

Homework3

Prajay Pandiri

2025-09-25

Libraries

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --  
## v dplyr      1.1.4      v readr      2.1.5  
## v forcats    1.0.1      v stringr   1.5.2  
## v ggplot2    4.0.0      v tibble    3.3.0  
## v lubridate  1.9.4      v tidyr     1.3.1  
## v purrr      1.1.0
```

```
## -- Conflicts ----- tidyverse_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag()     masks stats::lag()  
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(colorfindr)
```

Reading in the data

```
data = read.csv('homework3_data.csv')
```

Should the company commit to redesigning the web site based on the criteria for success that they laid out?

My Recommendation is that the company should commit to redesigning the website based on the criteria that it would increase sales.

Getting colors

```
stripe <- get_colors("stripe.png")  
palette <- make_palette(stripe)
```



palette

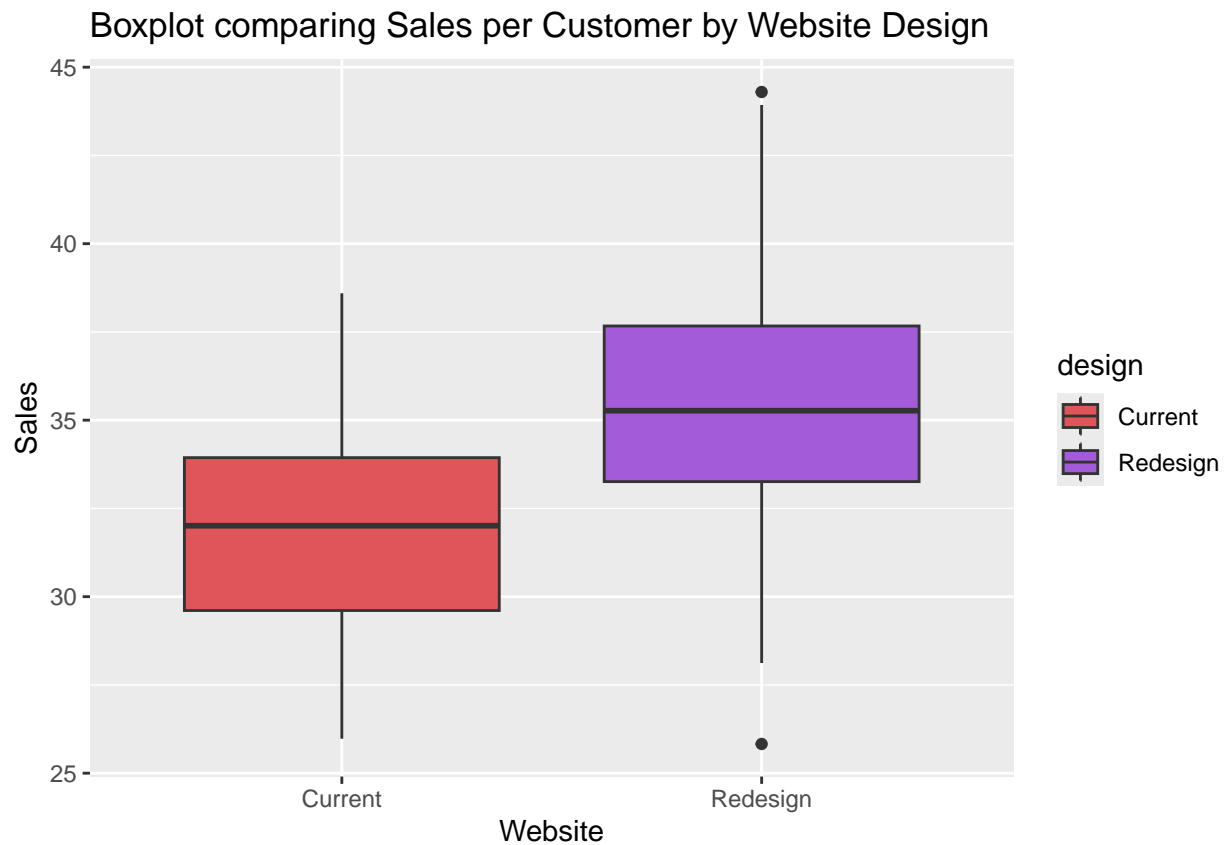
```
## [1] "#FFFFFF" "#11243E" "#E05559" "#CECED4" "#A35BDA" "#A0D1F8" "#EF8461"
## [8] "#5384ED" "#757F94" "#B65EB2"
```

I am choosing colors #E05559 and #A35BDA from Stripe color palette.

Graphical Representation 1

Boxplot

```
data %>%
  mutate(design = factor(design,
                          labels = c("Current", "Redesign"))) %>%
  ggplot(aes(x = design, y = sales, fill = design)) +
  geom_boxplot() +
  scale_fill_manual(values = c("#E05559", "#A35BDA")) +
  labs(title = "Boxplot comparing Sales per Customer by Website Design",
       x = "Website", y = "Sales")
```

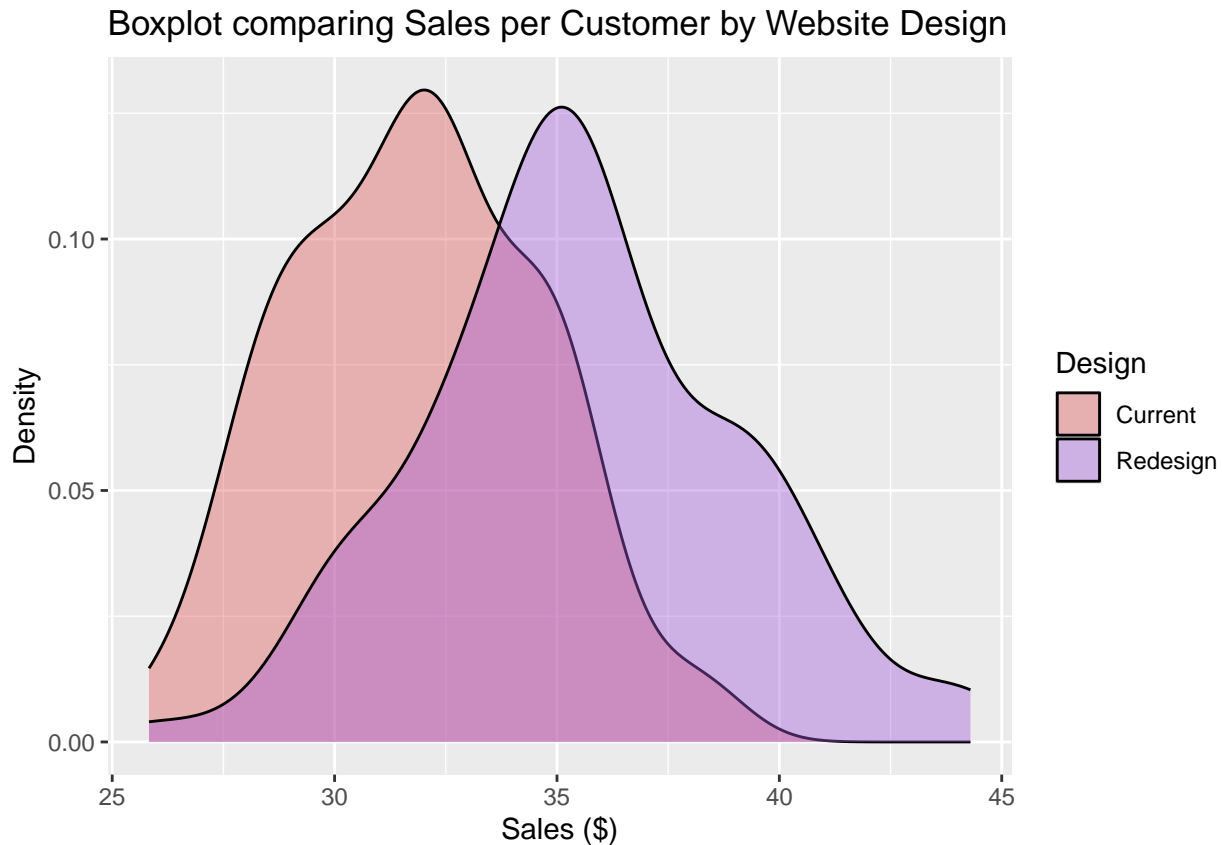


The boxplot distribution shows that the distribution of sales for the redesigned website is in general higher than the one for the current website.

Graphical Representation 2

Histogram

```
data %>%
  mutate(design = factor(design,
                          labels = c("Current", "Redesign"))) %>%
  ggplot(aes(x = sales, fill = design)) +
  geom_density(alpha = 0.4) +
  scale_fill_manual(values = c("#E05559", "#A35BDA")) +
  labs(title = "Boxplot comparing Sales per Customer by Website Design", x = "Sales ($)", y = "Density")
```



The density plot suggest that the redesign leads to higher typical sales per customer when compared to the current design similar to the boxplot.

Estimation of how much sales will increase/decrease if the redesign is done;

```
mean_diff <- data %>%
  mutate(design = factor(design,
    labels = c("Current", "Redesign"))) %>%
  summarise(
    mean_current = mean(sales[design == "Current"]),
    mean_redesign = mean(sales[design == "Redesign"]),
    mean_diff = mean_redesign - mean_current
  )
mean_diff
```

```
##   mean_current mean_redesign mean_diff
## 1      31.84819      35.51309  3.664904
```

The mean difference in sales shows that sales will increase by approximately 3.66 if the redesign is done.

Addressing the question of whether the redesign will lead to an average increase in sales of at least \$1.80 per customer

```
cur_sales <- data$sales[data$design == 0]
re_sales <- data$sales[data$design == 1]
t.test(re_sales, cur_sales,
  alternative = "greater",
  mu = 1.80,
```

```
var.equal = FALSE)
```

```
##  
## Welch Two Sample t-test  
##  
## data: re_sales and cur_sales  
## t = 4.1499, df = 186.01, p-value = 2.528e-05  
## alternative hypothesis: true difference in means is greater than 1.8  
## 95 percent confidence interval:  
## 2.922037 Inf  
## sample estimates:  
## mean of x mean of y  
## 35.51309 31.84819
```

This means we are 95% confident that the true difference in means is at least \$2.92 which is greater than \$1.80.