# STA160 Project Code

Hayat Botan

2024-12-03

```
data160 <-
read.csv("C:\\Users\\19162\\Downloads\\ObesityDataSet_raw_and_data_sinthetic.
csv")

names(data160)
```

```
##  [1] "Gender"                          "Age"
##  [3] "Height"                          "Weight"
##  [5] "family_history_with_overweight" "FAVC"
##  [7] "FCVC"                            "NCP"
##  [9] "CAEC"                            "SMOKE"
## [11] "CH2O"                            "SCC"
## [13] "FAF"                             "TUE"
## [15] "CALC"                            "MTRANS"
## [17] "NObeyesdad"
```

```
head(data160)
```

```
##   Gender Age Height Weight family_history_with_overweight FAVC FCVC NCP
## 1 Female  21   1.62   64.0                             yes   no    2   3
## 2 Female  21   1.52   56.0                             yes   no    3   3
## 3   Male  23   1.80   77.0                             yes   no    2   3
## 4   Male  27   1.80   87.0                              no   no    3   3
## 5   Male  22   1.78   89.8                              no   no    2   1
## 6   Male  29   1.62   53.0                              no  yes    2   3
##        CAEC SMOKE CH2O SCC FAF TUE       CALC                 MTRANS
## 1 Sometimes    no    2  no   0   1         no Public_Transportation
## 2 Sometimes   yes    3 yes   3   0  Sometimes Public_Transportation
## 3 Sometimes    no    2  no   2   1 Frequently Public_Transportation
## 4 Sometimes    no    2  no   2   0 Frequently               Walking
## 5 Sometimes    no    2  no   0   0  Sometimes Public_Transportation
## 6 Sometimes    no    2  no   0   0  Sometimes            Automobile
##           NObeyesdad
## 1       Normal_Weight
## 2       Normal_Weight
## 3       Normal_Weight
## 4  Overweight_Level_I
## 5 Overweight_Level_II
## 6       Normal_Weight
```

```
summary(data160)
```

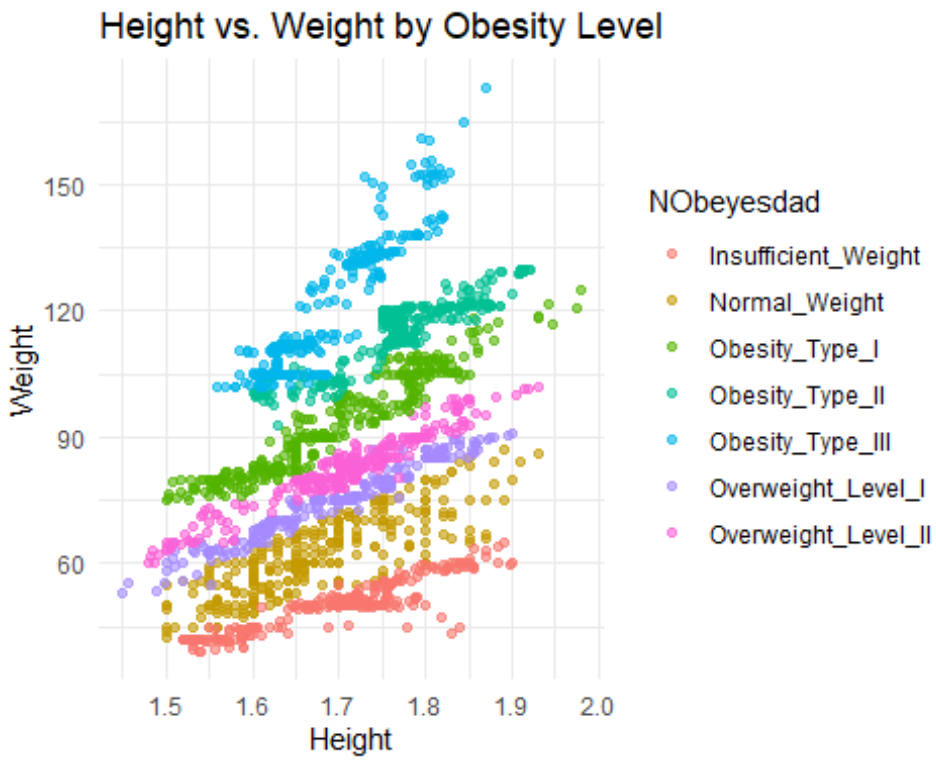
```
##     Gender               Age            Height          Weight
##  Length:2111        Min.   :14.00   Min.   :1.450   Min.   : 39.00
##  Class :character   1st Qu.:19.95   1st Qu.:1.630   1st Qu.: 65.47
##  Mode  :character   Median :22.78   Median :1.700   Median : 83.00
##                     Mean   :24.31   Mean   :1.702   Mean   : 86.59
##                     3rd Qu.:26.00   3rd Qu.:1.768   3rd Qu.:107.43
##                     Max.   :61.00   Max.   :1.980   Max.   :173.00
##  family_history_with_overweight     FAVC                FCVC
##  Length:2111                    Length:2111        Min.   :1.000
##  Class :character               Class :character   1st Qu.:2.000
##  Mode  :character               Mode  :character   Median :2.386
##                                                    Mean   :2.419
##                                                    3rd Qu.:3.000
##                                                    Max.   :3.000
##       NCP            CAEC              SMOKE                CH2O
##  Min.   :1.000   Length:2111        Length:2111        Min.   :1.000
##  1st Qu.:2.659   Class :character   Class :character   1st Qu.:1.585
##  Median :3.000   Mode  :character   Mode  :character   Median :2.000
##  Mean   :2.686                                         Mean   :2.008
##  3rd Qu.:3.000                                         3rd Qu.:2.477
##  Max.   :4.000                                         Max.   :3.000
##      SCC                 FAF              TUE              CALC
##  Length:2111        Min.   :0.0000   Min.   :0.0000   Length:2111
##  Class :character   1st Qu.:0.1245   1st Qu.:0.0000   Class :character
##  Mode  :character   Median :1.0000   Median :0.6253   Mode  :character
##                     Mean   :1.0103   Mean   :0.6579
##                     3rd Qu.:1.6667   3rd Qu.:1.0000
##                     Max.   :3.0000   Max.   :2.0000
##     MTRANS           NObeyesdad
##  Length:2111        Length:2111
##  Class :character   Class :character
##  Mode  :character   Mode  :character
##
##
##
```

```r
#height vs. weight by obesity level

library(ggplot2)
```

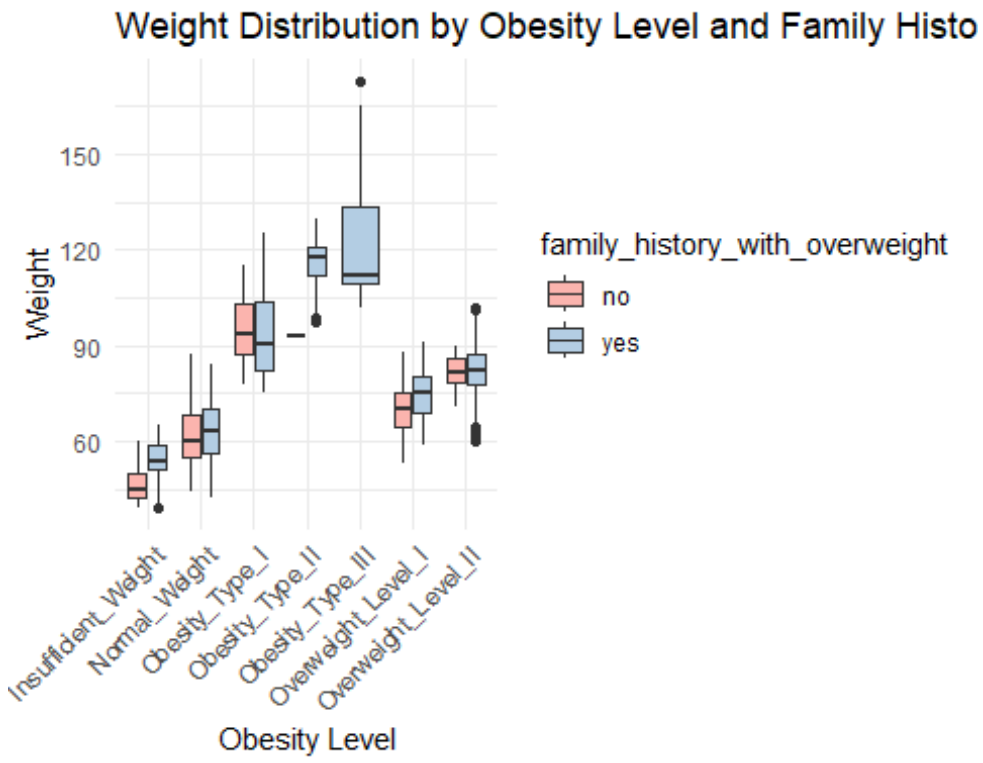
```
## Warning: package 'ggplot2' was built under R version 4.3.3
```

```r
ggplot(data160, aes(x = Height, y = Weight, color = NObeyesdad)) +
  geom_point(alpha = 0.6) +
  labs(title = "Height vs. Weight by Obesity Level", x = "Height", y =
"Weight") +
  theme_minimal()
```




```r
#weight distribution by obesity level and family history

ggplot(data160, aes(x = NObeyesdad, y = Weight, fill =
family_history_with_overweight)) +
  geom_boxplot() +
  labs(title = "Weight Distribution by Obesity Level and Family History",
       x = "Obesity Level",
       y = "Weight") +
  scale_fill_brewer(palette = "Pastel1") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```




```
#Kruskal-Wallis & Dunn's Test

library(FSA)

## Warning: package 'FSA' was built under R version 4.3.3

## ## FSA v0.9.5. See citation('FSA') if used in publication.
## ## Run fishR() for related website and fishR('IFAR') for related book.

variables <- c("Weight", "FAF", "CH2O", "NCP")

for (var in variables) {
  cat("\nKruskal-Wallis Test for", var, "\n")
  kruskal_results <- kruskal.test(data160[[var]] ~ data160$NObeyesdad)
  print(kruskal_results)
}

##
## Kruskal-Wallis Test for Weight
##
##  Kruskal-Wallis rank sum test
##
## data:  data160[[var]] by data160$NObeyesdad
## Kruskal-Wallis chi-squared = 1822.2, df = 6, p-value < 2.2e-16
##
##
## Kruskal-Wallis Test for FAF
```

```
##
##  Kruskal-Wallis rank sum test
##
## data:  data160[[var]] by data160$NObeyesdad
## Kruskal-Wallis chi-squared = 88.42, df = 6, p-value < 2.2e-16
##
##
## Kruskal-Wallis Test for CH2O
##
##  Kruskal-Wallis rank sum test
##
## data:  data160[[var]] by data160$NObeyesdad
## Kruskal-Wallis chi-squared = 94.747, df = 6, p-value < 2.2e-16
##
##
## Kruskal-Wallis Test for NCP
##
##  Kruskal-Wallis rank sum test
##
## data:  data160[[var]] by data160$NObeyesdad
## Kruskal-Wallis chi-squared = 230.95, df = 6, p-value < 2.2e-16
```

```
dunn_weight <- dunnTest(Weight ~ NObeyesdad, data = data160, method =
"bonferroni")
```

```
## Warning: NObeyesdad was coerced to a factor.
```

```
cat("\nDunn's Test Results for Weight (Post Hoc)\n")
```

```
##
## Dunn's Test Results for Weight (Post Hoc)
```

```
print(dunn_weight)
```

```
## Dunn (1964) Kruskal-Wallis multiple comparison

##   p-values adjusted with the Bonferroni method.

##                                    Comparison           Z       P.unadj
## 1        Insufficient_Weight - Normal_Weight  -5.3130167  1.078252e-07
## 2       Insufficient_Weight - Obesity_Type_I -21.8328868 1.130506e-105
## 3             Normal_Weight - Obesity_Type_I -16.5121660  2.999226e-61
## 4      Insufficient_Weight - Obesity_Type_II -30.8838472 1.968368e-209
## 5            Normal_Weight - Obesity_Type_II -25.8822486 1.055249e-147
## 6           Obesity_Type_I - Obesity_Type_II -10.5054883  8.149990e-26
## 7     Insufficient_Weight - Obesity_Type_III -32.2323844 6.211332e-228
## 8           Normal_Weight - Obesity_Type_III -27.1538361 2.281236e-162
## 9          Obesity_Type_I - Obesity_Type_III -11.5134875  1.128217e-30
## 10        Obesity_Type_II - Obesity_Type_III  -0.7313106  4.645895e-01
## 11  Insufficient_Weight - Overweight_Level_I -11.3482436  7.566744e-30
## 12        Normal_Weight - Overweight_Level_I  -6.1046868  1.030025e-09
## 13       Obesity_Type_I - Overweight_Level_I  10.1541228  3.176537e-24
```

```
## 14      Obesity_Type_II - Overweight_Level_I  19.7935963  3.380388e-87
## 15     Obesity_Type_III - Overweight_Level_I  20.9408227  2.274969e-97
## 16 Insufficient_Weight - Overweight_Level_II -15.7704811  4.966995e-56
## 17       Normal_Weight - Overweight_Level_II -10.5877909  3.395096e-26
## 18      Obesity_Type_I - Overweight_Level_II   5.4503069  5.028298e-08
## 19     Obesity_Type_II - Overweight_Level_II  15.2720711  1.173848e-52
## 20    Obesity_Type_III - Overweight_Level_II  16.3232462  6.745135e-60
## 21  Overweight_Level_I - Overweight_Level_II  -4.4948042  6.963398e-06
##            P.adj
## 1   2.264328e-06
## 2  2.374062e-104
## 3   6.298374e-60
## 4  4.133572e-208
## 5  2.216023e-146
## 6   1.711498e-24
## 7  1.304380e-226
## 8  4.790595e-161
## 9   2.369256e-29
## 10  1.000000e+00
## 11  1.589016e-28
## 12  2.163052e-08
## 13  6.670729e-23
## 14  7.098815e-86
## 15  4.777435e-96
## 16  1.043069e-54
## 17  7.129703e-25
## 18  1.055943e-06
## 19  2.465082e-51
## 20  1.416478e-58
## 21  1.462314e-04

dunn_physical_activity <- dunnTest(FAF ~ NObeyesdad, data = data160, method =
"bonferroni")

## Warning: NObeyesdad was coerced to a factor.

cat("\nDunn's Test Results for Physical Activity (Post Hoc)\n")

##
## Dunn's Test Results for Physical Activity (Post Hoc)

print(dunn_physical_activity)

## Dunn (1964) Kruskal-Wallis multiple comparison
##   p-values adjusted with the Bonferroni method.

##                                    Comparison          Z      P.unadj
## 1        Insufficient_Weight - Normal_Weight  1.4517192 1.465797e-01
## 2       Insufficient_Weight - Obesity_Type_I  4.5632451 5.036895e-06
## 3             Normal_Weight - Obesity_Type_I  3.0883068 2.013006e-03
## 4      Insufficient_Weight - Obesity_Type_II  3.5188381 4.334411e-04
```

```
## 5            Normal_Weight - Obesity_Type_II  2.0836976 3.718768e-02
## 6           Obesity_Type_I - Obesity_Type_II -0.9297225 3.525148e-01
## 7     Insufficient_Weight - Obesity_Type_III  8.4639603 2.584479e-17
## 8           Normal_Weight - Obesity_Type_III  7.0713254 1.534608e-12
## 9          Obesity_Type_I - Obesity_Type_III  4.2500128 2.137583e-05
## 10        Obesity_Type_II - Obesity_Type_III  4.9883365 6.090142e-07
## 11  Insufficient_Weight - Overweight_Level_I  3.1726077 1.510765e-03
## 12        Normal_Weight - Overweight_Level_I  1.7408563 8.170877e-02
## 13       Obesity_Type_I - Overweight_Level_I -1.2705706 2.038814e-01
## 14      Obesity_Type_II - Overweight_Level_I -0.3334247 7.388137e-01
## 15     Obesity_Type_III - Overweight_Level_I -5.2977270 1.172531e-07
## 16 Insufficient_Weight - Overweight_Level_II  4.7250073 2.301071e-06
## 17       Normal_Weight - Overweight_Level_II  3.3146227 9.176688e-04
## 18      Obesity_Type_I - Overweight_Level_II  0.3806754 7.034441e-01
## 19     Obesity_Type_II - Overweight_Level_II  1.2538292 2.099041e-01
## 20    Obesity_Type_III - Overweight_Level_II -3.6767548 2.362199e-04
## 21  Overweight_Level_I - Overweight_Level_II  1.5778737 1.145946e-01
##           P.adj
## 1  1.000000e+00
## 2  1.057748e-04
## 3  4.227312e-02
## 4  9.102262e-03
## 5  7.809413e-01
## 6  1.000000e+00
## 7  5.427405e-16
## 8  3.222676e-11
## 9  4.488925e-04
## 10 1.278930e-05
## 11 3.172607e-02
## 12 1.000000e+00
## 13 1.000000e+00
## 14 1.000000e+00
## 15 2.462315e-06
## 16 4.832249e-05
## 17 1.927105e-02
## 18 1.000000e+00
## 19 1.000000e+00
## 20 4.960617e-03
## 21 1.000000e+00
```

```
dunn_water_intake <- dunnTest(CH2O ~ NObeyesdad, data = data160, method =
"bonferroni")
```

```
## Warning: NObeyesdad was coerced to a factor.
```

```
cat("\nDunn's Test Results for Water Intake (Post Hoc)\n")
```

```
##
## Dunn's Test Results for Water Intake (Post Hoc)
```

```
print(dunn_water_intake)
```

```
## Dunn (1964) Kruskal-Wallis multiple comparison
##   p-values adjusted with the Bonferroni method.

##                                       Comparison         Z      P.unadj
## 1        Insufficient_Weight - Normal_Weight  0.681965 4.952611e-01
## 2       Insufficient_Weight - Obesity_Type_I -4.963759 6.914161e-07
## 3             Normal_Weight - Obesity_Type_I -5.763644 8.231698e-09
## 4      Insufficient_Weight - Obesity_Type_II -1.014655 3.102701e-01
## 5            Normal_Weight - Obesity_Type_II -1.725982 8.435062e-02
## 6           Obesity_Type_I - Obesity_Type_II  4.005739 6.182381e-05
## 7     Insufficient_Weight - Obesity_Type_III -6.944755 3.791174e-12
## 8           Normal_Weight - Obesity_Type_III -7.757494 8.662390e-15
## 9          Obesity_Type_I - Obesity_Type_III -2.208414 2.721546e-02
## 10        Obesity_Type_II - Obesity_Type_III -6.049292 1.454833e-09
## 11  Insufficient_Weight - Overweight_Level_I -3.765785 1.660264e-04
## 12        Normal_Weight - Overweight_Level_I -4.510715 6.460960e-06
## 13       Obesity_Type_I - Overweight_Level_I  1.047330 2.949475e-01
## 14      Obesity_Type_II - Overweight_Level_I -2.818832 4.819874e-03
## 15     Obesity_Type_III - Overweight_Level_I  3.132844 1.731213e-03
## 16 Insufficient_Weight - Overweight_Level_II -2.436778 1.481879e-02
## 17       Normal_Weight - Overweight_Level_II -3.163415 1.559300e-03
## 18      Obesity_Type_I - Overweight_Level_II  2.460960 1.385659e-02
## 19     Obesity_Type_II - Overweight_Level_II -1.459986 1.442940e-01
## 20    Obesity_Type_III - Overweight_Level_II  4.520557 6.167721e-06
## 21  Overweight_Level_I - Overweight_Level_II  1.350816 1.767544e-01
##           P.adj
## 1  1.000000e+00
## 2  1.451974e-05
## 3  1.728657e-07
## 4  1.000000e+00
## 5  1.000000e+00
## 6  1.298300e-03
## 7  7.961465e-11
## 8  1.819102e-13
## 9  5.715246e-01
## 10 3.055150e-08
## 11 3.486554e-03
## 12 1.356802e-04
## 13 1.000000e+00
## 14 1.012174e-01
## 15 3.635547e-02
## 16 3.111945e-01
## 17 3.274529e-02
## 18 2.909885e-01
## 19 1.000000e+00
## 20 1.295221e-04
## 21 1.000000e+00

dunn_snacking <- dunnTest(NCP ~ NObeyesdad, data = data160, method =
"bonferroni")
```

```
## Warning: NObeyesdad was coerced to a factor.
```

```
cat("\nDunn's Test Results for Snacking Patterns (Post Hoc)\n")
```

```
##
## Dunn's Test Results for Snacking Patterns (Post Hoc)
```

```
print(dunn_snacking)
```

```
## Dunn (1964) Kruskal-Wallis multiple comparison
##   p-values adjusted with the Bonferroni method.

##                                      Comparison           Z      P.unadj
## 1        Insufficient_Weight - Normal_Weight   3.7948453 1.477355e-04
## 2       Insufficient_Weight - Obesity_Type_I  11.0842082 1.496688e-28
## 3             Normal_Weight - Obesity_Type_I   7.2159553 5.355665e-13
## 4      Insufficient_Weight - Obesity_Type_II   6.9674615 3.227107e-12
## 5            Normal_Weight - Obesity_Type_II   3.1848861 1.448110e-03
## 6           Obesity_Type_I - Obesity_Type_II  -3.9400462 8.146593e-05
## 7     Insufficient_Weight - Obesity_Type_III   0.9458078 3.442466e-01
## 8           Normal_Weight - Obesity_Type_III  -3.0020335 2.681827e-03
## 9          Obesity_Type_I - Obesity_Type_III -10.6125210 2.606205e-26
## 10        Obesity_Type_II - Obesity_Type_III  -6.3108010 2.775949e-10
## 11  Insufficient_Weight - Overweight_Level_I   7.3888063 1.481528e-13
## 12        Normal_Weight - Overweight_Level_I   3.6337082 2.793768e-04
## 13       Obesity_Type_I - Overweight_Level_I  -3.4239747 6.171238e-04
## 14      Obesity_Type_II - Overweight_Level_I   0.4715527 6.372461e-01
## 15     Obesity_Type_III - Overweight_Level_I   6.7530042 1.448147e-11
## 16 Insufficient_Weight - Overweight_Level_II  10.0297121 1.128452e-23
## 17       Normal_Weight - Overweight_Level_II   6.3109627 2.773050e-10
## 18      Obesity_Type_I - Overweight_Level_II  -0.6149137 5.386117e-01
## 19     Obesity_Type_II - Overweight_Level_II   3.1717518 1.515225e-03
## 20    Obesity_Type_III - Overweight_Level_II   9.5105640 1.896316e-21
## 21  Overweight_Level_I - Overweight_Level_II   2.6842417 7.269453e-03
##            P.adj
## 1  3.102446e-03
## 2  3.143044e-27
## 3  1.124690e-11
## 4  6.776924e-11
## 5  3.041031e-02
## 6  1.710784e-03
## 7  1.000000e+00
## 8  5.631836e-02
## 9  5.473031e-25
## 10 5.829493e-09
## 11 3.111209e-12
## 12 5.866913e-03
## 13 1.295960e-02
## 14 1.000000e+00
## 15 3.041110e-10
## 16 2.369748e-22
```

```
## 17 5.823405e-09
## 18 1.000000e+00
## 19 3.181972e-02
## 20 3.982264e-20
## 21 1.526585e-01
```