NYT_API

Pricilla

2025-10-12

Get Book data from NYT API for the last 4 months. Does a high rank translate to a longer time on the best selling list?

```
#  Read API key from GitHub file
api_key_url <- "https://raw.githubusercontent.com/prnakyazze94/Data_607/refs/heads/main/NYT_API"
api_key <- readLines(api_key_url, warn = FALSE)
api_key <- trimws(api_key[1])  # ensure clean string, remove any spaces/newlines

#  NYT Books API setup
base_url <- "https://api.nytimes.com/svc/books/v3/lists"
list_name <- "hardcover-fiction"  # You can change this to another NYT list name

#  Function to fetch books for a given date
get_books_for_date <- function(date) {
  url <- paste0(base_url, "/", date, "/", list_name, ".json")
  res <- GET(url, query = list("api-key" = api_key))

  if (status_code(res) != 200) {
    warning(paste("Failed for date:", date))
    return(NULL)
  }

  data <- fromJSON(content(res, as = "text"), flatten = TRUE)
  books <- data$results$books
  if (is.null(books)) return(NULL)

  df <- data.frame(
    title = books$title,
    author = books$author,
    publisher = books$publisher,
    rank = books$rank,
    weeks_on_list = books$weeks_on_list,
    stringsAsFactors = FALSE
  )

  return(df)
}

# --- Get top books for multiple dates this year
current_year <- year(Sys.Date())
dates <- seq.Date(from = Sys.Date(), to = Sys.Date() - 120, by = "-7 days")  # last 4 months
dates <- format(dates, "%Y-%m-%d")
```

```r
all_books <- lapply(dates, get_books_for_date)
```

```
## Warning in FUN(X[[i]], ...): Failed for date: 2025-09-11

## Warning in FUN(X[[i]], ...): Failed for date: 2025-09-04

## Warning in FUN(X[[i]], ...): Failed for date: 2025-08-28

## Warning in FUN(X[[i]], ...): Failed for date: 2025-08-21

## Warning in FUN(X[[i]], ...): Failed for date: 2025-08-14

## Warning in FUN(X[[i]], ...): Failed for date: 2025-08-07

## Warning in FUN(X[[i]], ...): Failed for date: 2025-07-31

## Warning in FUN(X[[i]], ...): Failed for date: 2025-07-24

## Warning in FUN(X[[i]], ...): Failed for date: 2025-07-17

## Warning in FUN(X[[i]], ...): Failed for date: 2025-07-10

## Warning in FUN(X[[i]], ...): Failed for date: 2025-07-03

## Warning in FUN(X[[i]], ...): Failed for date: 2025-06-26

## Warning in FUN(X[[i]], ...): Failed for date: 2025-06-19
```

```r
all_books <- bind_rows(all_books)

# --- Keep top 40 unique books
top_books <- all_books %>%
  distinct(title, .keep_all = TRUE) %>%
  arrange(rank) %>%
  head(40)

# --- Print results
print(top_books)
```

```
##                                title                                author
## 1                           ALCHEMISED                              SenLinYu
## 2                THE SECRET OF SECRETS                            Dan Brown
## 3                   THE HALLMARKED MAN                     Robert Galbraith
## 4                       TOURIST SEASON                        Brynne Weaver
## 5                     THE LAST LETTER                       Rebecca Yarros
## 6         THE PRIMAL OF BLOOD AND BONE                 Jennifer L. Armentrout
## 7            AMONG THE BURNING FLOWERS                      Samantha Shannon
## 8                           CLOWN TOWN                          Mick Herron
## 9                     ONE DARK WINDOW                        Rachel Gillig
```

```
## 10                  LOVER FORBIDDEN                            J.R. Ward
## 11               TWO TWISTED CROWNS                          Rachel Gillig
## 12          ASSISTANT TO THE VILLAIN                  Hannah Nicole Maehrer
## 13              THIS INEVITABLE RUIN                          Matt Dinniman
## 14     CLIVE CUSSLER: THE IRON STORM                           Jack Du Brul
## 15                   APOSTLE'S COVE                   William Kent Krueger
## 16                   CIRCLE OF DAYS                             Ken Follett
## 17                       ONYX STORM                         Rebecca Yarros
## 18                   THE ACADEMY Elin Hilderbrand and Shelby Cunningham
## 19                      ATMOSPHERE                    Taylor Jenkins Reid
## 20                ON WINGS OF BLOOD                          Briar Boleyn
## 21                       ANATHEMA                             Keri Lake
## 22                      MY FRIENDS                       Fredrik Backman
## 23                   WILD REVERENCE                          Rebecca Ross
## 24                     NEVER FLINCH                         Stephen King
## 25                WHAT WE CAN KNOW                           Ian McEwan
## 26                   BROKEN COUNTRY                    Clare Leslie Hall
## 27            BILLION-DOLLAR RANSOM James Patterson and Duane Swierczynski
## 28     TOM CLANCY: TERMINAL VELOCITY                       M.P. Woodward
## 29          GREAT BIG BEAUTIFUL LIFE                         Emily Henry
## 30 THE LONELINESS OF SONIA AND SUNNY                         Kiran Desai
## 31                 THE CORRESPONDENT                       Virginia Evans
## 32                       PLAY NICE                      Rachel Harrison
## 33                    FORGET ME NOT                    Stacy Willingham
## 34                       KATABASIS                           R.F. Kuang
## 35                THE COLOR OF DEATH    Trey Gowdy with Christopher Greyson
## 36            THE PUMPKIN SPICE CAFÉ                       Laurie Gilmore
## 37                THE WEDDING PEOPLE                       Alison Espach
## 38                           BUCKEYE                        Patrick Ryan
## 39                  FRAMED IN DEATH                            J.D. Robb
## 40              DUNGEON CRAWLER CARL                       Matt Dinniman
##           publisher rank weeks_on_list
## 1           Del Rey    1             1
## 2         Doubleday    2             3
## 3        Mulholland    2             1
## 4          Slowburn    3             1
## 5             Amara    3             1
## 6          Blue Box    4             1
## 7        Bloomsbury    4             1
## 8        Soho Crime    4             1
## 9             Orbit    5             1
## 10          Gallery    5             1
## 11            Orbit    6             1
## 12       Red Tower     6             1
## 13              Ace    7             1
## 14          Putnam     7             1
## 15            Atria    7             1
## 16    Grand Central    8             1
## 17        Red Tower    8            35
## 18    Little, Brown    9             2
## 19       Ballantine    9            16
## 20             MIRA    9             2
## 21            Bloom   10             1
## 22            Atria   10            20
```

```
## 23         Saturday   10            2
## 24         Scribner   10           13
## 25            Knopf   11            1
## 26 Simon & Schuster   11           26
## 27    Little, Brown   11            2
## 28           Putnam   11            1
## 29          Berkley   11           17
## 30          Hogarth   12            1
## 31            Crown   12            1
## 32          Berkley   12            1
## 33         Minotaur   12            1
## 34   Harper Voyager   13            5
## 35         Fox News   13            3
## 36    HarperCollins   13            4
## 37             Holt   13           39
## 38     Random House   14            4
## 39     St. Martin's   14            3
## 40              Ace   14            4
```

```
# Save to CSV
write.csv(top_books, "nyt_top40_books.csv", row.names = FALSE)
```

```
df_selected <- top_books %>%
  select(title, author, publisher, rank, weeks_on_list)
# Display in a clean table
kable(df_selected, caption = "Top NYT Books - Selected Columns")
```

Table 1: Top NYT Books - Selected Columns

title	author	publisher	rank	weeks_on_list
ALCHEMISED	SenLinYu	Del Rey	1	1
THE SECRET OF SECRETS	Dan Brown	Doubleday	2	3
THE HALLMARKED MAN	Robert Galbraith	Mulholland	2	1
TOURIST SEASON	Brynne Weaver	Slowburn	3	1
THE LAST LETTER	Rebecca Yarros	Amara	3	1
THE PRIMAL OF BLOOD AND BONE	Jennifer L. Armentrout	Blue Box	4	1
AMONG THE BURNING FLOWERS	Samantha Shannon	Bloomsbury	4	1
CLOWN TOWN	Mick Herron	Soho Crime	4	1
ONE DARK WINDOW	Rachel Gillig	Orbit	5	1
LOVER FORBIDDEN	J.R. Ward	Gallery	5	1
TWO TWISTED CROWNS	Rachel Gillig	Orbit	6	1
ASSISTANT TO THE VILLAIN	Hannah Nicole Maehrer	Red Tower	6	1
THIS INEVITABLE RUIN	Matt Dinniman	Ace	7	1
CLIVE CUSSLER: THE IRON STORM	Jack Du Brul	Putnam	7	1
APOSTLE'S COVE	William Kent Krueger	Atria	7	1
CIRCLE OF DAYS	Ken Follett	Grand Central	8	1
ONYX STORM	Rebecca Yarros	Red Tower	8	35

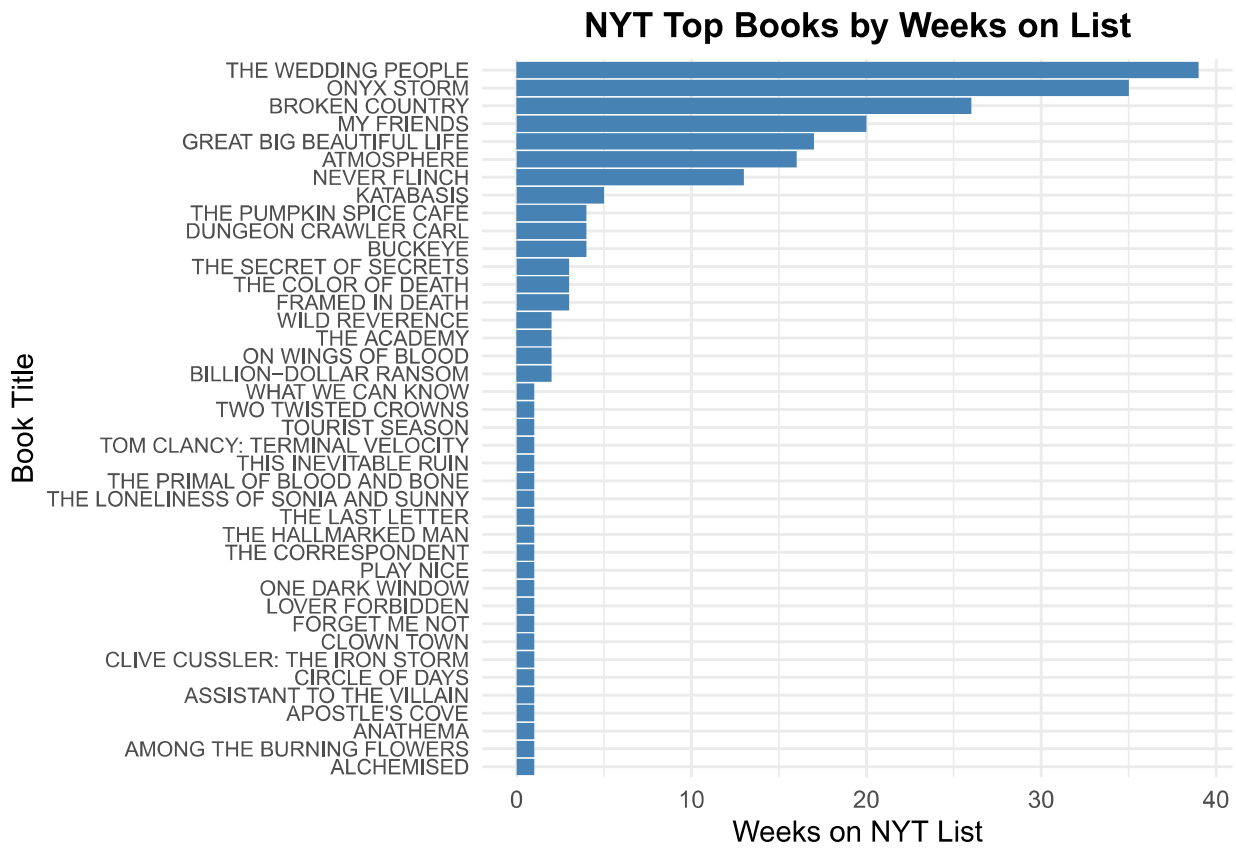
title	author	publisher	rank	weeks_on_list
THE ACADEMY	Elin Hilderbrand and Shelby Cunningham	Little, Brown	9	2
ATMOSPHERE	Taylor Jenkins Reid	Ballantine	9	16
ON WINGS OF BLOOD	Briar Boleyn	MIRA	9	2
ANATHEMA	Keri Lake	Bloom	10	1
MY FRIENDS	Fredrik Backman	Atria	10	20
WILD REVERENCE	Rebecca Ross	Saturday	10	2
NEVER FLINCH	Stephen King	Scribner	10	13
WHAT WE CAN KNOW	Ian McEwan	Knopf	11	1
BROKEN COUNTRY	Clare Leslie Hall	Simon & Schuster	11	26
BILLION-DOLLAR RANSOM	James Patterson and Duane Swierczynski	Little, Brown	11	2
TOM CLANCY: TERMINAL VELOCITY	M.P. Woodward	Putnam	11	1
GREAT BIG BEAUTIFUL LIFE	Emily Henry	Berkley	11	17
THE LONELINESS OF SONIA AND SUNNY	Kiran Desai	Hogarth	12	1
THE CORRESPONDENT	Virginia Evans	Crown	12	1
PLAY NICE	Rachel Harrison	Berkley	12	1
FORGET ME NOT	Stacy Willingham	Minotaur	12	1
KATABASIS	R.F. Kuang	Harper Voyager	13	5
THE COLOR OF DEATH	Trey Gowdy with Christopher Greyson	Fox News	13	3
THE PUMPKIN SPICE CAFÉ	Laurie Gilmore	HarperCollins	13	4
THE WEDDING PEOPLE	Alison Espach	Holt	13	39
BUCKEYE	Patrick Ryan	Random House	14	4
FRAMED IN DEATH	J.D. Robb	St. Martin's	14	3
DUNGEON CRAWLER CARL	Matt Dinniman	Ace	14	4

Plot NYT Top Books by Weeks on List

```
# Arrange books by descending weeks_on_list
plot_data <- df_selected %>%
  arrange(desc(weeks_on_list))

# Create horizontal bar plot
ggplot(plot_data, aes(x = reorder(title, weeks_on_list), y = weeks_on_list)) +
  geom_col(fill = "steelblue") +
  coord_flip() +
  labs(
    title = "NYT Top Books by Weeks on List",
    x = "Book Title",
    y = "Weeks on NYT List"
  ) +
  theme_minimal() +
  theme(
```

```
    plot.title = element_text(hjust = 0.5, face = "bold"),
    axis.text.y = element_text(size = 8)
  )
```



The Wedding people remained top sellers for over 35 weeks on the NYT best selling list. While almost 18 titles out of 40 were only on the list for less than 4 weeks.

```
# from top_books data frame get top 15 `top_books`
top_15_books <- top_books %>%
  arrange(rank) %>%                # sort by rank ascending
  select(rank, title) %>%          # keep only rank and title
  head(15)                         # get top 15

# View results
print(top_15_books)
```

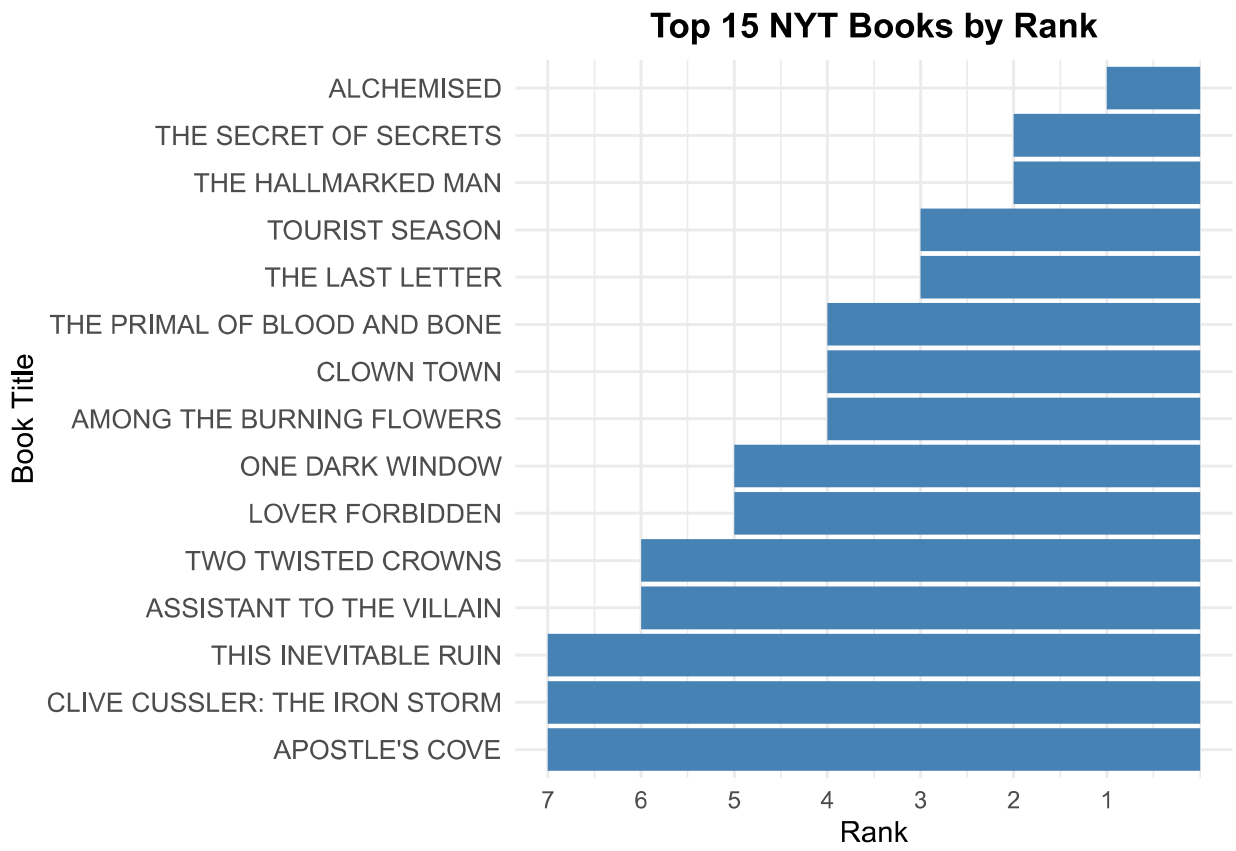
```
##    rank                          title
## 1     1                     ALCHEMISED
## 2     2          THE SECRET OF SECRETS
## 3     2             THE HALLMARKED MAN
## 4     3                 TOURIST SEASON
## 5     3               THE LAST LETTER
## 6     4  THE PRIMAL OF BLOOD AND BONE
## 7     4     AMONG THE BURNING FLOWERS
## 8     4                     CLOWN TOWN
## 9     5               ONE DARK WINDOW
```

```
## 10    5               LOVER FORBIDDEN
## 11    6            TWO TWISTED CROWNS
## 12    6      ASSISTANT TO THE VILLAIN
## 13    7          THIS INEVITABLE RUIN
## 14    7 CLIVE CUSSLER: THE IRON STORM
## 15    7                APOSTLE'S COVE
```

Top 15 books by rank over the last 4 months.

```r
# Prepare top 15 books
top_15_books <- top_books %>%
  arrange(rank) %>%
  select(rank, title) %>%
  head(15)

# Horizontal bar plot
ggplot(top_15_books, aes(x = reorder(title, -rank), y = rank)) +
  geom_col(fill = "steelblue") +
  coord_flip() +  # horizontal bars
  scale_y_reverse(breaks = 1:15) + # reverse y-axis so rank 1 is at top
  labs(
    title = "Top 15 NYT Books by Rank",
    x = "Book Title",
    y = "Rank"
  ) +
  theme_minimal() +
  theme(
    plot.title = element_text(hjust = 0.5, face = "bold"),
    axis.text.y = element_text(size = 10)
  )
```

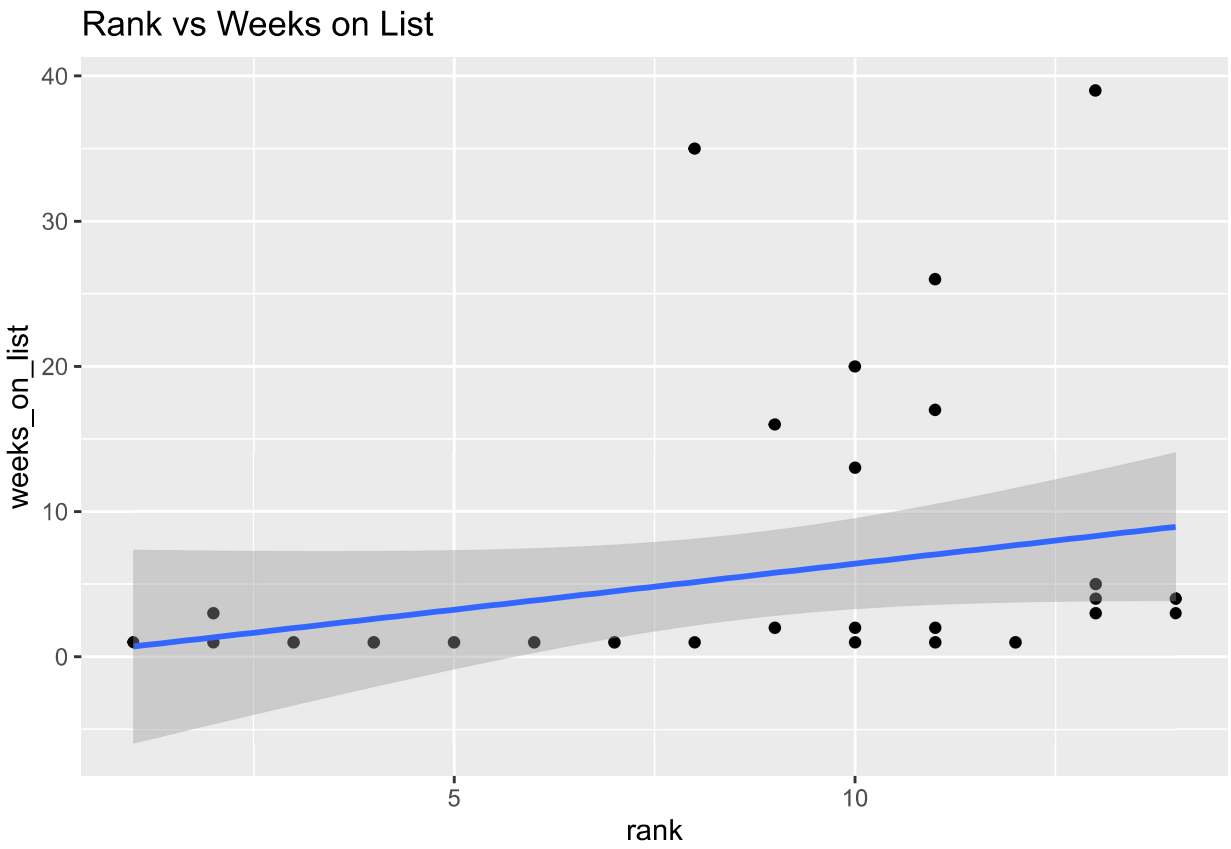


Relationship Between Rank and Weeks on List

Do books that debut at 1 stay longer on the list?

```
ggplot(top_books, aes(x = rank, y = weeks_on_list)) +
  geom_point() +
  geom_smooth(method = "lm") +
  labs(title = "Rank vs Weeks on List")
```

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



A single data point is showing a specific rank and the corresponding number of weeks it spent on a list.

The data is quite scattered, meaning there isn't a perfect or very strong relationship.

Noticeable points exist far from the main cluster, especially those with high "Weeks on List" (For example one point near Rank 12 is close to 40 Weeks on List).

Trend Line (Blue Line) is a linear regression line (or a similar smoothing line) that attempts to summarize the general trend in the data.

The line has a positive slope, indicating a positive correlation between Rank and Weeks on List.

In the context of typical rankings (where Rank 1 is the best), this positive slope suggests that as the Rank gets worse (increases from 1 to 12), the Weeks on List tends to increase.

This is an unusual finding for a typical ranking where higher rank numbers should usually mean less time on a list. It suggests that items with lower (worse) ranks tend to stay on the list longer, or perhaps the 'Rank' variable is measuring something where a higher number is actually better, or the relationship is non-linear and not well-captured by the straight line.

In summary just going by the data as it is, the image graphically explores the relationship between a list Rank and the number of Weeks on List, suggesting a weak positive trend, but with significant variability in the data.

Count how many titles spent how long on best selling list.

```
# Count how many titles per weeks_on_list
cluster_counts <- top_books %>%
  group_by(weeks_on_list) %>%
  summarise(num_titles = n())
```

```r
# View results
print(cluster_counts)
```

```
## # A tibble: 12 x 2
##    weeks_on_list num_titles
##            <int>      <int>
##  1             1         22
##  2             2          4
##  3             3          3
##  4             4          3
##  5             5          1
##  6            13          1
##  7            16          1
##  8            17          1
##  9            20          1
## 10            26          1
## 11            35          1
## 12            39          1
```

Most of the titles where on the best selling list for 1 to 3 weeks.

Average time on the Top selling weeks list.

```r
# Ensure weeks_on_list is numeric
top_books <- top_books %>%
  mutate(weeks_on_list = as.numeric(as.character(weeks_on_list)))

# Calculate average weeks on list
average_weeks <- mean(top_books$weeks_on_list, na.rm = TRUE)

# Print result
print(average_weeks)
```

```
## [1] 5.55
```

Summary of Weeks on List

The analysis of the weeks_on_list data shows that most books are relatively new to the bestseller list. Out of all titles, 22 books have appeared on the list for only one week, indicating frequent turnover and strong competition among recent releases. A smaller number of books 4 titles for 2 weeks, and 3 titles each for 3 and 4 weeks demonstrate moderate staying power. Only a handful of books have shown long-term popularity, remaining on the list for 13, 16, 17, 20, and 26 weeks respectively. Overall, this suggests that while the bestseller list changes often, a few standout titles maintain consistent reader interest over time. The Average time a book stays on the best seller List is 5.55 weeks.