

# Homework3

Sahithi Uppalapati

2025-10-01

## Reccomendation

The company should proceed with redesigning the web site based on the criteria of success laid out in the analysis below (more detailed final reccomendation at the end)

## Problem statement and Intro

```
data <- read_csv("homework3_data.csv")
```

```
## Rows: 200 Columns: 4
## -- Column specification -----
## Delimiter: ","
## dbf (4): sales, design, items, nps
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
head(data)
```

```
## # A tibble: 6 x 4
##   sales design items  nps
##   <dbl> <dbl> <dbl> <dbl>
## 1  32.6     0     2     4
## 2  35.4     0     4     5
## 3  30.9     0     3     4
## 4  35.5     0     3     6
## 5  32.1     0     2     6
## 6  31.6     0     1     4
```

This data has columns titled sales, design, items, and nps. We are trying to make a recommendation to the company: Should they commit to redesigning the web site based on the criteria for success that they laid out?

##Tiktok Compnay Colors

```
#getting custom tiktok colors
tiktok_color_palette <- get_colors("~/Downloads/tiktok.png")

head(tiktok_color_palette)
```

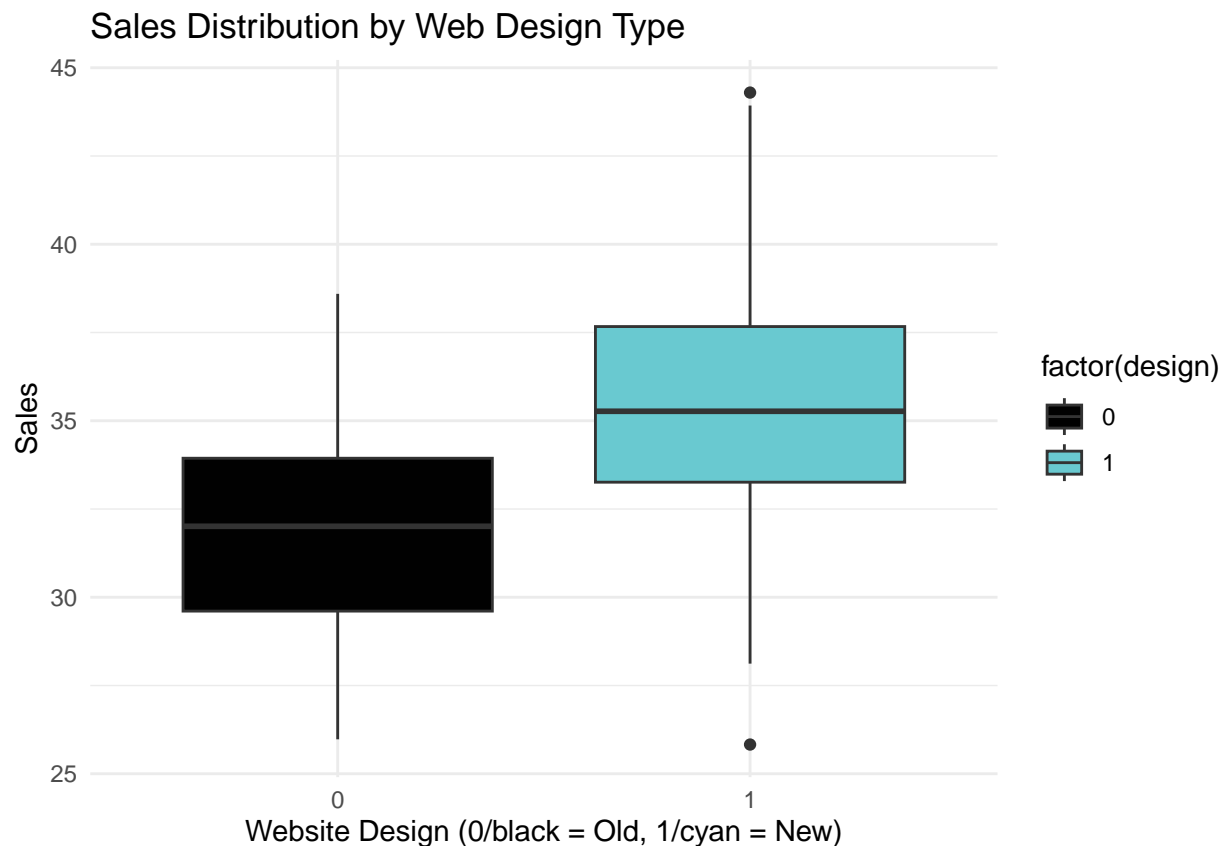
```
## # A tibble: 6 x 3
##   col_hex col_freq col_share
##   <chr>    <int>    <dbl>
## 1 #010101  118853  0.695
## 2 #FFFFFF   21626  0.126
## 3 #69C9D0    4536  0.0265
## 4 #EE1D52    4051  0.0237
## 5 #000000    3731  0.0218
## 6 #020202    1158  0.00677
```

```
tiktok_colors <- c("#010101", "#FFFFFF", "#69C9D0", "#EE1D52")
```

I will be using Tiktok as the companys color pallete.

##Graphical Visuals

```
ggplot(data, aes(x = factor(design), y = sales, fill = factor(design))) +
  geom_boxplot() +
  scale_fill_manual(values = c("0" = tiktok_colors[1], "1" = tiktok_colors[3])) +
  labs(x = "Website Design (0/black = Old, 1/cyan = New)",
       y = "Sales",
       color = "Website Design (0 = Old, 1 = New)",
       title = "Sales Distribution by Web Design Type") +
  theme_minimal()
```

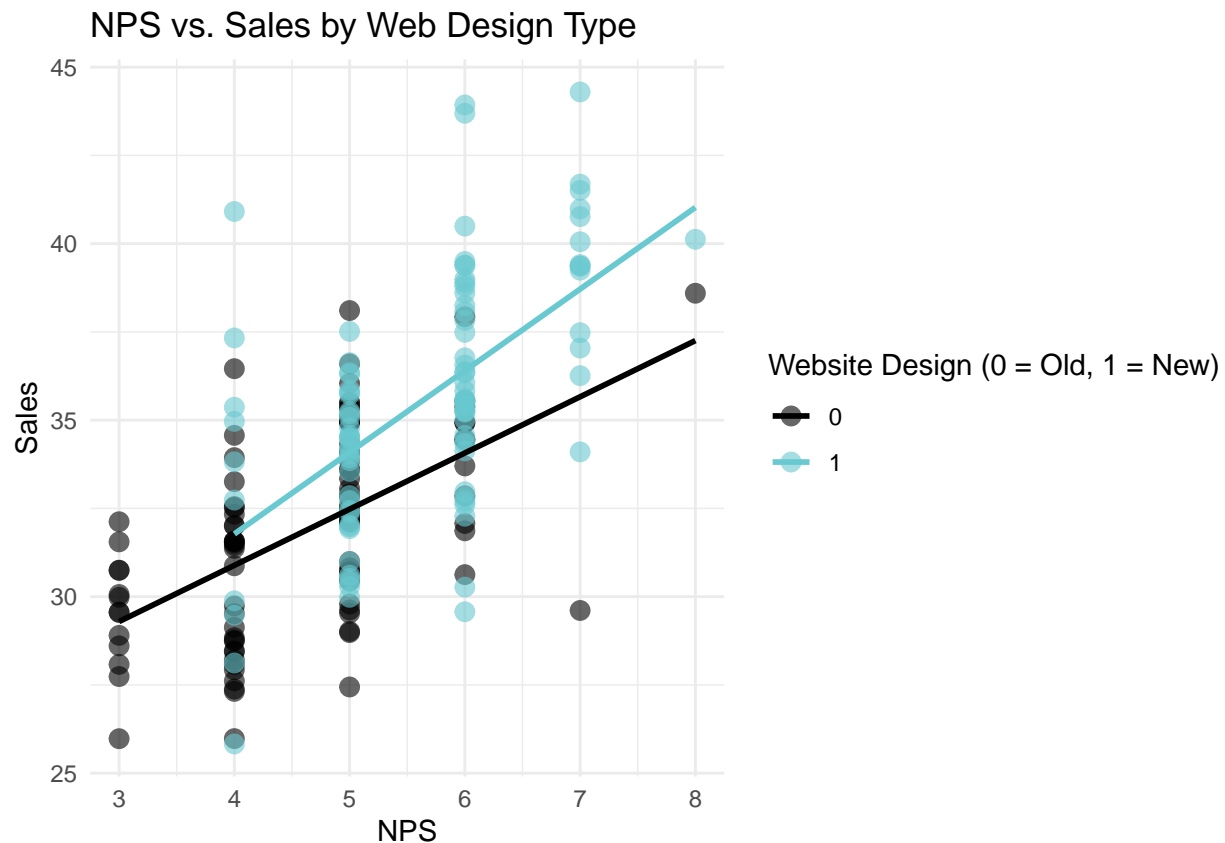


This box plot displays the distribution of sales with the old and new web designs. The boxplot shows that the median sales actually increased with the new design (1/cyan), and the middle 50% of sales shifted

upwards in comparison to the old website. The central tendency clearly favors the new design showing that customers spend more with the new website. This supports the analysis that the redesign of the website increases sales.

```
ggplot(data, aes(x = nps, y = sales, color = factor(design))) +
  geom_point(alpha = 0.6, size = 3) +
  geom_smooth(method = "lm", se = FALSE) + # add regression lines to the graph
  scale_color_manual(values = c("0" = tiktok_colors[1], "1" = tiktok_colors[3])) +
  labs(x = "NPS",
       y = "Sales",
       color = "Website Design (0 = Old, 1 = New)",
       title = "NPS vs. Sales by Web Design Type") +
  theme_minimal()
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```



Through this scatter graph, we can see that sales increase with NPS. This goes to show that happier customers will spend more money. Since nps is indicating how likely the customer is to recommend the web site, we can see that as the nps score gets higher, the sales tend to increase for redesign of website 1 rather than the old website. The regression line for the new design is consistently above the old website design (black). This also goes to support the analysis that the new design raises average sales.

## Sales Estimate

```
#run a regression
model <- lm(sales ~ design, data = data)
summary(model)

##
## Call:
## lm(formula = sales ~ design, data = data)
##
## 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
```

Here we can see that the intercept is 31.85 which is what the old design average sales would estimate. This means customers spent on average 31.85 with the old website. We can see the slope coefficient in this case (design) is 3.66. This would mean that we increase \$3.66 more with the new website. The p-value seems to be < 0.001 too which means the difference is in fact statistically significant.

```
# see how confident we are that sales will increase
confint(model)
```

```
##              2.5 %    97.5 %
## (Intercept) 31.226168 32.470212
## design      2.780801  4.549008
```

Just to be sure, I ran a confidence interval test. We can see that we are 95% confident that the redesign increases sales between \$2.78 and \$4.55 per customer. The company will commit to redesign if the sales increase on average above \$1.80. According to our results above, the estimate for the new redesign is \$3.66, which is above the threshold.

## Final Recommendation

The graphical visuals show that sales increase with the new design. The regression model suggests that the redesign increases sales by \$3.66 per customer. Since this increase is above the company's threshold of \$1.80 and the effect is statistically significant, I would suggest that the company does the website redesign.

## **Alternative Statement**

Even though the regression and graphical analysis show that the redesign increases sales, the improvement could be due to other factors like increased sales due to seasonality or customer demographics. Due to these confounding variables, the old design could still increase sales and would be a good choice to avoid extra efforts.