

# final work

[Code ▼](#)

*Oksana Rebrik*

## Introduction

The database contains information on the evaluation of homework students philologists on the course of programming on the Python. Since the work will be graded by various TAs, I would like to know which of them evaluates more strictly and who is softer, since a biased assessment greatly influences the learning process.

## Load libraries

I load libraries for further work which are responsible for uploading data, connecting functions, data manipulation, and its visualization.

[Hide](#)

```
# uplading data
library(haven)
library(readr)

# data manipulation
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##   filter, lag
```

```
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
```

[Hide](#)

```
library(psych)

# visualizing data
library(ggplot2)
```

```
##
## Attaching package: 'ggplot2'
```

```
## The following objects are masked from 'package:psych':  
##  
##      %+%, alpha
```

[Hide](#)

```
library(knitr)  
library(tidyquant)
```

```
## Loading required package: lubridate
```

```
##  
## Attaching package: 'lubridate'
```

```
## The following object is masked from 'package:base':  
##  
##      date
```

```
## Loading required package: PerformanceAnalytics
```

```
## Loading required package: xts
```

```
## Loading required package: zoo
```

```
##  
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':  
##  
##      as.Date, as.Date.numeric
```

```
##  
## Attaching package: 'xts'
```

```
## The following objects are masked from 'package:dplyr':  
##  
##      first, last
```

```
##  
## Attaching package: 'PerformanceAnalytics'
```

```
## The following object is masked from 'package:graphics':  
##  
##      legend
```

```
## Loading required package: quantmod
```

```
## Loading required package: TTR
```

```
## Version 0.4-0 included new data defaults. See ?getSymbols.
```

```
## Loading required package: tidyverse
```

```
## — Attaching packages —
```

```
— tidyverse 1.2.1 —
```

```
## ✓ tibble 1.4.2      ✓ purrr 0.2.5  
## ✓ tidyr 0.7.2       ✓ stringr 1.3.1  
## ✓ tibble 1.4.2      ✓ forcats 0.3.0
```

```
## — Conflicts — tid  
yverse_conflicts() —  
## X ggplot2::%+%( ) masks psych::%+%( )  
## X ggplot2::alpha( ) masks psych::alpha( )  
## X lubridate::as.difftime( ) masks base::as.difftime( )  
## X lubridate::date( ) masks base::date( )  
## X dplyr::filter( ) masks stats::filter( )  
## X xts::first( ) masks dplyr::first( )  
## X lubridate::intersect( ) masks base::intersect( )  
## X dplyr::lag( ) masks stats::lag( )  
## X xts::last( ) masks dplyr::last( )  
## X lubridate::setdiff( ) masks base::setdiff( )  
## X lubridate::union( ) masks base::union( )
```

Hide

```
library(sjPlot)
```

```
## Warning in checkMatrixPackageVersion(): Package version inconsistency detected.  
## TMB was built with Matrix version 1.2.11  
## Current Matrix version is 1.2.14  
## Please re-install 'TMB' from source or ask CRAN for a binary version of 'TMB' matching  
## CRAN's 'Matrix' package
```

[Hide](#)

```
library(stargazer)
```

```
##  
## Please cite as:
```

```
## Hlavac, Marek (2018). stargazer: Well-Formatted Regression and Summary Statistics Tables.
```

```
## R package version 5.2.1. https://CRAN.R-project.org/package=stargazer
```

[Hide](#)

```
library(gridExtra)
```

```
##  
## Attaching package: 'gridExtra'
```

```
## The following object is masked from 'package:dplyr':  
##  
## combine
```

[Hide](#)

```
library(rockchalk)
```

```
##  
## Attaching package: 'rockchalk'
```

```
## The following object is masked from 'package:quantmod':  
##  
## addLines
```

```
## The following objects are masked from 'package:PerformanceAnalytics':  
##  
## kurtosis, skewness
```

```
## The following object is masked from 'package:dplyr':  
##  
## summarize
```

[Hide](#)

```
library(ggpubr)
```

```
## Loading required package: magrittr
```

```
##  
## Attaching package: 'magrittr'
```

```
## The following object is masked from 'package:purrr':  
##  
##   set_names
```

```
## The following object is masked from 'package:tidyr':  
##  
##   extract
```

[Hide](#)

```
#making default theme  
theme_set(theme_tq())  
  
#tests  
library(stats)  
library(car)
```

```
## Loading required package: carData
```

```
##  
## Attaching package: 'car'
```

```
## The following object is masked from 'package:purrr':  
##  
##   some
```

```
## The following object is masked from 'package:psych':  
##  
##   logit
```

```
## The following object is masked from 'package:dplyr':  
##  
##   recode
```

[Hide](#)

```
library(pwr)
```

```
#source
```

```
source("functions.R")
```

## Theme

Later, I will use this theme instead of the default one.

[Hide](#)

```
theme_set(theme_tq())
```

```
theme_update(plot.title = element_text(hjust = 0.5))
```

## Generate data

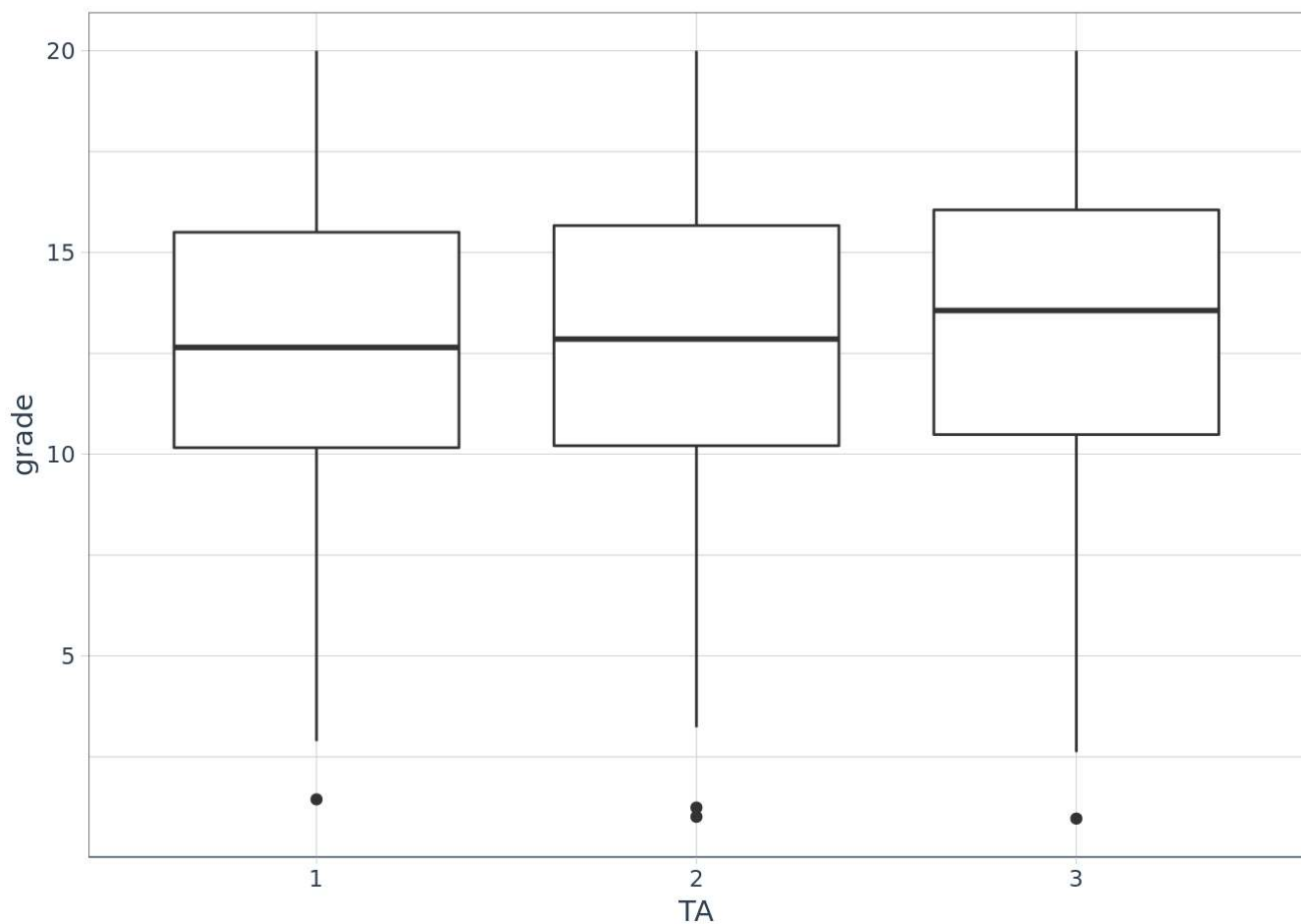
[Hide](#)

```
set.seed(1)
```

```
marks_norm <- data.frame(student = c(1:1000),  
                          grade = my_sample(1000, 13, 4, 0, 20),  
                          TA = sample(1:3, 1000, replace = TRUE))
```

```
marks_norm$TA <- ordered(marks_norm$TA, levels = c("1", "2", "3"))
```

```
ggplot(marks_norm)+  
  geom_boxplot(aes(TA, grade))
```



Hide

```
describe(marks_norm)
```

```
##      vars    n  mean    sd median trimmed   mad  min  max  range
## student    1 1000 500.50 288.82 500.50  500.50 370.65  1.00 1000  999.00
## grade      2 1000  12.88   3.99  12.86   12.96   4.13  0.97   20   19.03
## TA*        3 1000   1.98   0.81   2.00    1.97   1.48  1.00    3    2.00
##
##      skew kurtosis   se
## student  0.00    -1.20  9.13
## grade    -0.22    -0.32  0.13
## TA*      0.05    -1.49  0.03
```

First data will be generated from normal distribution with large sample, to meet assumptions, but it does not have any difference in means. Mean grade is near 13, data is negatively skewed and platykurtic, but a few.

Hide

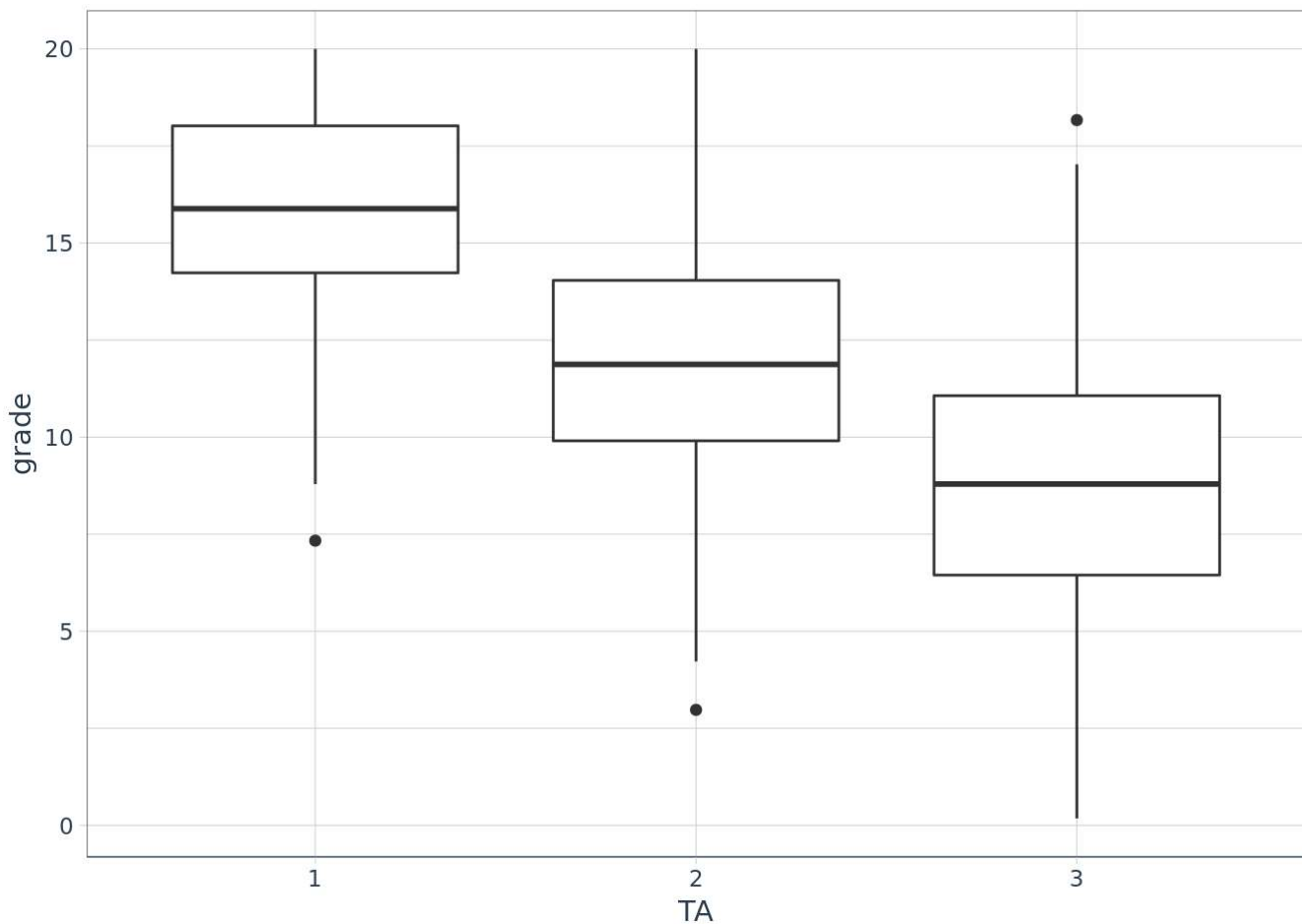
```

set.seed(1)
TA1 <- data.frame(student = c(1:300),
                  grade = my_sample(300, 16, 3, 0, 20),
                  TA = 1)
TA2 <- data.frame(student = c(301:600),
                  grade = my_sample(300, 12, 3, 0, 20),
                  TA = 2)
TA3 <- data.frame(student = c(601:900),
                  grade = my_sample(300, 9, 3, 0, 20),
                  TA = 3)
marks_diff_TA <- rbind(TA1, TA2, TA3)

marks_diff_TA$TA <- ordered(marks_diff_TA$TA, levels = c("1", "2", "3"))

ggplot(marks_diff_TA)+
  geom_boxplot(aes(TA, grade))

```



Hide

```
describe(marks_diff_TA)
```



```
##          vars    n  mean    sd median trimmed   mad  min max  range  skew
## student     1 900 450.50 259.95 450.50  450.50 333.58 1.00 900 899.00  0.00
## grade       2 900  12.25   4.20  12.24   12.32   4.50 0.18  20  19.82 -0.17
## TA*         3 900   2.00   0.82   2.00    2.00   1.48 1.00   3   2.00  0.00
##          kurtosis  se
## student     -1.20 8.67
## grade       -0.59 0.14
## TA*        -1.50 0.03
```

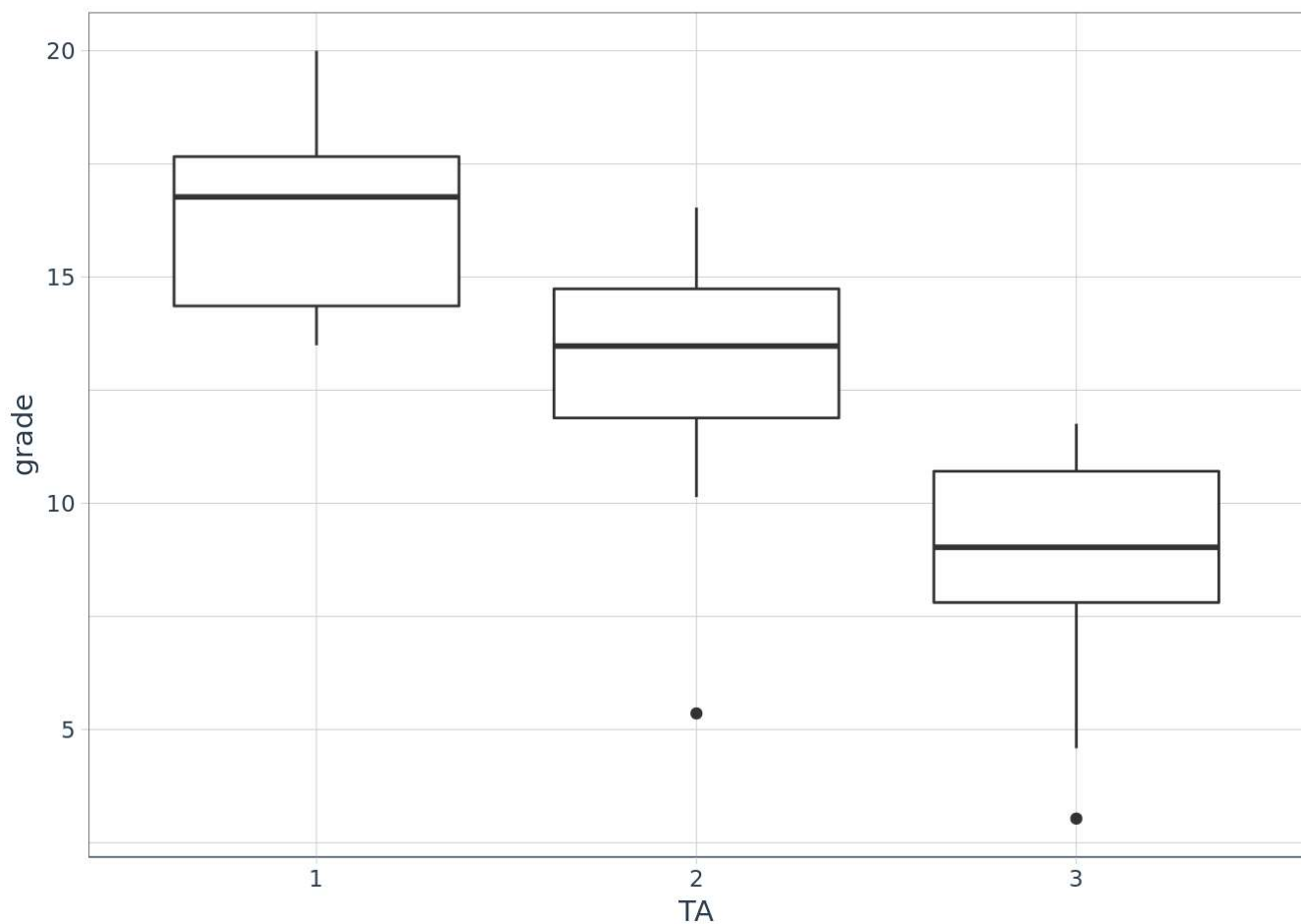
Next piece of data generated with difference in means, but it is normal, groups have same sizes and same standard deviation. Mean grade for firsts TA is 16, and 12 and 9 for second and third respectively. For general sample median is 12, SD is 4.2. *(If we multiply 4.2 by 10...)*

Hide

```
set.seed(1)
TA1 <- data.frame(student = c(1:10),
                  grade = my_sample(10, 16, 3, 0, 20),
                  TA = 1)
TA2 <- data.frame(student = c(11:20),
                  grade = my_sample(10, 12, 3, 0, 20),
                  TA = 2)
TA3 <- data.frame(student = c(21:30),
                  grade = my_sample(10, 9, 3, 0, 20),
                  TA = 3)
marks_small <- rbind(TA1, TA2, TA3)

marks_small$TA <- ordered(marks_small$TA, levels = c("1", "2", "3"))

ggplot(marks_small)+
  geom_boxplot(aes(TA, grade))
```


[Hide](#)

```
describe(marks_small)
```

```
##      vars  n  mean   sd median trimmed  mad  min max range  skew
## student   1 30 15.50 8.80  15.50   15.50 11.12 1.00 30 29.00  0.00
## grade     2 30 12.55 4.19  13.33   12.82  4.65 3.03 20 16.97 -0.42
## TA*       3 30  2.00 0.83   2.00    2.00  1.48 1.00  3  2.00  0.00
##      kurtosis  se
## student   -1.32 1.61
## grade     -0.54 0.76
## TA*       -1.60 0.15
```

This data generated for emulation of small sample size. There is only 30 observations.

[Hide](#)

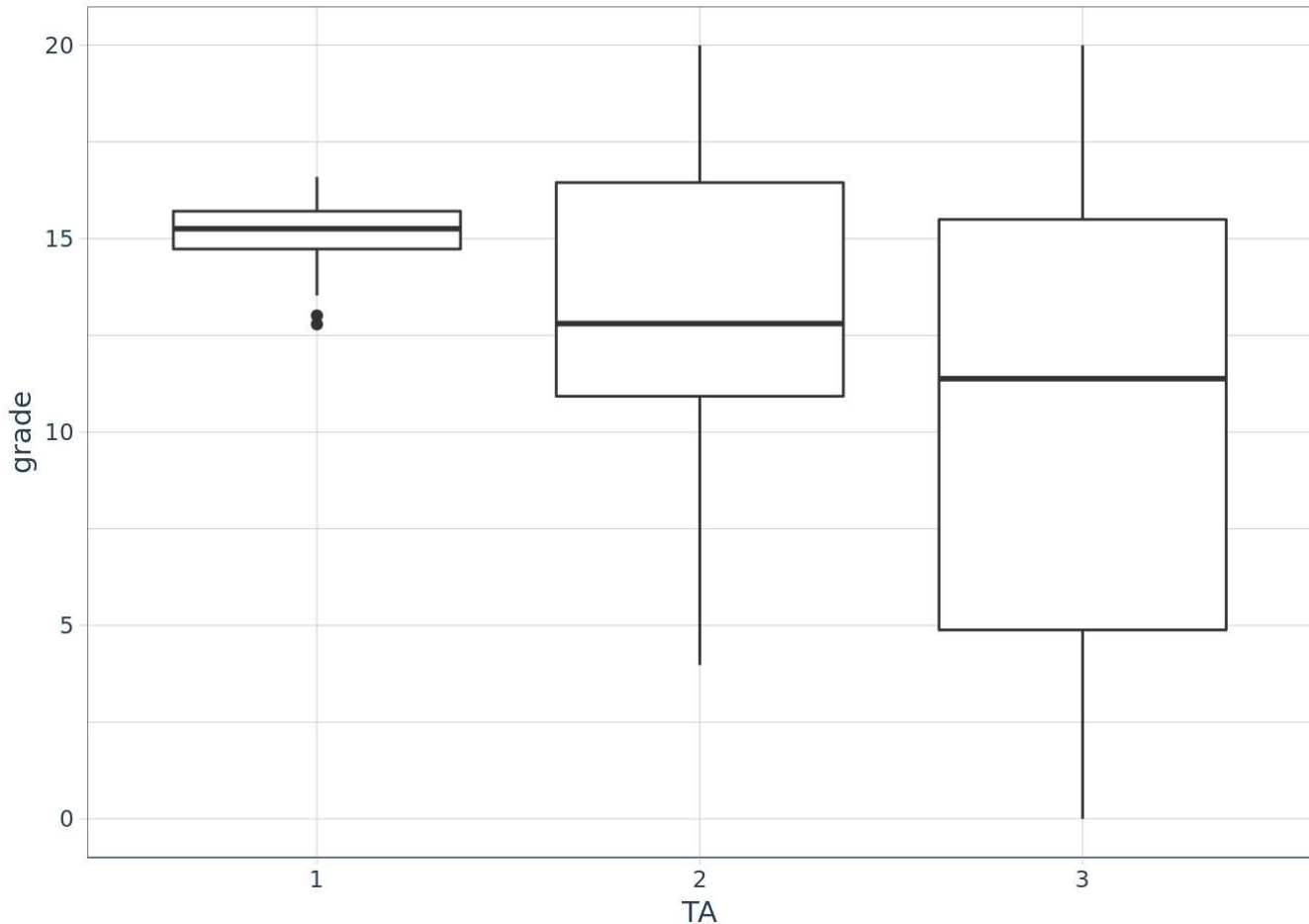
```

set.seed(1)
TA1 <- data.frame(student = c(1:34),
                  grade = my_sample(34, 15, 1, 0, 20),
                  TA = 1)
TA2 <- data.frame(student = c(35:67),
                  grade = my_sample(33, 13, 5, 0, 20),
                  TA = 2)
TA3 <- data.frame(student = c(68:100),
                  grade = my_sample(33, 10, 9, 0, 20),
                  TA = 3)
marks_sd <- rbind(TA1, TA2, TA3)

marks_sd$TA <- ordered(marks_sd$TA, levels = c("1", "2", "3"))

ggplot(marks_sd)+
  geom_boxplot(aes(TA, grade))

```



Hide

```
describeBy(marks_sd, group = "TA")
```

```
##
## Descriptive statistics by group
## group: 1
##      vars  n  mean   sd median trimmed   mad   min  max range  skew
## student    1 34 17.50 9.96  17.50   17.50 12.60   1.00 34.0 33.00  0.00
## grade      2 34 15.12 0.90  15.26   15.19  0.75 12.79 16.6  3.81 -0.73
## TA*        3 34  1.00 0.00   1.00    1.00  0.00   1.00  1.0  0.00   NaN
##      kurtosis   se
## student    -1.31 1.71
## grade       0.28 0.15
## TA*         NaN 0.00
## -----
## group: 2
##      vars  n  mean   sd median trimmed   mad   min  max range  skew
## student    1 33 51.00 9.67  51.0    51.0 11.86 35.00  67 32.00  0.00
## grade      2 33 13.23 4.05  12.8    13.3  4.51  3.98  20 16.02 -0.18
## TA*        3 33  2.00 0.00   2.0     2.0  0.00  2.00   2  0.00   NaN
##      kurtosis   se
## student    -1.31 1.68
## grade     -0.66 0.71
## TA*         NaN 0.00
## -----
## group: 3
##      vars  n  mean   sd median trimmed   mad   min  max range  skew
## student    1 33 84.00 9.67  84.00   84.00 11.86  68 100   32  0.00
## grade      2 33 10.67 6.88  11.38   10.81  9.28   0  20   20 -0.08
## TA*        3 33  3.00 0.00   3.00    3.00  0.00   3   3    0  NaN
##      kurtosis   se
## student    -1.31 1.68
## grade     -1.36 1.20
## TA*         NaN 0.00
```

This data for not equal variances and small difference between groups due to increased SD for second and third TA.

Hide

```
remove(TA1, TA2, TA3)
# they won't be used later
```

## ANOVA

ANOVA or Analysis of Variance, is most common test for checking difference in means, if there are more than two groups. There is also a Tukey test, but it is less friendly and harder to interpret. In our data we have three groups.

### Assumptions

Before making ANOVA itself, I check its assumptions:

## Normality

As far as we do have less than 5k observations, I check the normality of residuals.

$(H_0)$ : the data is distributed normally;  
 $(H_A)$ : the data is not distributed normally.

Hide

```
set.seed(1)
model = aov(marks_norm$grade ~ marks_norm$TA)
residuals = model$residuals
shapiro.test(residuals)
```

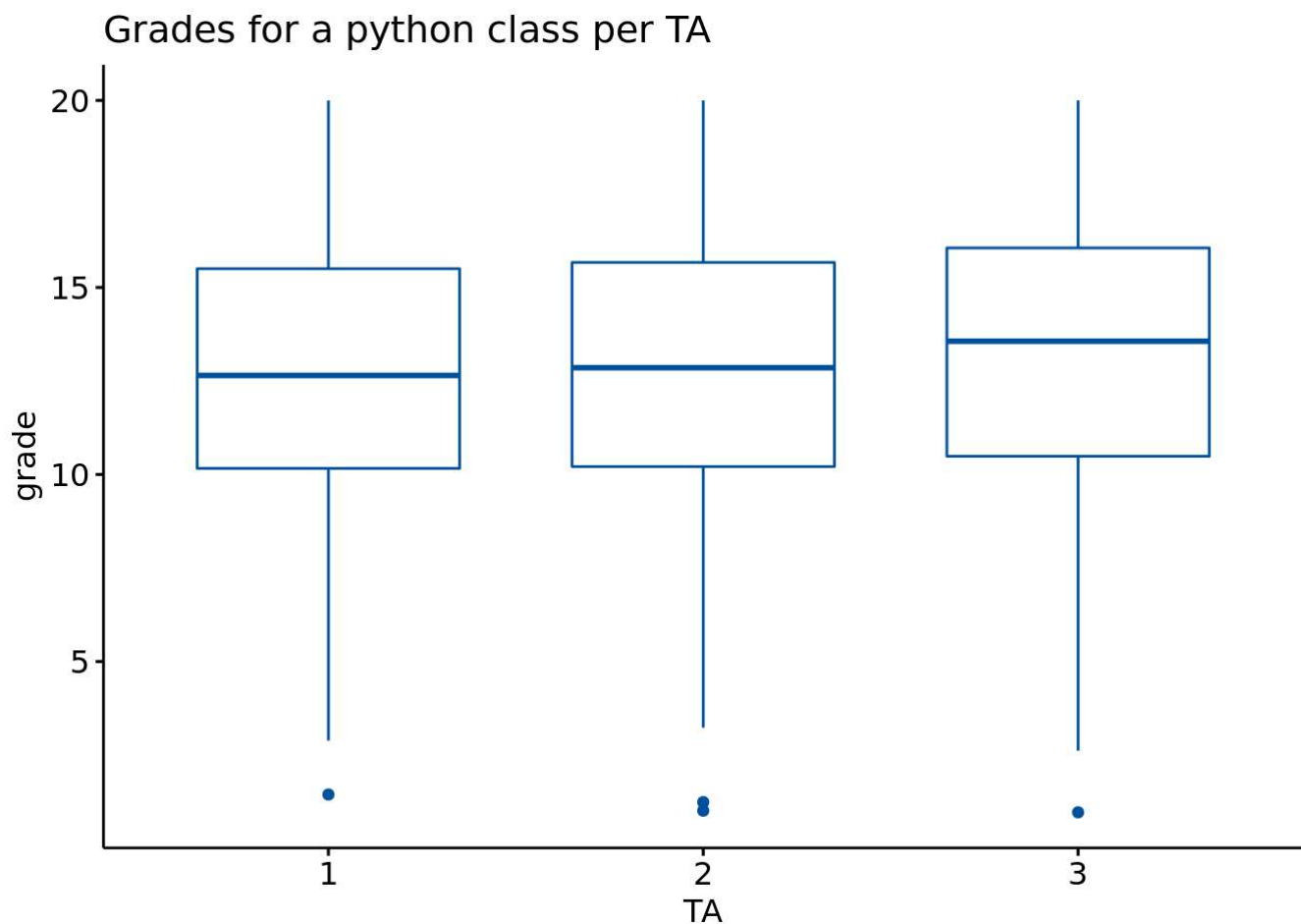
```
##
##  Shapiro-Wilk normality test
##
## data:  residuals
## W = 0.98793, p-value = 2.496e-07
```

P-value is less than 0.05 ( $2.496e-07$ ). So, we reject null hypothesis about normality of the distribution.

Also, we create box-plots for checking availability of outliers.

Hide

```
ggpubr::ggboxplot(marks_norm, x = "TA", y = "grade",
  color = "TA", palette = c("#00519d", "#00519d", "#00519d"),
  xlab = "TA",
  ylab = "grade",
  title = "Grades for a python class per TA") +
  theme(legend.position="none")
```



We can see that there are outliers in all categories. It is a good reason for running non-parametric test.

Homogeneity of variances

Next step is checking homogeneity with the help of Levene's test. I choose this one because variable is not distributed normally. And it is less sensitive.

$(H_0)$ : variances are equal between categories;  
 $(H_A)$ : variances are not equal.

Hide

```
set.seed(1)
leveneTest(marks_norm$grade ~ marks_norm$TA)
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value Pr(>F)
## group  2   0.303 0.7387
##      997
```

Since p-value is more than 0.05 (\* 0.7387\*), I do not reject the null hypothesis and meet second assumption.

Test itself (finally!)

Now I can apply Kruskal-Wallis to our data with following hypotheses:

$(H_0)$ : the difference between the means of grades among the groups of different TAs is equal to zero;  
 $(H_A)$ : the difference between the means of grades among the groups of different TAs is not equal to zero.

Hide

```
kruskal.test(grade ~ TA, data = marks_norm)
```

```
##  
##  Kruskal-Wallis rank sum test  
##  
## data:  grade by TA  
## Kruskal-Wallis chi-squared = 2.5612, df = 2, p-value = 0.2779
```

P-value is more than 0.05 – 0.2779. Therefore, I can not reject the null hypothesis, and I prove that there is no difference between TAs grading.

However, I would like to use parametric test (just in case) with the correction on equal variances.

Hide

```
norm_anova = oneway.test(grade ~ TA, data = marks_norm, var.equal = TRUE)  
norm_anova
```

```
##  
##  One-way analysis of means  
##  
## data:  grade and TA  
## F = 1.1296, num df = 2, denom df = 997, p-value = 0.3236
```

As the result of large F-statistics ( $F(2, 197) = 1.1296$ ) and p-value more than 0.05 (0.3236), I can not confirm (one more time) rejection of null hypothesis. ##### Post-hoc test

As in our case, I have deal with equal variances, I choose pairwise t-test with bonferroni correction.

Hide

```
pairwise.t.test(marks_norm$grade, marks_norm$TA, adjust = "bonferroni")
```

```
##
## Pairwise comparisons using t tests with pooled SD
##
## data: marks_norm$grade and marks_norm$TA
##
## 1 2
## 2 0.89 -
## 3 0.51 0.51
##
## P value adjustment method: holm
```

Now I can say that grades is different for all TAs because p-value for each sell is really small.

Hide

```
pwr.anova.test(k = 3, n = 660, f = 0, sig.level = 0.05, power = NULL)
```

```
##
## Balanced one-way analysis of variance power calculation
##
## k = 3
## n = 660
## f = 0
## sig.level = 0.05
## power = 0.05
##
## NOTE: n is number in each group
```

I do not really understand, how to calculate the exact value for effect size, however, according to Jacob Cohen, it can be used as 0 for no effect, 0.1 for small effect, 0.25 and 0.5, for medium and strong respectively. (Kind of similar to correlation)

For data without effect of TAs, power is really small, even for large sample.

## Data 2

Normality of data 2

$(H_0)$ : the data is distributed normally;  
 $(H_A)$ : the data is not distributed normally.

Hide



```
set.seed(1)
model = aov(marks_diff_TA$grade ~ marks_diff_TA$TA)
residuals = model$residuals
shapiro.test(residuals)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  residuals
## W = 0.99745, p-value = 0.176
```

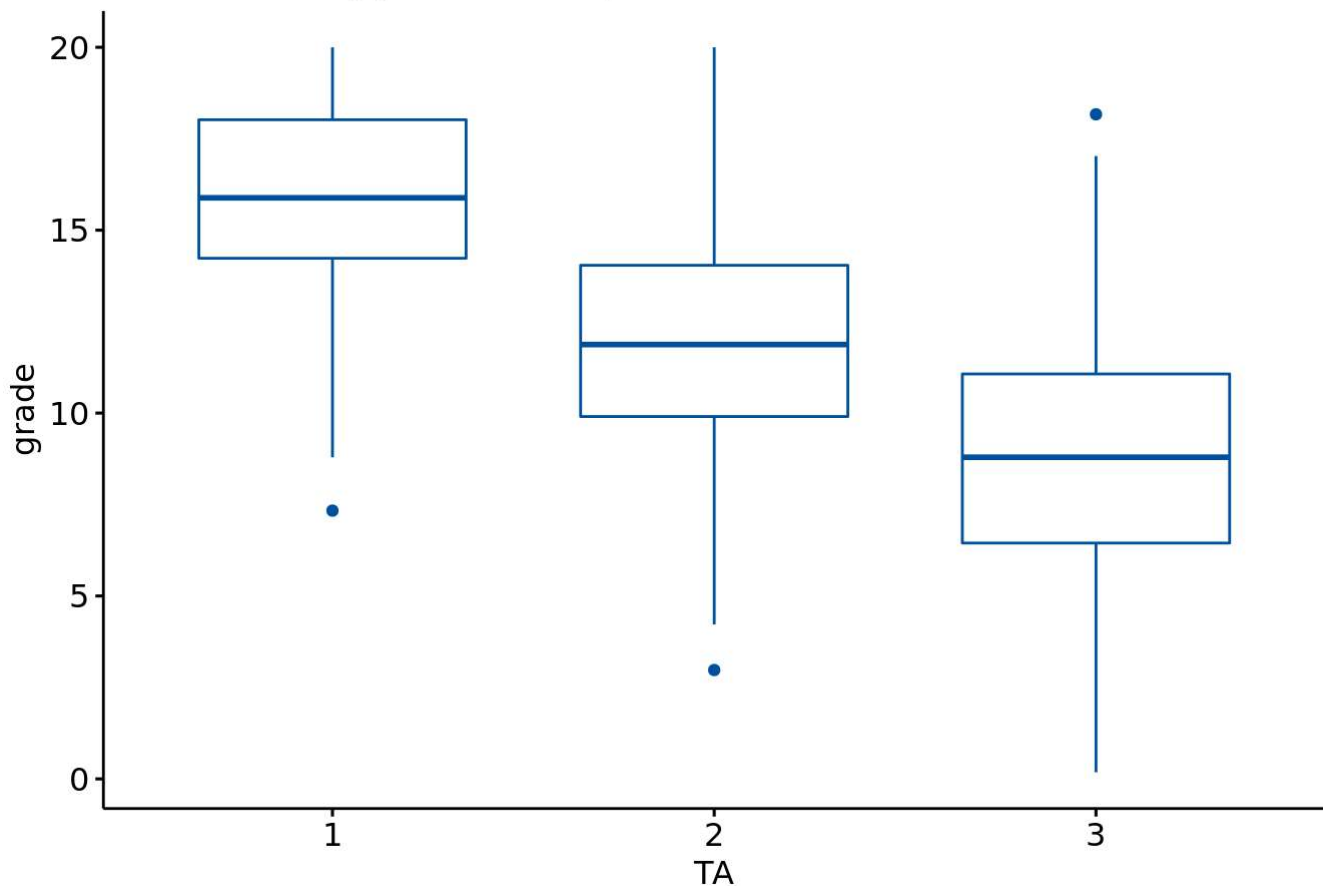
P-value is more than 0.05 (*0.176*). So, I can not reject null hypothesis about normality of the distribution. Our data is distributed normally.

Also, I create boxplots for checking availability of outliers.

Hide

```
ggpubr::ggboxplot(marks_diff_TA, x = "TA", y = "grade",
  color = "TA", palette = c("#00519d", "#00519d", "#00519d"),
  xlab = "TA",
  ylab = "grade",
  title = "Grades for a python class per TA") +
  theme(legend.position="none")
```

Grades for a python class per TA



There are outliers in two categories.

Homogeneity of variances

$(H_0)$ : variances are equal between categories;  
 $(H_A)$ : variances are not equal.

Hide

```
set.seed(1)
leveneTest(marks_diff_TA$grade ~ marks_diff_TA$TA)
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value  Pr(>F)
## group  2  5.9337 0.002754 **
##      897
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Since p-value is less than 0.05 (0.002754), I reject the null hypothesis and violate second assumption.

Test

$(H_0)$ : the difference between the means of grades among the groups of different TAs is equal to zero;  
 $(H_A)$ : the difference between the means of grades among the groups of different TAs is not equal to zero.

Anova with correction of not equal variances.

Hide

```
diff_anova = oneway.test(grade ~ TA, data = marks_diff_TA, var.equal = FALSE)
diff_anova
```

```
##
## One-way analysis of means (not assuming equal variances)
##
## data: grade and TA
## F = 448.68, num df = 2.00, denom df = 592.95, p-value < 2.2e-16
```

As the result of large F-statistics ( $F(2, 592) = 448.68$ ) and p-value less than 0.05 ( $2.2e-16$ ), I confirm rejection of null hypothesis. For now, however, I cannot say which groups do differ from each other. That means that it is the time to run post-hoc test.

Post-hoc test

As in our case, I have deal with not equal variances, we choose pairwise t-test with bonferroni correction.

Hide

```
pairwise.t.test(marks_diff_TA$grade, marks_diff_TA$TA, adjust = "bonferroni")
```

```
##
## Pairwise comparisons using t tests with pooled SD
##
## data: marks_diff_TA$grade and marks_diff_TA$TA
##
## 1      2
## 2 <2e-16 -
## 3 <2e-16 <2e-16
##
## P value adjustment method: holm
```

Now I can say that grades is different for all TAs because p-value for each sell is really small.

Hide

```
pwr.anova.test(k = 3, n = 300, f = 0.5, sig.level = 0.05, power = NULL)
```

```
##
##      Balanced one-way analysis of variance power calculation
##
##          k = 3
##          n = 300
##          f = 0.5
##      sig.level = 0.05
##          power = 1
##
## NOTE: n is number in each group
```

As well there is huge difference between groups, i will run power test with  $f = 0.5$ , and for sample of 900, power is 1.

For data with strong effect of TAs, power is maximum, even for medium sample.

## Data small

Normality of small data

Hide

```
set.seed(1)
model = aov(marks_small$grade ~ marks_small$TA)
residuals = model$residuals
shapiro.test(residuals)
```

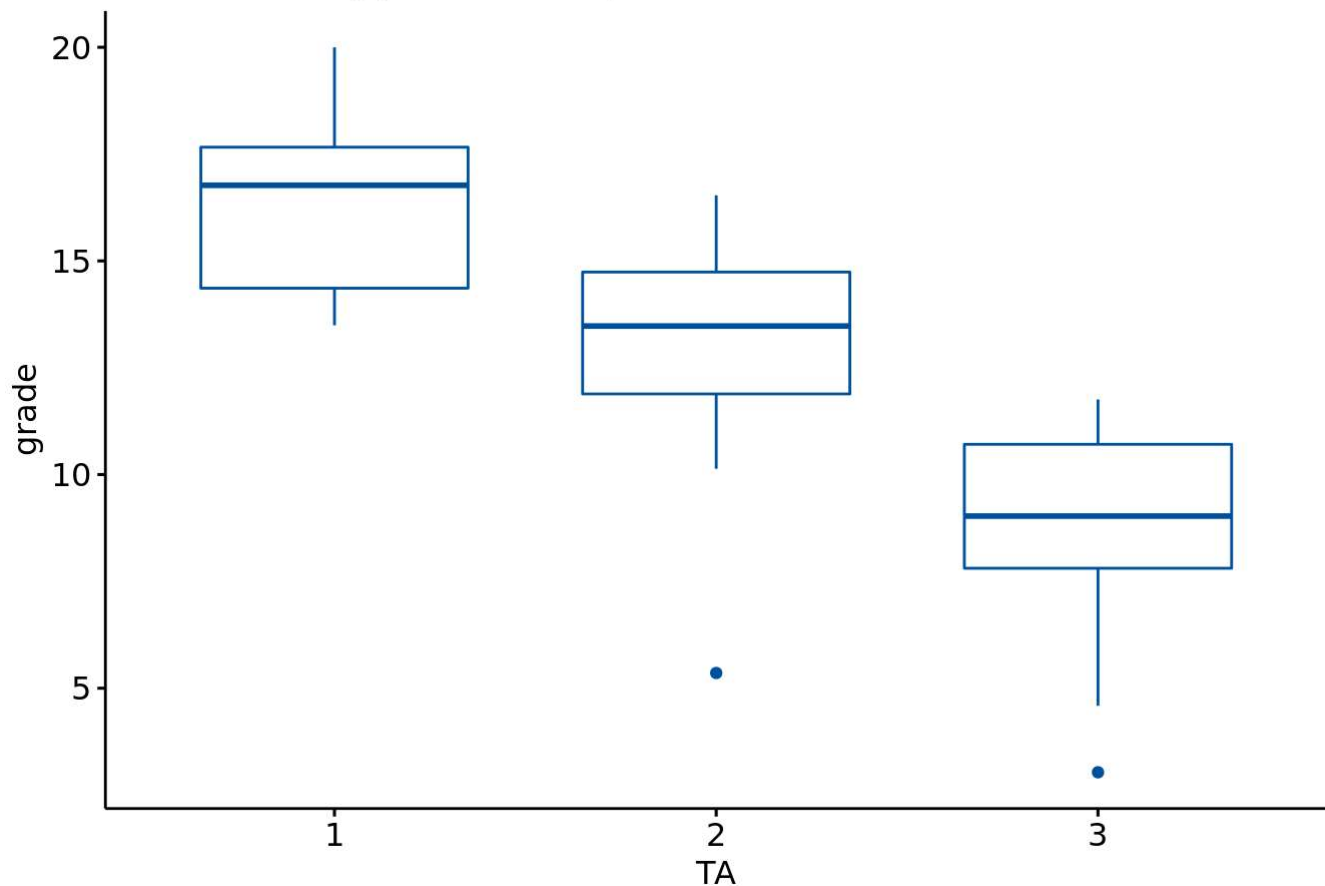
```
##
##  Shapiro-Wilk normality test
##
## data:  residuals
## W = 0.93855, p-value = 0.08314
```

P-value is more than 0.05 (*0.08314*). So, we can not reject null hypothesis. Our data is distributed normally.

Hide

```
ggpubr::ggboxplot(marks_small, x = "TA", y = "grade",
  color = "TA", palette = c("#00519d", "#00519d", "#00519d"),
  xlab = "TA",
  ylab = "grade",
  title = "Grades for a python class per TA") +
  theme(legend.position="none")
```

Grades for a python class per TA



There are outliers in two categories.

Homogeneity of variances

Hide

```
set.seed(1)
leveneTest(marks_small$grade ~ marks_small$TA)
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value Pr(>F)
## group 2  0.1802 0.8361
##      27
```

Since p-value is more than 0.05 (*0.8361*), I reject the null hypothesis and meet assumption about homogeneity of variances.

Test

Hide

```
small_anova = oneway.test(grade ~ TA, data = marks_small, var.equal = TRUE)
small_anova
```

```
##
## One-way analysis of means
##
## data: grade and TA
## F = 19.221, num df = 2, denom df = 27, p-value = 6.447e-06
```

As the result of small F-statistics ( $F(2, 27) = 19.221$ ) and p-value less than 0.05  $6.447e-06$ , I reject the null hypothesis. There is significant difference between means.

Post-hoc test

However, let`s check pairwise t-test

Hide

```
pairwise.t.test(marks_small$grade, marks_small$TA, adjust = "bonferroni")
```

```
##
## Pairwise comparisons using t tests with pooled SD
##
## data: marks_small$grade and marks_small$TA
##
##      1      2
## 2 0.0080 -
## 3 3.8e-06 0.0051
##
## P value adjustment method: holm
```

there is difference between each TA.

Hide

```
pwr.anova.test(k = 3, n = 10, f = 0.5, sig.level = 0.05, power = NULL)
```

```
##
##      Balanced one-way analysis of variance power calculation
##
##              k = 3
##              n = 10
##              f = 0.5
##      sig.level = 0.05
##      power = 0.6352398
##
## NOTE: n is number in each group
```

As well there is huge difference between groups, i will run power test with  $f = 0.5$ , and for sample of 300, power is 0.63. That is not really good, so we have to increase groups sizes in real experiment.

# Data not really different

Normality of data with less difference

Hide

```
set.seed(1)
model = aov(marks_sd$grade ~ marks_sd$TA)
residuals = model$residuals
shapiro.test(residuals)
```

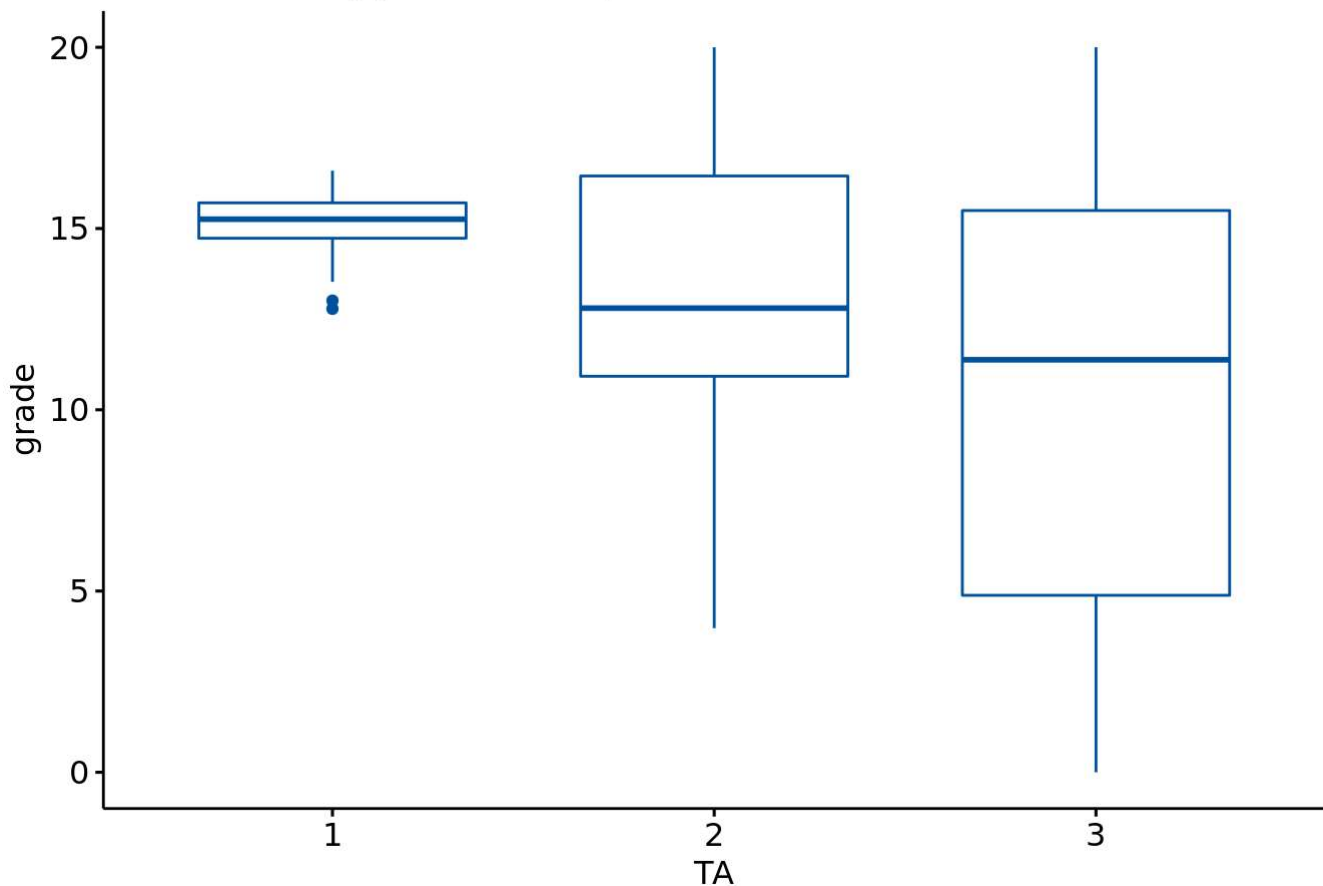
```
##
##  Shapiro-Wilk normality test
##
## data:  residuals
## W = 0.95843, p-value = 0.003115
```

P-value is less than 0.05 (*0.003115*). Our data is not distributed normally.

Hide

```
ggpubr::ggboxplot(marks_sd, x = "TA", y = "grade",
  color = "TA", palette = c("#00519d", "#00519d", "#00519d"),
  xlab = "TA",
  ylab = "grade",
  title = "Grades for a python class per TA") +
  theme(legend.position="none")
```

Grades for a python class per TA



There are outliers in one categories.

Homogeneity of variances

Hide

```
set.seed(1)
leveneTest(marks_sd$grade ~ marks_sd$TA)
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value    Pr(>F)
## group 2  37.992 6.532e-13 ***
##      97
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Since p-value is less than 0.05 (\* 6.532e-13\*), I don't the null hypothesis and violate assumption about homogeneity of variances.

Test

Hide

```
small_anova = oneway.test(grade ~ TA, data = marks_sd, var.equal = FALSE)
small_anova
```



```
##
## One-way analysis of means (not assuming equal variances)
##
## data: grade and TA
## F = 9.8436, num df = 2.000, denom df = 45.256, p-value = 0.0002822
```

As the result of small F-statistics ( $F(2, 45) = 9.8436$ ) and p-value less than 0.05 0.0002822, I reject the null hypothesis. There is significant difference between means.

Post-hoc test

However, let`s check pairwise t-test

Hide

```
pairwise.t.test(marks_sd$grade, marks_sd$TA, adjust = "bonferroni")
```

```
##
## Pairwise comparisons using t tests with pooled SD
##
## data: marks_sd$grade and marks_sd$TA
##
## 1      2
## 2 0.09653 -
## 3 0.00045 0.05278
##
## P value adjustment method: holm
```

There is difference between first and third TA.

Hide

```
pwr.anova.test(k = 3, n = 30, f = 0.25, sig.level = 0.05, power = NULL)
```

```
##
## Balanced one-way analysis of variance power calculation
##
## k = 3
## n = 30
## f = 0.25
## sig.level = 0.05
## power = 0.5396133
##
## NOTE: n is number in each group
```

As well there is difference between two groups, i will run power test with  $f = 0.25$ , and for sample of 90, power is 0.54. If we not able to influence on difference, our only chance to produce good data analysis is to increase sample size.

# Power analysis

Last but not least. I have found and edit code for power analysis of ANOVA.

Hide

```
# Plot sample size curves for detecting correlations of
# various sizes.

# range of effect size
f <- seq(0.1, 0.5, 0.01)
nf <- length(f)

# power values
p <- seq(0.4, 0.9, 0.1)
np <- length(p)

# obtain sample sizes
samsize <- array(numeric(nf * np), dim=c(nf, np))
for (i in 1:np){
  for (j in 1:nf){
    result <- pwr.anova.test(k = 3, n = NULL, f = f[j],
      sig.level = 0.05, power = p[i])
    samsize[j, i] <- ceiling(result$n)
  }
}

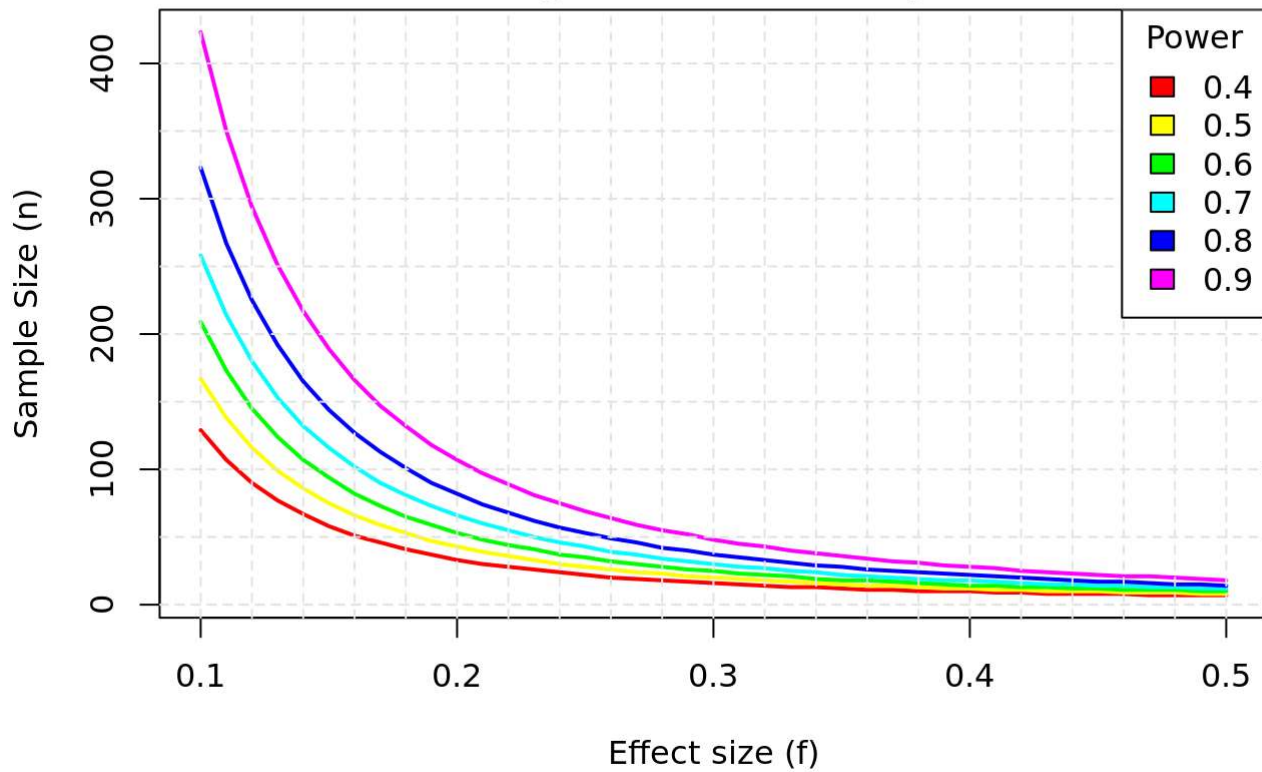
# set up graph
xrange <- range(f)
yrange <- round(range(samsize))
colors <- rainbow(length(p))
plot(xrange, yrange, type = "n",
  xlab = "Effect size (f)",
  ylab = "Sample Size (n)")

# add power curves
for (i in 1:np){
  lines(f, samsize[,i], type = "l", lwd = 2, col = colors[i])
}

# add annotation (grid lines, title, Legend)
abline(v = 0, h = seq(0, yrange[2], 50), lty = 2, col = "grey89")
abline(h = 0, v = seq(xrange[1], xrange[2], 0.02), lty = 2, col = "grey89")
title("Sample Size Estimation for ANOVA\n
  Sig = 0.05 (One-way)")
legend("topright", title = "Power", as.character(p), fill = colors)
```

## Sample Size Estimation for ANOVA

Sig = 0.05 (One-way)



Hide

```
# all credits to  
# https://www.statmethods.net/stats/power.html
```

Sample size is crucial of ANOVA tests with small effect because for effect size near 0.1 for appropriate power sample should be more than 200. On medium effect size, size of sample does not play the leading role and for power from 0.6 to 0.9 sample differ on 30-40. For Large effect size even tiny sample provide appropriate power.