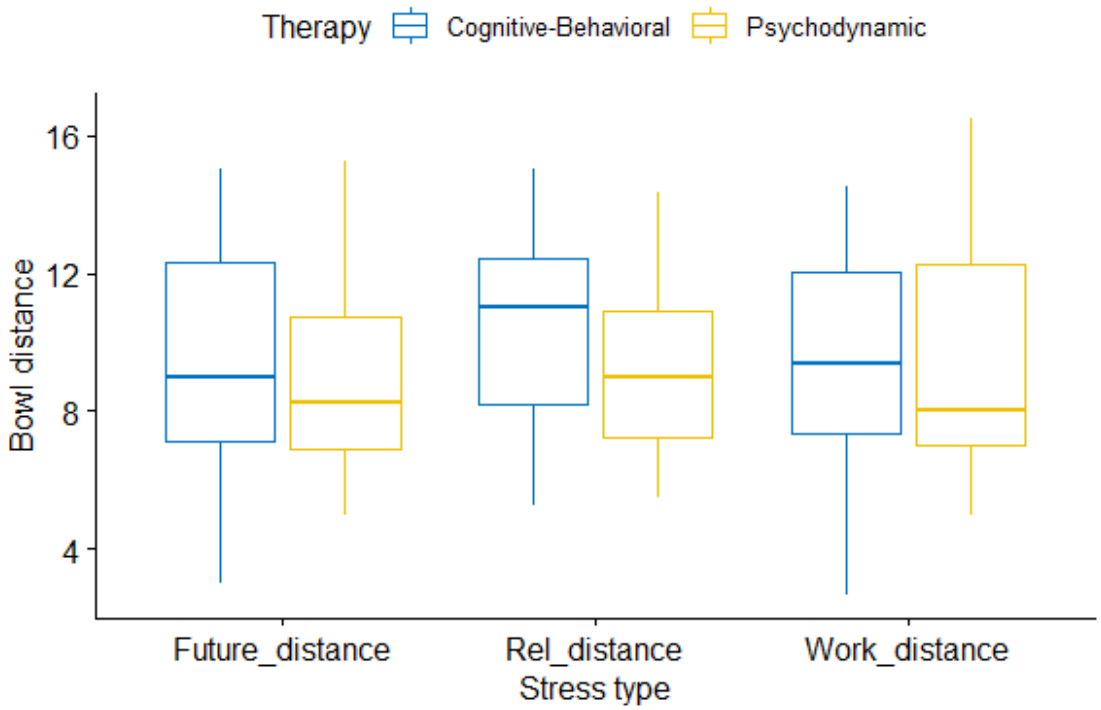
Robust ANOVA

```
## Call:
## t2way(formula = stress_grapes ~ stress_type * Therapy, data =
d_Stress_type_mnar)
##
##                       value p.value
## stress_type          7.0622   0.037
## Therapy             48.5656   0.001
## stress_type:Therapy  1.2481   0.543
```

The number of grapes eaten between therapies and stress type are pretty much same. We provide this conclusion because the p-value we received from ANOVA is greater than 0.05. This is the same result as original?complete and MCAR dataset.

Distance of the bowl after exposed to stress

By using the same method as before we check whether there was any noticeable difference in the distance of the bowl moved to the two therapies participants after being exposed to each of the stresses.



The graphs shows that there is no noticeable difference in the distance of the bowl moved after being exposed to all three kinds of stresses in both the treatments.

Assumptions

All the assumptions are satisfied.

Two-way mixed ANOVA test

Since the assumption are satisfied, we use the normal ANOVA model for the analysis

```
##                      Df Sum Sq Mean Sq F value Pr(>F)
## Therapy               1   15.9  15.888   1.838  0.177
## stress_type           2   14.4   7.216   0.835  0.436
## Therapy:stress_type   2   17.5   8.740   1.011  0.366
## Residuals           183 1581.9   8.644
```

From the ANOVA table we can conclude that both therapy and stress type are not statistically significant. These results would lead us to believe that changing therapies or tress type, will not impact significantly the mean distance of the bowl. This is the same conclusion as complete/original data and MCAR data.

Conclusion

We conclude that eating habits of Psychodynamic and Cognitive Behavioral are different. Means of overall grapes eaten by two therapy populations are different. Similar results were concluded from MCAR data and MNAR data.

From two-way mixed ANOVA we concluded that for both Cognitive Behavioral therapy and Psychodynamic therapy there was no statistical significant interaction between therapy and number of grapes eaten in each of stress types. We conclude that the mean of the number of grapes eaten when exposed to relationship stress, work, stress and future stress are all same for both the therapies. Also there was no difference in the distance of the bowl moved for both the therapy participants after being exposed to each of the stresses. Similar results are concluded fo both MCAR data and MNAR data.

There was also no relation seen between number of grapes and bowl distance in each of the stress types and therapies.

References

1. https://en.wikipedia.org/wiki/Mixed-design_analysis_of_variance
2. https://teachpsychscience.org/index.php/design-data-which-therapy-is-best-for-treating-eating-disorders-mixed-design-and-mixed-anova/432/research-methods/
3. https://www.datanovia.com/en/lessons/mixed-anova-in-r/
4. https://cran.r-project.org/web/packages/WRS2/vignettes/WRS2.pdf