JMP_defects2

2023-10-25

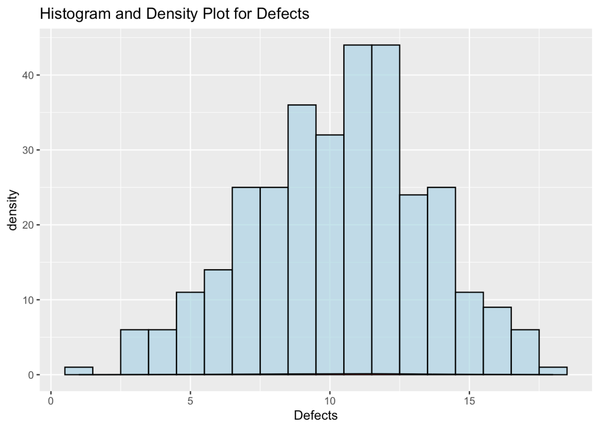
```
library(haven)
library(tidyverse)
```

```
## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.2 ──
## ✔ ggplot2 3.4.0      ✔ purrr   1.0.0
## ✔ tibble  3.1.8      ✔ dplyr   1.0.10
## ✔ tidyr   1.2.1      ✔ stringr 1.5.0
## ✔ readr   2.1.3      ✔ forcats 0.5.2
## ── Conflicts ────────────────────────────────────────── tidyverse_conflicts() ──
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()    masks stats::lag()
```

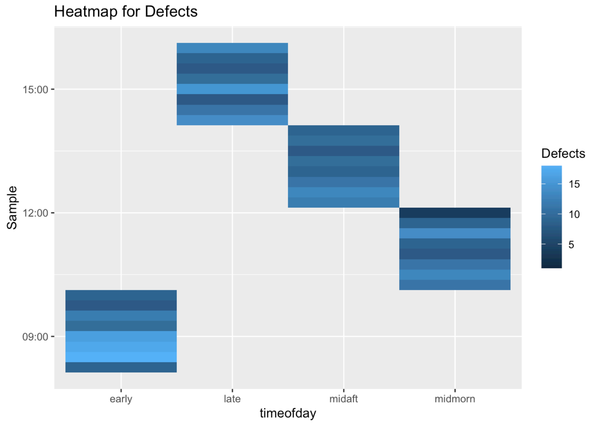
```
library(plotly)
```

```
##
## Attaching package: 'plotly'
##
## The following object is masked from 'package:ggplot2':
##
##     last_plot
##
## The following object is masked from 'package:stats':
##
##     filter
##
## The following object is masked from 'package:graphics':
##
##     layout
```

```
# Read the data from the CSV file
defects <- read.csv("defects.csv")

title <- 'JMP Case Study for Defective Parts'
title2 <- 'Sampling for comparison of defects by time of day'

defects$Sample <- strptime(defects$Sample, "%H:%M")
defects$Sample <- as.POSIXct(defects$Sample)

defects <- defects %>%
  mutate(timeofday = case_when(
    between(Sample, as.POSIXct("2023-10-24 08:00:00"), as.POSIXct("2023-10-24 10:00:00")) ~ 'early',
    between(Sample, as.POSIXct("2023-10-24 10:15:00"), as.POSIXct("2023-10-24 12:00:00")) ~ 'midmorn',
    between(Sample, as.POSIXct("2023-10-24 12:15:00"), as.POSIXct("2023-10-24 14:00:00")) ~ 'midaft',
    between(Sample, as.POSIXct("2023-10-24 14:15:00"), as.POSIXct("2023-10-24 16:00:00")) ~ 'late',
    TRUE ~ NA_character_
  ))

# Summary statistics and univariate analysis
summary_stats <- summary(defects$Defects)

# Histogram and density plot
ggplot(defects, aes(x = Defects)) +
  geom_histogram(binwidth = 1, fill = "lightblue", color = "black", alpha = 0.7) +
  geom_density(alpha = 0.5, fill = "#FF6666") +
  labs(title = "Histogram and Density Plot for Defects")
```




```
# Calculate mean by timeofday
defects_summary <- defects %>%
  group_by(timeofday) %>%
  summarize(mean_defects = mean(Defects))

# Heatmap
ggplot(defects, aes(x = timeofday, y = Sample, fill = Defects)) +
  geom_tile() +
  labs(title = "Heatmap for Defects")
```




```
# ANOVA for defects
model <- glm(Day ~ timeofday, data = defects, family = gaussian)
summary(model)
```

```
##
## Call:
## glm(formula = Day ~ timeofday, family = gaussian, data = defects)
##
## Deviance Residuals:
##    Min      1Q  Median      3Q     Max
##   -4.5    -2.5     0.0     2.5     4.5
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)       5.500e+00  3.232e-01   17.02   <2e-16 ***
## timeofdaylate    -3.049e-15  4.570e-01    0.00        1
## timeofdaymidaft  -2.274e-15  4.570e-01    0.00        1
## timeofdaymidmorn -3.090e-15  4.570e-01    0.00        1
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 8.35443)
##
##     Null deviance: 2640  on 319  degrees of freedom
## Residual deviance: 2640  on 316  degrees of freedom
## AIC: 1593.4
##
## Number of Fisher Scoring iterations: 2
```