Homework 3 Company Design Analysis

Mason Reed

2025-10-01

```
library(ggplot2)
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 lubridate 1.9.4     v tibble    3.3.0
## v purrr     1.1.0     v tidyr     1.3.1
## -- 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 error
```

```
library(colorfindr)
customer_sales = read.csv('homework3_data.csv')
```

"The finance people at the company decided that they would only commit to the full redesign if the data showed that the redesign would lead to an average increase in sales of at least $1.80 per customer."

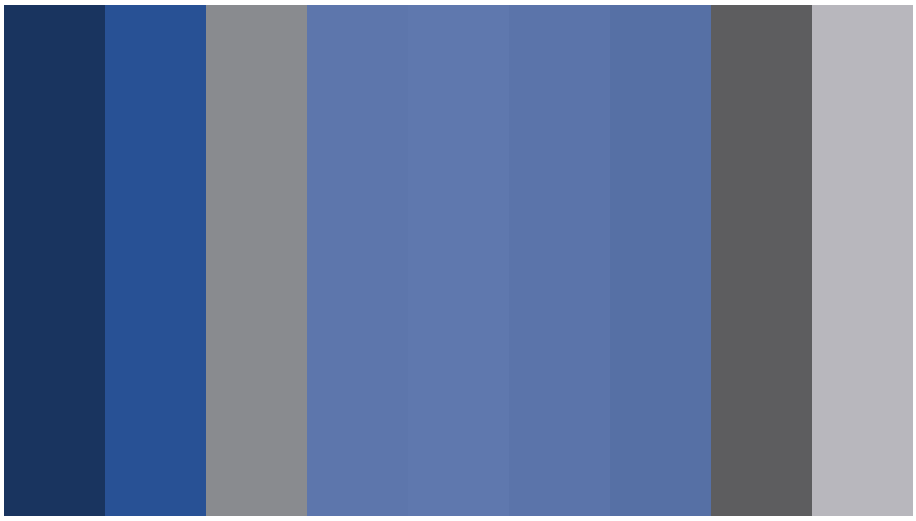
1. My recommendation to the company would be to go ahead with the redesign as analysis on the past redesign seems to support that a new design would increase sales per customer by at least $1.80, with the past redesign having an estimated increase of 3.66 in sale value per customer. The distribution of sale value per customer in the new design seems to be centered at a higher average sale value as well, seen in the boxplot and histogram. The boxplot displays this difference better, although the histogram shows the right shift in the distribution of sale value as well. It's possible that the data suggests that a redesign would increase sales despite the redesign actualy not increasing sales, with possible reasons being confounding variables between the new design and old design, impactful differences in the new redesign and past redesign, and poor data on the customer sales.

To investigate whether or not a redesign will have an expected average increase in sales of at least $1.80 we can compare the distribution of sales per customer, as well as compare means and run a hypothesis/regression test to see if there is a significant difference between the effect of the design sales value per customer.

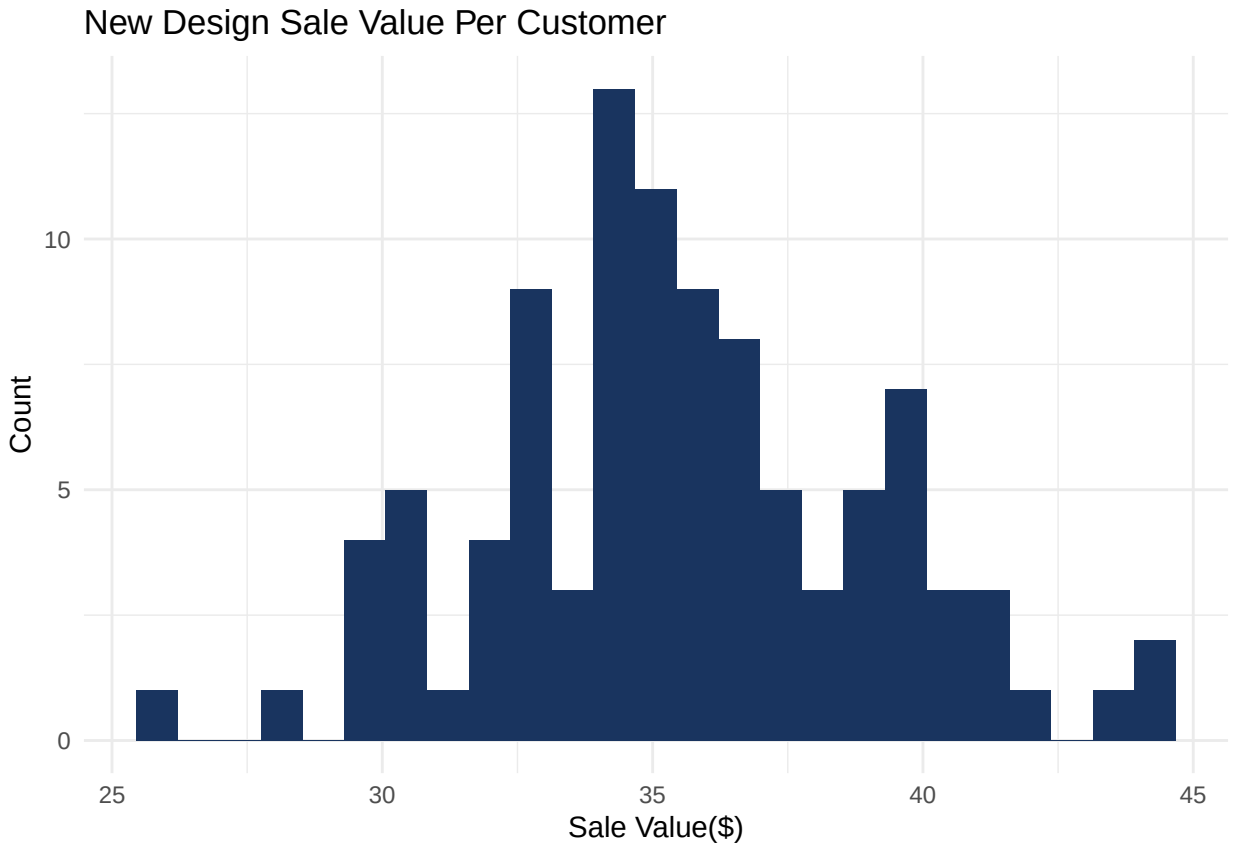
```
# Separating Old And New Design Sales Data
sales_design_old = filter(customer_sales, design == 0)
sales_design_new = filter(customer_sales, design == 1)
```

2. & 3. : Visualizations With Citadel Colors

```
citadel_colordat = get_colors('citadelwebpage.png')
citadel_palette = make_palette(citadel_colordat[1:100,])
```

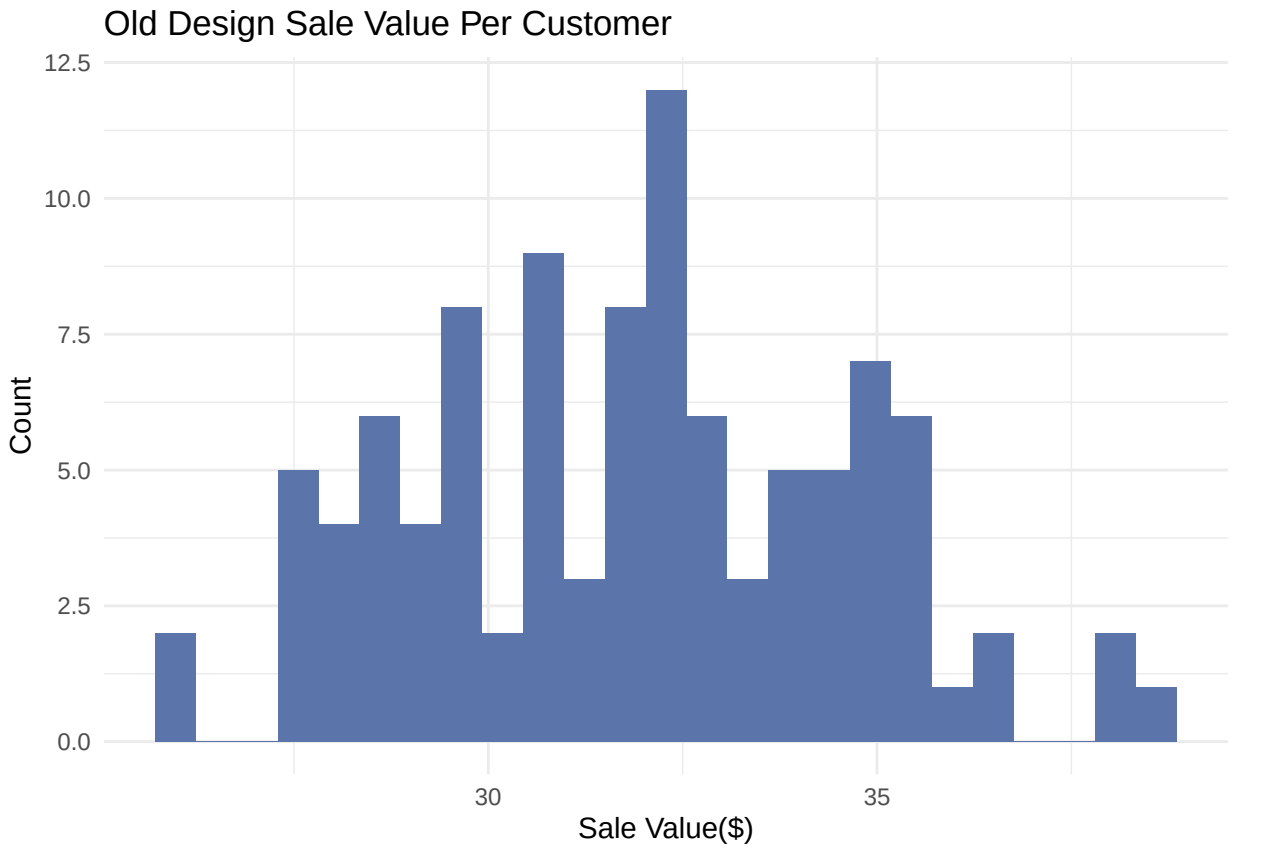


```
ggplot() + geom_histogram(data = sales_design_new, aes(x = sales),
                  fill = "#19345F", bins = 25) +
  labs(title = "New Design Sale Value Per Customer",
       x = "Sale Value($)",
       y = "Count") +
  theme_minimal()
```

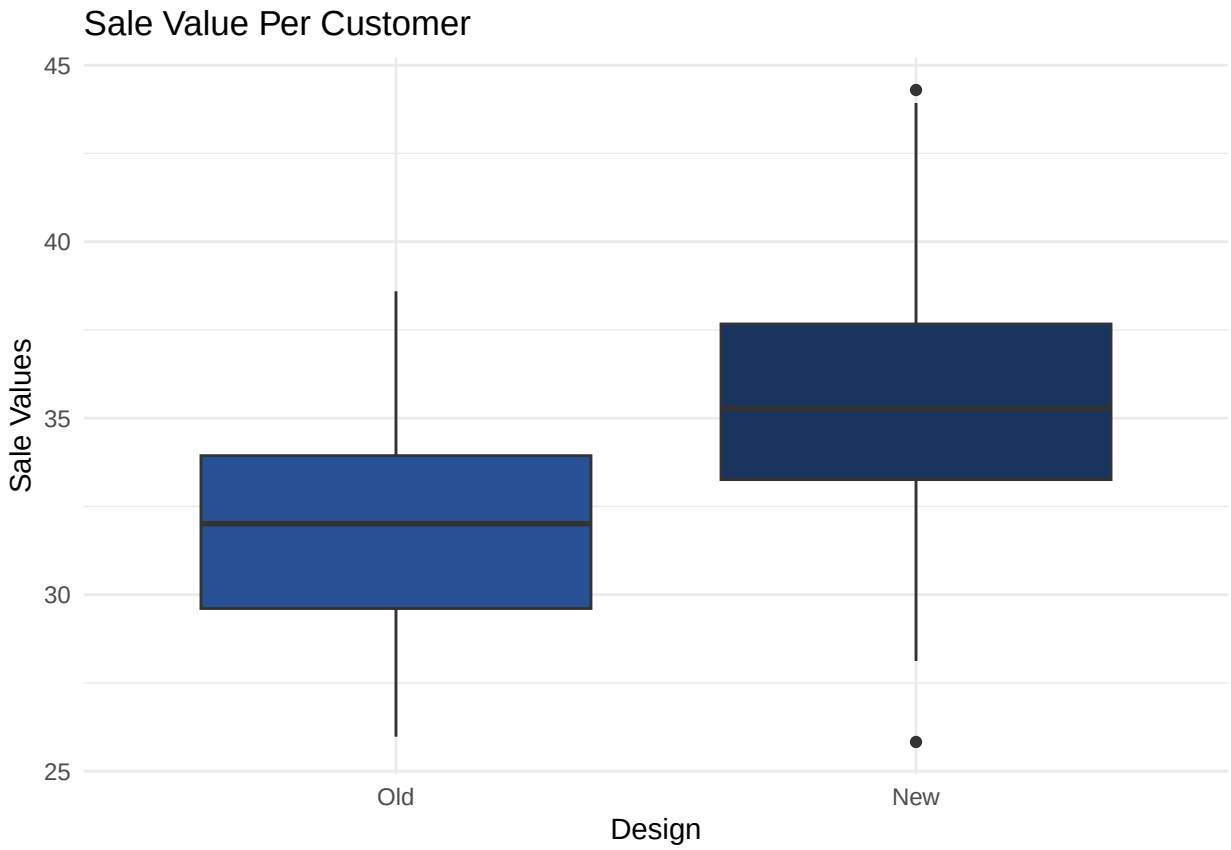



```
ggplot() + geom_histogram(data = sales_design_old, aes(x = sales),
                  fill = "#5B74AA", bins = 25) +
  labs(title = "Old Design Sale Value Per Customer",
       x = "Sale Value($)",
```

```
      y = "Count") +
  theme_minimal()
```




```
ggplot(customer_sales, aes(x = factor(design), y = sales, fill = factor(design))) +
  geom_boxplot() +
  scale_fill_manual(values = c("0" = "#285195", "1" = "#19345F")) +
  scale_x_discrete(labels = c("0" = "Old", "1" = "New")) +
  labs(title = "Sale Value Per Customer",
       x = "Design",
       y = "Sale Values") +
  theme_minimal() + theme(legend.position = 'none')
```



2. Looking At Mean NPS And Regression Based Off Design

```
design_lm = lm(sales ~ design, data = customer_sales)

summary(design_lm)
```

```
##
## Call:
## lm(formula = sales ~ design, data = customer_sales)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -9.6855 -2.2511 -0.0673  2.1205  8.7838
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  31.8482     0.3154 100.969  < 2e-16 ***
## design        3.6649     0.4483   8.175 3.49e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.17 on 198 degrees of freedom
## Multiple R-squared:  0.2523, Adjusted R-squared:  0.2486
## F-statistic: 66.83 on 1 and 198 DF,  p-value: 3.489e-14
```

```
nps_design1mean = mean(sales_design_new$nps)
design2_npsmean = mean(sales_design_old$nps)
```

The mean nps score for the new design is 5.616 which is greater than the old nps score of 4.604.

The new design seems to increase sale value per customer by $3.66 based off the regression on sale value by design. It has low p-value as well which suggests there is enough data to support the difference between the new design sales and old design sales.

This estimated $3.66 increase from the new design is greater than the $ 1.80 the company wants to see, which supports the redesign.

4. Alternative Statement / Possible Faults

Although the regression and distribution plots suggest that the redesign increased sales, there is not a true 100% certainty that the redesign actually increased sales or that the new redesign would increase sales. The suggested $3.66 increase could be a result of other factors that happened when the old design existed that didn't when the new design was implemented. If the old design was in a period where the economy was in a recession and the new design was in a standard or healthier market this could affect consumers ability to spend. Pricing changes between the time the new and old design were implemented could increase sale values as well. General growth and awareness of the company from other marketing could increase sales.