# Effect of Potassium Fertilizer Rates on Soybean Trifoliate Concentration at Four different Location

Faith Oyewale OLABISI

## Introduction

This study was conducted at four different locations in one year to assess the effect of potassium (K) fertilizer rates on soybean trifoliate K concentration at R4.

Each study was a randomized complete block design (RCBD) with K rate as the only treatment factor with levels 0, 50, 100, and 150 lbs K20/ac.

Note: The experiment work analysed in this project is a simulated experiment.

```
library(tidyverse)
library(car)
library(knitr)
library(readxl)
library(openxlsx)
library(lme4)
library(broom.mixed)
library(ggthemes)
library(emmeans)
library(multcomp)
library(multcompView)
```

```
soy_data <- read_excel('data/SoybeanWorkshop.xlsx', sheet = 'R')

view(soy_data)
```

## Data Insight and Exploration

```r
soy_data |>
  glimpse()
```

```
Rows: 16
Columns: 8
$ Plot        <dbl> 105, 203, 305, 401, 102, 207, 306, 402, 106, 202, 307, 404,~
$ TRT         <dbl> 1, 1, 1, 1, 2, 2, 2, 2, 4, 4, 4, 4, 5, 5, 5, 5
$ Stage       <chr> "R4", "R4", "R4", "R4", "R4", "R4", "R4", "R4", "R4", "R4",~
$ Krate_lbac <dbl> 0, 0, 0, 0, 50, 50, 50, 50, 100, 100, 100, 100, 150, 150, 1~
$ A1.Kpct     <dbl> 2.63, 2.76, 2.79, 2.67, 2.48, 2.87, 3.04, 2.67, 2.83, 3.17,~
$ O2.Kpct     <dbl> 1.17, 1.22, 1.32, 1.14, 1.60, 1.28, 1.30, 1.34, 1.85, 1.74,~
$ S1.Kpct     <dbl> 3.21, 2.90, 3.33, 3.13, 3.24, 3.28, 3.11, 3.35, 3.13, 3.10,~
$ W1.Kpct     <dbl> 0.74, 0.82, 1.05, 1.05, 0.95, 1.17, 1.25, 1.49, 1.21, 1.71,~
```

```r
soy_data |>
  str()
```

```
tibble [16 x 8] (S3: tbl_df/tbl/data.frame)
 $ Plot      : num [1:16] 105 203 305 401 102 207 306 402 106 202 ...
 $ TRT       : num [1:16] 1 1 1 1 2 2 2 2 4 4 ...
 $ Stage     : chr [1:16] "R4" "R4" "R4" "R4" ...
 $ Krate_lbac: num [1:16] 0 0 0 0 50 50 50 50 100 100 ...
 $ A1.Kpct   : num [1:16] 2.63 2.76 2.79 2.67 2.48 2.87 3.04 2.67 2.83 3.17 ...
 $ O2.Kpct   : num [1:16] 1.17 1.22 1.32 1.14 1.6 1.28 1.3 1.34 1.85 1.74 ...
 $ S1.Kpct   : num [1:16] 3.21 2.9 3.33 3.13 3.24 3.28 3.11 3.35 3.13 3.1 ...
 $ W1.Kpct   : num [1:16] 0.74 0.82 1.05 1.05 0.95 1.17 1.25 1.49 1.21 1.71 ...
```

```r
soy_data |>
  filter(
    Krate_lbac == 50 | Krate_lbac == 100
  )
```

```
# A tibble: 8 x 8
   Plot   TRT Stage Krate_lbac A1.Kpct O2.Kpct S1.Kpct W1.Kpct
  <dbl> <dbl> <chr>      <dbl>   <dbl>   <dbl>   <dbl>   <dbl>
1   102     2 R4            50    2.48    1.6     3.24    0.95
2   207     2 R4            50    2.87    1.28    3.28    1.17
```

```
3    306       2 R4             50     3.04     1.3      3.11     1.25
4    402       2 R4             50     2.67     1.34     3.35     1.49
5    106       4 R4            100     2.83     1.85     3.13     1.21
6    202       4 R4            100     3.17     1.74     3.1      1.71
7    307       4 R4            100     2.95     1.5      3.16     1.25
8    404       4 R4            100     2.79     1.19     2.95     1.53
```

```
soy_data |>
  head(n = 10) |>
  kable(
    caption = 'First Ten rows of the data'
  )
```

Table 1: First Ten rows of the data

| Plot | TRT | Stage | Krate_lbac | A1.Kpct | O2.Kpct | S1.Kpct | W1.Kpct |
|---|---|---|---|---|---|---|---|
| 105 | 1 | R4 | 0 | 2.63 | 1.17 | 3.21 | 0.74 |
| 203 | 1 | R4 | 0 | 2.76 | 1.22 | 2.90 | 0.82 |
| 305 | 1 | R4 | 0 | 2.79 | 1.32 | 3.33 | 1.05 |
| 401 | 1 | R4 | 0 | 2.67 | 1.14 | 3.13 | 1.05 |
| 102 | 2 | R4 | 50 | 2.48 | 1.60 | 3.24 | 0.95 |
| 207 | 2 | R4 | 50 | 2.87 | 1.28 | 3.28 | 1.17 |
| 306 | 2 | R4 | 50 | 3.04 | 1.30 | 3.11 | 1.25 |
| 402 | 2 | R4 | 50 | 2.67 | 1.34 | 3.35 | 1.49 |
| 106 | 4 | R4 | 100 | 2.83 | 1.85 | 3.13 | 1.21 |
| 202 | 4 | R4 | 100 | 3.17 | 1.74 | 3.10 | 1.71 |

```
soy_data |>
  summary() |>
  kable(
    caption = 'Summary of the data'
  )
```

Table 2: Summary of the data

| | Plot | TRT | Stage | Krate_lbac | A1.Kpct | O2.Kpct | S1.Kpct | W1.Kpct |
|---|---|---|---|---|---|---|---|---|
| | Min. :102.0 | Min. :1.00 | Length:16 | Min. : 0.0 | Min. :2.480 | Min. :1.140 | Min. :2.900 | Min. :0.740 |
| | 1st Qu.:177.2 | 1st Qu.:1.75 | Class :character | 1st Qu.: 37.5 | 1st Qu.:2.737 | 1st Qu.:1.265 | 1st Qu.:3.087 | 1st Qu.:1.050 |

| Plot | TRT | Stage | Krate_lbac | A1.Kpct | O2.Kpct | S1.Kpct | W1.Kpct |
|---|---|---|---|---|---|---|---|
| Median :255.5 | Median :3.00 | Mode :character | Median : 75.0 | Median :2.810 | Median :1.405 | Median :3.145 | Median :1.250 |
| Mean :254.1 | Mean :3.00 | NA | Mean : 75.0 | Mean :2.825 | Mean :1.446 | Mean :3.147 | Mean :1.312 |
| 3rd Qu.:330.5 | 3rd Qu.:4.25 | NA | 3rd Qu.:112.5 | 3rd Qu.:2.942 | 3rd Qu.:1.570 | 3rd Qu.:3.217 | 3rd Qu.:1.575 |
| Max. :407.0 | Max. :5.00 | NA | Max. :150.0 | Max. :3.170 | Max. :1.940 | Max. :3.350 | Max. :1.980 |

```
soy_data |>
  summarise(
    mean_A1 = mean(A1.Kpct),
    mean_O2 = mean(O2.Kpct),
    mean_S1 = mean(S1.Kpct)
  ) |>
  kable(
    caption = 'Mean of the three Plots'
  )
```

Table 3: Mean of the three Plots

| mean_A1 | mean_O2 | mean_S1 |
|---|---|---|
| 2.825 | 1.445625 | 3.146875 |

```
soy_data |>
  names()
```

```
[1] "Plot"       "TRT"        "Stage"      "Krate_lbac" "A1.Kpct"
[6] "O2.Kpct"    "S1.Kpct"    "W1.Kpct"
```

### Data Preparation for Analysis

In this chunk, we wrangled the data to best fit to perform data analysis and randomized complete block design

```r
soy_data_final <- soy_data |>
  ## Convert the Krate from lbs/ac to Kg per hectare
  ## Represent the plot between 100 and 400 with a value of 1 to 4
  mutate(
    krate_kgha = Krate_lbac*0.453592/0.4044686,
    krate_kgha = round(krate_kgha, 0),
    Rep = case_when(
      Plot > 100 & Plot < 200 ~ 1,
      Plot > 200 & Plot < 300 ~ 2,
      Plot > 300 & Plot < 400 ~ 3,
      Plot > 400 ~ 4
    )
  ) |>
  ## Gather the location columns into a single column called 'location.k'
  gather(
    Location.k, K_pct,
    A1.Kpct, O2.Kpct, S1.Kpct, W1.Kpct
  ) |>
  ## Separate the location.k column into two different column
  ## called location and kname
  separate(
    Location.k,
    into = c('Location', 'Kname')
  ) |>
  ## select the columns that will be needed to perform the analysis
  dplyr::select(
    Location, Rep, krate_kgha, K_pct
  )
```

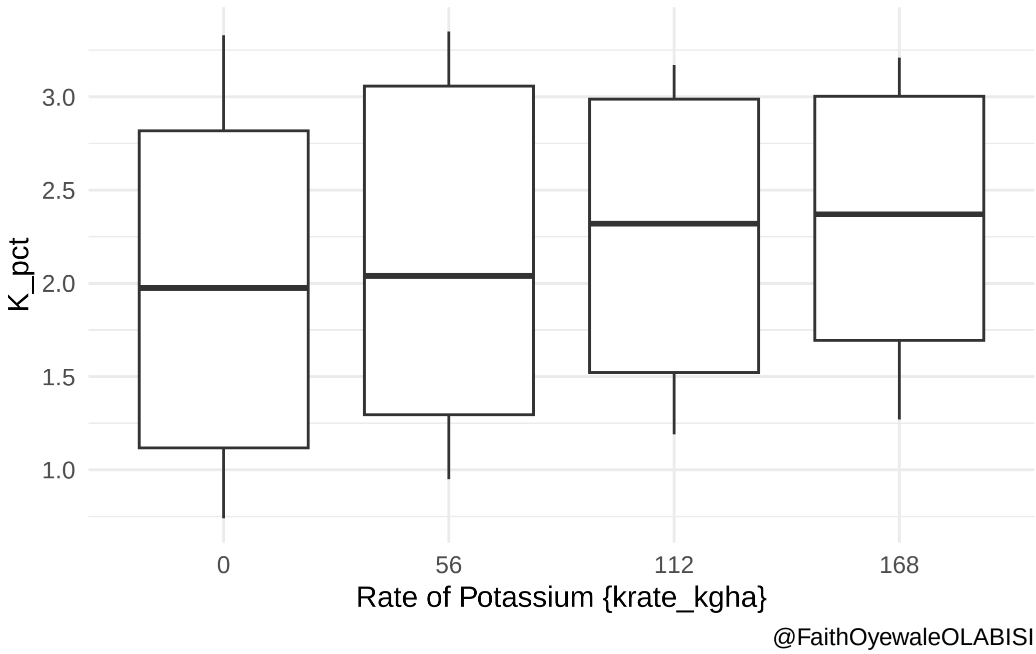
```r
soy_data_final |>
  ## Group the data by the location of the site.
  group_by(
    Location
  )|>
  summarise(
    meanK_pct = mean(K_pct),
  )
```

```
# A tibble: 4 x 2
  Location meanK_pct
  <chr>        <dbl>
1 A1            2.82
```

```
2 O2              1.45
3 S1              3.15
4 W1              1.31
```

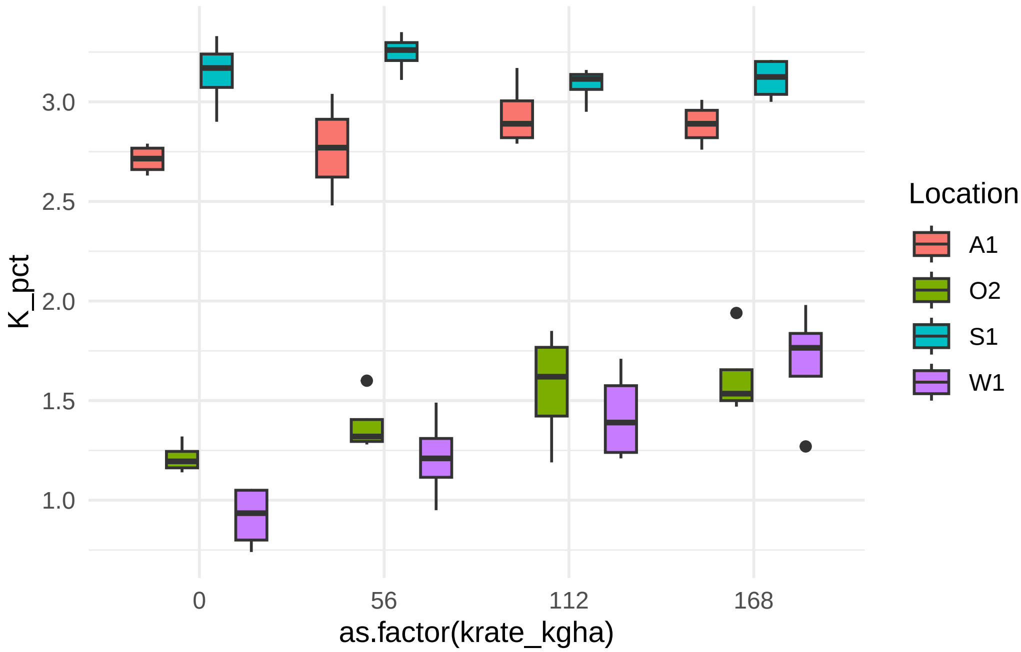
### Data Visualization

```r
## Visualize the rate of potassium application on the various site
soy_data_final |>
  ggplot(aes(as.factor(krate_kgha), K_pct))+
  geom_boxplot()+
  labs(
     x = "Rate of Potassium {krate_kgha}",
     y = 'K_pct',
     caption = '@FaithOyewaleOLABISI'
  )+
  theme_minimal()
```

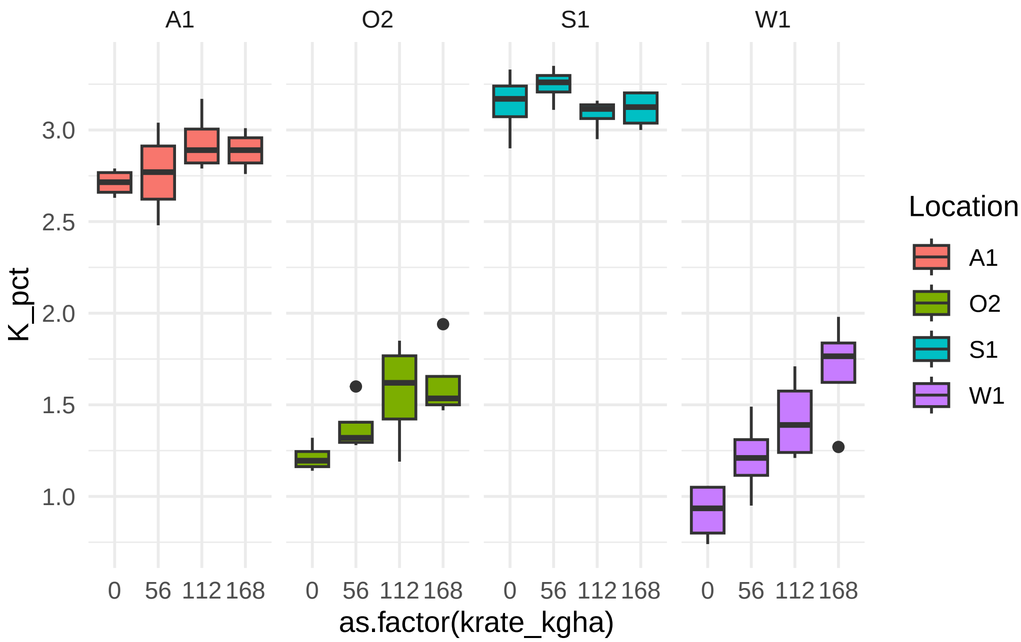



```r
## Visualize the rate of potassium application on the various site or location
soy_data_final |>
  ggplot(aes(as.factor(krate_kgha), K_pct))+
```

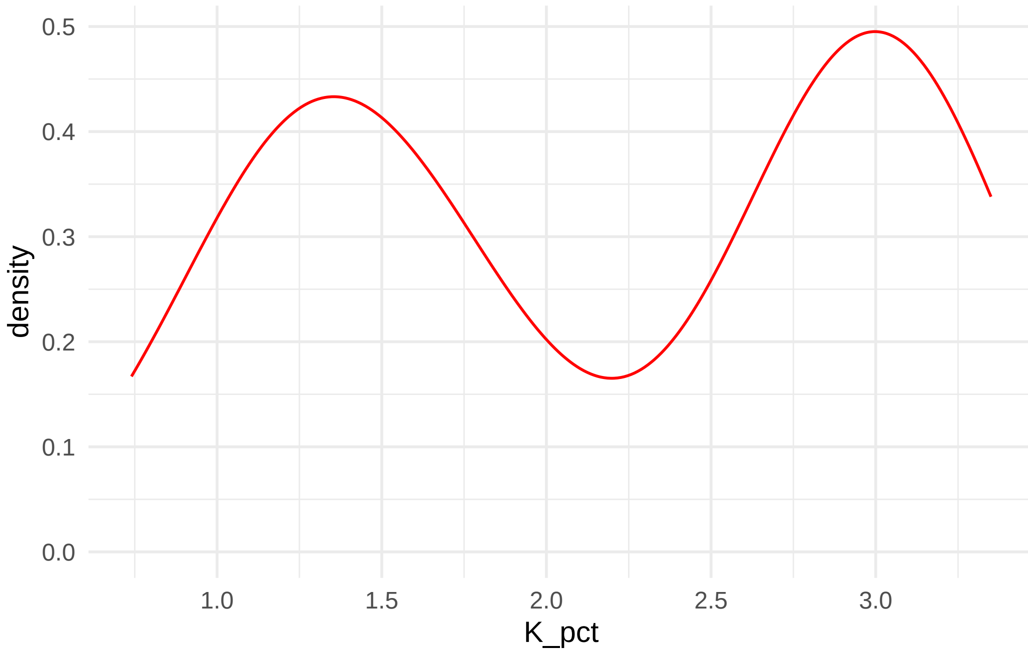
```
  geom_boxplot(aes(fill = Location))+
  theme_minimal()
```




```
soy_data_final |>
  ggplot(aes(as.factor(krate_kgha), K_pct))+
  geom_boxplot(aes(fill = Location))+
  facet_grid(.~Location)+
  theme_minimal()
```




```
soy_data_final |>
  ggplot(aes(K_pct))+
  geom_density(color = 'red')+
  theme_minimal()
```




```
# Creating krate as categorical variable(factor) and other columns
# also to factor
soy_data_final_1 <- soy_data_final |>
  mutate(
    fkrate_kgha = factor(krate_kgha),
    frep = factor(Rep),
    Location = factor(Location)
  ) |>
  filter(
    Location == 'W1' | frep != '1'
  )
```

### Randomized Complete Block Design (RCBD) Model

```
# Changing type of matrix restriction
options(contrasts = c('contr.sum', 'contr.poly'))

# Running the Model
soyk_mod <- lmer(K_pct ~ Location*fkrate_kgha + (1|Location/frep),
                 data = soy_data_final_1)
```

```
Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, : Model is near
 - Rescale variables?
```

```r
soyk_mod
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: K_pct ~ Location * fkrate_kgha + (1 | Location/frep)
   Data: soy_data_final_1
REML criterion at convergence: 17.9372
Random effects:
 Groups        Name        Std.Dev.
 frep:Location (Intercept) 0.09027
 Location      (Intercept) 0.15565
 Residual                  0.15851
Number of obs: 52, groups:  frep:Location, 13; Location, 4
Fixed Effects:
          (Intercept)              Location1              Location2
              2.17417                0.68583               -0.79333
            Location3           fkrate_kgha1           fkrate_kgha2
              0.96917               -0.17375               -0.01708
         fkrate_kgha3  Location1:fkrate_kgha1  Location2:fkrate_kgha1
              0.06125                0.05375                0.01958
Location3:fkrate_kgha1  Location1:fkrate_kgha2  Location2:fkrate_kgha2
              0.15042                0.01708               -0.05708
Location3:fkrate_kgha2  Location1:fkrate_kgha3  Location2:fkrate_kgha3
              0.12042                0.04875                0.03458
Location3:fkrate_kgha3
             -0.13458
optimizer (nloptwrap) convergence code: 0 (OK) ; 0 optimizer warnings; 1 lme4 warnings
```

```r
Anova(soyk_mod, type = 3)
```

```
Analysis of Deviance Table (Type III Wald chisquare tests)

Response: K_pct
                       Chisq Df Pr(>Chisq)
(Intercept)          657.961  1  < 2.2e-16 ***
Location              96.881  3  < 2.2e-16 ***
fkrate_kgha           25.995  3  9.561e-06 ***
Location:fkrate_kgha  30.987  9  0.0002975 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

## Model Assumption

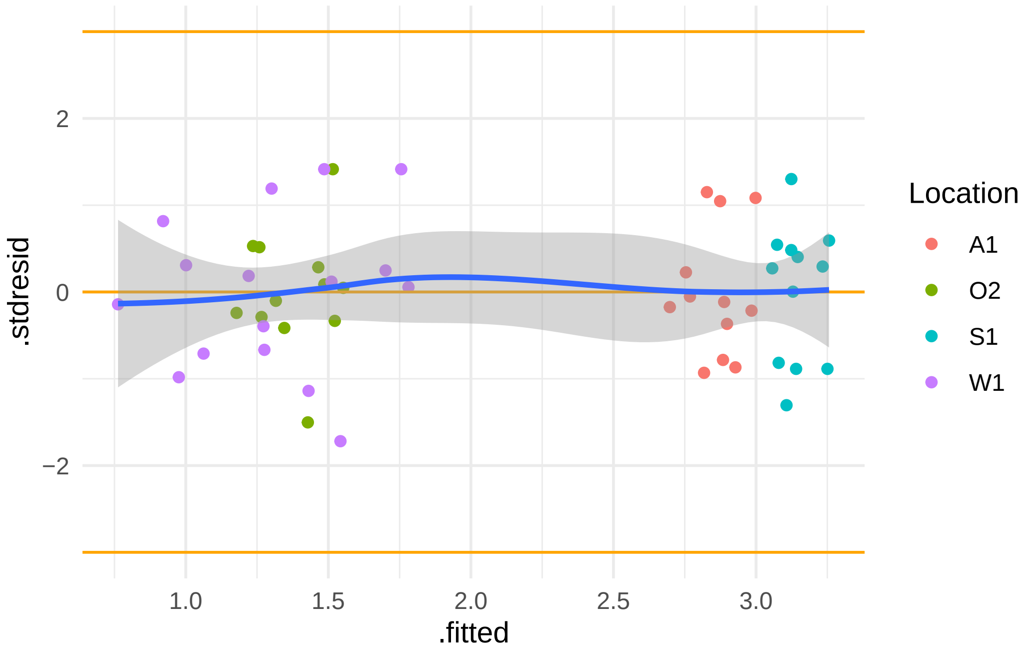
```r
## Extracting residual information

soy_data_aug <- augment(soyk_mod)

## Adding pearson standardized residuals

soy_data_aug$.stdresid <- resid(soyk_mod, type = 'pearson',
                                scale = 1)
```

### Within-group errors are iid ~ N(0, var2)

On this plot, looking for:
- Spread around zero
- Homogeneity (no patterns)
- Outliers (>3 or <-3)

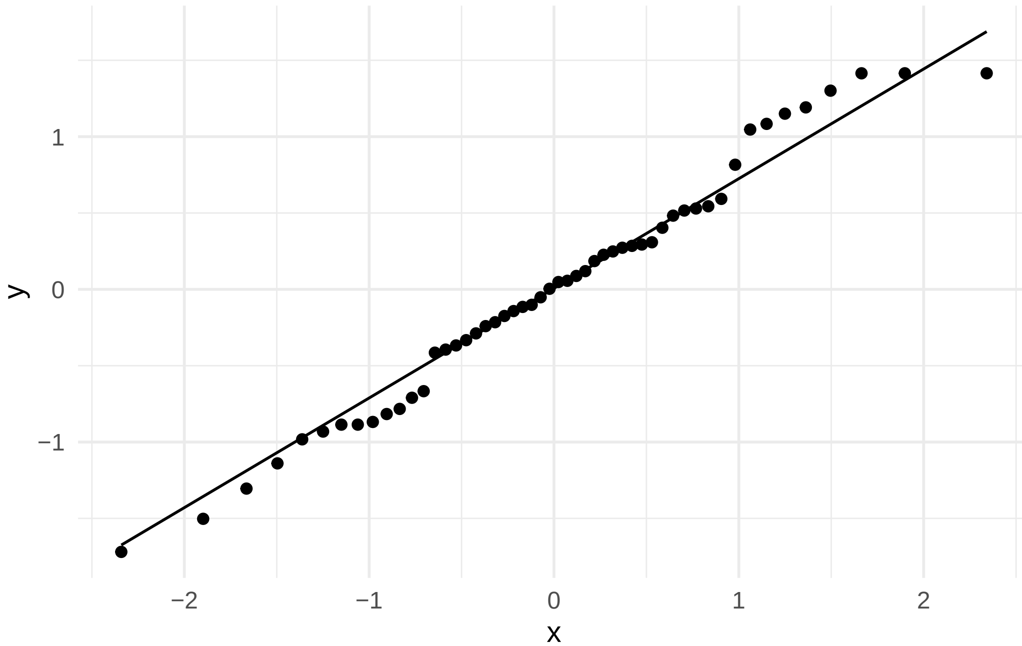
```r
soy_data_aug |>
  ggplot(aes(.fitted, .stdresid))+
  geom_point(aes(color = Location))+
  geom_hline(yintercept = c(-3, 0, 3), color = 'orange')+
  geom_smooth()+
  theme_minimal()
```

```
`geom_smooth()` using method = 'loess' and formula = 'y ~ x'
```




On this plot, looking for normality (points on top of line).

```
soy_data_aug |>
  ggplot(aes(sample = .stdresid))+
  stat_qq()+
  stat_qq_line()+
  theme_minimal()
```




**Random effects are iid ~ N(0,var1)**

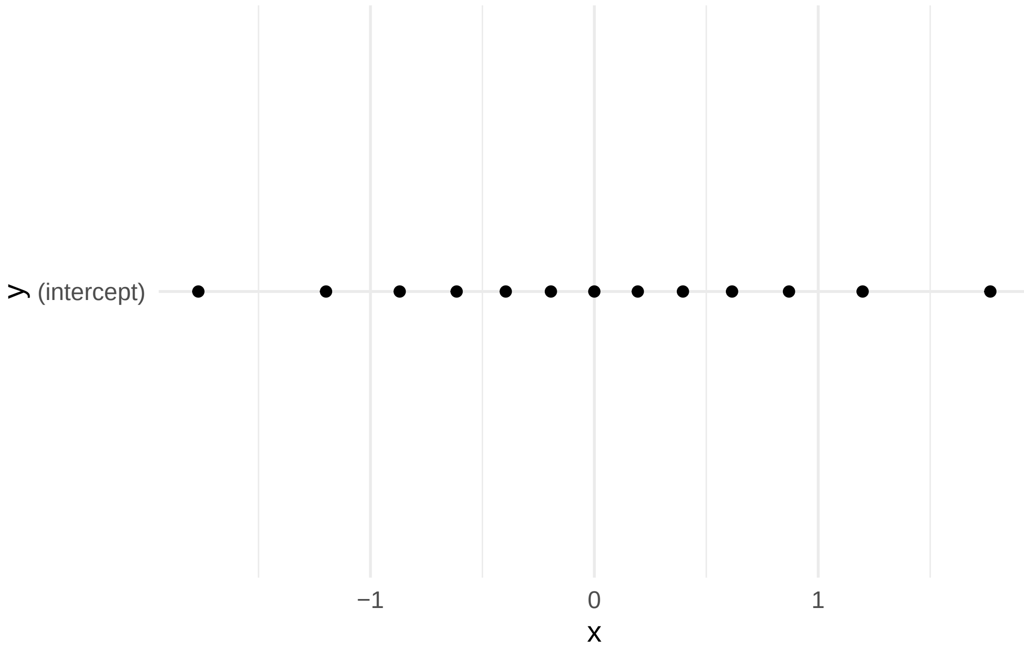
On this plot, looking for normality.

```
randeff_LocRep <- ranef(soyk_mod)[[1]]

randeff_LocRep |>
  ggplot(aes(sample = '(intercept)'))+
  stat_qq()+
  stat_qq_line()+
  theme_minimal()
```

```
Warning in stat_qq(): All aesthetics have length 1, but the data has 13 rows.
i Please consider using `annotate()` or provide this layer with data containing
  a single row.
```

```
Warning in stat_qq_line(): All aesthetics have length 1, but the data has 13 rows.
i Please consider using `annotate()` or provide this layer with data containing
  a single row.
```

```
Warning: Computation failed in `stat_qq_line()`.
Caused by error in `(1 - h) * qs[i]`:
! non-numeric argument to binary operator
```



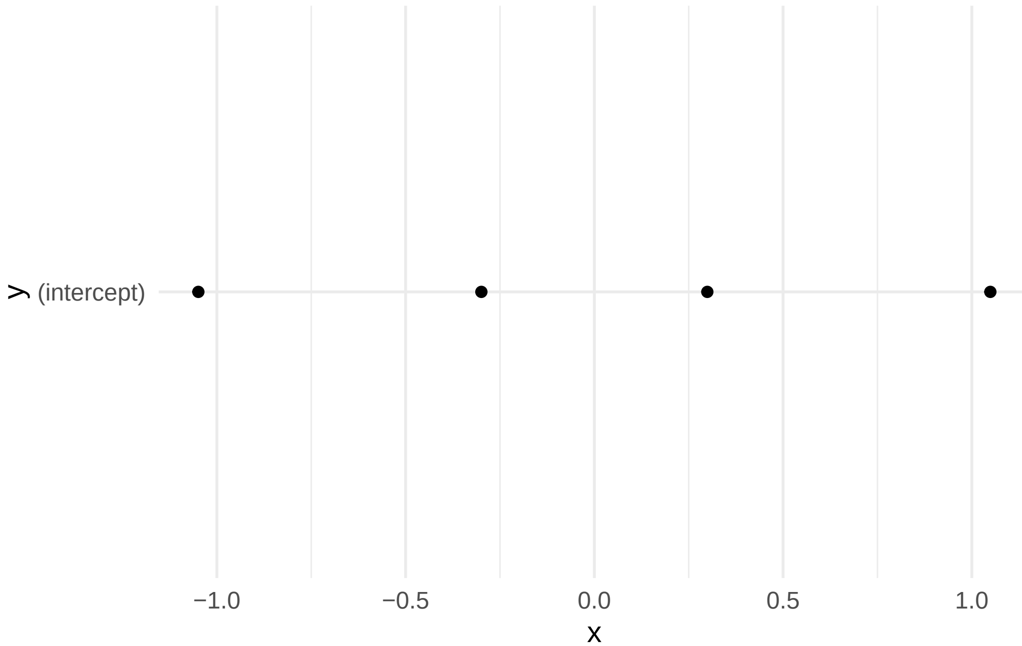

```
randeff_rep <- ranef(soyk_mod)[[2]]

randeff_rep |>
  ggplot(aes(sample = '(intercept)'))+
  stat_qq()+
  stat_qq_line()+
  theme_minimal()
```

```
Warning in stat_qq(): All aesthetics have length 1, but the data has 4 rows.
i Please consider using `annotate()` or provide this layer with data containing
  a single row.
```

```
Warning in stat_qq_line(): All aesthetics have length 1, but the data has 4 rows.
i Please consider using `annotate()` or provide this layer with data containing
  a single row.
```

```
Warning: Computation failed in `stat_qq_line()`.
```

```
Caused by error in `(1 - h) * qs[i]`:
! non-numeric argument to binary operator
```




#### Extracting model means and pairwise comparisons

```
soyk_means_LocRate <- emmeans(soyk_mod,
                              ~ fkrate_kgha | Location)

soyk_means_LocRate
```

```
Location = A1:
 fkrate_kgha emmean    SE  df lower.CL upper.CL
 0            2.740 0.188 309    2.370     3.11
 56           2.860 0.188 309    2.490     3.23
 112          2.970 0.188 309    2.600     3.34
 168          2.870 0.188 309    2.500     3.24

Location = O2:
 fkrate_kgha emmean    SE  df lower.CL upper.CL
 0            1.227 0.188 309    0.857     1.60
```

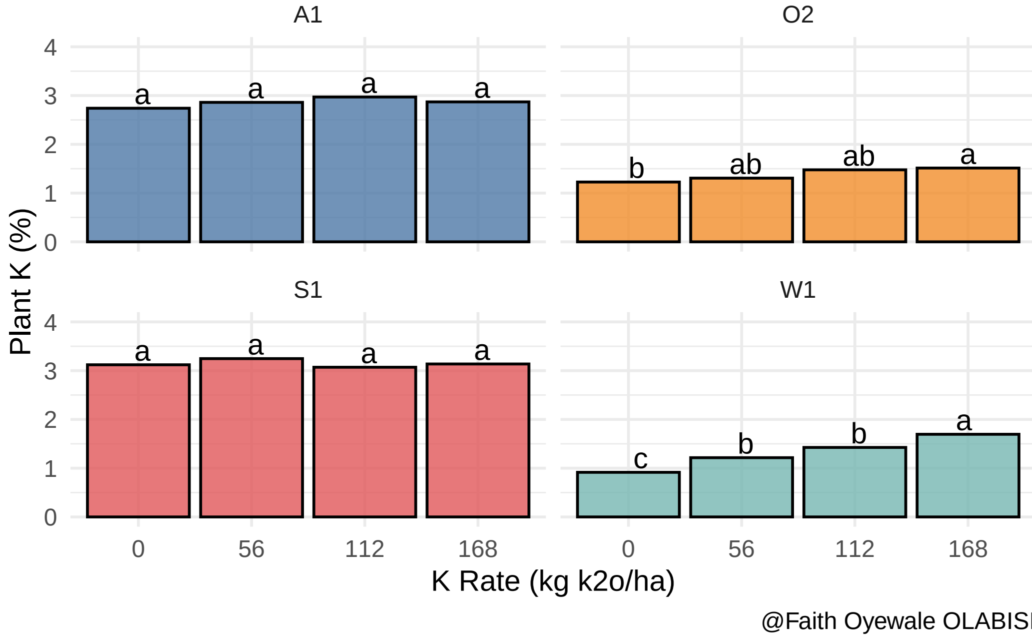
```
 56           1.307 0.188 309    0.937     1.68
 112          1.477 0.188 309    1.107     1.85
 168          1.513 0.188 309    1.144     1.88

Location = S1:
 fkrate_kgha emmean    SE  df lower.CL upper.CL
 0            3.120 0.188 309    2.750     3.49
 56           3.247 0.188 309    2.877     3.62
 112          3.070 0.188 309    2.700     3.44
 168          3.137 0.188 309    2.767     3.51

Location = W1:
 fkrate_kgha emmean    SE  df lower.CL upper.CL
 0            0.915 0.180 467    0.560     1.27
 56           1.215 0.180 467    0.860     1.57
 112          1.425 0.180 467    1.070     1.78
 168          1.695 0.180 467    1.340     2.05

Degrees-of-freedom method: kenward-roger
Confidence level used: 0.95
```

```
soyk_pwc_LocRate<- cld(soyk_means_LocRate,
                       adjust="none",
                       Letters=letters,
                       reversed=T)

soyk_pwc_LocRate <- soyk_pwc_LocRate |>
  as.data.frame()

soyk_pwc_LocRate
```

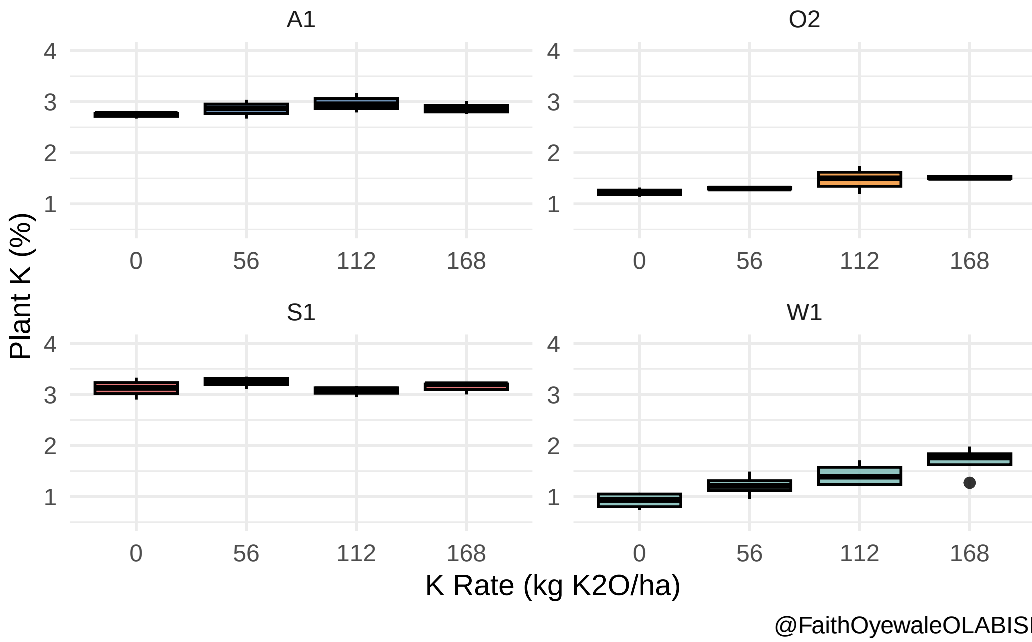
```
Location = A1:
 fkrate_kgha   emmean        SE     df  lower.CL upper.CL .group
 112         2.970000 0.1879335 309.42 2.6002107 3.339789  a
 168         2.870000 0.1879335 309.42 2.5002107 3.239789  a
 56          2.860000 0.1879335 309.42 2.4902107 3.229789  a
 0           2.740000 0.1879335 309.42 2.3702107 3.109789  a

Location = O2:
 fkrate_kgha   emmean        SE     df  lower.CL upper.CL .group
 168         1.513333 0.1879335 309.42 1.1435440 1.883123  a
 112         1.476667 0.1879335 309.42 1.1068774 1.846456  ab
```

```
 56          1.306667 0.1879335 309.42 0.9368774 1.676456  ab
 0           1.226667 0.1879335 309.42 0.8568774 1.596456   b

Location = S1:
 fkrate_kgha   emmean        SE     df  lower.CL upper.CL .group
 56          3.246667 0.1879335 309.42 2.8768774 3.616456  a
 168         3.136667 0.1879335 309.42 2.7668774 3.506456  a
 0           3.120000 0.1879335 309.42 2.7502107 3.489789  a
 112         3.070000 0.1879335 309.42 2.7002107 3.439789  a

Location = W1:
 fkrate_kgha   emmean        SE     df  lower.CL upper.CL .group
 168         1.695000 0.1804059 467.10 1.3404923 2.049508  a
 112         1.425000 0.1804059 467.10 1.0704923 1.779508   b
 56          1.215000 0.1804059 467.10 0.8604923 1.569508   b
 0           0.915000 0.1804059 467.10 0.5604923 1.269508    c

Degrees-of-freedom method: kenward-roger
Confidence level used: 0.95
significance level used: alpha = 0.05
NOTE: If two or more means share the same grouping symbol,
      then we cannot show them to be different.
      But we also did not show them to be the same.
```

```r
soyk_pwc_LocRate |>
  ggplot(aes(fkrate_kgha, emmean))+
  geom_bar(aes(fill = Location), color = 'black',
           stat = 'identity', alpha = 0.8)+
  geom_text(aes(label = .group), nudge_y = 0.3,
            show.legend = F)+
  labs(
    x = 'K Rate (kg k2o/ha)',
    y = 'Plant K (%)',
    caption = '@Faith Oyewale OLABISI'
  )+
  theme_minimal()+
  scale_fill_tableau()+
  theme(legend.position = 'none')+
  scale_y_continuous(limits = c(0, 4))+
  facet_wrap(~Location)
```




```r
soy_data_final_1 |>
  ggplot(aes(fkrate_kgha, K_pct))+
  geom_boxplot(aes(fill = Location), color = 'black', alpha = .8)+
  facet_wrap(~Location, scales = 'free')+
  scale_y_continuous(limits = c(.5, 4))+
  labs(
    x = 'K Rate (kg K2O/ha)',
    y = 'Plant K (%)',
    caption = '@FaithOyewaleOLABISI'
  )+
  theme_minimal()+
  scale_fill_tableau()+
  theme(legend.position = 'none')
```




## Acknowledgement

Kansas State University 2019 AGSA R Workshop - Introduction to R, taught by @leombastos.