Client Report (Placeholder)

Parham Pishrobat, Sarah Masri, Johnson Chen

2024-02-27

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.4
## v forcats   1.0.0     v stringr   1.5.0
## v ggplot2   3.4.4     v tibble    3.2.1
## v lubridate 1.9.2     v tidyr     1.3.0
## v purrr     1.0.1
## -- 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 errors
```

```
library(lubridate)
library(rlang)
```

```
##
## Attaching package: 'rlang'
##
## The following objects are masked from 'package:purrr':
##
##     %@%, flatten, flatten_chr, flatten_dbl, flatten_int, flatten_lgl,
##     flatten_raw, invoke, splice
```

```
source("../code/EDA.R")
source("../code/ITSA.R")
source("../code/GEE.R")
source("../code/LMEM.R")
```

Introduction

- Background of the study.
- Objective(s) of the study.
- Statistical questions to answer.

Concerns around sugar consumption and its health implications have prompted an array of interventions aimed at modifying consumer behaviours in relation to sugary beverages. The current study investigates the effectiveness of various strategies to encourage consumers to choose zero-calorie beverages over sugary

alternatives. In particular, the research question focuses on the impact of two types of visual presentations of calorie content through posters that highlight either the calorie content or the physical activity required to burn these calories. Furthermore, the effectiveness of price discounts, both independently and in conjunction with explanatory messaging, as a means to influence consumer choices is explored.

Data Description and Summaries

- Data collection method.
- Study design.
- Sample size.
- Variables measured.
- Missing data.

The data has been collected data from cafeterias and convenience shops within three hospital sites, denoted by A, B, C, after conducting the interventions. Hospitals A is urban and has two cafeterias and two convenience shops. Hospital B is also urban setting and but has only one cafeteria. And finally hospital C is suburban setting, and has one cafeteria and one convenience shop. Both interventions (delivery of discount or the messaging type) and data collection (recording the sales) was automatic at site A and by trained personnel in sites B and C. In the context of this study, sugary beverages include regular soft drinks and iced teas that are sweetened with natural sugars like sucrose and corn syrup, and zero-calorie beverages include diet soft drinks and teas, and water. Note that other beverage types like juices, milk, coffee, and fountain-dispensed drinks are excluded due to categorization challenges.

The study adopts an interrupted time-series multi-site quasi-experimental design to assess the outcomes of five distinct interventions on the purchase patterns of bottled sugary and zero-calorie beverages. The interventions consist of two price discounts and three calorie messaging strategies, each designed to influence consumer purchasing behaviour towards healthier beverage options. The price interventions involved a 10% discount on zero-calorie beverages, with one intervention additionally providing explanatory messaging about the discount. The calorie messaging interventions varied in their approach, providing information on the caloric content of sugary beverages, the physical activity required to burn off these calories, and a combination of both strategies.

The primary outcome of interest is daily sales of bottled sugary and zero-calorie beverages over a span of 30 weeks, from October 27 to May 23. The study periods include baseline data collection phases, intervention phases for both price discounts and calorie messaging, and washout periods to assess the persistence of intervention effects.

(INCOMPLETE)

```
# summarize_data(beverage_sales)$numerical
# unique(beverage_sales$Site)
# unique(beverage_sales$Intervention)
```

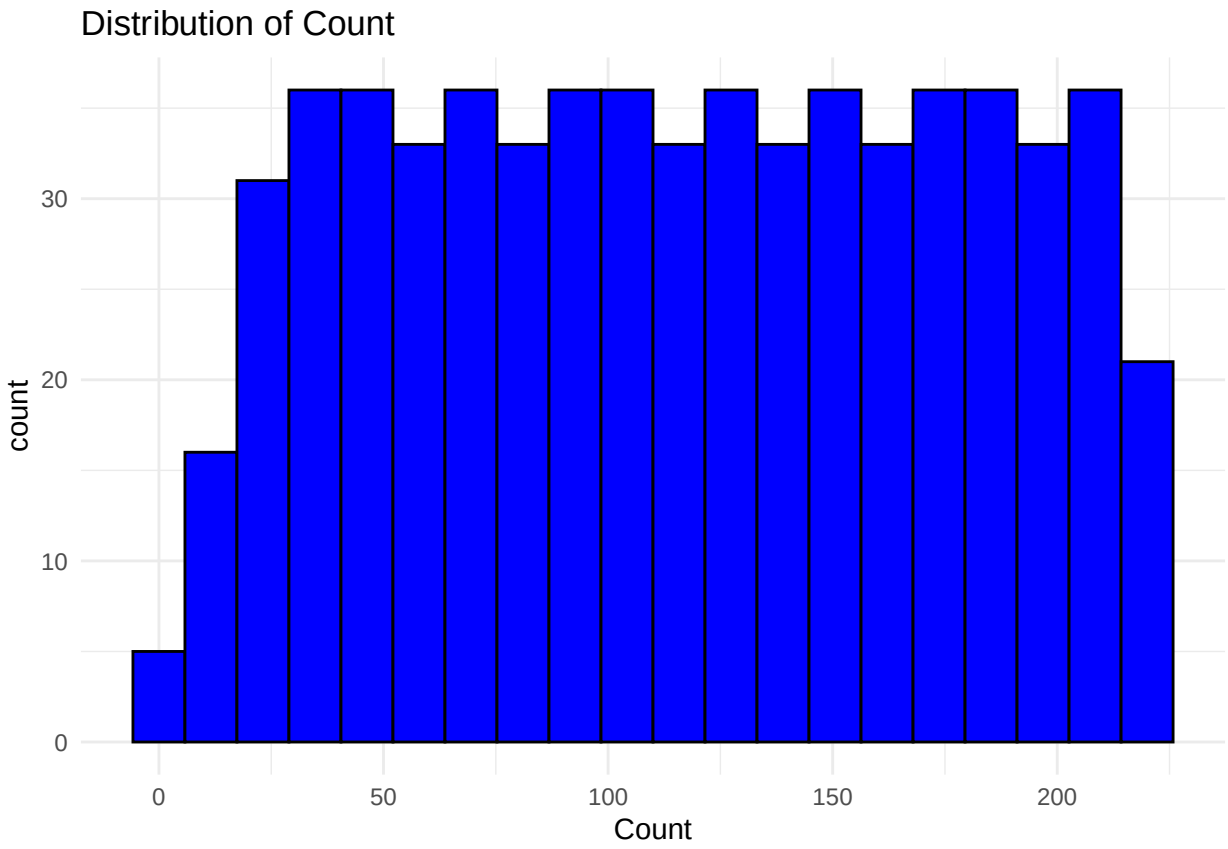
3. Exploratory Analysis

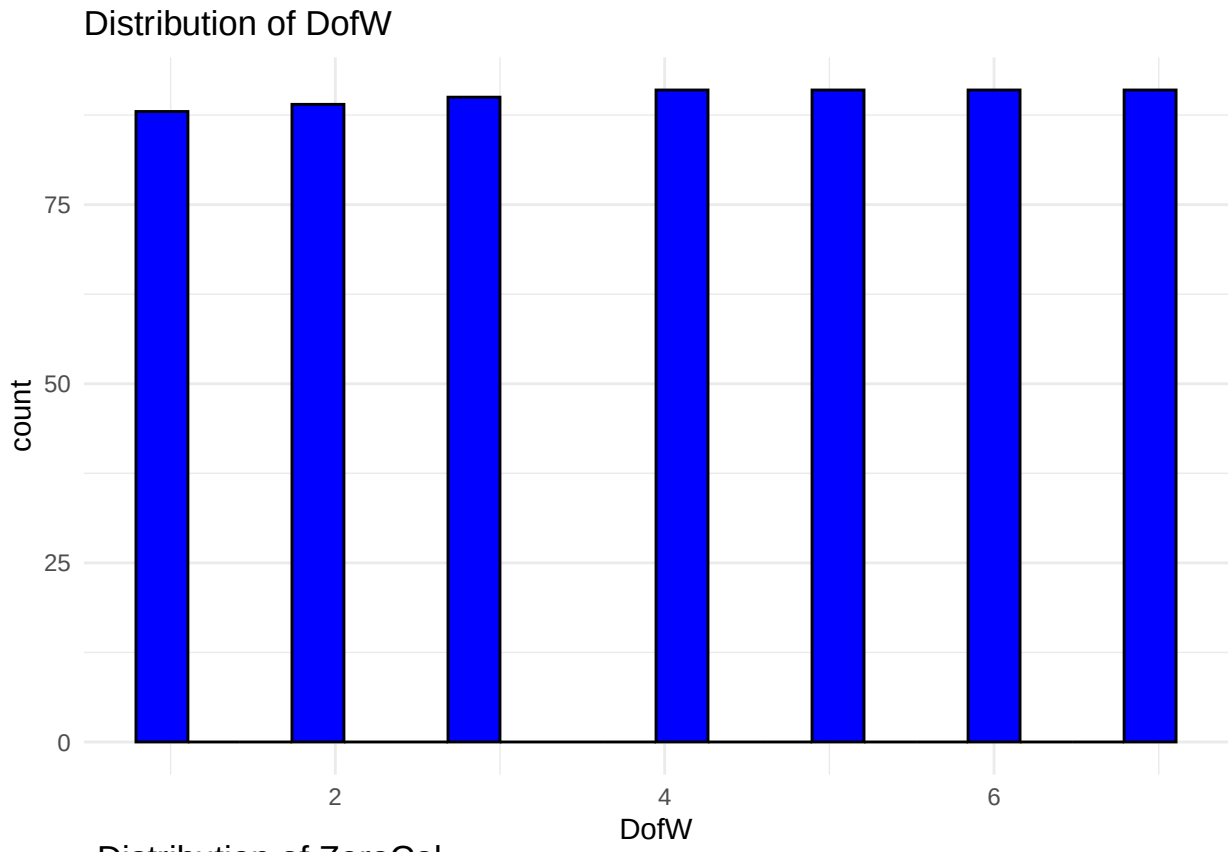
NOTE: we will figure out the order and title of these sub sections at the end

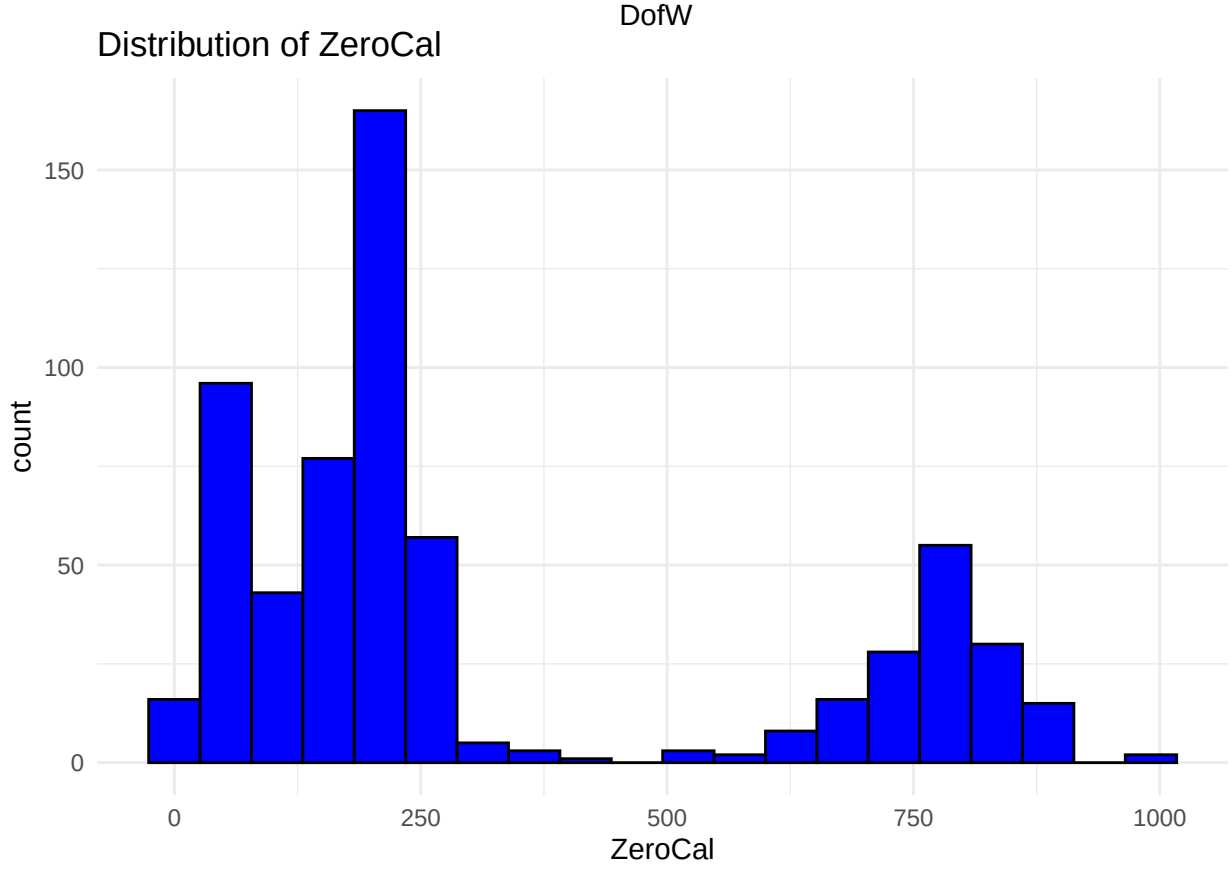
histogram and scatter plots, boxplot (Par; almost done)

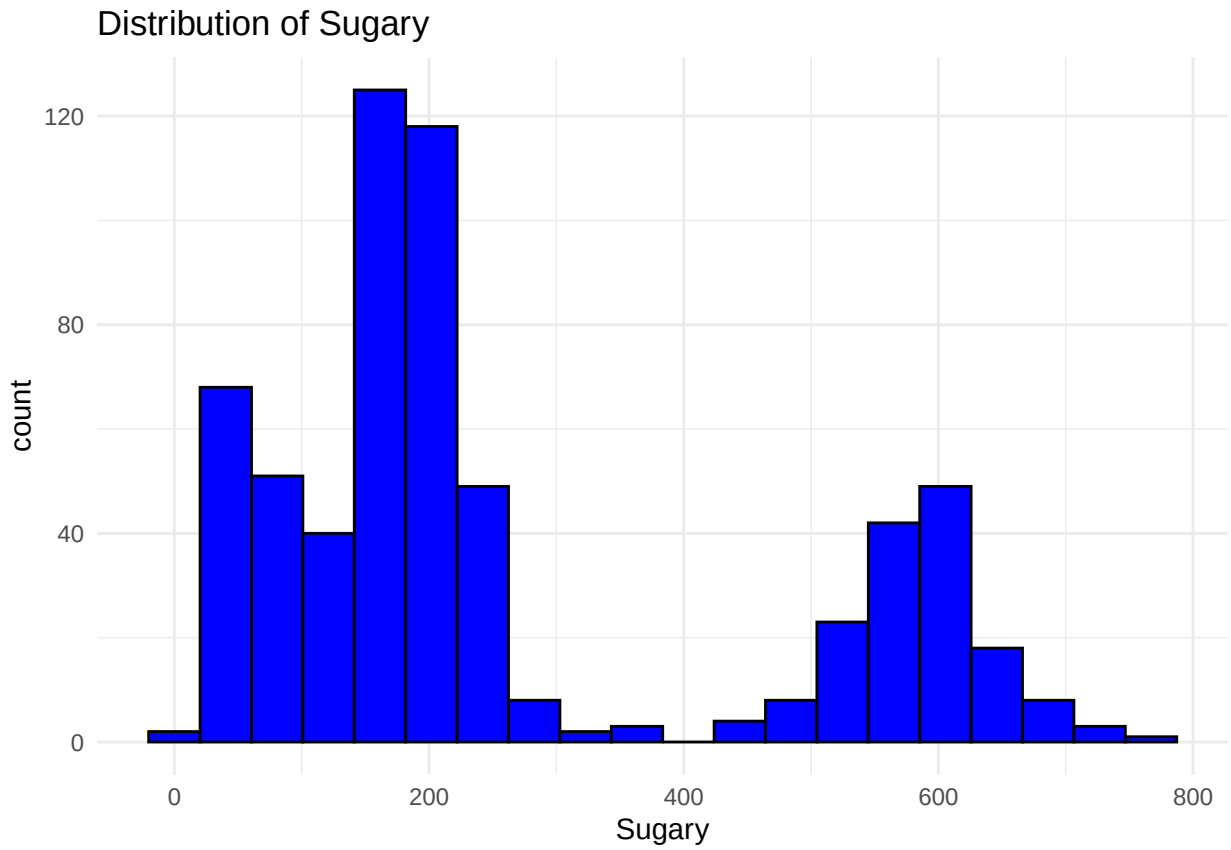
```
plot_histograms(beverage_sales)
```

```
## Warning: ‘aes_string()‘ was deprecated in ggplot2 3.0.0.
## i Please use tidy evaluation idioms with ‘aes()‘.
## i See also ‘vignette("ggplot2-in-packages")‘ for more information.
## This warning is displayed once every 8 hours.
## Call ‘lifecycle::last_lifecycle_warnings()‘ to see where this warning was
## generated.
```

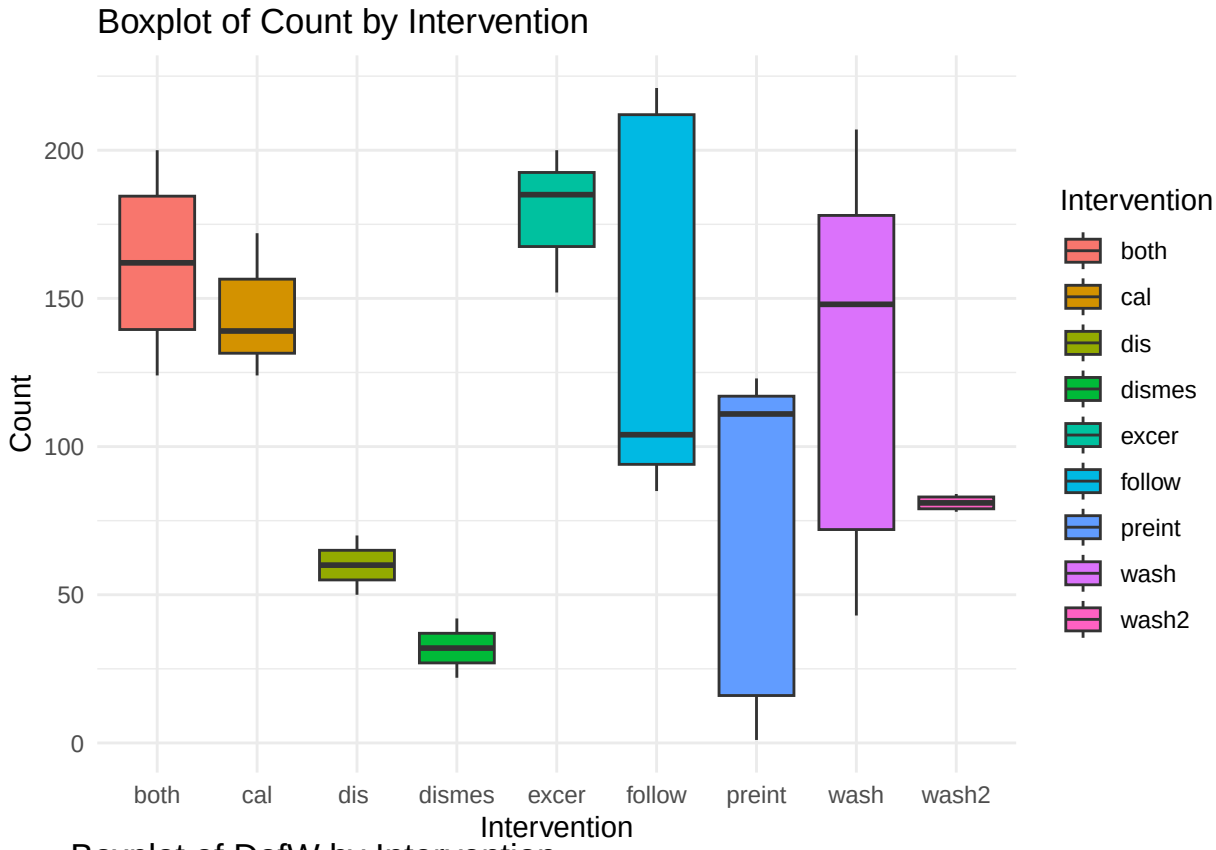


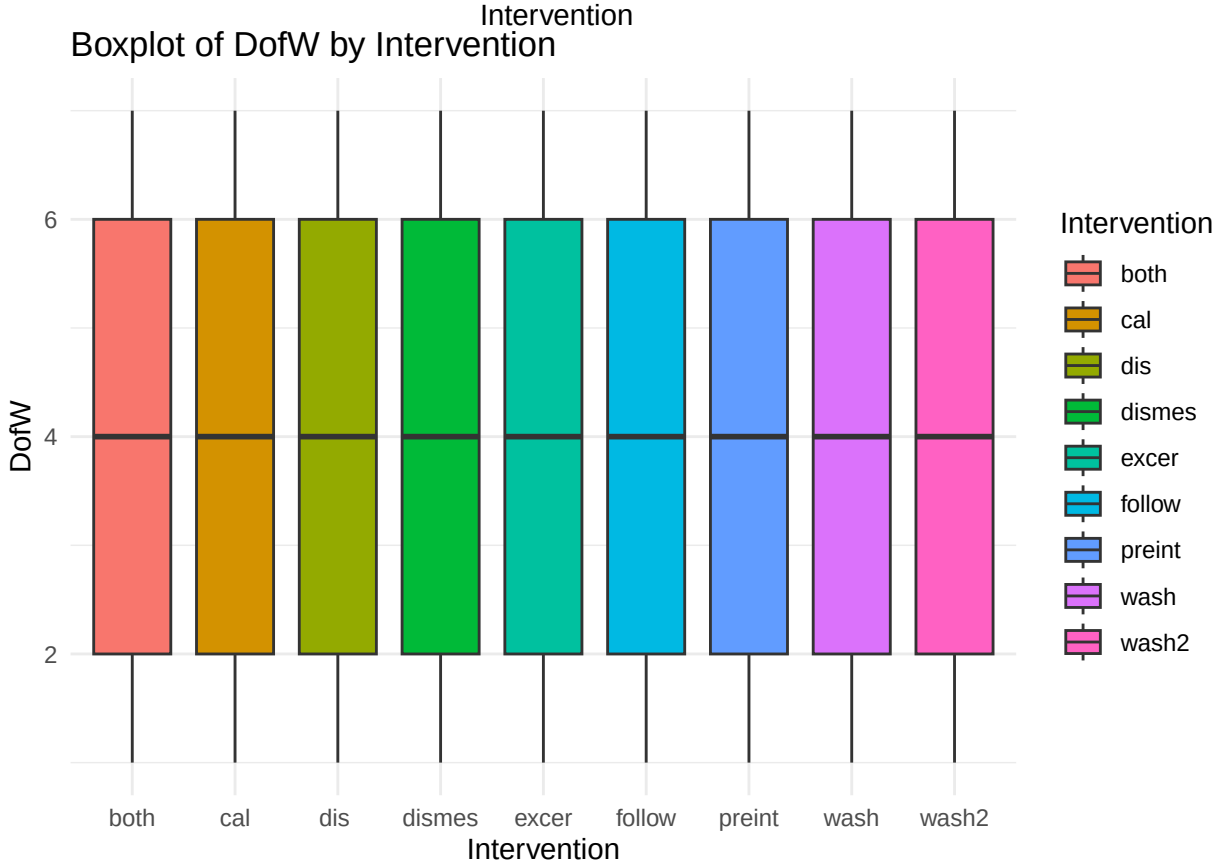

Distribution of DofW
count
75
50
25
0
2
4
6
DofW

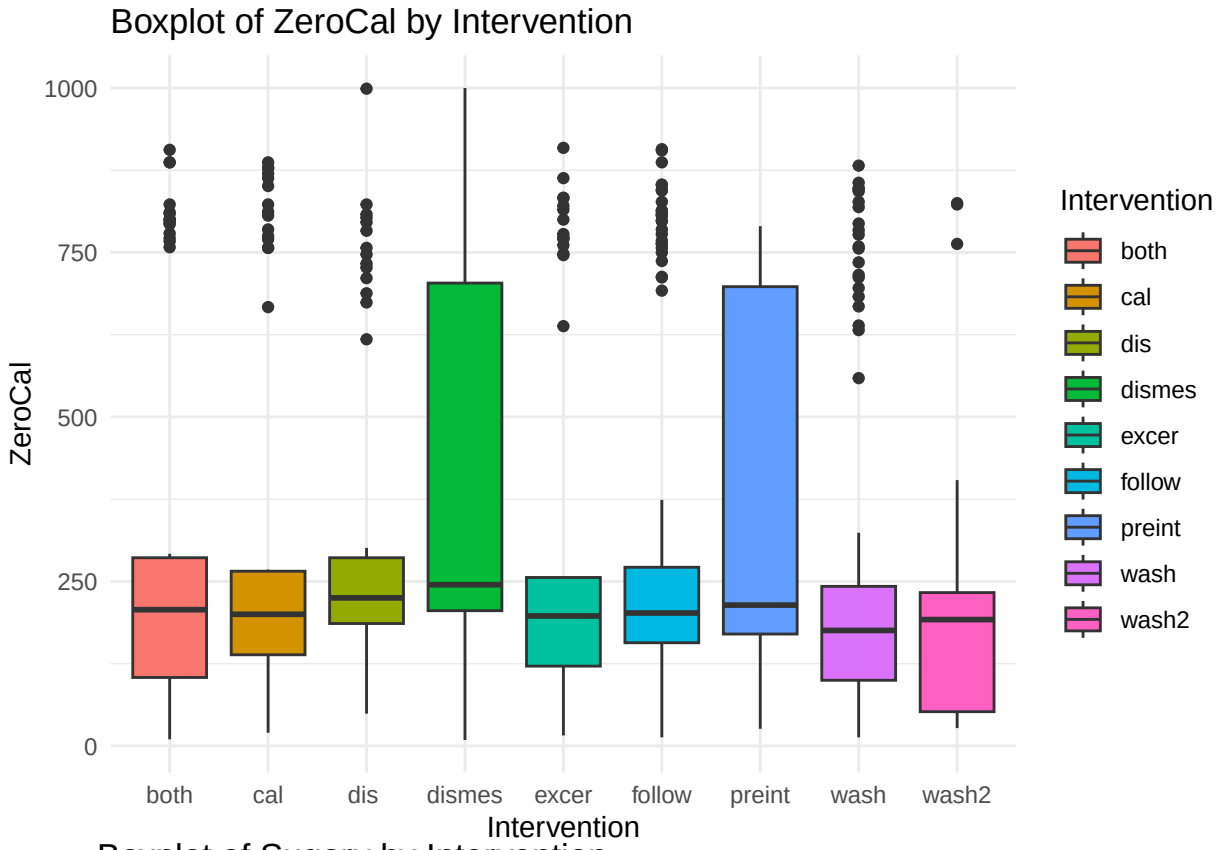

Distribution of ZeroCal
count
150
100
50
0
0
250
500
750
1000
ZeroCal

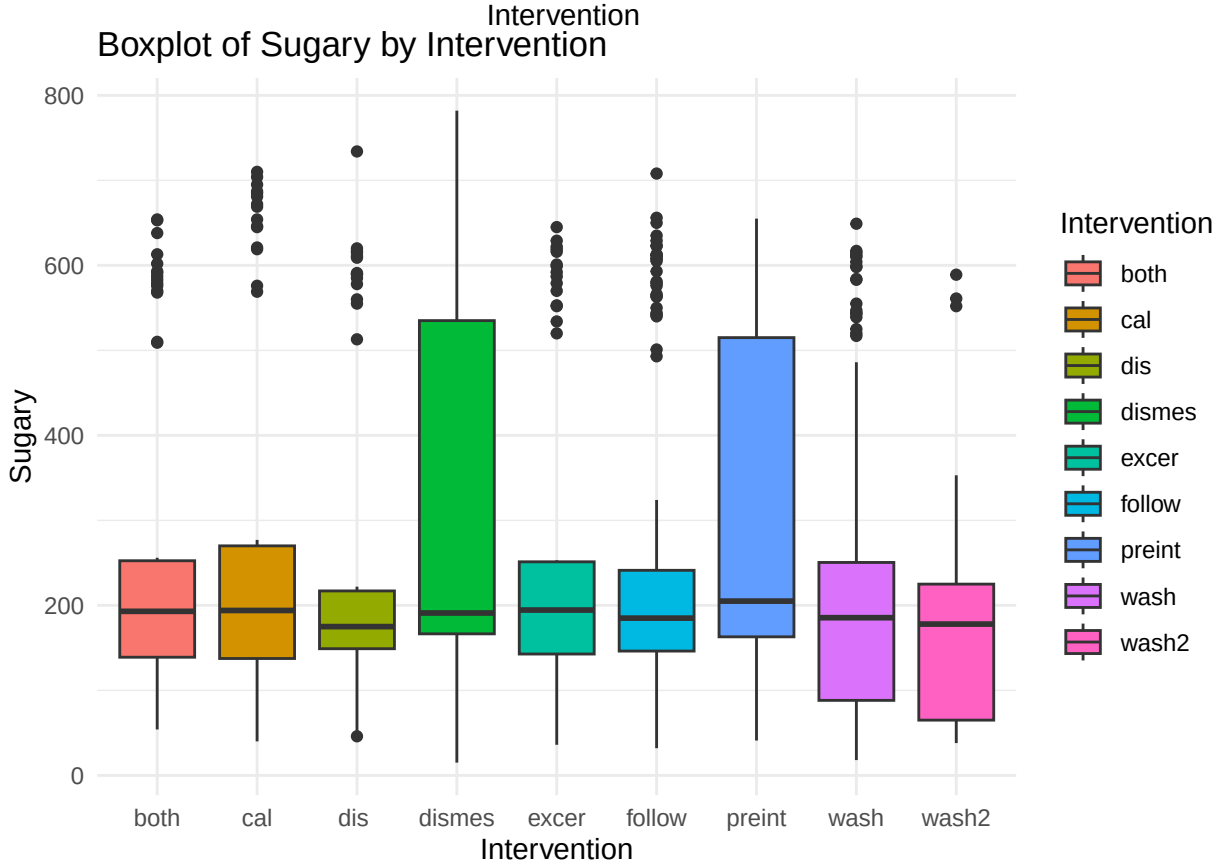



```
boxplots_by_category(beverage_sales, "Intervention")
```

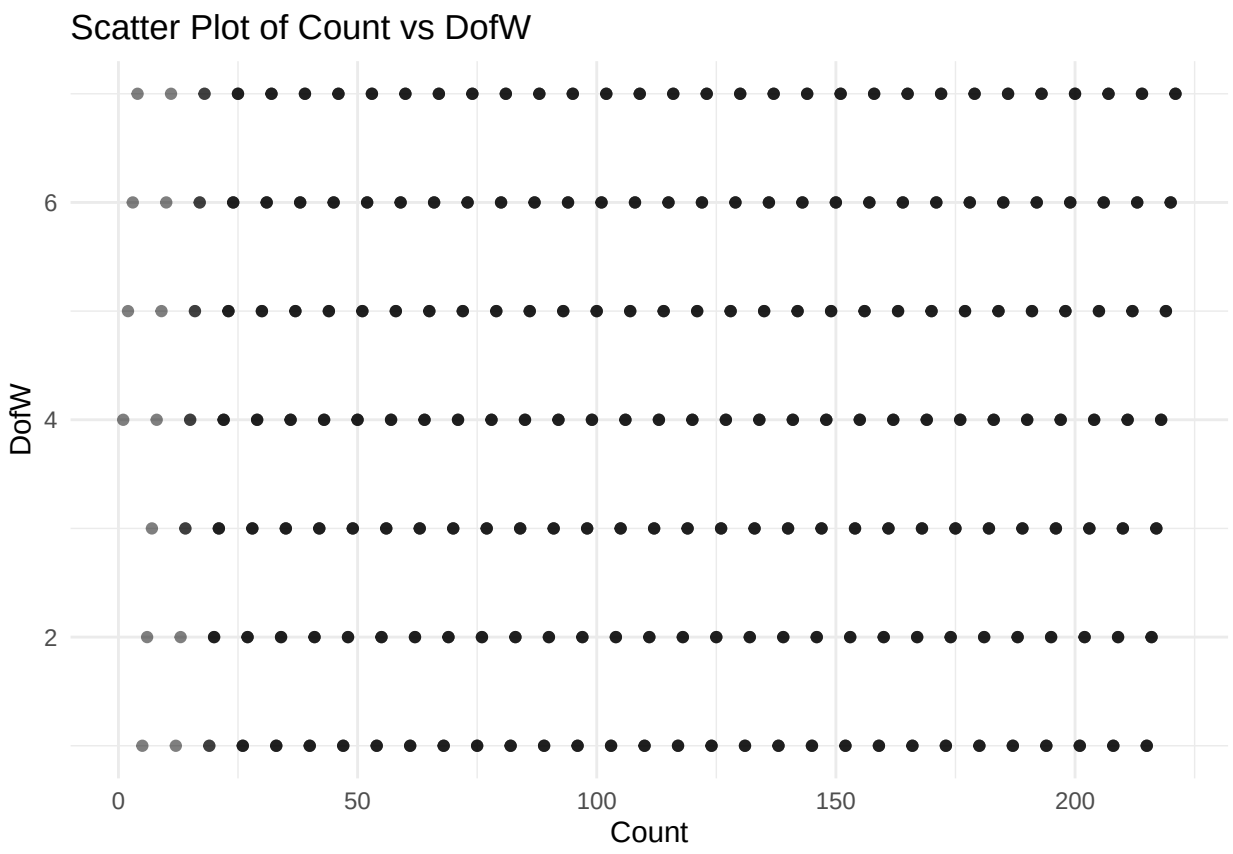

Boxplot of Count by Intervention
Count
200
150
100
50
0
both
cal
dis
dismes
excer
follow
preint
wash
wash2
Intervention
Intervention
both
cal
dis
dismes
excer
follow
preint
wash
wash2

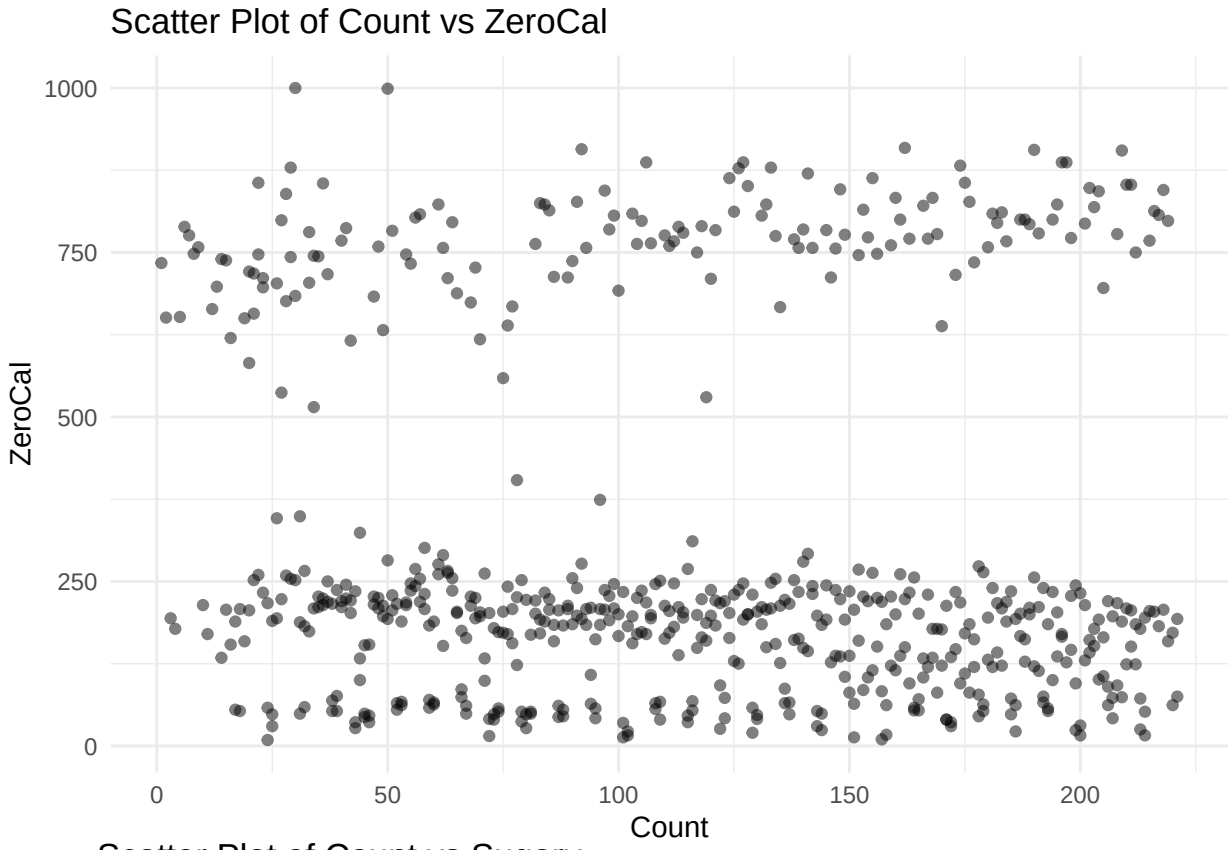

Boxplot of DofW by Intervention
DofW
6
4
2
both
cal
dis
dismes
excer
follow
preint
wash
wash2
Intervention
Intervention
both
cal
dis
dismes
excer
follow
preint
wash
wash2

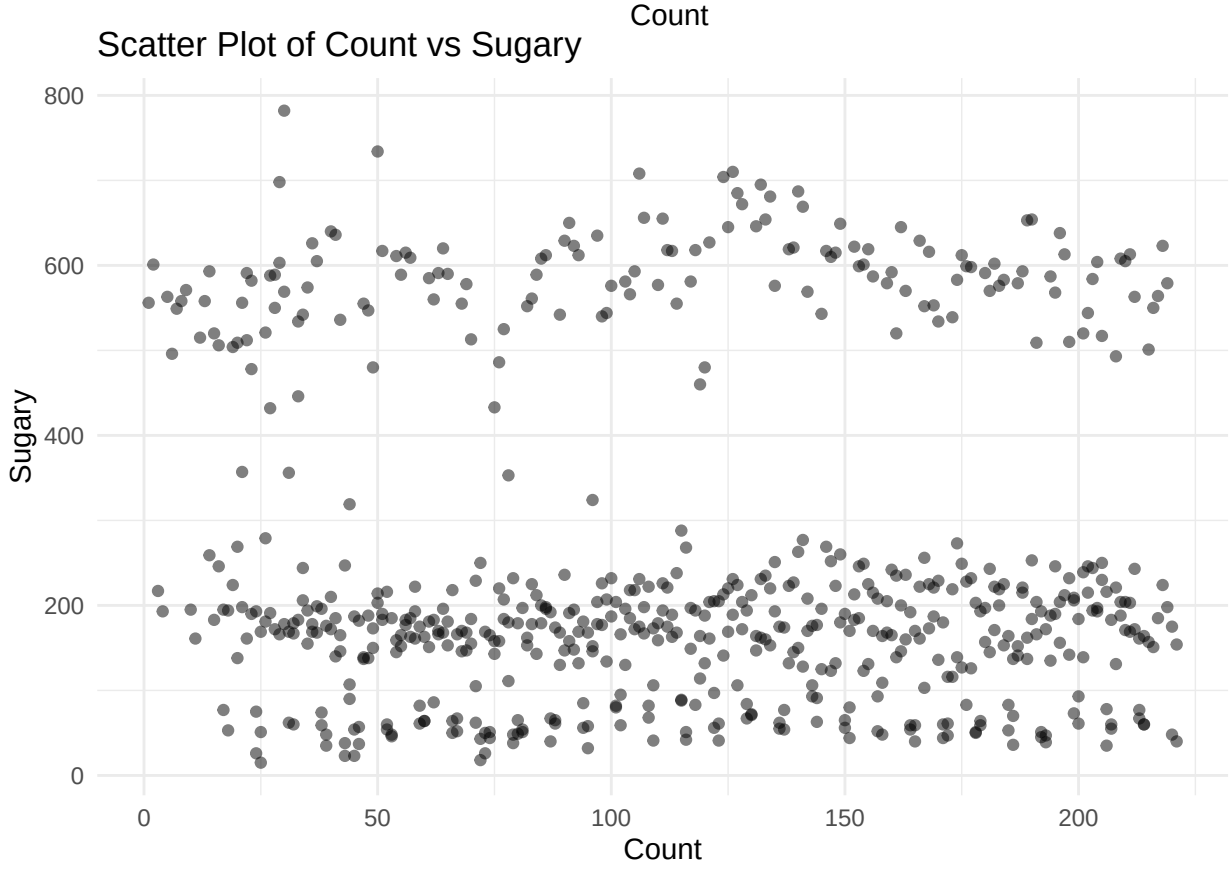

Boxplot of ZeroCal by Intervention
ZeroCal
1000
750
500
250
0
both
cal
dis
dismes
excer
follow
preint
wash
wash2
Intervention
Intervention
both
cal
dis
dismes
excer
follow
preint
wash
wash2

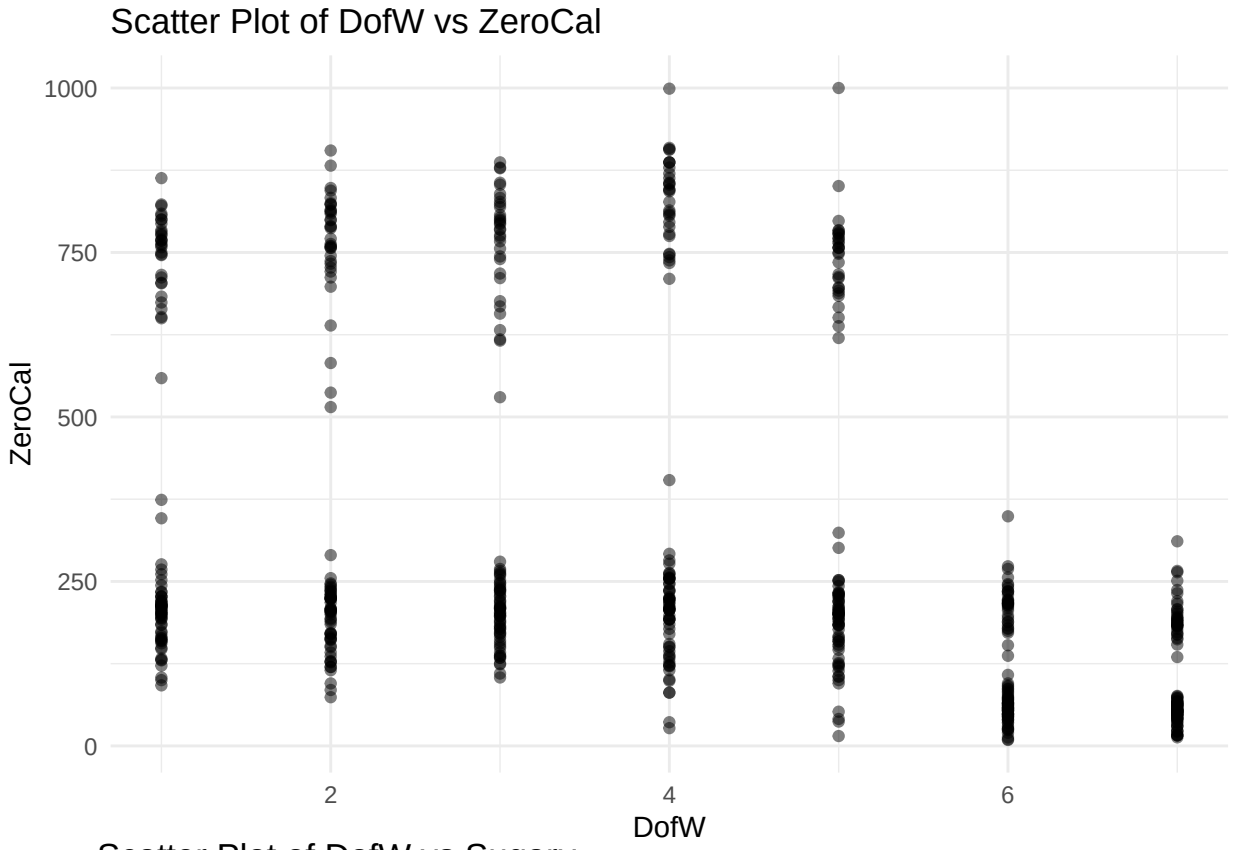

Boxplot of Sugary by Intervention
Sugary
800
600
400
200
0
both
cal
dis
dismes
excer
follow
preint
wash
wash2
Intervention
Intervention
both
cal
dis
dismes
excer
follow
preint
wash
wash2

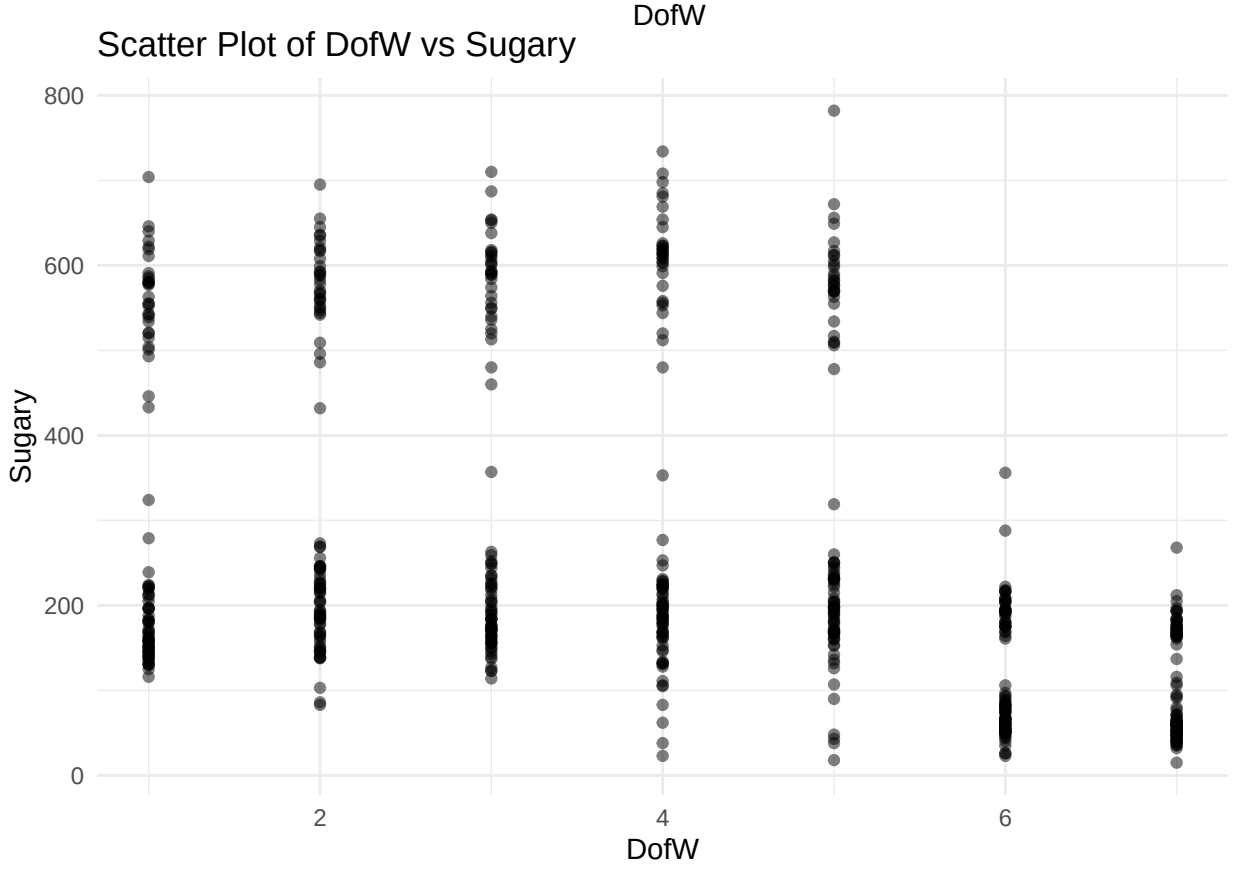
```
scatter_plots(beverage_sales)
```

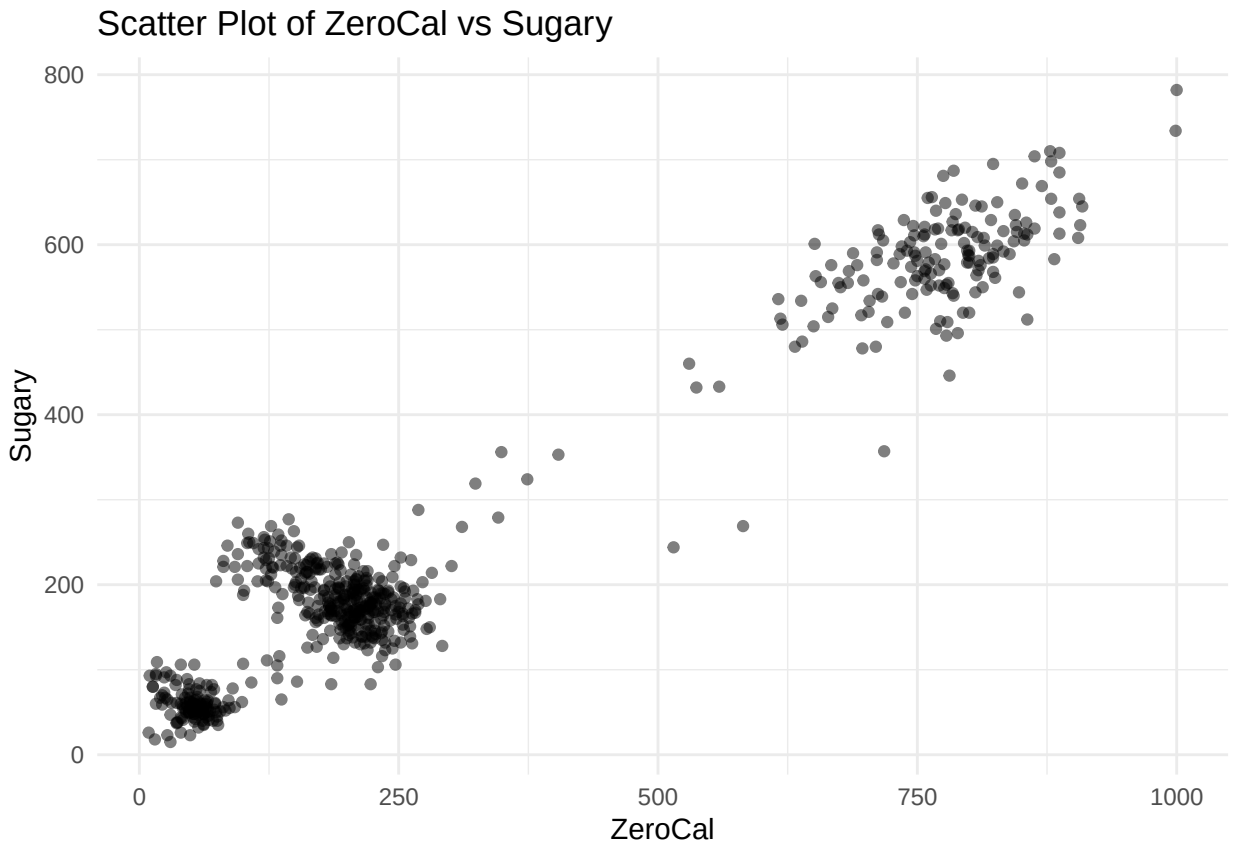



Scatter Plot of Count vs ZeroCal
ZeroCal
1000
750
500
250
0
0
50
100
150
200
Count


Scatter Plot of Count vs Sugary
Sugary
800
600
400
200
0
0
50
100
150
200
Count

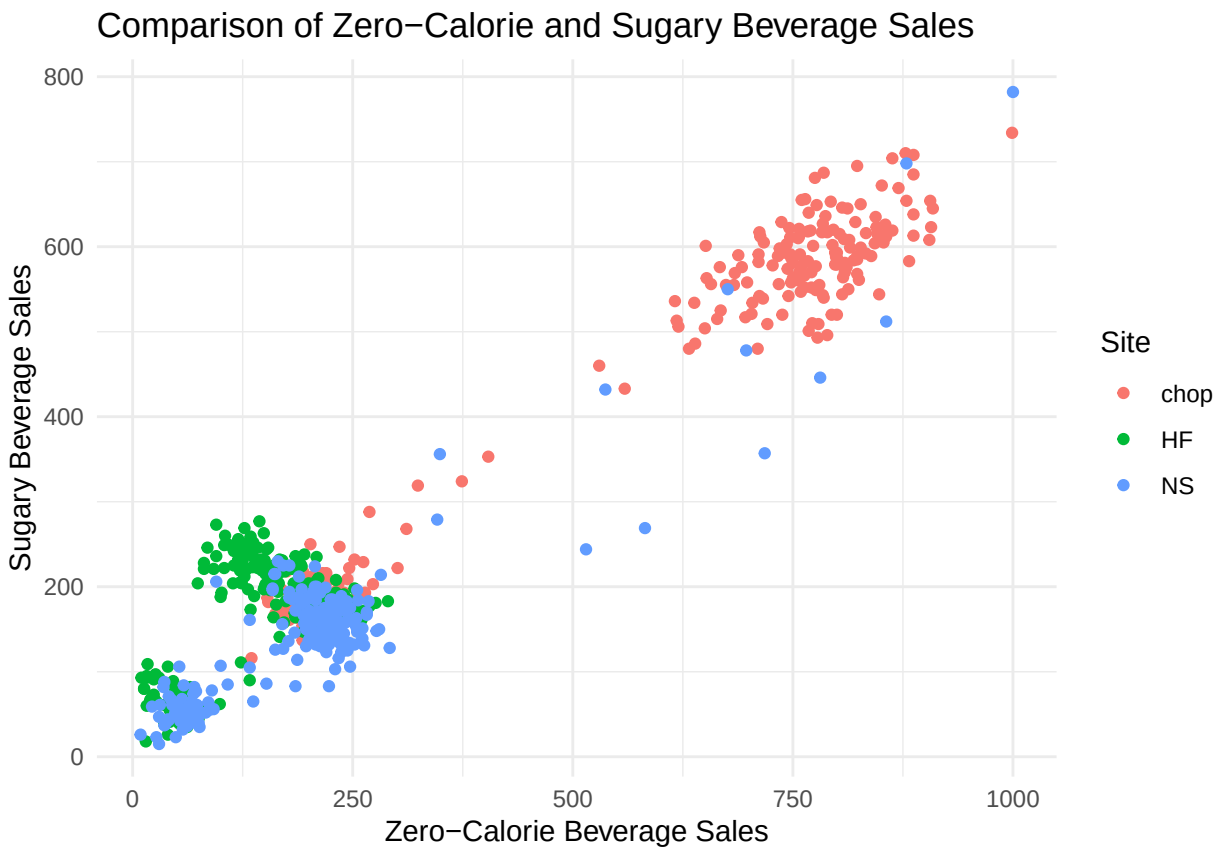

Scatter Plot of DofW vs ZeroCal
ZeroCal
1000
750
500
250
0
2
4
6
DofW


Scatter Plot of DofW vs Sugary
Sugary
800
600
400
200
0
2
4
6
DofW




```
plot_beverage_sales_comparison(beverage_sales, "ZeroCal", "Sugary", "Site")
```

```
## Warning: Removed 9 rows containing missing values (‘geom_point()‘).
```



correlation plot (Johnson)

imbalance/missing value (Sarah) visdat

line plot for trajectory (Par)

4. Formal Analysis

```
# handle_missing_data(beverage_sales)$MissingOverview
```

ITSA (Par)

GEE (Sarah)

LMEM (Johnson)

5. Conclusions

- Recommendations to the clients.

6. References

- Properly formatted citations.

7. Statistical Appendix

- Mathematical formulas.
- Additional tables/figures.