S2 - Read and data plotting from the Manuscript Marreiros et al. 2022

Joao Marreiros, Mae Goder-Goldberger, Eduardo Paixao

7/21/2021

Brief description of the script

This R markdown document reads, summarizes and plots data concerning the analysis of the Manot core engraved incisions, included in the manuscript Marreiros et al. 2022.

The document contains:

1. Tables
2. Figures (data analysis)

This R project and respective scripts follow the procedures described by Marwick et al. 2017.

To compile this markdown document do not delete or move files from their original folders. Please note that most of the tables and figures in this file do not match the numbering in the original manuscript.

For any questions, comments and inputs, please contact:

João Marreiros, joao.marreiros@rgzm.de

Load data into R project

Imported files are in: '../analysis/raw_data'

Figures are saved in: '../analysis/plots'

Tables are saved in: '../analysis/derived_data'

Load libraries

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --

## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.6     v dplyr   1.0.7
## v tidyr   1.1.4     v stringr 1.4.0
## v readr   2.1.1     v forcats 0.5.1
```

```
## -- Conflicts ---------------------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(knitr)
library(kableExtra)
```

```
##
## Attaching package: 'kableExtra'
```

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

```
library(GGally)
```

```
## Registered S3 method overwritten by 'GGally':
##   method from
##   +.gg   ggplot2
```

```
library(doBy)
```

```
##
## Attaching package: 'doBy'
```

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

```
library(ggfortify)
library(tools)
library(rstatix)
```

```
##
## Attaching package: 'rstatix'
```

```
## The following object is masked from 'package:stats':
##
##     filter
```

```
library(ggpubr)
```

Import dataset

```
db <- read_csv("../raw_data/incisionsangle.csv", na = c("", "NA"), quoted_na = TRUE)
```

```
## Warning: The ‘quoted_na‘ argument of ‘read_csv()‘ is deprecated as of readr
## 2.0.0.
```

```
## Rows: 114 Columns: 5
```

```
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (1): site
## dbl (4): aoi, incision, angleid, angle
```

```
##
## i Use ‘spec()‘ to retrieve the full column specification for this data.
## i Specify the column types or set ‘show_col_types = FALSE‘ to quiet this message.
```

```
data_file <- list.files("../raw_data/", pattern = "\\.csv$", full.names = TRUE)
```

Show dataset:

```
str(db)
```

```
## spec_tbl_df [114 x 5] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
##  $ site    : chr [1:114] "Manot" "Manot" "Manot" "Manot" ...
##  $ aoi     : num [1:114] 1 1 1 1 1 1 1 1 1 1 ...
##  $ incision: num [1:114] 4 6 6 6 5 5 4 5 8 4 ...
##  $ angleid : num [1:114] 1 2 3 4 5 6 7 8 9 10 ...
##  $ angle   : num [1:114] 168 162 168 158 165 ...
##  - attr(*, "spec")=
##   .. cols(
##   ..   site = col_character(),
##   ..   aoi = col_double(),
##   ..   incision = col_double(),
##   ..   angleid = col_double(),
##   ..   angle = col_double()
##   .. )
##  - attr(*, "problems")=<externalptr>
```

Import and summarize data

```
# compute descriptive statistics

nminmaxmeanmedsd <- function(x){
    y <- x[!is.na(x)]
    n_test <- length(y)
    min_test <- min(y)
    max_test <- max(y)
    mean_test <- mean(y)
    med_test <- median(y)
    sd_test <- sd(y)
    out <- c(n_test, min_test, max_test, mean_test, med_test, sd_test)
```

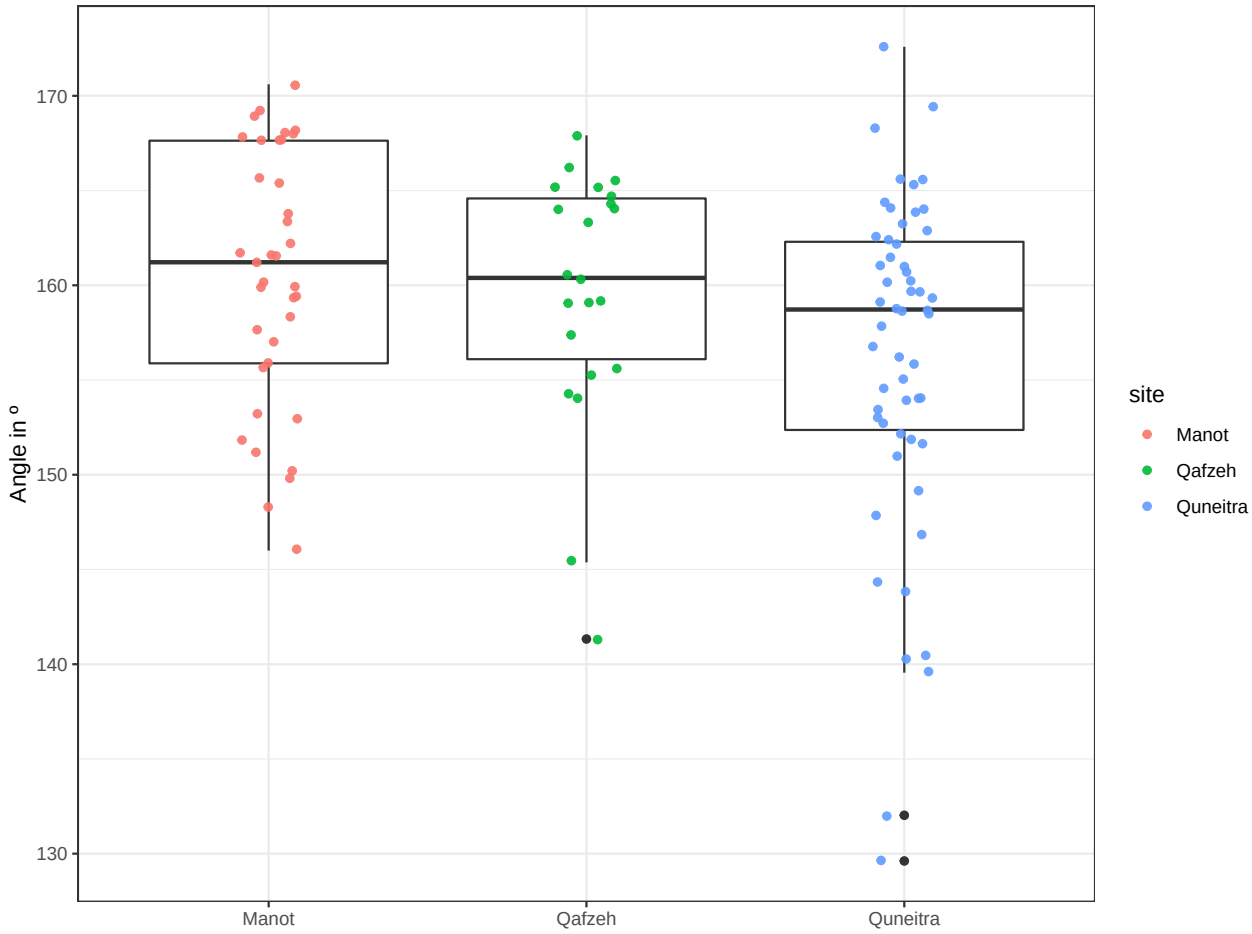
```
    names(out) <- c("n", "min", "max", "mean", "median", "sd")
    return(out)
}

num.var <- 5:length(db)

anglestats <- summaryBy(.~aoi + site + incision, data=db[c("site", "incision", names(db)[num.var])], FU

anglestats
```

```
## # A tibble: 25 x 8
##    site   incision angle.n angle.min angle.max angle.mean angle.median angle.sd
##    <chr>     <dbl>   <dbl>     <dbl>     <dbl>      <dbl>        <dbl>    <dbl>
##  1 Manot         1      12      148.      168.       156.         158.     5.86
##  2 Manot         2       7      153.      171.       164.         168.     7.07
##  3 Manot         3       3      146.      164.       156.         159.     9.27
##  4 Manot         4       7      156.      168.       163.         163.     4.81
##  5 Manot         5       3      162.      168.       165.         165.     2.99
##  6 Manot         6       3      158.      168.       162.         162.     5.11
##  7 Manot         7       1      152.      152.       152.         152.    NA
##  8 Manot         8       1      166.      166.       166.         166.    NA
##  9 Qafzeh        1       1      154.      154.       154.         154.    NA
## 10 Qafzeh        2       2      160.      165.       163.         163.     3.64
## # ... with 15 more rows
```

```
write_csv(anglestats, "../summary_stats/anglestats.csv")
```

Plot data

```
# converting to categorical
db$aoi = as.factor(db$aoi)
# db$incision = as.factor(db$incision)

sitesangle <- ggplot(data = db, aes(x = site, y = angle)) +
           geom_boxplot() +
           theme_bw()+scale_fill_grey(start = 0.8, end = 1) +
           labs( x = "", y = "Angle in º") +
           geom_jitter(aes(colour = site), alpha=0.9, position=position_jitter(w=0.1,h=0.1))

print(sitesangle)
```

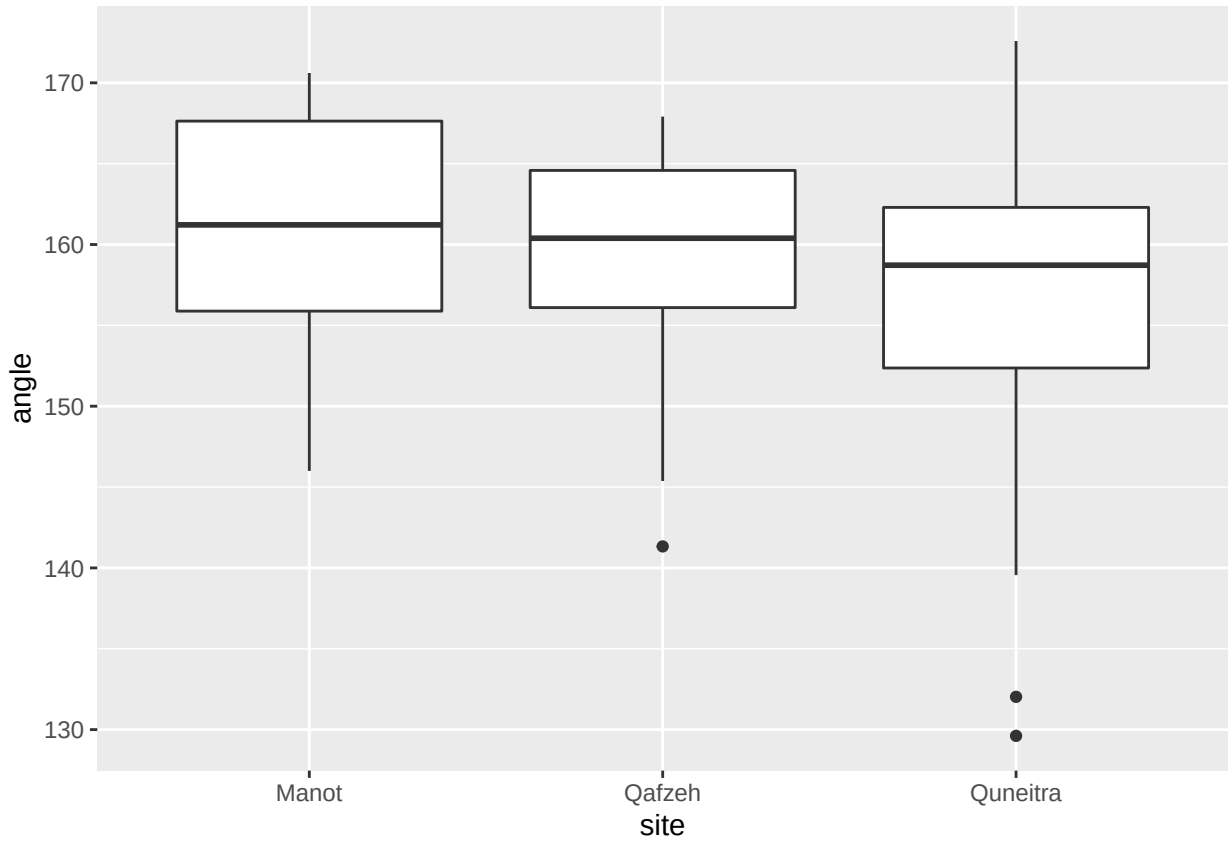



```
ggsave("../plots/sitesangle.png")
```

```
## Saving 8.5 x 6.5 in image
```

Run ANOVA for comparison

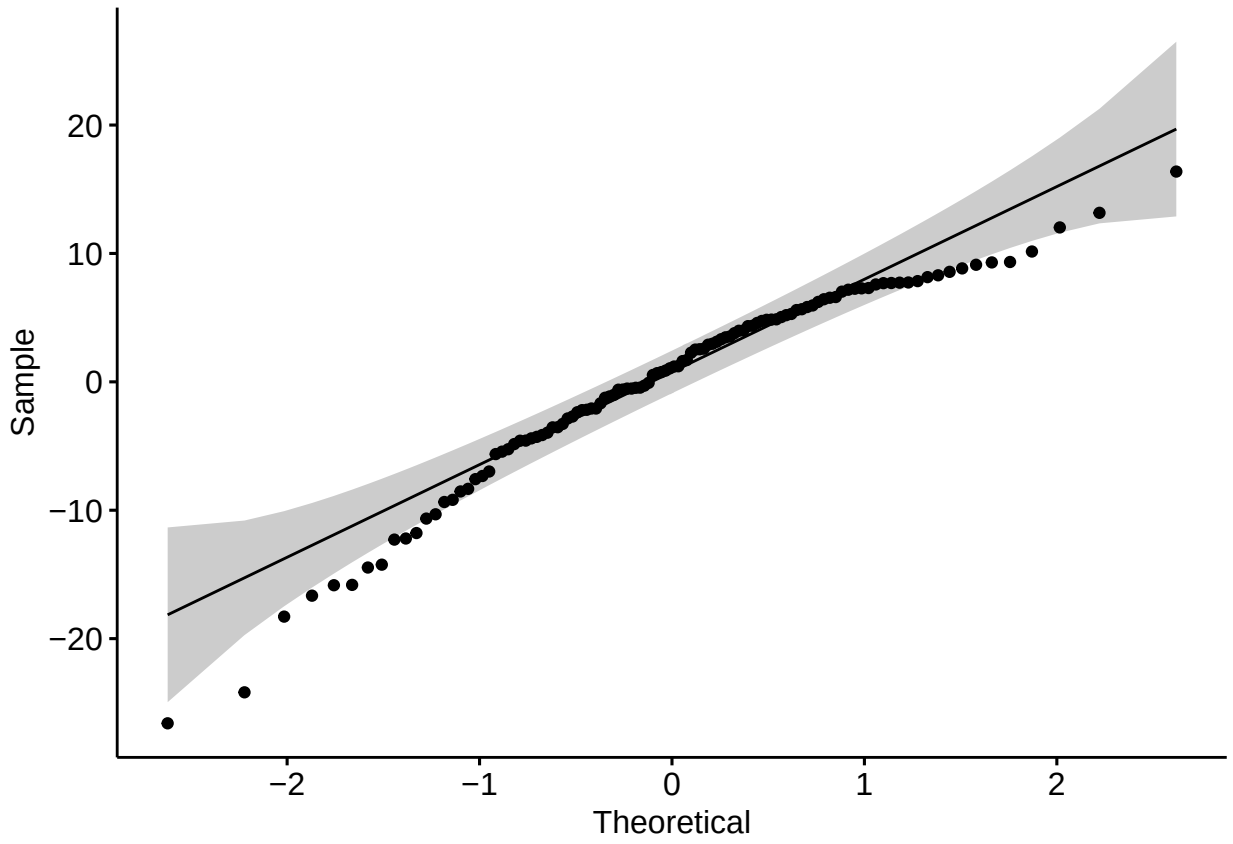
```
# brief visualization of the data from the 3 artifacts

ggplot(data = db, aes(x= site, y= angle)) +
  geom_boxplot()
```




```
# check for outliers

db %>%
  group_by(site) %>%
  identify_outliers(angle)
```

```
## # A tibble: 3 x 7
##   site     aoi   incision angleid angle is.outlier is.extreme
##   <chr>    <fct>    <dbl>   <dbl> <dbl> <lgl>      <lgl>
## 1 Qafzeh   1           10       1  141. TRUE       FALSE
## 2 Quneitra 2            2       2  132. TRUE       FALSE
## 3 Quneitra 2            7      32  130. TRUE       FALSE
```

```
# there are no extreme outliers

# check normality

model <- lm(angle ~ site, data = db)

# QQ plot of residuals

ggqqplot(residuals(model))
```



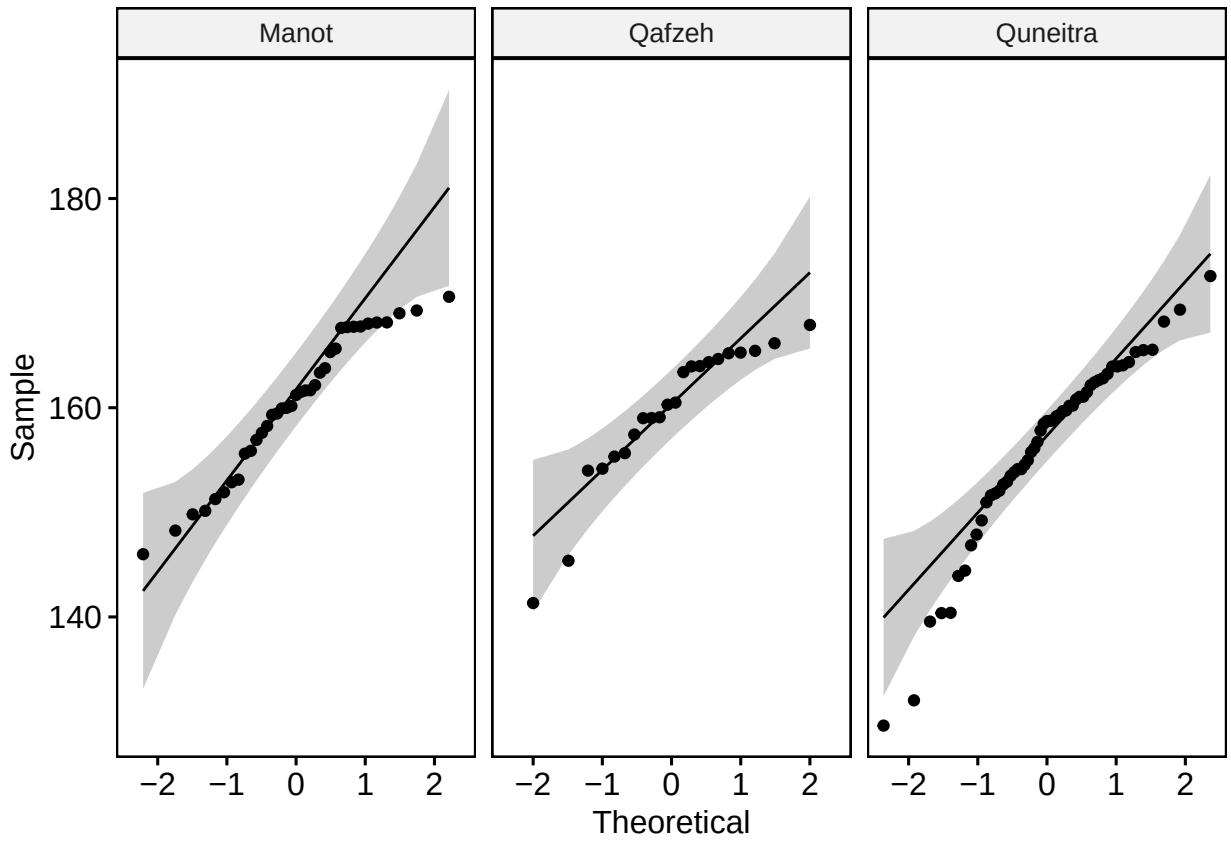

```
shapiro_test(residuals(model))
```

```
## # A tibble: 1 x 3
##   variable         statistic  p.value
##   <chr>                <dbl>    <dbl>
## 1 residuals(model)     0.947 0.000198
```

```
db %>%
  group_by(site) %>%
  shapiro_test(angle)
```

```
## # A tibble: 3 x 4
##   site     variable statistic       p
##   <chr>    <chr>        <dbl>   <dbl>
## 1 Manot    angle        0.949 0.0893
## 2 Qafzeh   angle        0.876 0.0101
## 3 Quneitra angle        0.934 0.00483
```

```
ggqqplot(db, "angle", facet.by = "site")
```




```
ang.aov <- db %>% anova_test(angle ~ site)
```

```
## Coefficient covariances computed by hccm()
```

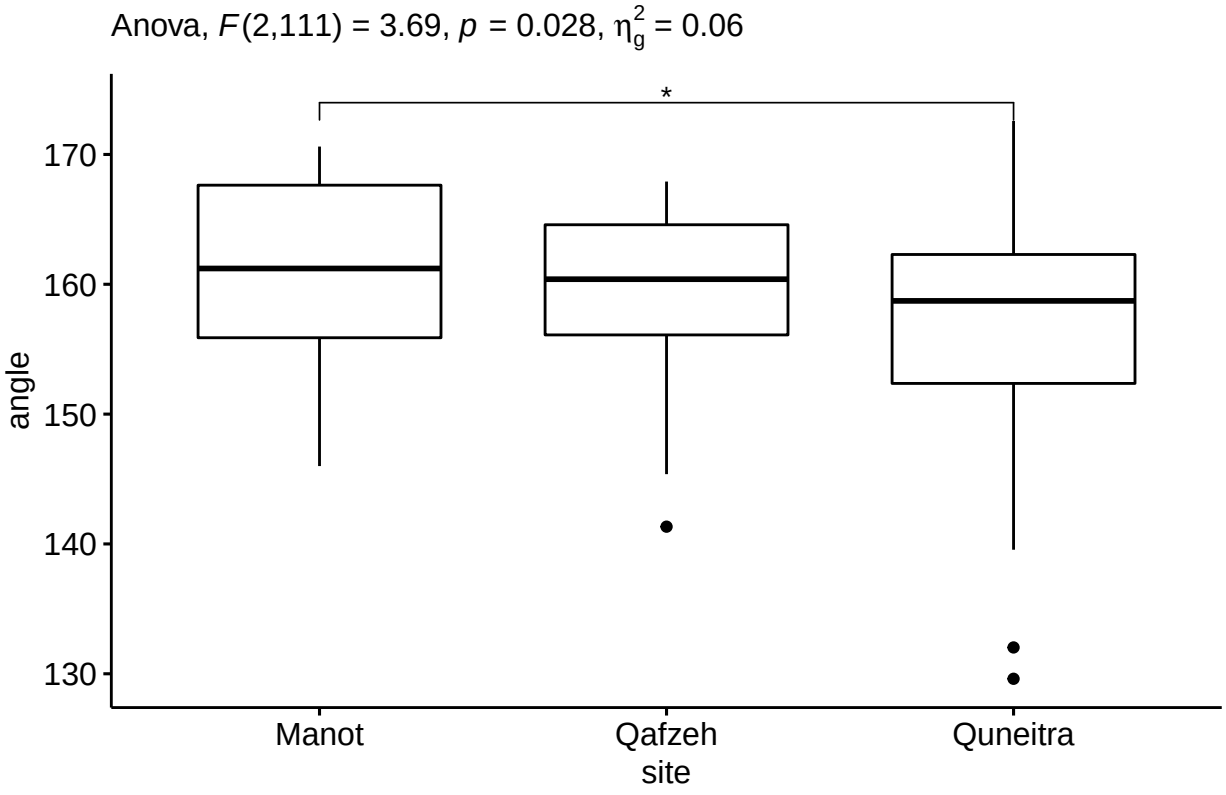
```
ang.aov
```

```
## ANOVA Table (type II tests)
##
##   Effect DFn DFd     F     p p<.05   ges
## 1   site   2 111 3.686 0.028     * 0.062
```

```
# post-hoc test

ang.pwc <- db %>% tukey_hsd(angle ~ site)
ang.pwc
```

```
## # A tibble: 3 x 9
##   term  group1 group2 null.value estimate conf.low conf.high  p.adj p.adj.signif
## * <chr> <chr>  <chr>       <dbl>    <dbl>    <dbl>     <dbl>  <dbl> <chr>
## 1 site  Manot  Qafzeh          0   -0.844    -5.84     4.16  0.915  ns
## 2 site  Manot  Qunei~          0   -4.25     -8.20    -0.301 0.0318 *
## 3 site  Qafzeh Qunei~          0   -3.41     -8.09     1.28  0.2    ns
```

```r
# plot the results

ang.pwc.plot <- ang.pwc %>% add_xy_position(x = "site")


ggboxplot(db, x = "site", y = "angle") +
  stat_pvalue_manual(ang.pwc.plot, hide.ns = TRUE) +
  labs(
    subtitle = get_test_label(ang.aov, detailed = TRUE),
    caption = get_pwc_label(ang.pwc.plot)
    )
```



Kruskal-Wallis test

```r
ang.kruskal <- db %>% kruskal_test(angle ~ site)
ang.kruskal
```

```
## # A tibble: 1 x 6
##   .y.       n statistic    df      p method
## * <chr> <int>     <dbl> <int>  <dbl> <chr>
## 1 angle   114      6.24     2 0.0441 Kruskal-Wallis
```

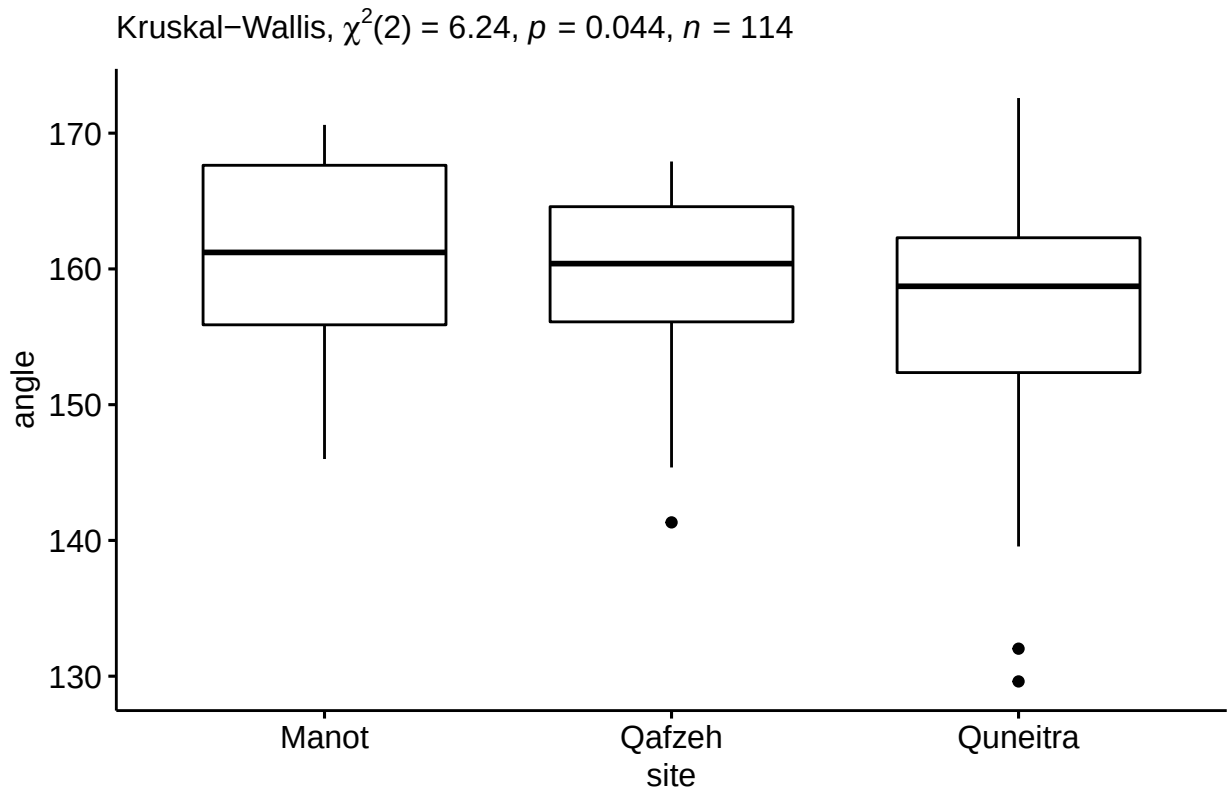
```
db %>% kruskal_effsize(angle ~ site)
```

```
## # A tibble: 1 x 5
##   .y.       n effsize method  magnitude
## * <chr> <int>   <dbl> <chr>   <ord>
## 1 angle   114  0.0382 eta2[H] small
```

```
# comparison between pairs

ang.will <- db %>%
  wilcox_test(angle ~ site, p.adjust.method = "bonferroni")
ang.will
```

```
## # A tibble: 3 x 9
##   .y.   group1 group2      n1    n2 statistic     p p.adj p.adj.signif
## * <chr> <chr>  <chr>    <int> <int>     <dbl> <dbl> <dbl> <chr>
## 1 angle Manot  Qafzeh      37    22       437 0.646 1     ns
## 2 angle Manot  Quneitra    37    55      1292 0.029 0.087 ns
## 3 angle Qafzeh Quneitra    22    55       769 0.065 0.196 ns
```

```
ang.will <- ang.will %>% add_xy_position(x = "site")

ggboxplot(db, x = "site", y = "angle") +
  stat_pvalue_manual(ang.will, hide.ns = TRUE) +
  labs(
    subtitle = get_test_label(ang.kruskal, detailed = TRUE),
    caption = get_pwc_label(ang.will)
    )
```




```
ggsave("../plots/anova.png")
```

```
## Saving 6.5 x 4.5 in image
```

End of Script