

# Analyse

Baptiste Crinière-Boizet

## Contents

<b>Analyse on raw data</b>	<b>1</b>
Descriptive statistics . . . . .	1
<b>Bi-variate analysis</b>	<b>2</b>
Correlation . . . . .	2
Boxplot & wilcoxon test . . . . .	2
Linear model . . . . .	3
Linear model with log transformation . . . . .	5
<b>Partial data</b>	<b>7</b>

```
# Library
library(ggsci)
library(tidyverse)
library(ggpubr)

# Data management
data <- openxlsx::read.xlsx("/home/baptiste.crinier/Document/PB_cell_CB/Data/Tableau data 170723.xlsx")
names(data) <- c("ID", "Group", "Cell_count", "Cell_tot", "Astros", "Area", "Prop", "Density")
data <- data %>%
  dplyr::mutate(ID = ID %>% factor) %>%
  dplyr::mutate(Group = Group %>% factor)
```

## Analyse on raw data

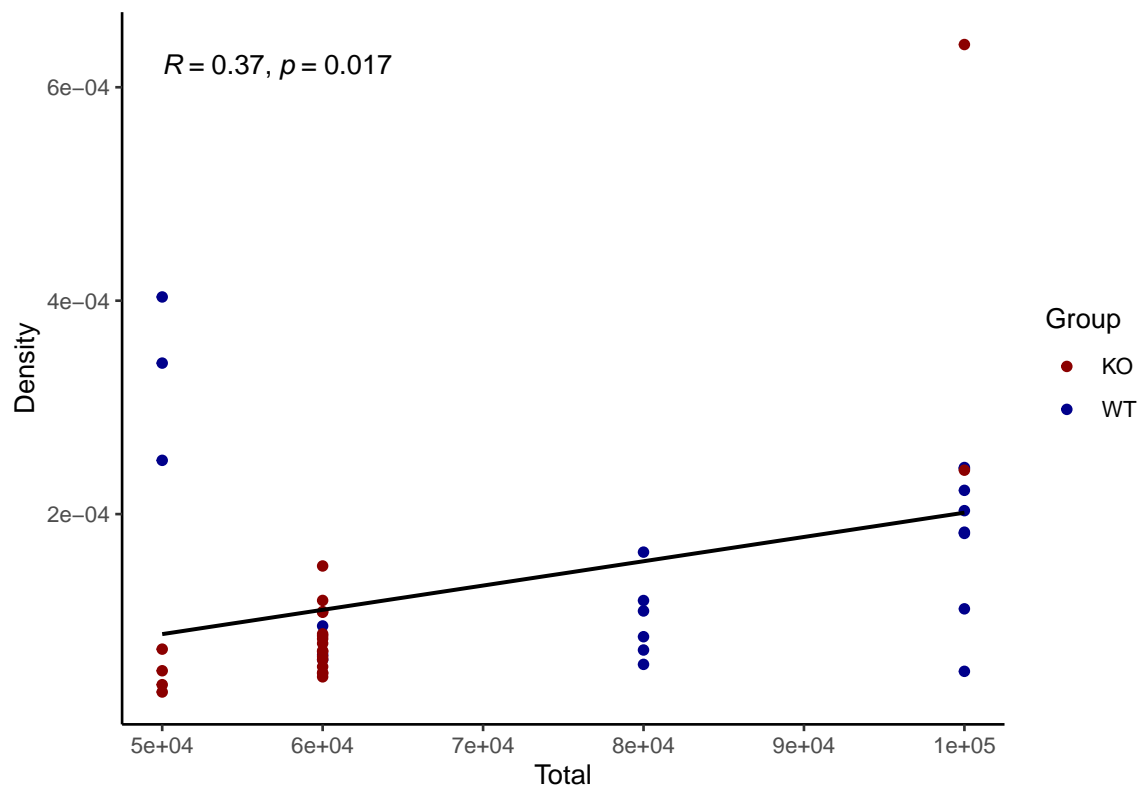
### Descriptive statistics

```
tab <- data %>% dplyr::select(-ID) %>%
  furniture::table1("Number of cells" = Cell_count,
                    "Total cell count" = Cell_tot,
                    Area,
                    "Ratio counted/plated" = Prop,
                    Density,
                    splitby = ~ Group,
                    digits = 5)
```

## Bi-variate analysis

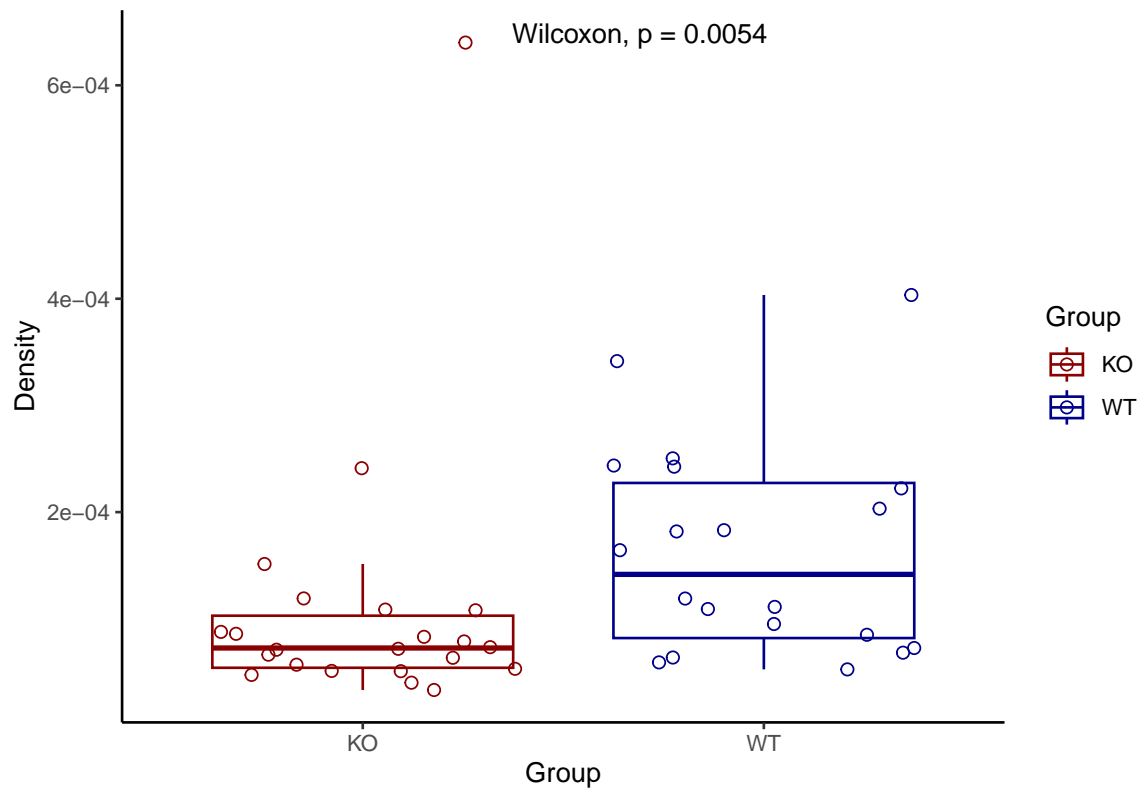
### Correlation

```
data %>%  
  ggplot(aes(x = Cell_tot, y = Density))+  
  geom_point(aes(color = Group), size = 1.5)+  
  geom_smooth(method = "lm", se = F, color = "black", size = 0.75)+  
  theme_classic()+  
  stat_cor()+  
  scale_color_manual(values = c("darkred", "darkblue"))+  
  labs(x = "Total")
```



### Boxplot & wilcoxon test

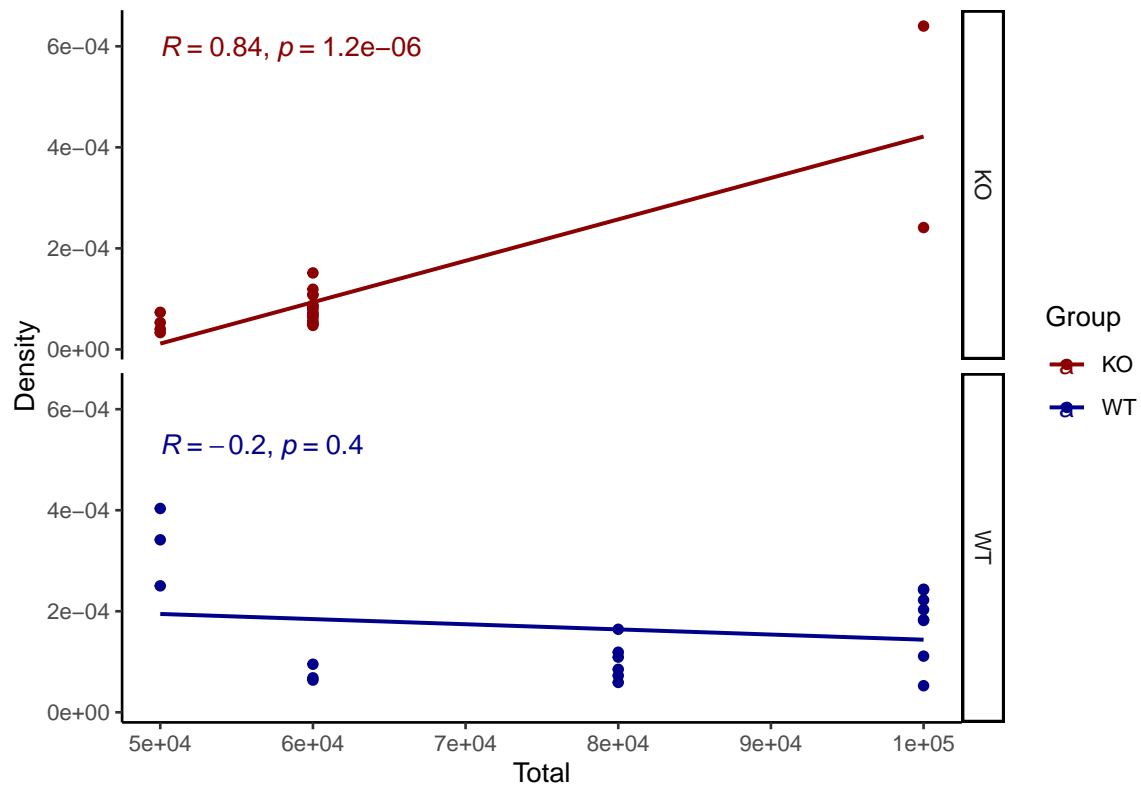
```
data %>%  
  ggplot(aes(x = Group, y = Density))+  
  geom_boxplot(aes(color = Group), outlier.shape = NA)+  
  geom_jitter(aes(color = Group), size = 2, shape = 1)+  
  stat_compare_means(label.x = 1.5)+  
  theme_classic()+  
  scale_color_manual(values = c("darkred", "darkblue"))
```



## Linear model

### Data visualisation

```
data %>%
  ggplot(aes(x = Cell_tot, y = Density, color = Group))+
  geom_point()+
  geom_smooth(method = "lm", se = FALSE, size = 0.75)+
  stat_cor()+
  facet_grid(rows = vars(Group))+
  theme_classic()+
  scale_color_manual(values = c("darkred", "darkblue"))+
  labs(x = "Total")
```



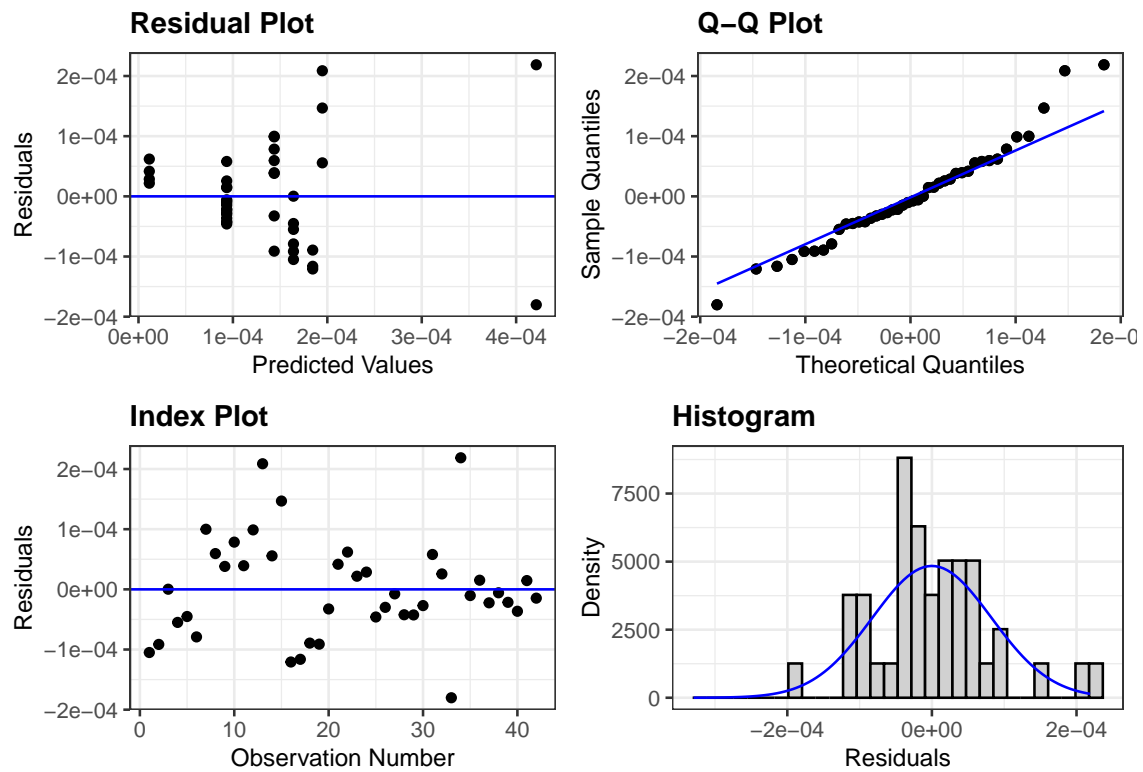
## Model

```
model <- lm(Density ~ Cell_tot*Group, data = data)
car::Anova(model)
```

```
## Anova Table (Type II tests)
##
## Response: Density
##              Sum Sq Df F value    Pr(>F)
## Cell_tot      4.4260e-08  1  6.0399  0.01866 *
## Group          2.2930e-09  1  0.3129  0.57921
## Cell_tot:Group 2.0012e-07  1 27.3084 6.546e-06 ***
## Residuals      2.7847e-07 38
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

## Assumptions

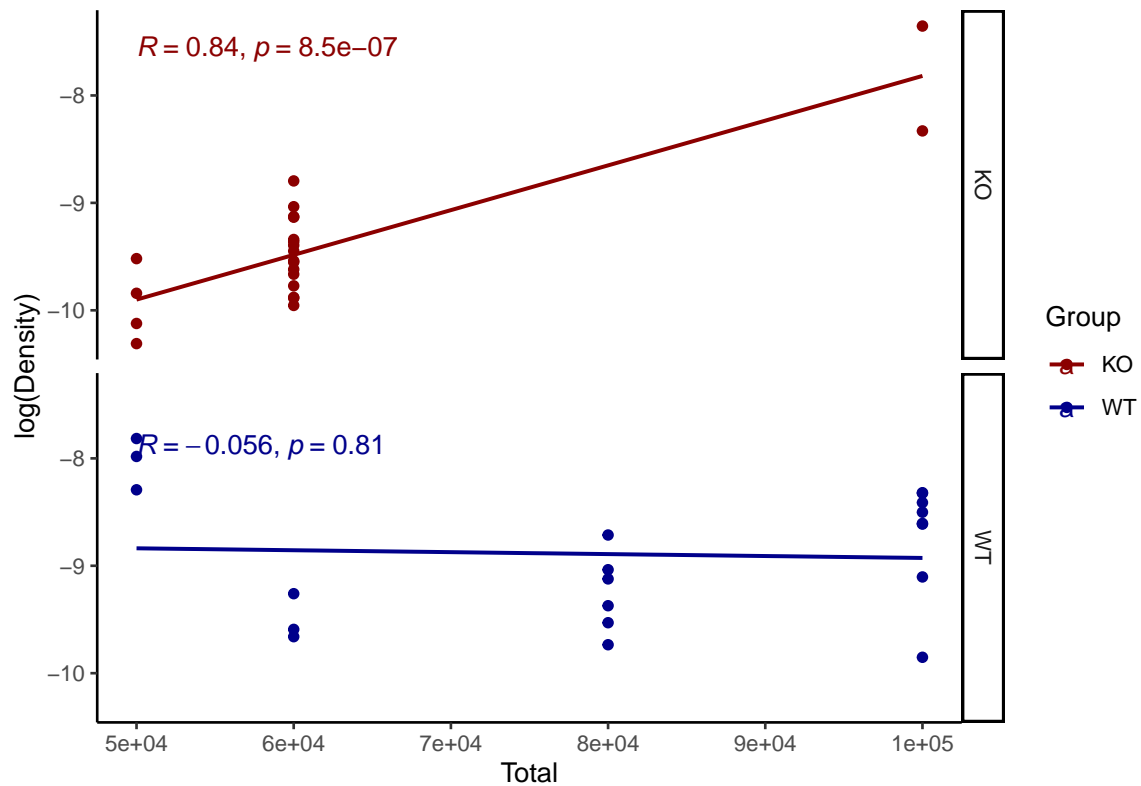
```
ggResidpanel::resid_panel(model)
```



## Linear model with log transformation

### Data visualisation

```
data %>%
  ggplot(aes(x = Cell_tot, y = log(Density), color = Group))+
  geom_point()+
  geom_smooth(method = "lm", se = FALSE, size = 0.75)+
  stat_cor()+
  facet_grid(rows = vars(Group))+
  theme_classic()+
  scale_color_manual(values = c("darkred", "darkblue"))+
  labs(x = "Total")
```



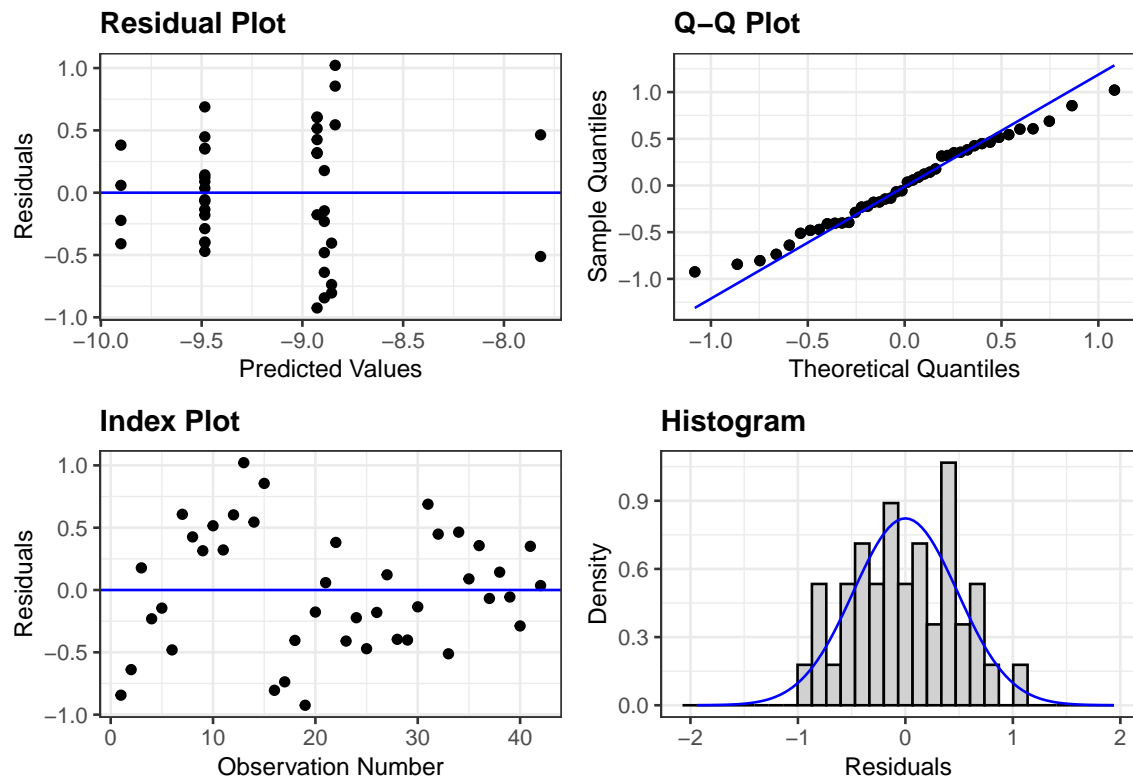
## Model

```
model <- lm(log(Density) ~ Cell_tot*Group, data = data)
car::Anova(model)
```

```
## Anova Table (Type II tests)
##
## Response: log(Density)
##           Sum Sq Df F value    Pr(>F)
## Cell_tot      1.6946  1  6.6753 0.0137466 *
## Group          0.6125  1  2.4129 0.1286321
## Cell_tot:Group 4.4472  1 17.5186 0.0001622 ***
## Residuals      9.6465 38
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

## Assumptions

```
ggResidpanel::resid_panel(model)
```



## Partial data

```
# data_bis <- data %>%
#   dplyr::filter(!(ID %in% c("WT3.2", "K-8.1MBPQK11", "K-8.1MBPQK12")))
```