MA_HW1

Eduard Petrosyan

2024-02-25

For this analytics, I have chosen to focus on athletic footwear producs; more specifically, I have chosen the innovation Adidas Adizero Adios Pro Evo 1. Let's briefly introduce this shoe type. It is demonstrating cutting-edge technologyy in the realm of distance racing and marathons and Ethiopian runner Tigst Assefa's record-breaking performance at the Berlin Marathon, clocking 2:11:53, showcased the effectiveness of the Evo 1. Weighing 25 percent less than the Nike Vaporfly and featuring a 4.9 oz. The Evo 1 introduces innovative features, such as a translucent lightweight upper and a super-thin rubber outsole. A proprietary manufacturing process for the Lightstrike Pro foam further enhances its appeal. However, it has an issue about its cost which is $500, a price higher than the average sport shoes.

On the other hand, similar in its target customers, Nike's Vaporfly series, an earlier innovation that in its period revolutionized distance running, have become a subject of interesting discussions. Featuring carbon plates and springy midsole foam, the Vaporfly series has demonstrated a real and substantial advantage, as highlighted by The New York Times analysis. The Zoom Vaporfly 4% and ZoomX Vaporfly Next%, in particular, have set a benchmark, providing runners with a 4 to 5 percent edge over average shoes and a 2 to 3 percent advantage over the next-fastest popular shoe. This notable advantage has not only stirred discussions among professionals and amateurs alike but has also prompted other brands like Brooks, Saucony, New Balance, Hoka One One, and Asics to introduce similar shoes to the market. This historical context sets the stage for analyzing the market dynamics and potential adoption patterns of the Adidas Adizero Adios Pro Evo 1, using some availables statistics of its Nike counterpart.

```
df <- read_excel("shoes.xlsx", sheet = "Data")
```

At this part of my project, I wanted to approximate the market dynamics of Nike's Vaporfly shoes, because of the inherent difficulty in directly accessing granular data specific to Vaportfly sales. The dataset that is being analyzed represents Nike's worldwide revenue from 2016 to 2021 within the Running category, which is a pragmatic alternative. When taking the narrow part of the data on Running category, where Vaporfly shoes are the primary focus, I could understand the insights into their market impact and adoption patterns.

I wasn't able to obtain detailed revenue figures at the individual shoe model level because of the lack of necessary statistics. To approximate it, I turned to the broader dataset by utilizing the revenue withing the Running category as a reliable proxy for Vaporfly's market presence. In this way, I was able to overcome the limitations in direct data access and use the available data to make assumptions about the unavailable dataset.

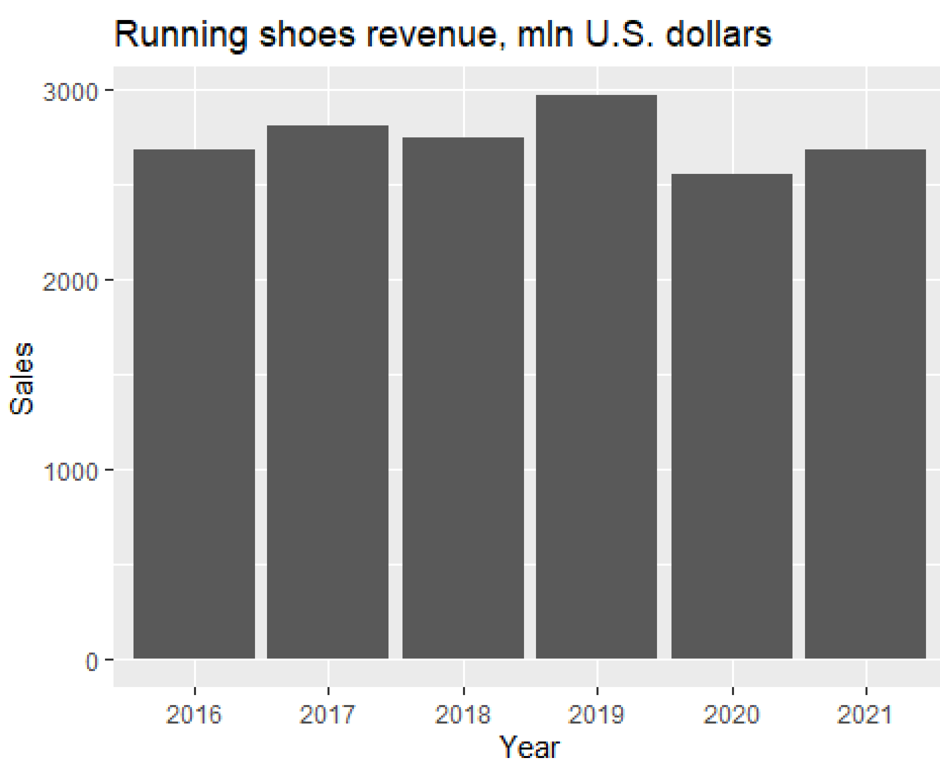
It is important to note, that I used specific source to emphasize, that in my assignment, the data that demonstrated Nike's overall revenue worldwide, is modified to demonstrate Nike specifically shoe revenue using each year's percent share from source.

```r
running_df <- df[, c("Year", "Running")]
percentages <- c(61.1, 61.4, 61.2, 66.2, 66.7, 67.4)

running_df$Running <- running_df$Running * (percentages / 100)

colnames(running_df) <- c("Year", "Sales")
print(running_df)
```

```
## # A tibble: 6 × 2
##   Year  Sales
##   <chr> <dbl>
## 1 2016  2689.
## 2 2017  2810.
## 3 2018  2752.
## 4 2019  2971.
## 5 2020  2555.
## 6 2021  2687.
```

This is an approximate Nike's revenue from running shoes worldwide from 2016 to 2021. Now, let's visualize this data.

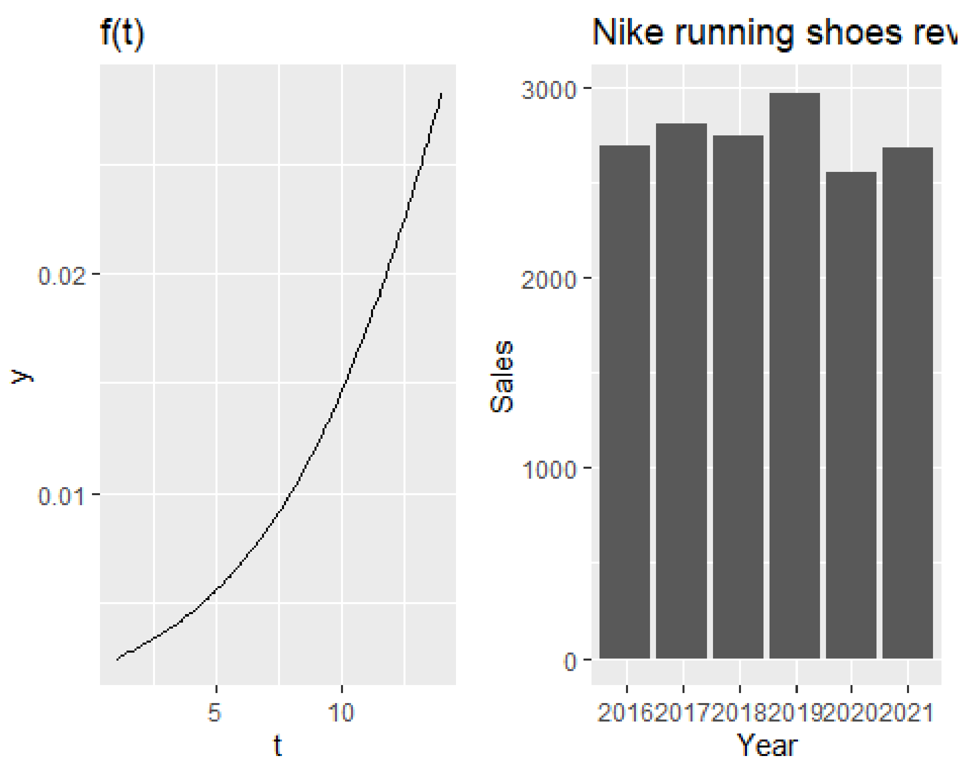
```r
ggplot(data = running_df, aes(x = Year, y = Sales)) +
geom_bar(stat = 'identity') +
ggtitle('Running shoes revenue, mln U.S. dollars')
```



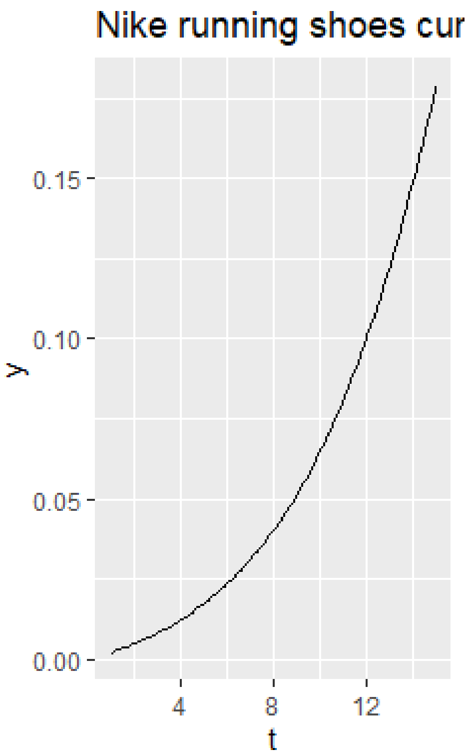
Let's now define the bass model functions.

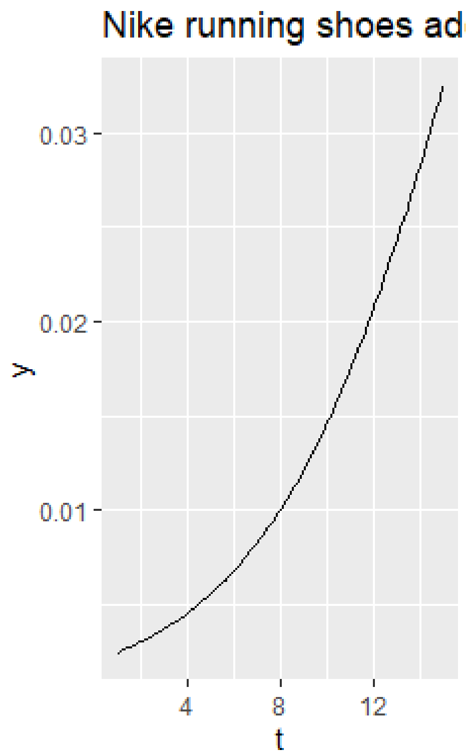
```r
bass.f <- function(t, p, q) {
  ((p + q)^2 / p) * exp(-(p + q) * t) / (1 + (q / p) * exp(-(p + q) * t))^2
}
```

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

time_ad = ggplot(data.frame(t = c(1:14)), aes(t)) +
stat_function(fun = bass.f, args = c(p=0.002, q=0.21)) +
labs(title = 'f(t)')

sm_sales = ggplot(data = running_df, aes(x = Year, y = Sales)) +
geom_bar(stat = 'identity') +
ggtitle('Nike running shoes revenue, mln U.S. dollars')

ggarrange(time_ad, sm_sales)
```




```r
cum_ad = ggplot(data.frame(t = c(1, 15)), aes(t)) +
stat_function(fun = bass.F, args = c(p=0.002, q=0.21)) +
labs(title = 'Nike running shoes cumulative adoptions')


time_ad = ggplot(data.frame(t = c(1, 15)), aes(t)) +
stat_function(fun = bass.f, args = c(p=0.002, q=0.21)) +
labs(title = 'Nike running shoes adoptions at time t')
ggarrange(cum_ad, time_ad)
```





Starting parameters for innovation (p) and imitation (q) coefficients do not result in accurately depiction of what our sales distribution look like, so we move on to estimation of the parameters. We will try the estimation via 2 methods and try to identify the best from them.

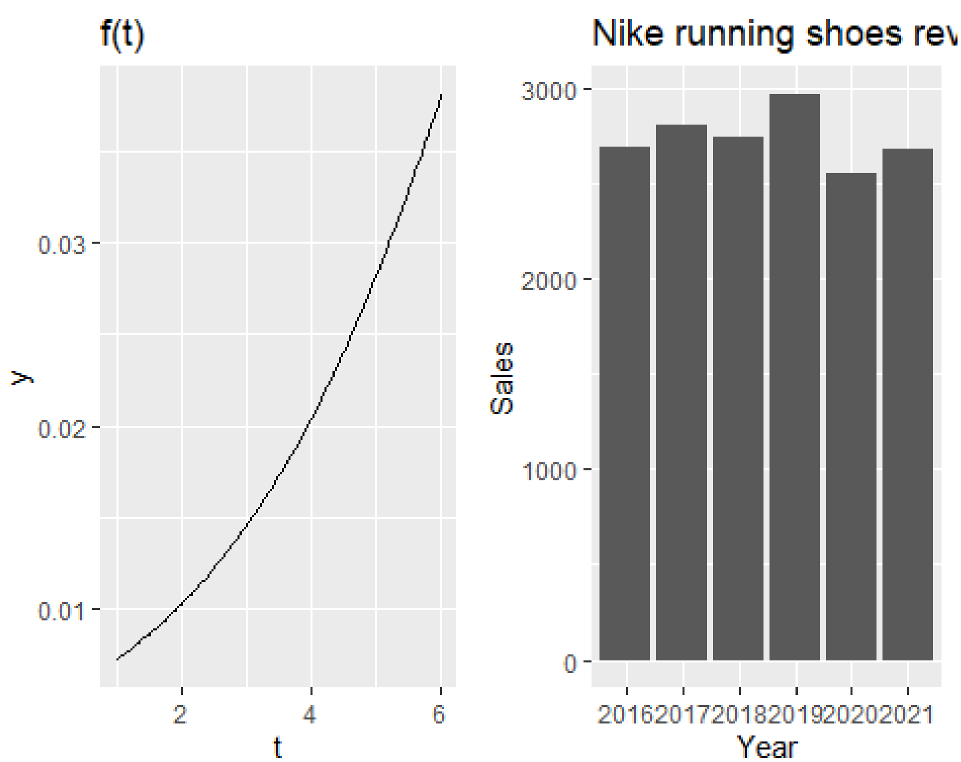
1) We will use nls() - Non-linear Least Squares for first approach.

```
sales = running_df$Sales
t = 1:length(sales)
bass_m = nls(sales ~ m*(((p+q)^2/p)*exp(-(p+q)*t))/
(1+(q/p)*exp(-(p+q)*t))^2,
start=c(list(m=sum(sales),p=0.02,q=0.4)))

summary(bass_m)

##
## 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 3.794e+04  1.350e+04   2.811   0.0673 .
## p 6.782e-02  1.919e-02   3.533   0.0386 *
## q 1.250e-01  7.995e-02   1.563   0.2160
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 153.4 on 3 degrees of freedom
##
## Number of iterations to convergence: 7
## Achieved convergence tolerance: 3.474e-07
```

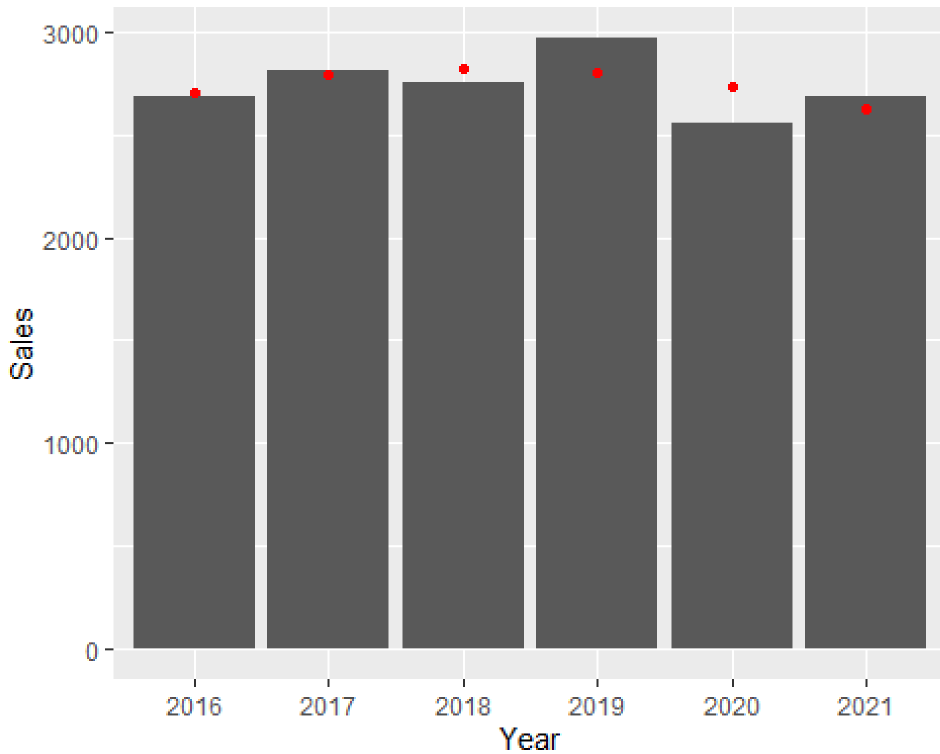
The summary demonstrates, that the potential market size (m) is estimated at approximately 37,940, indicating the anticipated number of adopters. The coefficients for innovators (p) and imitators (q) are estimated at 0.068 and 0.125. Innovation rate is approximately two times less than imitation rate, which indicates that the rate at which new adopters independently embrace the innovation (p) is roughly half the rate at which individuals imitate others in their decision to adopt (q). The model's fit to the data is reflected in the residual standard error of 153.4.

Now, let's visualize with those parameters.

```
time_ad = ggplot(data.frame(t = c(1:6)), aes(t)) +
stat_function(fun = bass.f, args = c(p=0.005, q=0.369)) +
labs(title = 'f(t)')

ggarrange(time_ad, sm_sales)
```




```
running_df$pred_sales = bass.f(1:6, p = 0.068, q = 0.125)*37940
ggplot(data = running_df, aes(x = Year, y = Sales)) +
geom_bar(stat = 'identity') +
geom_point(mapping = aes(x=Year, y=pred_sales), color = 'red')
```



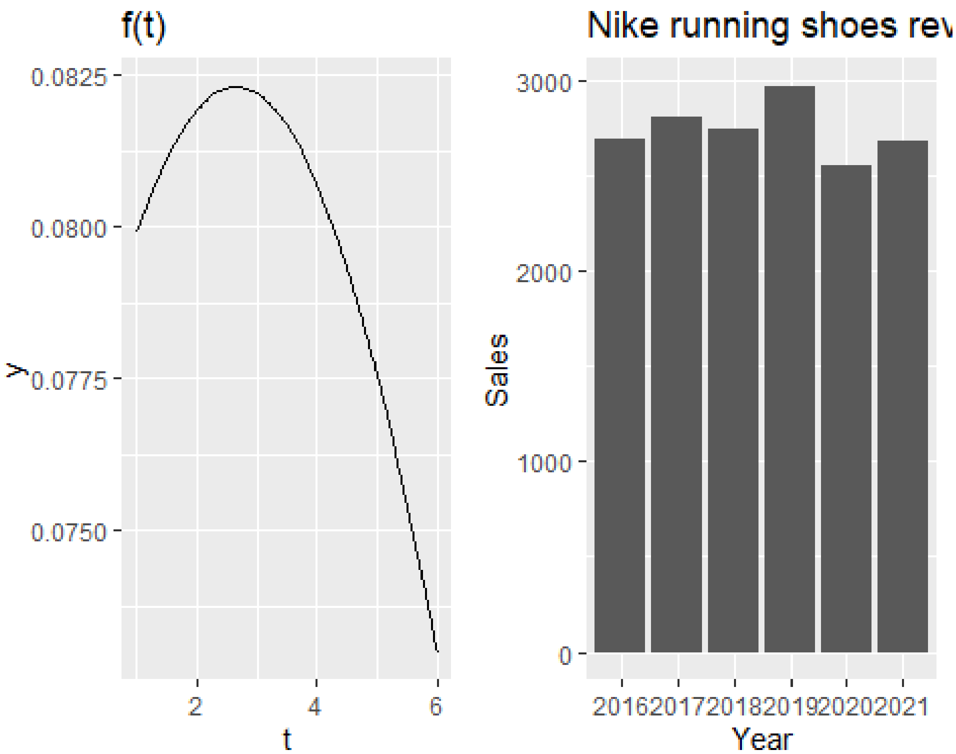
The red dots demonstrate our estimation for corresponding years, which are pretty accurate and close to actual values. Now, let's try other approach for parameter estimation using diffusion library.

```
diff_m = diffusion(sales)
p=round(diff_m$w,4)[1]
q=round(diff_m$w,4)[2]
m=round(diff_m$w,4)[3]
diff_m

## bass model
##
## Parameters:
##                                  Estimate p-value
## p - Coefficient of innovation      0.0763      NA
## q - Coefficient of imitation       0.1328      NA
## m - Market potential           34435.3336      NA
##
## sigma: 109.2018
```

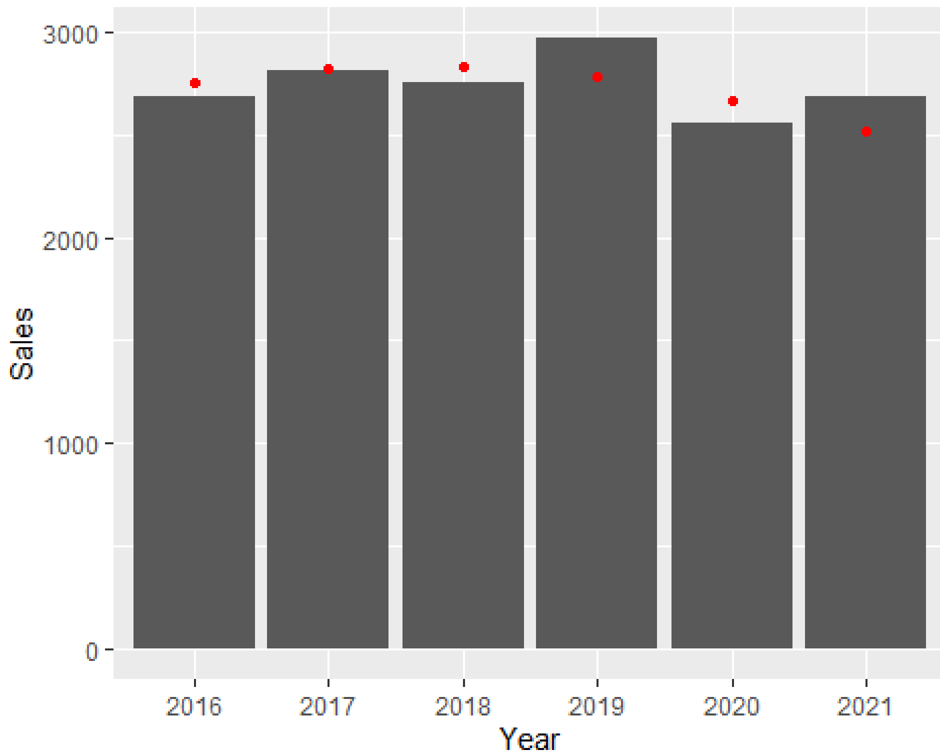
For this approach, we see, that summary demonstrates, the potential market size (m) is estimated at approximately 34435, indicating the anticipated number of adopters. The coefficients for innovators (p) and imitators (q) are close in values to previous approach estimations, 0.076 and 0.133. The model's fit to the data is reflected in the residual standard error of 109.2018, which is less than previous approach's sigma, which implies, that this method may more accurately fit to our sales distribution.

```
time_ad = ggplot(data.frame(t = c(1:6)), aes(t)) +
stat_function(fun = bass.f, args = c(p, q)) +
labs(title = 'f(t)')
```

```r
ggarrange(time_ad, sm_sales)
```




```r
running_df$pred_sales = bass.f(1:6, p , q )*m
ggplot(data = running_df, aes(x = Year, y = Sales)) +
geom_bar(stat = 'identity') +
geom_point(mapping = aes(x=Year, y=pred_sales), color = 'red')
```

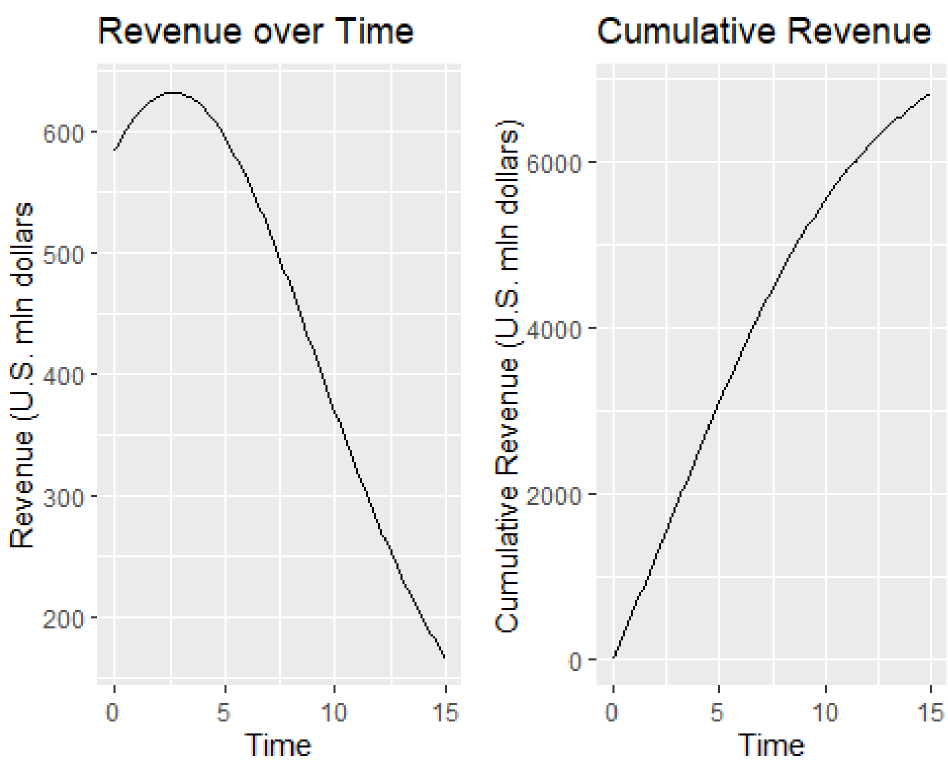


From visualizations we also see, that the predictions in average are more fitted to actual sales. Due to points above, let's continue our work with last approach's resulted parameters.

From research, I found out, that Nike actually dominates the global sneaker market with a 65.9% market share, while Adidas has a 14.7% market share. So we will change the market (m) for making predictions of the diffusion of the innovation.

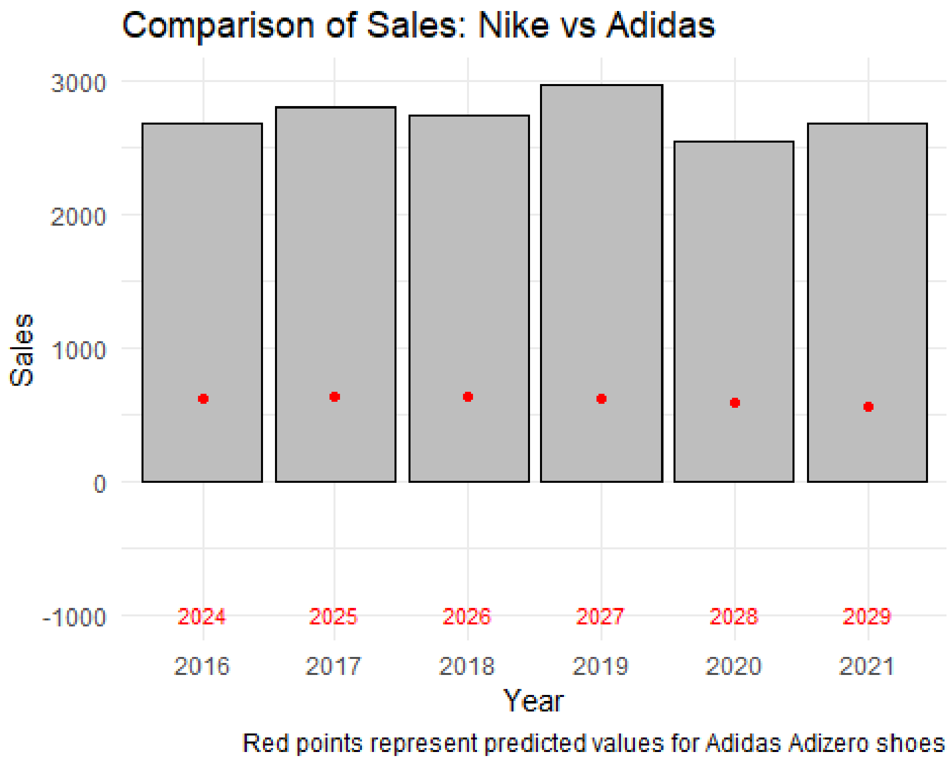
```
M = 14.7/65.9 * m

bass_f <- function(t, p, q, M) {
  ((p + q)^2 / p) * exp(-(p + q) * t) / (1 + (q / p) * exp(-(p + q) * t))^2 * M
}

bass_F <- function(t, p, q, M) {
  (1 - exp(-(p + q) * t)) / (1 + (q / p) * exp(-(p + q) * t)) * M
}

time_values <- seq(0, 15, by = 0.1)

f_values <- bass_f(time_values, p, q, M)
F_values <- bass_F(time_values, p, q, M)

data_f <- data.frame(time = time_values, f = f_values)
data_F <- data.frame(time = time_values, F = F_values)

plot_f <- ggplot(data_f, aes(x = time, y = f)) +
  geom_line() +
  labs(title = "Revenue over Time", x = "Time", y = "Revenue (U.S. mln dollars")
```

```
plot_F <- ggplot(data_F, aes(x = time, y = F)) +
  geom_line() +
  labs(title = "Cumulative Revenue   over Time", x = "Time", y = "Cumulative Rev
enue (U.S. mln dollars)")

ggarrange(plot_f, plot_F)
```



Here, we can observe the probable diffusion of the innovation shoes of Adidas. Through first graph, we can observe, that the starting revenue will be already high and after 2-3 years will reach its peak, and shrink over time.

```
running_df$pred_Year <- seq(2024, 2029)

running_df$pred_sales <- bass.f(1:6, p, q) * M

ggplot(data = running_df, aes(x = Year, y = Sales)) +
  geom_bar(stat = 'identity', fill = 'gray', color = 'black') +
  geom_point(mapping = aes(x = Year, y = pred_sales), color = 'red') +
  labs(title = 'Comparison of Sales: Nike vs Adidas',
       x = 'Year',
       y = 'Sales',
       caption = 'Red points represent predicted values for Adidas Adizero shoes
') +
  theme_minimal() +
  annotate("text", x = running_df$Year, y = rep(-1000, nrow(running_df)),
           label = running_df$pred_Year, color = "red", size = 3, hjust = 0.5)
```



Through this graph, we can see, how different are revenue values after start of production of Adidas shoes (red points) and Nike shoes (gray bars). The start of production of Nike Vaporfly was in 2016, so the above graph is useful in its interpretation, as we find also how the revenues could be for Adidas Adizero shoes for same period, as we have statistics about Adidas. Our estimated market potential for Adidas Adizero shoes is 7681 mln U.S. dollars.

For final part, we will try to estimate the potential market share of our product. From research, I found out, that in 2020, Adidas, in the global athletic footwear market accounted for a 15.1% market share. Also, Adidas has a 14.7% market share in global sneaker market. So, we can say that approximately 15% of market share is reached by Adidas. From other source, I found , that The Athletic Footwear Market size is estimated at USD 116.82 billion in 2024, and is expected to reach USD 146.48 billion by 2029. Hence, approximately, in the start, the product will have potential market 116.82* 15.1/100 = 17.64 billion U.S. dollars for 2024 period and for 2029 period it will be calculated in same way approximately 22.12 billion U.S. dollars. If we assume, that for the next 5 years, no structural breaks will occur, and the trend will be linear, we can say, that each year from 2024 the market share will be increased by (22.12 - 17.64)/5 = 0.9 billion dollars

```
innovators_percentage <- 0.025
early_adopters_percentage <- 0.135
early_majority_percentage <- 0.34
late_majority_percentage <- 0.34
laggards_percentage <- 0.16

#Estimated Market share for 2024
m <- 17.64

innovators_share <- m * innovators_percentage
```

```r
early_adopters_share <- m * early_adopters_percentage
early_majority_share <- m * early_majority_percentage
late_majority_share <- m * late_majority_percentage
laggards_share <- m * laggards_percentage

result_df <- data.frame(
  Innovators_Share = innovators_share,
  Early_Adopters_Share = early_adopters_share,
  Early_Majority_Share = early_majority_share,
  Late_Majority_Share = late_majority_share,
  Laggards_Share = laggards_share
)

#Market shares in billions for each category
print(result_df)
```

```
##   Innovators_Share Early_Adopters_Share Early_Majority_Share
## 1            0.441               2.3814               5.9976
##   Late_Majority_Share Laggards_Share
## 1              5.9976         2.8224
```

Sources: https://time.com/collection/best-inventions-2023/6324416/adidas-adizero-adios-pro-evo-1/ https://www.statista.com/statistics/240975/athletic-footwwear-wholesale-sales-in-the-us/ https://www.nytimes.com/interactive/2019/12/13/upshot/nike-vaporfly-next-percent-shoe-estimates.html https://runrepeat.com/nike-shoes-statistics https://gitnux.org/nike-vs-adidas-statistics/#:~:text=Nike’s%20global%20brand%20value%20in,has%20a%2014.7%25%20market%20share. https://www.mordorintelligence.com/industry-reports/athletic-footwear-market