SOM - Read and plot data from the manuscript Paixao et al.

Paixao et al. Using Mechanical experiments to study Ground Stone Tools use: exploring the formation of

28/09/2020

Brief description of the script

This R markdown socument reads, summarizes and plots data for the manuscript Paixao et al. The document contais:

- 1. Manuscript tables
- 2. Manuscript figures (data plots)
- 3. Supplementary material, including extra tables and figures (data plots)

This R project and respective script follows the proceadures described by Marwick et al. 2017. To compile this markdown document do not delete or move files from their original folders. Please note that the tables and figures in this file do not match the numbering in the original manuscript.

For any questions, comments and inputs, please contact:

Joao Marreiros, marreiros@rgzm.de, or Eduardo Paixao, paixao@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'

1) Load libraries and datasets

```
# Load required libraries
library(tidyverse)
## -- Attaching packages ------ tidyverse 1.3.0 --
## v ggplot2 3.3.2
                     v purrr
                               0.3.4
## v tibble 3.0.3
                     v dplyr
                               1.0.2
## v tidyr
          1.1.2
                     v stringr 1.4.0
## v readr
          1.3.1
                     v forcats 0.5.0
## Warning: package 'tibble' was built under R version 4.0.2
## Warning: package 'dplyr' was built under R version 4.0.2
```

```
## -- Conflicts ------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()
                    masks stats::lag()
library(utils)
library(knitr)
library(janitor)
##
## Attaching package: 'janitor'
## The following objects are masked from 'package:stats':
##
##
       chisq.test, fisher.test
library(kableExtra)
##
## Attaching package: 'kableExtra'
## The following object is masked from 'package:dplyr':
##
##
      group_rows
library(GGally)
## Warning: package 'GGally' was built under R version 4.0.2
## Registered S3 method overwritten by 'GGally':
##
    method from
    +.gg ggplot2
library(doBy)
## Warning: package 'doBy' was built under R version 4.0.2
## Attaching package: 'doBy'
## The following object is masked from 'package:dplyr':
##
##
      order_by
library(ggpubr)
## Warning: package 'ggpubr' was built under R version 4.0.2
```

```
# Load datasets
gisdata <- read.csv("../analysis/raw_data/gisdata.csv")</pre>
confocaldata <- read_csv("../analysis/raw_data/confocaldata.csv")</pre>
## Parsed with column specification:
## cols(
##
     .default = col_double(),
##
    Name = col_character(),
##
     `Created on` = col_character(),
##
     sample = col_character(),
##
    motion = col_character(),
     workedmaterial = col_character(),
##
     `Studiable type` = col_character(),
     `Axis name - X` = col_character(),
##
##
     `Axis name - Y` = col_character(),
##
     `Axis name - Z` = col_character(),
     `Layer type - Z` = col_character(),
     `Lengthscale anisotropy Sfrax epLsar` = col_character(),
##
     `Lengthscale anisotropy NewEplsar` = col_character()
##
## )
## See spec(...) for full column specifications.
In this study, two datasets are used:
  1) gisdata.csv: dataset for the QGIS analysis
str(gisdata)
## 'data.frame':
                   197 obs. of 11 variables:
## $ sample : chr "id2-5" "id2-5" "id3-3" "id3-3" ...
## $ parameter: chr "tri" "tri" "tri" "tri" ...
## $ motion : chr "Impact" "Impact" "Impact" "Impact" ...
## $ material : chr "Flint" "Flint" "Flint" "Flint" ...
           : int 0 1 0 1 2 3 4 0 1 0 ...
## $ id
## $ elev_min : num 0 0.01 0 0.01 0.02 0.03 0.04 0 0.01 0 ...
## $ elev_max : num 0.01 0.02 0.01 0.02 0.03 0.04 0.05 0.01 0.02 0.01 ...
## $ nparts : int 1007 1010 131 165 47 10 1 11 12 6 ...
## $ npoints : int 67487 59648 14961 16143 3142 423 7 6559 1248 8538 ...
## $ perimeter: num 485.2 446 160.6 177.5 32.8 ...
              : num 204.49 17.48 83.65 38.12 1.83 ...
## $ area
  2) confocaldata.csv: dataset for the Confocal microscopy surface texture analysis
str(confocaldata)
## tibble [25 x 54] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
## $ Name
                                         : chr [1:25] "Lime2-5_LSM_50x075_suf1_Topo > Leveled (LS-plane
## $ Created on
                                         : chr [1:25] "6/24/2020 12:03:05 PM" "6/24/2020 12:21:59 PM" "
                                         : chr [1:25] "id2-5" "id2-5" "id2-5" "id3-3" ...
## $ sample
```

```
: num [1:25] 0.836 4.619 2.321 5.348 3.491 ...
## $ Sq
## $ Ssk
                                      : num [1:25] 0.997 0.107 0.142 -0.491 0.201 ...
## $ Sku
                                       : num [1:25] 8.63 4.48 3.11 5.97 3.46 ...
## $ Sp
                                      : num [1:25] 4.8 15.02 7.46 23 11.88 ...
## $ Sv
                                      : num [1:25] 3.49 15.58 7.45 25.58 11.11 ...
## $ Sz
                                       : num [1:25] 8.3 30.6 14.9 48.6 23 ...
## $ Sa
                                       : num [1:25] 0.572 3.244 1.819 3.887 2.699 ...
## $ Smr
                                      : num [1:25] 0.464 0.497 0.448 0.18 0.354 ...
## $ Smc
                                      : num [1:25] 0.775 5.691 2.944 5.799 4.436 ...
## $ Sxp
                                       : num [1:25] 1.53 10.61 4.3 11.65 6.74 ...
## $ Sal
                                      : num [1:25] 13.5 18.9 20.1 18.7 22.9 ...
## $ Str
                                      : num [1:25] 0.371 0.416 0.592 0.468 0.803 ...
                                      : num [1:25] 149 65 150 51 124 ...
## $ Std
## $ Sdq
                                      : num [1:25] 0.328 1.153 0.688 1.126 0.897 ...
## $ Sdr
                                      : num [1:25] 4.36 20.02 16.47 31.23 24.6 ...
## $ Vm
                                      : num [1:25] 0.0866 0.3378 0.1331 0.3133 0.2114 ...
## $ Vv
                                      : num [1:25] 0.861 6.029 3.078 6.113 4.648 ...
                                       : num [1:25] 0.0866 0.3378 0.1331 0.3133 0.2114 ...
## $ Vmp
## $ Vmc
                                      : num [1:25] 0.528 3.111 2.097 3.932 3.146 ...
## $ Vvc
                                      : num [1:25] 0.769 5.335 2.824 5.33 4.303 ...
## $ Vvv
                                      : num [1:25] 0.0923 0.694 0.2534 0.7826 0.3444 ...
## $ Maximum depth of furrows
                                    : num [1:25] 4.56 20.63 7.65 25.68 13.88 ...
## $ Mean depth of furrows
                                      : num [1:25] 0.962 4.63 2.49 5.112 3.932 ...
## $ Mean density of furrows
                                      : num [1:25] 4523 3830 4509 2286 2201 ...
## $ First direction
                                      : num [1:25] 1.50e+02 6.36e+01 2.66e-03 9.00e+01 9.00e+01 ...
## $ Second direction
                                      : num [1:25] 180 45 154 45 135 ...
## $ Third direction
                                      : num [1:25] 141.3 56.2 63.5 51.2 123.7 ...
                                       : num [1:25] 23.5 26.2 77.8 33.1 73.8 ...
## $ Isotropy
## $ Lengthscale anisotropy Sfrax epLsar: chr [1:25] "***** "***** "***** "0.000493204" ...
## $ Lengthscale anisotropy NewEplsar : chr [1:25] "*****" "*****" "*****" "0.017686885" ...
## $ Fractal complexity Asfc : num [1:25] 8.66 23.18 30.55 37.79 44.11 ... ## $ Scale of max complexity Smfc : num [1:25] 1.71e+06 1.81e+08 2.71e+06 1.17e
                                     : num [1:25] 1.71e+06 1.81e+08 2.71e+06 1.17e+01 1.52e+01 ...
                                      : num [1:25] 0.449 2.496 0.388 0.544 0.254 ...
## $ HAsfc9
## $ HAsfc81
                                      : num [1:25] 0.659 3.446 0.481 0.701 0.487 ...
## - attr(*, "spec")=
##
    .. cols(
##
   .. Name = col_character(),
```

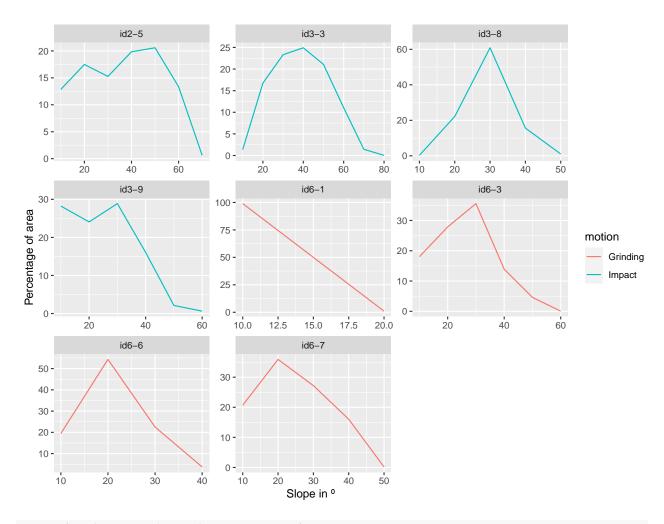
```
##
          `Created on` = col_character(),
##
          sample = col_character(),
     . .
          motion = col character(),
##
     . .
          workedmaterial = col_character(),
##
##
          `Studiable type` = col_character(),
     . .
          `Axis name - X` = col_character(),
##
##
          `Axis length - X` = col double(),
     . .
          `Axis size - X` = col_double(),
##
          `Axis spacing - X` = col_double(),
##
     . .
##
          `Axis name - Y` = col_character(),
##
          `Axis length - Y` = col_double(),
          `Axis size - Y` = col_double(),
##
          `Axis spacing - Y` = col_double(),
##
     . .
##
          `Axis name - Z` = col_character(),
     . .
##
          `Layer type - Z` = col_character(),
          `Axis length - Z` = col_double(),
##
     . .
##
          `Axis size - Z` = col_double(),
##
          `Axis spacing - Z` = col double(),
     . .
##
          `NM-points ratio - Z` = col_double(),
          Sq = col_double(),
##
     . .
##
          Ssk = col_double(),
##
          Sku = col_double(),
     . .
          Sp = col_double(),
##
##
          Sv = col double(),
     . .
##
          Sz = col_double(),
          Sa = col_double(),
##
     . .
##
          Smr = col_double(),
          Smc = col_double(),
##
     . .
##
          Sxp = col_double(),
##
          Sal = col_double(),
##
          Str = col_double(),
     . .
##
          Std = col_double(),
##
          Sdq = col_double(),
     . .
##
          Sdr = col_double(),
##
          Vm = col double(),
     . .
##
          Vv = col_double(),
     . .
##
     . .
          Vmp = col double(),
##
          Vmc = col_double(),
##
          Vvc = col_double(),
     . .
##
          Vvv = col_double(),
##
          `Maximum depth of furrows` = col_double(),
     . .
##
          `Mean depth of furrows` = col_double(),
          `Mean density of furrows` = col_double(),
##
     . .
##
          `First direction` = col_double(),
##
          `Second direction` = col_double(),
          `Third direction` = col_double(),
##
     . .
##
          Isotropy = col_double(),
     . .
##
          `Lengthscale anisotropy Sfrax epLsar` = col_character(),
     . .
##
          `Lengthscale anisotropy NewEplsar` = col_character(),
          `Fractal complexity Asfc` = col_double(),
##
##
          `Scale of max complexity Smfc` = col_double(),
     . .
##
     . .
          HAsfc9 = col double(),
##
          HAsfc81 = col_double()
     . .
##
     ..)
```

GIS analysis, Terrain analysis for Slope and TRI

Slope

```
# Compute proportions for perimeter and area groupped by sample and parameter
slope <- filter(gisdata, parameter == "slope")</pre>
id2.5 \leftarrow filter(slope, sample == "id2-5")
id3.3 <- filter(slope, sample == "id3-3")
id3.8 <- filter(slope, sample == "id3-8")
id3.9 <- filter(slope, sample == "id3-9")
id6.1 <- filter(slope, sample == "id6-1")</pre>
id6.3 <- filter(slope, sample == "id6-3")</pre>
id6.6 <- filter(slope, sample == "id6-6")
id6.7 <- filter(slope, sample == "id6-7")
id2.5 <- id2.5 %>%
 group_by(sample) %>%
 mutate(
    areaperc = area / sum(area) * 100,
    perimperc = perimeter / sum(perimeter) * 100)
id3.3 <- id3.3 %>%
  group_by(sample) %>%
 mutate(
    areaperc = area / sum(area) * 100,
    perimperc = perimeter / sum(perimeter) * 100)
id3.8 <- id3.8 %>%
  group_by(sample) %>%
 mutate(
    areaperc = area / sum(area) * 100,
    perimperc = perimeter / sum(perimeter) * 100)
id3.9 <- id3.9 %>%
  group_by(sample) %>%
  mutate(
    areaperc = area / sum(area) * 100,
    perimperc = perimeter / sum(perimeter) * 100)
id6.1 <- id6.1 %>%
  group_by(sample) %>%
 mutate(
    areaperc = area / sum(area) * 100,
    perimperc = perimeter / sum(perimeter) * 100)
id6.3 <- id6.3 %>%
  group_by(sample) %>%
    areaperc = area / sum(area) * 100,
  perimperc = perimeter / sum(perimeter) * 100)
```

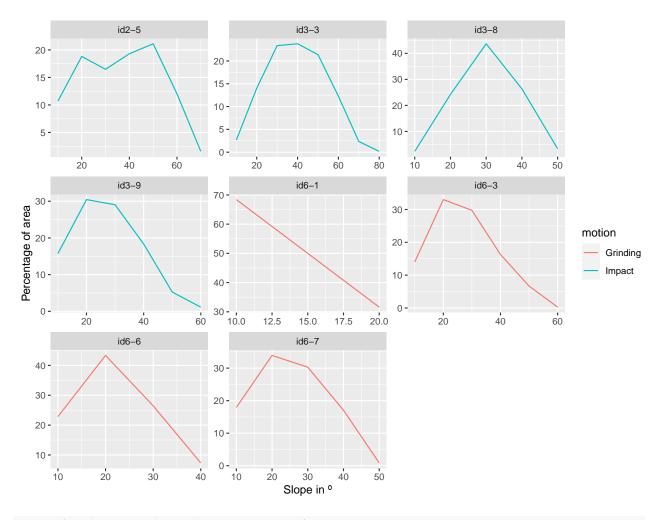
```
id6.6 <- id6.6 %>%
  group_by(sample) %>%
 mutate(
   areaperc = area / sum(area) * 100,
    perimperc = perimeter / sum(perimeter) * 100)
id6.7 <- id6.7 %>%
  group_by(sample) %>%
  mutate(
    areaperc = area / sum(area) * 100,
    perimperc = perimeter / sum(perimeter) * 100)
newslope <- do.call("rbind", list(id2.5, id3.3, id3.8, id3.9, id6.1, id6.3, id6.6, id6.7))
# save output
write_csv(newslope,"../analysis/derived_data/newslope.csv")
# Plot data
areaP <- ggplot(newslope, aes(x = elev_max, y = areaperc, colour = motion)) +</pre>
  geom_line() +
 facet_wrap(~sample, scale = "free") +
 ylab("Percentage of area") +
 xlab("Slope in º")
areaP
```



ggsave("../analysis/plots/slopearea.png")

Saving 8.5×6.5 in image

```
perimP <- ggplot(newslope, aes(x = elev_max, y = perimperc, colour = motion)) +
    geom_line() +
    facet_wrap(~sample, scale = "free") +
    ylab("Percentage of area") +
    xlab("Slope in º")</pre>
```



ggsave("../analysis/plots/slopeperim.png")

Saving 8.5×6.5 in image

TRI (Terrain roughness index)

```
tri <- filter(gisdata, parameter == "tri")

id2.5 <- filter(tri, sample == "id2-5")

id3.3 <- filter(tri, sample == "id3-3")

id3.8 <- filter(tri, sample == "id3-8")

id3.9 <- filter(tri, sample == "id3-9")

id6.1 <- filter(tri, sample == "id6-1")

id6.3 <- filter(tri, sample == "id6-3")

id6.6 <- filter(tri, sample == "id6-6")

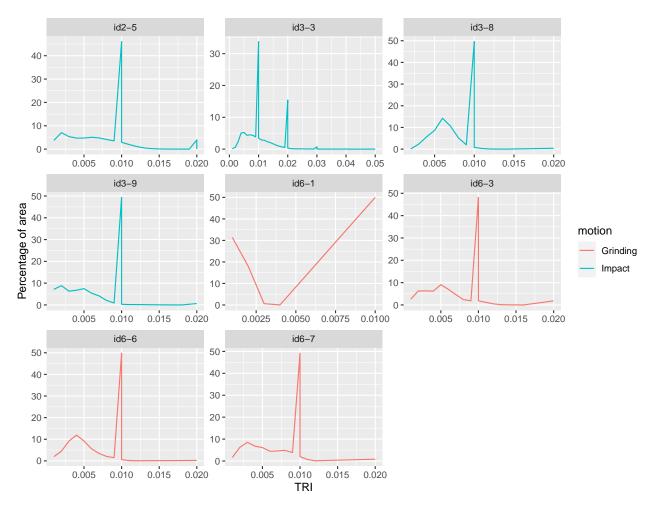
id6.7 <- filter(tri, sample == "id6-7")

id2.5 <- id2.5 %>%

    group_by(sample) %>%
    mutate(
```

```
areaperc = area / sum(area) * 100,
    perimperc = perimeter / sum(perimeter) * 100)
id3.3 <- id3.3 %>%
  group_by(sample) %>%
  mutate(
    areaperc = area / sum(area) * 100,
    perimperc = perimeter / sum(perimeter) * 100)
id3.8 <- id3.8 %>%
  group_by(sample) %>%
  mutate(
    areaperc = area / sum(area) * 100,
    perimperc = perimeter / sum(perimeter) * 100)
id3.9 <- id3.9 %>%
  group_by(sample) %>%
  mutate(
    areaperc = area / sum(area) * 100,
    perimperc = perimeter / sum(perimeter) * 100)
id6.1 <- id6.1 %>%
  group_by(sample) %>%
  mutate(
    areaperc = area / sum(area) * 100,
    perimperc = perimeter / sum(perimeter) * 100)
id6.3 <- id6.3 %>%
  group_by(sample) %>%
  mutate(
    areaperc = area / sum(area) * 100,
    perimperc = perimeter / sum(perimeter) * 100)
id6.6 <- id6.6 %>%
  group_by(sample) %>%
  mutate(
    areaperc = area / sum(area) * 100,
    perimperc = perimeter / sum(perimeter) * 100)
id6.7 <- id6.7 %>%
  group_by(sample) %>%
  mutate(
    areaperc = area / sum(area) * 100,
    perimperc = perimeter / sum(perimeter) * 100)
newtri <- do.call("rbind", list(id2.5, id3.3, id3.8, id3.9, id6.1, id6.3, id6.6, id6.7))
# save output
write_csv(newtri,"../analysis/derived_data/newtri.csv")
# Plot data
```

```
areaP <- ggplot(newtri, aes(x = elev_max, y = areaperc, colour = motion)) +
   geom_line() +
   facet_wrap(~sample, scale = "free") +
   ylab("Percentage of area") +
   xlab("TRI")
areaP</pre>
```

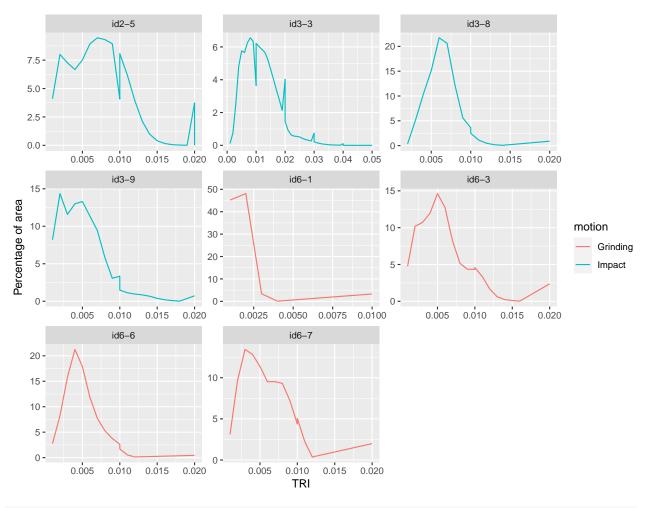


ggsave("../analysis/plots/triarea.png")

Saving 8.5×6.5 in image

```
perimP <- ggplot(newtri, aes(x = elev_max, y = perimperc, colour = motion)) +
   geom_line() +
   facet_wrap(~sample, scale = "free") +
   ylab("Percentage of area") +
   xlab("TRI")

perimP</pre>
```



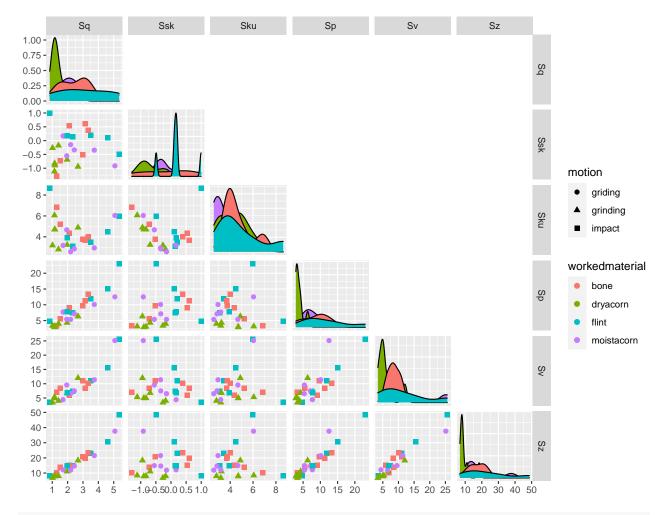
```
ggsave("../analysis/plots/perimtri.png")
```

Saving 8.5 x 6.5 in image

Import, summarize and plot all Confocal data

```
lower = list(continuous = wrap("points", alpha = 1, size = 2)),
upper = list(continuous = "blank"),
legend = c(2,1)
) +

theme(legend.position = "right") +
labs(fill = "Micro polish type")
```



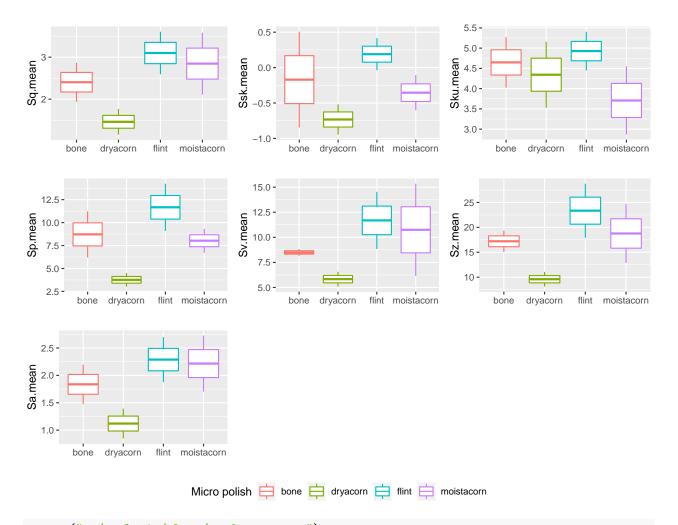
ggsave("../analysis/plots/confocal.png")

Saving 8.5×6.5 in image

```
# 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)</pre>
```

```
sd_test <- sd(y)</pre>
    out <- c(n_test, min_test, max_test, mean_test, med_test, sd_test)</pre>
    names(out) <- c("n", "min", "max", "mean", "median", "sd")</pre>
    return(out)
}
positions <-c(1:5,21:27)
df <- confocaldata %>%
  select(positions)
## Note: Using an external vector in selections is ambiguous.
## i Use `all_of(positions)` instead of `positions` to silence this message.
## i See <https://tidyselect.r-lib.org/reference/faq-external-vector.html>.
## This message is displayed once per session.
num.var <- 5:length(df)</pre>
confostats <- summaryBy(.~sample+workedmaterial, data=df[c("sample", "motion", "workedmaterial", names(d
write_csv(confostats, "../analysis/derived_data/confostats.csv")
# Plot confostats for the ISO 25178 Height parameters
# select parameter from datase
heightconfostats <- select(confostats, sample, workedmaterial, Sq. mean, Ssk. mean, Sku. mean, Sp. mean, Sv. mean,
p1 <- ggplot(heightconfostats, aes(x=workedmaterial, y=Sq.mean, colour=workedmaterial)) +
  geom_boxplot() +
  labs(x="", colour="Micro polish")
p2 <- ggplot(heightconfostats, aes(x=workedmaterial, y=Ssk.mean, colour=workedmaterial)) +
                                                                                               geom_boxpl
  labs(x="", colour="Micro polish")
p3 <- ggplot(heightconfostats, aes(x=workedmaterial, y=Sku.mean, colour=workedmaterial)) +
                                                                                               geom_boxpl
  labs(x="", colour="Micro polish")
p4 <- ggplot(heightconfostats, aes(x=workedmaterial, y=Sp.mean, colour=workedmaterial)) +
  geom_boxplot() +
  labs(x="", colour="Micro polish")
p5 <- ggplot(heightconfostats, aes(x=workedmaterial, y=Sv.mean, colour=workedmaterial)) +
  geom_boxplot() +
  labs(x="", colour="Micro polish")
p6 <- ggplot(heightconfostats, aes(x=workedmaterial, y=Sz.mean, colour=workedmaterial)) +
  geom_boxplot() +
  labs(x="", colour="Micro polish")
p7 <- ggplot(heightconfostats, aes(x=workedmaterial, y=Sa.mean, colour=workedmaterial)) +
  geom_boxplot() +
  labs(x="", colour="Micro polish")
ggarrange(p1, p2, p3, p4, p5, p6, p7, ncol = 3, nrow = 3, common.legend = TRUE, legend="bottom")
```



ggsave("../analysis/plots/confostats.png")

Saving 8.5×6.5 in image

End of script

sessionInfo()

```
## R version 4.0.0 Patched (2020-05-04 r78358)
## Platform: x86_64-apple-darwin17.0 (64-bit)
## Running under: macOS Catalina 10.15.6
##
## Matrix products: default
## BLAS: /Library/Frameworks/R.framework/Versions/4.0/Resources/lib/libRblas.dylib
## LAPACK: /Library/Frameworks/R.framework/Versions/4.0/Resources/lib/libRlapack.dylib
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## attached base packages:
```

```
## [1] stats
                 graphics grDevices utils
                                                datasets methods
##
## other attached packages:
   [1] ggpubr_0.4.0
                         doBy_4.6.7
                                           GGally_2.0.0
                                                            kableExtra_1.2.1
##
   [5] janitor_2.0.1
                         knitr_1.29
                                           forcats_0.5.0
                                                            stringr_1.4.0
  [9] dplyr 1.0.2
                         purrr_0.3.4
                                           readr 1.3.1
                                                            tidyr_1.1.2
##
                                           tidyverse_1.3.0
## [13] tibble 3.0.3
                         ggplot2_3.3.2
##
## loaded via a namespace (and not attached):
  [1] httr_1.4.2
                                                                  carData_3.0-4
                           jsonlite_1.7.1
                                               viridisLite_0.3.0
                                                                  cellranger_1.1.0
  [5] modelr_0.1.8
                           assertthat_0.2.1
                                               blob_1.2.1
   [9] yaml_2.2.1
                           pillar_1.4.6
                                                                  lattice_0.20-41
                                               backports_1.1.9
##
                           digest_0.6.25
                                               RColorBrewer_1.1-2 ggsignif_0.6.0
## [13] glue_1.4.2
## [17] rvest_0.3.6
                           snakecase_0.11.0
                                               colorspace_1.4-1
                                                                  cowplot_1.0.0
## [21] htmltools_0.5.0
                           Matrix_1.2-18
                                               plyr_1.8.6
                                                                  pkgconfig_2.0.3
## [25] broom_0.7.0
                           haven_2.3.1
                                               scales_1.1.1
                                                                  webshot_0.5.2
## [29] openxlsx_4.1.5
                           rio_0.5.16
                                               farver_2.0.3
                                                                  generics_0.0.2
## [33] car 3.0-9
                           ellipsis_0.3.1
                                               withr 2.2.0
                                                                  cli 2.0.2
                           crayon_1.3.4
                                               readxl_1.3.1
                                                                  evaluate_0.14
## [37] magrittr_1.5
## [41] fs 1.5.0
                           fansi_0.4.1
                                               MASS_7.3-52
                                                                  rstatix 0.6.0
## [45] xml2_1.3.2
                           foreign_0.8-80
                                               tools_4.0.0
                                                                  data.table_1.13.0
## [49] hms 0.5.3
                           lifecycle_0.2.0
                                               munsell_0.5.0
                                                                  reprex_0.3.0
## [53] zip_2.1.1
                           Deriv_4.0.1
                                               compiler_4.0.0
                                                                  rlang_0.4.7
## [57] grid 4.0.0
                           rstudioapi_0.11
                                               labeling 0.3
                                                                  rmarkdown 2.3
                                                                  reshape_0.8.8
                                               DBI 1.1.0
## [61] gtable_0.3.0
                           abind_1.4-5
                                               gridExtra_2.3
                                                                  lubridate_1.7.9
## [65] curl_4.3
                           R6_2.4.1
## [69] stringi_1.5.3
                           Rcpp_1.0.5
                                               vctrs_0.3.4
                                                                  dbplyr_1.4.4
## [73] tidyselect_1.1.0
                           xfun_0.17
```