Marketing Analytics

Elen Sukiasyan

2/26/2024

Homework 1

Cutting Plastic Lids

The Good Cup

1.Description

Every year, 500 billion disposable cups end up in landfills. Paper cups sound good in theory, but most use a plastic laminate as a sealant, which makes them difficult to recycle. This is why Choose Planet A's the Good Cup is glued together with a water-based coating, called aqueous, that makes it leak-proof, recyclable (up to seven times), and even biodegradable. "If you throw it in the water or the forest," Choose Planet A's co-founder and managing director Cyril Drouet says, "it will biodegrade and disappear." The cup also has a foldable paper lid, cutting the need for a plastic top. It will be available in the U.S. later this year. Hence, data is a country-wise.

2.Why I have chosen this topic

Choosing the topic of "The Good Cup" is not only a matter of interest but also a personal conviction stemming from a concern for our environment. Last year, my research delved into the impact of plastic pollution, particularly concentrating on how disposable items harm our planet, oceans, seas, animals and our overall surroundings. It was eye-opening to learn that even things we think are recyclable and are labeled as recyclable, aren't actually fully harmless. For example paper cups often have a thin layer of plastic inside that makes them hard to recycle properly.

But "The Good Cup" is different. This innovation made me believe that not everything is lost. It's designed to be leak-proof using a special water-based coating. This means, that it can be recycled up to seven times and even be biodegradable breaking down naturally over time.

So, I chose this topic because it represents a real solution to a big problem. And I believe, that "The Good Cup" is a step towards a cleaner future.

3.Statista innovation justification

I decided to focus on the usage of disposable cups and plates as a look-alike innovation because it aligns perfectly with my topic. The only difference about the topics is that the statista dataset has not only cups but also plates in research. However, the concept is completely the same, and statistically, it is related to my data too.

The dataset shows the usage of disposable cups and plates from 2011 to 2024, giving an opportunity to investigate the potential shift towards more sustainable alternatives, such as "The Good Cup." By examining the trends over time, my goal is to understand the readiness of consumers to accept eco-friendly solutions and the potential impact of innovations like "The Good Cup" on reducing plastic waste.

Hence, this dataset offers valuable insights about consumer behaviour and market dynamics, allowing me to estimate the market potential for sustainable alternatives and make informed predictions.

4.Estimating Bass Model

cleaning the data

```
library(readxl)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
# Read the Excel file, we are taking the second sheet as our excel file has 2 pages
data <- read_excel('/Users/elensukiasyan/Downloads/Statista_innovation.xlsx', sheet = 2)
```

```
## New names:
## * `` -> `...2`
## * `` -> `...3`
## * `` -> `...4`
```

```
# choosing the rows we need and assigning them column names that we want
data <- data[3:16, ]
colnames(data) <- c("Year", "Yes", "No", "don't know")

# Remove the last character from the Year column in the specified rows (as I had * from 2021 to 2024 an
rows_to_modify <- 11:14
data$Year[rows_to_modify] <- substr(data$Year[rows_to_modify], 1, nchar(data$Year[rows_to_modify]) - 1)
data
```

```
## # A tibble: 14 x 4
##    Year  Yes    No     `don't know`
##    <chr> <chr>  <chr>  <chr>
##  1 2011  208.04 99.99  3.54
##  2 2012  210.55 99.62  3.7
##  3 2013  211.48 101.31 3.27
##  4 2014  210.58 104.93 2.88
```

```
##  5 2015  216.19 101.48 3.07
##  6 2016  217.99 102.18 2.9
##  7 2017  218.61 103.44 3.1
##  8 2018  222.87 101.38 2.91
##  9 2019  218.74 109.1  2.43
## 10 2020  219.73 109.96 2.95
## 11 2021  221.44 110.75 2.81
## 12 2022  222.99 111.52 2.83
## 13 2023  224.53 112.29 2.84
## 14 2024  226.04 113.05 2.86
```

```
# changing characters to integers
data$Year <- as.integer(data$Year)
data$Yes <- as.integer(data$Yes)

# Calculating cumulative usage and making a new column by cumulative usage values
data$cumulative_usage <- cumsum(data$Yes)

head(data)
```

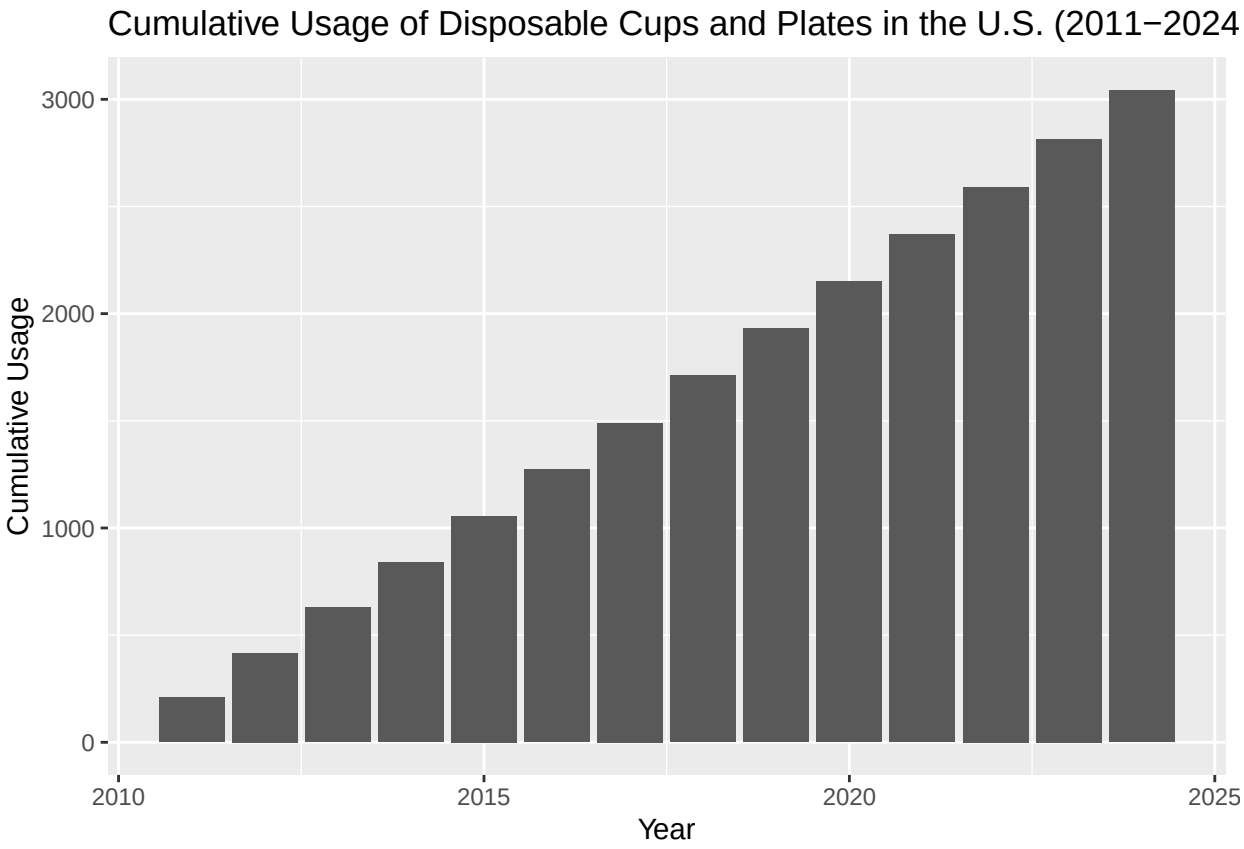
```
## # A tibble: 6 x 5
##    Year   Yes No     ‘don’t know‘ cumulative_usage
##   <int> <int> <chr>  <chr>                  <int>
## 1  2011   208 99.99  3.54                     208
## 2  2012   210 99.62  3.7                      418
## 3  2013   211 101.31 3.27                     629
## 4  2014   210 104.93 2.88                     839
## 5  2015   216 101.48 3.07                    1055
## 6  2016   217 102.18 2.9                     1272
```

```
colnames(data)
```

```
## [1] "Year"             "Yes"              "No"               "don’t know"
## [5] "cumulative_usage"
```

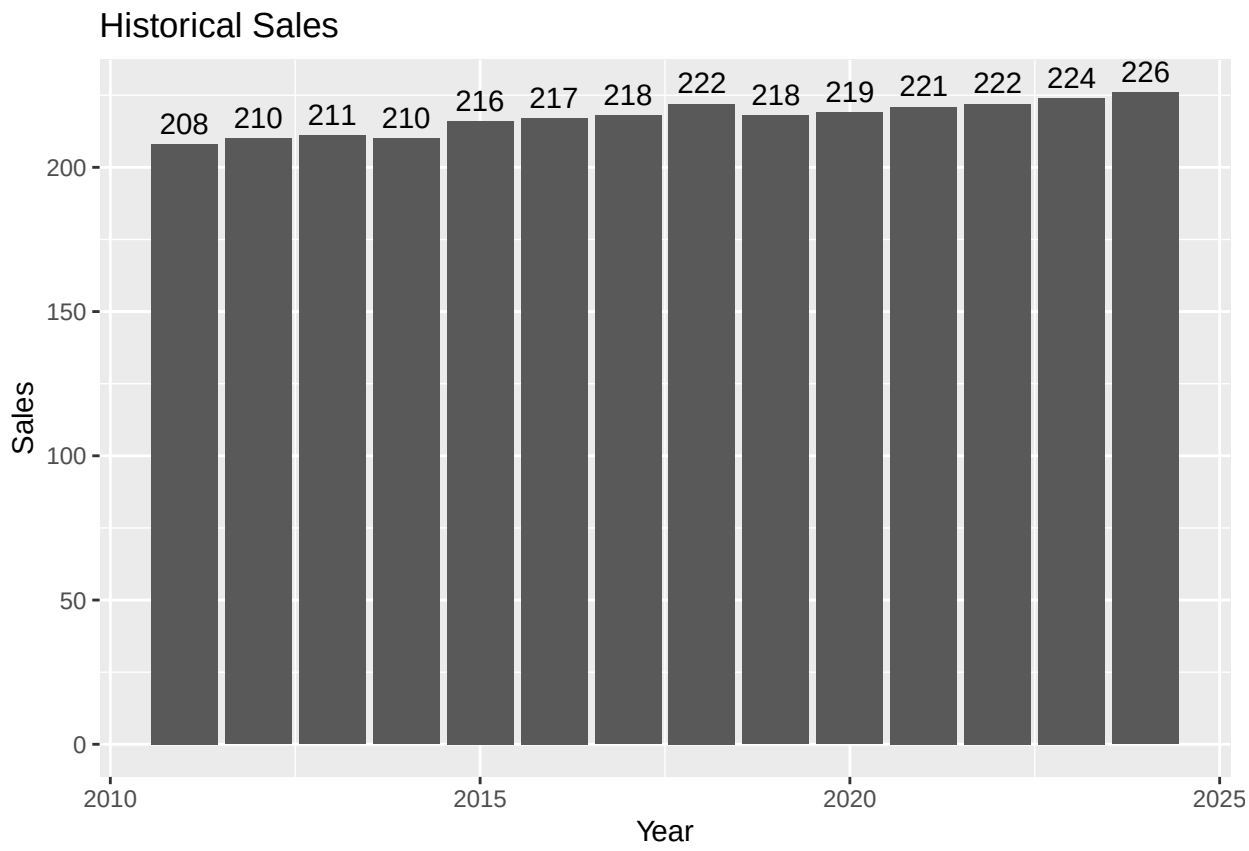
Printing cumulative plot

```
library(ggplot2)
cumulative_plot <- ggplot(data, aes(x = Year, y = cumulative_usage)) +
  geom_bar(stat = 'identity') +
  labs(title = 'Cumulative Usage of Disposable Cups and Plates in the U.S. (2011-2024)',
       x = 'Year',
       y = 'Cumulative Usage')
cumulative_plot
```



Plotting historical sales

```
ggplot(data, aes(x = Year, y = Yes)) +
  geom_bar(stat = 'identity') +
  geom_text(aes(label = Yes), vjust = -0.5) +  # Adding data labels above bars
  labs(x = "Year", y = "Sales", title = "Historical Sales")
```



bass.f is the fraction of the total market that adopts at time t bass.F is the fraction of the total market that has adopted up to and including time t q is imitation rate p is innovation rate

```
bass.F <- function(t,p,q){
(1-exp(-(p+q)*t))/
(1+(q/p)*exp(-(p+q)*t))
}

bass.f <- function(t,p,q){
((p+q)^2/p)*exp(-(p+q)*t)/(1+(q/p)*exp(-(p+q)*t))^2
}
```

Estimating Bass model

```
t <- 1:nrow(data)
sales <- data$Yes

# Fit the Bass model to the data
bass_fit <- nls(sales ~   m * (((p + q)^2 / p) * exp(-(p + q) * t)) / (1 + (q / p) * exp(-(p + q) * t))
                start = list(m = sum(sales), p= 0.02, q= 0.4))
summary(bass_fit)
```

```
##
## Formula: sales ~ m * (((p + q)^2/p) * exp(-(p + q) * t))/(1 + (q/p) *
```

```
##     exp(-(p + q) * t))^2
##
## Parameters:
##    Estimate Std. Error t value Pr(>|t|)
## m 2.106e+04  9.360e+03   2.250   0.0459 *
## p 9.791e-03  4.293e-03   2.281   0.0435 *
## q 1.855e-02  6.564e-03   2.826   0.0165 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.793 on 11 degrees of freedom
##
## Number of iterations to convergence: 9
## Achieved convergence tolerance: 1.102e-06
```

Extracting p and q

```
p <- coef(bass_fit)["p"]
q <- coef(bass_fit)["q"]
m <- coef(bass_fit)["m"]
p
```

```
##           p
## 0.009791109
```

```
q
```

```
##          q
## 0.01854869
```

```
m
```

```
##        m
## 21058.58
```

5.Make predictions of the diffusion of the innovation

```
library(diffusion)
sales <- data$Yes
diff_m <- diffusion(sales)

# Extract parameter estimates
p <- round(diff_m$w, 4)[1]
q <- round(diff_m$w, 4)[2]
m <- round(diff_m$w, 4)[3]
p
```

```
##      p
## 0.0127
```

```
q
```

```
##      q
## 0.0225
```

```
m
```

```
##        m
## 16315.74
```

6.Estimate the number of adopters by period. Here we are going to use 5th point's (method 2) parameter values as there are better fot to our data

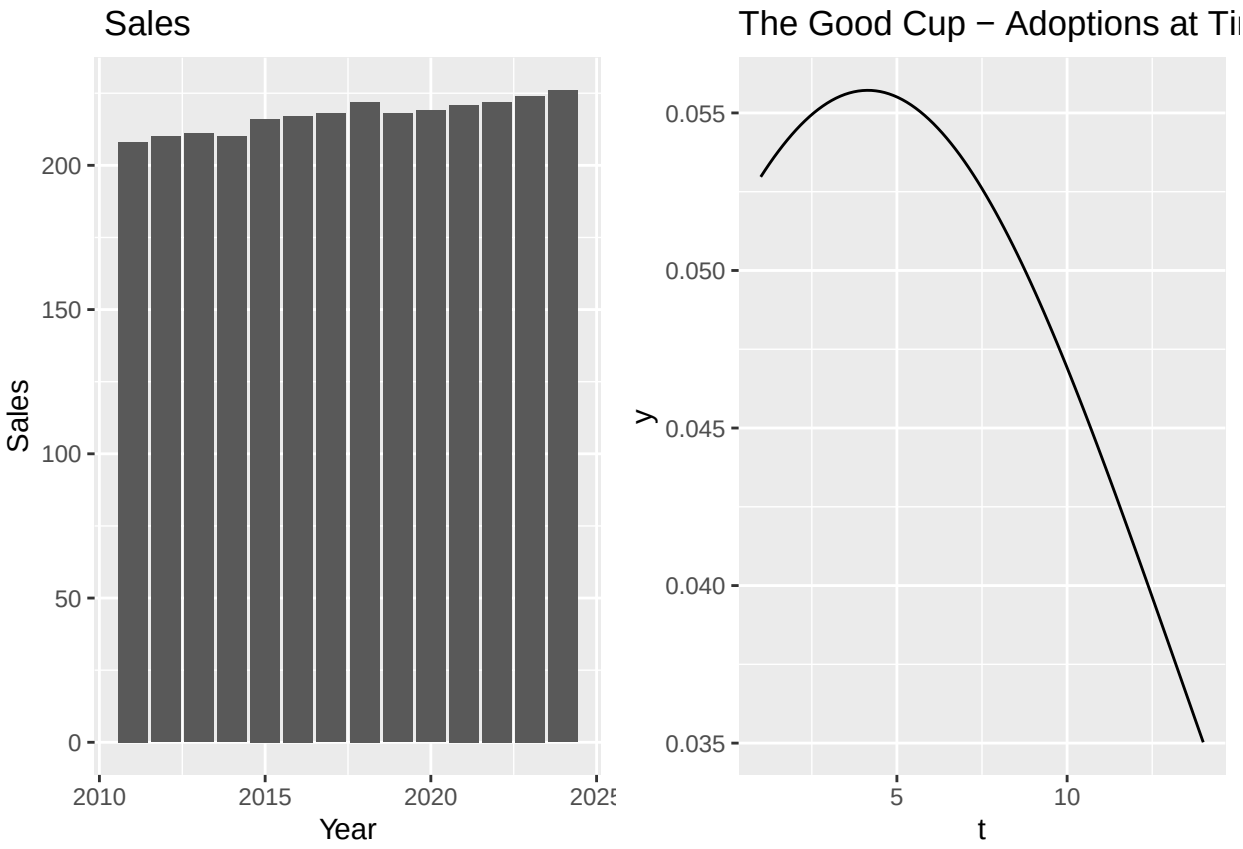
```
p
```

```
##      p
## 0.0127
```

```
q
```

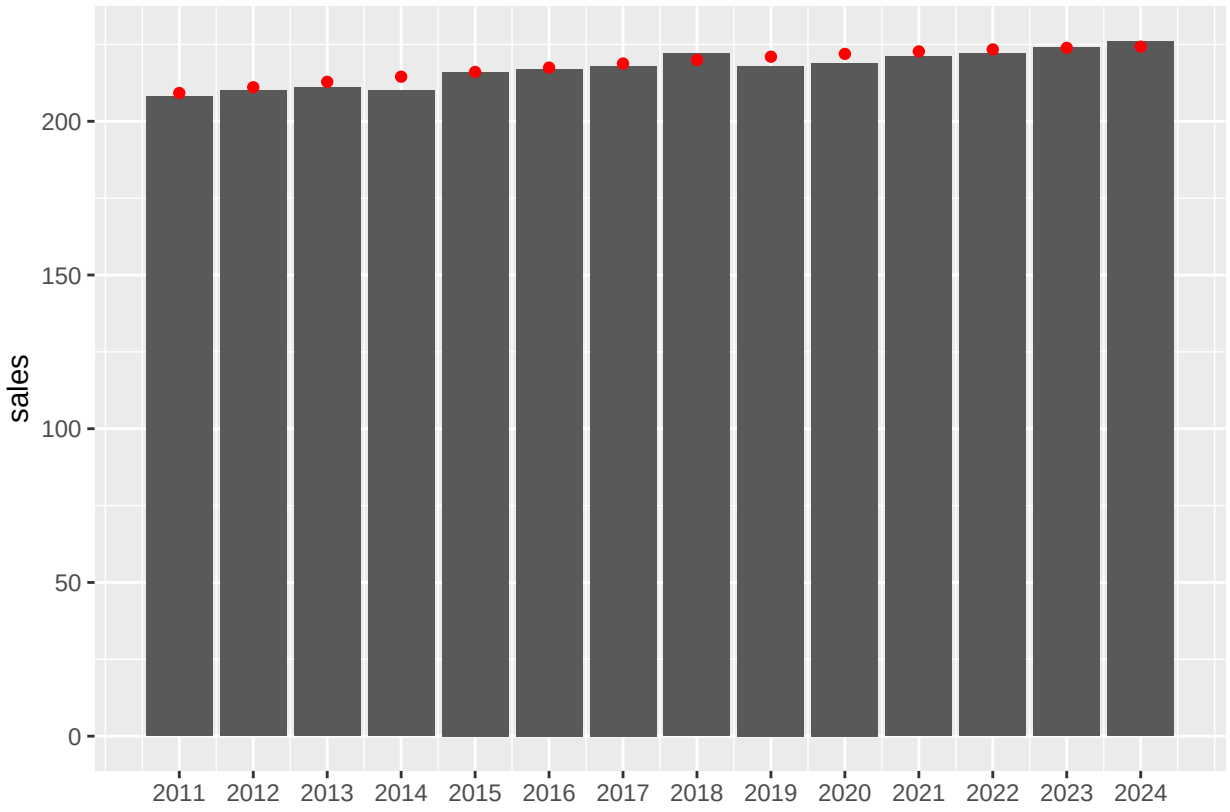
```
##      q
## 0.0225
```

```
library(ggpubr)
sales = ggplot(data, aes(x = Year, y = Yes)) +
  geom_bar(stat = 'identity') +
  labs(x = "Year", y = "Sales", title = " Sales")

# Plot adoptions at time t
time_ad <- ggplot(data.frame(t = c(1:14)), aes(t)) +
stat_function(fun = bass.f, args = c(p=0.05106, q=0.09259)) +
labs(title = 'The Good Cup - Adoptions at Time t')

# Arrange the plots side by side
ggarrange(sales, time_ad)
```



Estimated number of adopters by time

```
data$predicted = bass.f(1:14, p = 0.0127 , q = 0.0225 )*16315.74
ggplot(data , aes(x = Year, y = Yes)) +
geom_bar(stat = 'identity') +
geom_point(mapping = aes(x=Year, y=predicted), color = 'red')+
labs( x= ' ', y = 'sales')+scale_x_continuous(breaks = 2011:2024, labels = 2011:2024)
```


sales
200
150
100
50
0
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024