SGA Data Anaylsis

OP & PC

5/5/22

Introduction

Loading Libraries

```
library(dplyr)
library(ggplot2)
library(tinytex)
library(tidyverse)
library(readxl)
library(viridis)
library(ggpubr)
library(knitr)
```

Importing Data from Excel

```
raw_data <- read_excel(file.choose(), sheet = "Raw")
```

Identifying Special Characters in Column Names

```
raw_data_colnames <- raw_data
```

Checking for Special Characters in the Data

```
# if (length(i <- grep("\u0394", colnames(raw_data_colnames)))) {
#   cat("Special Characters appear the data frame", colnames(raw_data_colnames), "\n")
# }
```

Removing Special Characters From Column Names

```
if (length(i <- grep("\u0394", colnames(raw_data_colnames)))) {
 delta_replace <- gsub("\u0394", "del", colnames(raw_data_colnames))
}
if (length(i <- grep("\u0394", colnames(raw_data_colnames)))) {
  colnames(raw_data_colnames) <- delta_replace
}
options(dplyr.width = Inf)
raw_data_colnames
```

```
## # A tibble: 4 x 13
##   Time    `0% (Blank)` `WT 0%` `-1del 0%` `-2del 0%` `50% (Blank)` `WT 50%`
```

```
##   <chr>          <dbl>   <dbl>      <dbl>      <dbl>         <dbl>    <dbl>
## 1 11:23am        0.031   0.087      0.073      0.075         0.532    0.522
## 2 12:44pm        0.033   0.278      0.204      0.199         0.515    0.694
## 3 1:55pm         0.034   0.555      0.557      0.406         0.546    0.94
## 4 3:05pm         0.035   0.768      0.873      0.621        NA        1.51
##   `-1del 50%` `-2del 50%` `HI-50% (Blank)` `WT HI-50%` `-1del HI-50%`
##         <dbl>       <dbl>            <dbl>       <dbl>          <dbl>
## 1       0.461       0.553            0.494       0.586          0.508
## 2       0.462       0.479            0.496       0.695          0.659
## 3       0.461       0.475           NA           0.962          1.01
## 4       0.462       0.517           NA           1.16           1.24
##   `-2del HI-50%`
##            <dbl>
## 1          0.653
## 2          0.704
## 3          1.21
## 4          1.20
```

Averaging Blank Measurements for Removing Background Noise

```
raw_data_blank_mean <- raw_data_colnames
```

```
raw_data_blank_mean <- raw_data_colnames %>%
  select(contains("Blank")) %>%
  colMeans(na.rm = TRUE) %>%
  round(3)
```

Removing Blank Measurement from Table

```
raw_data_clean <- raw_data_colnames %>%
  select(-`0% (Blank)`, -`50% (Blank)`, -`HI-50% (Blank)`, -`Time`)
```

Removing Background Noise

```
for (a in 1:ncol(raw_data_clean)) {
 if (grepl(" 0%", colnames(raw_data_clean)[a])) {
   raw_data_clean[ , a] <- raw_data_clean[ , a] - as.numeric(raw_data_blank_mean[1])
 }
 if (grepl(" 50%", colnames(raw_data_clean)[a])) {
   raw_data_clean[ , a] <- raw_data_clean[ , a] - as.numeric(raw_data_blank_mean[2])
 }
 if (grepl("-50%", colnames(raw_data_clean)[a])) {
   raw_data_clean[ , a] <- raw_data_clean[ , a] - as.numeric(raw_data_blank_mean[3])
 }
}
options(dplyr.width = Inf)
raw_data_clean
```

```
## # A tibble: 4 x 9
##   `WT 0%` `-1del 0%` `-2del 0%` `WT 50%` `-1del 50%` `-2del 50%` `WT HI-50%`
##     <dbl>      <dbl>      <dbl>    <dbl>       <dbl>       <dbl>       <dbl>
## 1   0.054       0.04      0.042 -0.00900       -0.07      0.0220       0.091
## 2   0.245      0.171      0.166   0.163        -0.069     -0.0520       0.2
```

```
## 3   0.522      0.524      0.373  0.409        -0.07      -0.0560       0.467
## 4   0.735      0.84       0.588  0.979        -0.069     -0.0140       0.66
##   `-1del HI-50%` `-2del HI-50%`
##            <dbl>          <dbl>
## 1         0.0130          0.158
## 2         0.164           0.209
## 3         0.515           0.712
## 4         0.741           0.706
```

```
raw_data_clean_display <- raw_data_clean
raw_data_clean_display <- add_column(raw_data_clean_display, Timepoint = c(1, 2, 3, 4), .before = 1)
kable(raw_data_clean_display)
```

Timepoint	WT 0%	-1del 0%	-2del 0%	WT 50%	-1del 50%	-2del 50%	WT HI-50%	-1del HI-50%	-2del HI-50%
1	0.054	0.040	0.042	-0.009	-0.070	0.022	0.091	0.013	0.158
2	0.245	0.171	0.166	0.163	-0.069	-0.052	0.200	0.164	0.209
3	0.522	0.524	0.373	0.409	-0.070	-0.056	0.467	0.515	0.712
4	0.735	0.840	0.588	0.979	-0.069	-0.014	0.660	0.741	0.706

#Transforming the Data

```
trans_data <- stack(raw_data_clean) %>%
  mutate(Time = rep(c(1,2,3,4), times = 9))
colnames(trans_data) [1:2] <- c("OD", "Sample")
```

Graphing Cleaned Data

```
zero_plot <- trans_data %>%
  filter(grepl(" 0%", Sample)) %>%
  ggplot(aes(x = Time, color = Sample, y = OD))+
  geom_point(size = 1.5)+
  geom_line()+
  labs(title = "0% Serum")+
  theme(plot.title = element_text(hjust = 0.5))+
  scale_colour_manual(values = c("#c51b8a","#d95f0e","#2c7fb8"))+
  xlim(0,4)

fifty_plot <- trans_data %>%
  filter(grepl(" 50%", Sample)) %>%
  ggplot(aes(x = Time, color = Sample, y = OD))+
  geom_point(size = 1.5)+
  geom_line()+
  labs(title = "50% Serum")+
  theme(plot.title = element_text(hjust = 0.5))+
  scale_colour_manual(values = c("#c51b8a","#d95f0e","#2c7fb8"))+
  xlim(0,4)

hi_plot <- trans_data %>%
  filter(grepl("HI-50%", Sample)) %>%
  ggplot(aes(x = Time, color = Sample, y = OD))+
  geom_point(size = 1.5)+
```

```
  geom_line()+
  labs(title = "HI-50% Serum")+
  theme(plot.title = element_text(hjust = 0.5))+
  scale_colour_manual(values = c("#c51b8a","#d95f0e","#2c7fb8"))+
  xlim(0,4)

figure_1 <- ggarrange(zero_plot, fifty_plot, hi_plot,
                      labels = c("A", "B", "C"),
                      ncol = 2, nrow = 2)
figure_1
```