RWorksheet5_(Celestra, Caneso, Arenal)

2024-11-06

[EXTRACTING TV SHOWS REVIEWS]

1.

Libraries

```
library(rvest)
library(httr)
library(polite)
```

```
## Warning: package 'polite' was built under R version 4.4.2
```

```
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
```

```
library(stringr)
library(knitr)
```

```
## Warning: package 'knitr' was built under R version 4.4.2
```

IMDB URL

```
polite::use_manners(save_as = 'polite_scrape.R')
```

```
## v Setting active project to "C:/Users/canes/OneDrive/Documents/R_Analytics".
```

```
url <- "https://www.imdb.com/chart/toptv/?ref_=nv_tvv_250"

session <- bow(url,
user_agent = "Educational")
session
```

```
## <polite session> https://www.imdb.com/chart/toptv/?ref_=nv_tvv_250
##     User-agent: Educational
##     robots.txt: 35 rules are defined for 3 bots
##    Crawl delay: 5 sec
##   The path is scrapable for this user-agent
```

Title List

```
title_list <- scrape(session) %>%
html_nodes('h3.ipc-title__text') %>%
html_text

title_list_sub <- as.data.frame(title_list[2:26])
title_list_sub
```

```
##                       title_list[2:26]
## 1                      1. Breaking Bad
## 2                   2. Planet Earth II
## 3                      3. Planet Earth
## 4                  4. Band of Brothers
## 5                         5. Chernobyl
## 6                          6. The Wire
## 7        7. Avatar: The Last Airbender
## 8                     8. Blue Planet II
## 9                       9. The Sopranos
## 10     10. Cosmos: A Spacetime Odyssey
## 11                           11. Cosmos
## 12                       12. Our Planet
## 13                 13. Game of Thrones
## 14                              14. Bluey
## 15                15. The World at War
## 16 16. Fullmetal Alchemist Brotherhood
## 17                   17. Rick and Morty
## 18                              18. Life
## 19                   19. The Last Dance
## 20                20. The Twilight Zone
## 21                  21. The Vietnam War
## 22                         22. Sherlock
## 23                  23. Attack on Titan
## 24     24. Batman: The Animated Series
## 25                           25. Arcane
```

Split

```
colnames(title_list_sub) <- "ranks"

split_df <- strsplit(as.character(title_list_sub$ranks),".",fixed = T)
split_df <- data.frame(do.call(rbind,split_df))

colnames(split_df) <- c("ranks", "titles")
split_df
```

```
##    ranks                          titles
```

```
## 1      1                     Breaking Bad
## 2      2                  Planet Earth II
## 3      3                     Planet Earth
## 4      4                 Band of Brothers
## 5      5                       Chernobyl
## 6      6                       The Wire
## 7      7       Avatar: The Last Airbender
## 8      8                    Blue Planet II
## 9      9                     The Sopranos
## 10    10      Cosmos: A Spacetime Odyssey
## 11    11                           Cosmos
## 12    12                       Our Planet
## 13    13                  Game of Thrones
## 14    14                            Bluey
## 15    15                The World at War
## 16    16  Fullmetal Alchemist Brotherhood
## 17    17                    Rick and Morty
## 18    18                              Life
## 19    19                   The Last Dance
## 20    20                The Twilight Zone
## 21    21                  The Vietnam War
## 22    22                         Sherlock
## 23    23                  Attack on Titan
## 24    24      Batman: The Animated Series
## 25    25                           Arcane
```

Rating

```
rating_list <- scrape(session) %>%
html_nodes('span.ipc-rating-star--rating') %>%
html_text
```

Number of People who voted

```
people_list <- scrape(session) %>%
html_nodes('span.ipc-rating-star--voteCount') %>%
html_text
```

Episode and Year it was released

```
episode_list <- scrape(session) %>%
  html_nodes('span.sc-300a8231-7.eaXxft.cli-title-metadata-item:nth-of-type(2)') %>%
  html_text()


year_list <- scrape(session) %>%
  html_nodes('span.sc-300a8231-7.eaXxft.cli-title-metadata-item:nth-of-type(1)') %>%
  html_text()
```

User Reviews, Critic Reviews, and Popularity Rating

```r
home_link <- 'https://www.imdb.com/chart/toptv/'
main_page <- read_html(home_link)

links <- main_page %>%
  html_nodes("a.ipc-title-link-wrapper") %>%
  html_attr("href")

# Loop to get link of each show's page
show_data <- lapply(links, function(link) {
  complete_link <- paste0("https://imdb.com", link)

  #loop to get the link for user review page
  usrv_link <- read_html(complete_link)
  usrv_link_page <- usrv_link %>%
    html_nodes('a.isReview') %>%
    html_attr("href")

  #loop to get user reviews of each shows
  usrv <- read_html(paste0("https://imdb.com", usrv_link_page[1]))
  usrv_count <- usrv %>%
    html_nodes('[data-testid="tturv-total-reviews"]') %>%
    html_text()

  #loop to extract critic reviews
  critic <- usrv_link %>%
              html_nodes("span.score") %>%
              html_text()
  critic_df <- data.frame(Critic_Reviews = critic[2], stringsAsFactors = FALSE)

  #loop to extract pop rating
  pop_rating <- usrv_link %>%
              html_nodes('[data-testid="hero-rating-bar__popularity__score"]') %>%
              html_text()
  pop_rating_df <- data.frame(Popularity_Rating = pop_rating[2], stringsAsFactors = FALSE)

  return(data.frame(User_Reviews = usrv_count, Critic = critic_df, pop = pop_rating_df))
})

showss <- do.call(rbind, show_data)
```

FINAL DATA

```r
tv_shows <- data.frame(Ranking = split_df$ranks,
                       TV_Show_Title = split_df$titles,
                       Episodes = episode_list[1:25],
                       Release_Year = year_list[1:25],
                       Rating = rating_list[1:25],
                       No._of_Voter = people_list[1:25],
                       User_Review = showss$User_Reviews,
                       Popularity = showss$Popularity_Rating,
                       Critic_Reviews = showss$Critic_Reviews)
write.csv(tv_shows, "TV_Shows.csv", row.names = F)
```

```
#just for space
```

2. Top 5 TV Shows

```
urls <- c("https://www.imdb.com/title/tt0903747/reviews/?ref_=tt_urv_sm",
          "https://www.imdb.com/title/tt5491994/reviews/?ref_=tt_urv_sm",
          "https://www.imdb.com/title/tt0795176/reviews/?ref_=tt_urv_sm",
          "https://www.imdb.com/title/tt0185906/reviews/?ref_=tt_urv_sm",
          "https://www.imdb.com/title/tt7366338/reviews/?ref_=tt_urv_sm")

df <- list()

for(i in seq_along(urls)){

  sessions <- bow(urls[i], user_agent = "Educational")

#User name
userName <- scrape(sessions) %>%
html_nodes('a.ipc-link.ipc-link--base') %>%
html_text() %>%
  head(20)

#Date reviewed
dateReview <- scrape(sessions) %>%
html_nodes('li.ipc-inline-list__item.review-date') %>%
html_text() %>%
  head(20)

#User rating
userRating <- scrape(sessions) %>%
html_nodes('span.ipc-rating-star--rating') %>%
html_text() %>%
  head(20)

#Title of the review
reviewTitle <- scrape(sessions) %>%
html_nodes('h3.ipc-title__text') %>%
html_text() %>%
  head(20)

#Helpful
helpful <- scrape(sessions) %>%
html_nodes('span.ipc-voting__label') %>%
html_text()

#Unhelpful
unhelpful <- scrape(sessions) %>%
html_nodes('span.ipc-voting__label__count.ipc-voting__label__count--down') %>%
html_text()

#Data frame for the user reviews.
userReviews <- data.frame(user_Name = userName[1:20],
                          Date_Reviewed = dateReview[1:20],
```

```
                          user_Rating = userRating[1:20],
                          Review_Title = reviewTitle[1:20],
                          Helpful = helpful[1:20],
                          Unhelpful = unhelpful[1:20],
                          stringsAsFactors = F)

df[[i]] <- userReviews
}

df[[1]]
```

```
##             user_Name Date_Reviewed user_Rating
## 1             FiRE010   Jul 3, 2021          10
## 2           Permalink   Mar 6, 2019          10
## 3          bruhperson  Jul 29, 2021          10
## 4           Permalink  Feb 18, 2020          10
## 5        KinoKoopaKid   Nov 8, 2021          10
## 6           Permalink  May 30, 2019          10
## 7         jehuschultz   Dec 8, 2022          10
## 8           Permalink  Nov 15, 2019          10
## 9      Supermanfan-13  Jul 17, 2021          10
## 10          Permalink   Dec 8, 2022          10
## 11  manishsingh-03299  Sep 28, 2024           7
## 12          Permalink  Feb 11, 2014           4
## 13            Rob1331  Dec 13, 2017          10
## 14          Permalink  Nov 12, 2017          10
## 15           xpinerhd   Nov 8, 2021          10
## 16          Permalink  Jan 11, 2014          10
## 17 dhanushreddy-14919  Aug 11, 2021          10
## 18          Permalink  May 19, 2019          10
## 19        tushv-31482   May 4, 2021          10
## 20          Permalink  Jun 23, 2021          10
##
## 1
## 2
## 3
## 4
## 5
## 6                                                                                                  Those days a
## 7
## 8                                                                                                            I
## 9                                                                                                            A
## 10                                                                                                        Once
## 11                                                                                                My Review For
## 12                                                                                               Fantastic show
## 13                                                                                               Most Overrated
## 14                                                                                 Among the best and most ac
## 15                                                                                       By far the greatest
## 16 If you mix Scarface, Robin Hood and maybe Tyler Durden with enough meth - you'll get a mean cockta
## 17                                                                                                  in a ca
## 18                                                                  Since GOT is over, this is Officially the
## 19                                                                                                Every bit a
## 20
```

```
##    Helpful Unhelpful
## 1     <NA>      <NA>
## 2     <NA>      <NA>
## 3     <NA>      <NA>
## 4     <NA>      <NA>
## 5     <NA>      <NA>
## 6     <NA>      <NA>
## 7     <NA>      <NA>
## 8     <NA>      <NA>
## 9     <NA>      <NA>
## 10    <NA>      <NA>
## 11    <NA>      <NA>
## 12    <NA>      <NA>
## 13    <NA>      <NA>
## 14    <NA>      <NA>
## 15    <NA>      <NA>
## 16    <NA>      <NA>
## 17    <NA>      <NA>
## 18    <NA>      <NA>
## 19    <NA>      <NA>
## 20    <NA>      <NA>
```

```
df[[2]]
```

```
##             user_Name Date_Reviewed user_Rating
## 1        arjanhylkema   Nov 7, 2016          10
## 2           Permalink   Nov 5, 2016          10
## 3            Wentloog   Nov 5, 2016          10
## 4           Permalink   Nov 9, 2016          10
## 5       john-m-madsen   Nov 5, 2016          10
## 6           Permalink   Nov 8, 2016          10
## 7        thespookybuz  Nov 17, 2016          10
## 8           Permalink  Nov 13, 2016          10
## 9         pjdickinson   Nov 6, 2016          10
## 10          Permalink  Dec 31, 2016          10
## 11           dbijis33  Nov 19, 2016          10
## 12          Permalink  Dec 28, 2016           7
## 13     dhanrajjughead  May 19, 2019          10
## 14          Permalink  Oct 20, 2018          10
## 15        NeilBarnett  Sep 29, 2017          10
## 16          Permalink  Nov 22, 2016          10
## 17      salmanu-27386  Oct 12, 2017          10
## 18          Permalink   Dec 4, 2016          10
## 19 panagiotiskatsanos  Apr 23, 2020          10
## 20          Permalink   Jan 5, 2017          10
##
## 1
## 2                                                                                  At once awe-inspiring a
## 3                                                        Yet another masterpiece from BBC Nature & Davi
## 4
## 5
## 6                                                                                               Dangero
## 7                                                                                  Greatest documentary
## 8                                                                          Best thing on TV since las
```

```
## 9
## 10                                                                             One of the best documentaries I
## 11                                                                                        In times of climate
## 12
## 13                                                                    More irritated with IMDb for the bias than
## 14
## 15                                                                                      Should be required view
## 16                                                                                      What a Beautiful Planet
## 17 Like the first 'Planet Earth', does for nature and our planet as 'Walking with Dinosaurs' did with
## 18                                                                                    This masterpiece deserves
## 19                                                                                                    Absol
## 20                                                                             Peerless evocation of nature ac
##    Helpful Unhelpful
## 1     <NA>      <NA>
## 2     <NA>      <NA>
## 3     <NA>      <NA>
## 4     <NA>      <NA>
## 5     <NA>      <NA>
## 6     <NA>      <NA>
## 7     <NA>      <NA>
## 8     <NA>      <NA>
## 9     <NA>      <NA>
## 10    <NA>      <NA>
## 11    <NA>      <NA>
## 12    <NA>      <NA>
## 13    <NA>      <NA>
## 14    <NA>      <NA>
## 15    <NA>      <NA>
## 16    <NA>      <NA>
## 17    <NA>      <NA>
## 18    <NA>      <NA>
## 19    <NA>      <NA>
## 20    <NA>      <NA>
```

```r
df[[3]]
```

```
##          user_Name Date_Reviewed user_Rating
## 1     robert-kamer   Feb 8, 2007          10
## 2        Permalink  Nov 19, 2008          10
## 3         jim-1409   Jan 4, 2009          10
## 4        Permalink  Dec 15, 2006          10
## 5  ccthemovieman-1   Sep 1, 2007          10
## 6        Permalink  Aug 27, 2006          10
## 7         cmcoveos  Apr 30, 2006          10
## 8        Permalink  Jun 29, 2015           9
## 9         Loordssm  Jul 20, 2006          10
## 10       Permalink  Jan 28, 2009          10
## 11        ultimorn   Jun 1, 2015           7
## 12       Permalink   Oct 8, 2020           3
## 13     bob the moo   Dec 4, 2007          10
## 14       Permalink  Jan 15, 2007          10
## 15           alfeu  Jul 30, 2008          10
## 16       Permalink  Dec 25, 2017          10
## 17         Cabrone  Sep 14, 2009           9
```

```
## 18       Permalink  Jul 27, 2014          10
## 19        berndt65   Jan 4, 2023           9
## 20       Permalink  Sep 20, 2020          10
##                                                                  Review_Title
## 1                                                                 11 out of 10
## 2                                           A masterpiece of a documentary
## 3                                                      In A Word: Amazing
## 4      The most amazing achievement in natural history TV has ever given
## 5                                                    Simply put, stunning
## 6                             An amazing trip around our beautiful planet.
## 7  A visually impressive and memorable look at the world that we live in
## 8                                      Is it real? I mean, actual footagge?
## 9                                                            Beautiful
## 10                                          Are you kidding me people?
## 11                                    It doesn't get any better than this.
## 12                                          Only 4 Eps can touch my soul!
## 13                             Should be called "BBC - Yeah, animals suck"
## 14                                     Brilliant Documentary Series
## 15                              Explanation to those low-rating reviews...
## 16                                                   Truly Astonishing
## 17                                        The Greatest Series Ever
## 18                                         Absolutely Mindblowing!
## 19                                       Excellent documentary series
## 20                                Words fail me to describe such greatness
##    Helpful Unhelpful
## 1     <NA>      <NA>
## 2     <NA>      <NA>
## 3     <NA>      <NA>
## 4     <NA>      <NA>
## 5     <NA>      <NA>
## 6     <NA>      <NA>
## 7     <NA>      <NA>
## 8     <NA>      <NA>
## 9     <NA>      <NA>
## 10    <NA>      <NA>
## 11    <NA>      <NA>
## 12    <NA>      <NA>
## 13    <NA>      <NA>
## 14    <NA>      <NA>
## 15    <NA>      <NA>
## 16    <NA>      <NA>
## 17    <NA>      <NA>
## 18    <NA>      <NA>
## 19    <NA>      <NA>
## 20    <NA>      <NA>
```

```
df[[4]]
```

```
##             user_Name Date_Reviewed user_Rating
## 1             Rob1331  Sep 27, 2022          10
## 2            Permalink  Oct 14, 2001          10
## 3        sanderson777  Jan 18, 2002          10
## 4            Permalink  Apr 18, 2004          10
## 5         wildcatt268  Feb 13, 2003          10
```

```
## 6           Permalink  Jan 23, 2005          10
## 7             arjay24  Sep 16, 2004          10
## 8           Permalink   May 6, 2022          10
## 9           rbverhoef   Nov 4, 2019          10
## 10          Permalink   Nov 5, 2001          10
## 11         yodaschoda  Aug 25, 2004          10
## 12          Permalink  May 30, 2015           7
## 13 philip_vanderveken  Apr 10, 2021           5
## 14          Permalink   May 2, 2006          10
## 15     Supermanfan-13   Jun 3, 2019          10
## 16          Permalink  Jan 26, 2005          10
## 17          thiagoutp   May 3, 2022          10
## 18          Permalink  Oct 24, 2018           9
## 19         bsmith5552   Dec 7, 2002          10
## 20          Permalink  Nov 25, 2002          10
##                                                          Review_Title Helpful
## 1                                                        Incredible!!    <NA>
## 2                            Possibly the finest 10 hours ever created    <NA>
## 3                               One of the best war movies/series ever    <NA>
## 4                                                          Realistic    <NA>
## 5                                                          Excellent    <NA>
## 6                      One of, if not the best, mini series' ever made    <NA>
## 7              This series is so unbelievably realistic, so authentic.    <NA>
## 8                            One of the best mini-series ever created!    <NA>
## 9                                              Probably the best ever    <NA>
## 10                            Realistic WWII Drama With Warts Included    <NA>
## 11                                                   war, no frills    <NA>
## 12                                             You can't beat this....    <NA>
## 13                                                       Overrated??    <NA>
## 14                                          Not very realistic at all    <NA>
## 15                  Without Doubt, the Best Mini-Series Ever Recorded    <NA>
## 16                                                 Great Miniseries    <NA>
## 17 A series like this won't be made again (see below), so treasure it    <NA>
## 18                                             Share With Your Children    <NA>
## 19                                                Best Mini series ever    <NA>
## 20                            A-1, TOPS, the BOMB what else can I say?    <NA>
##    Unhelpful
## 1       <NA>
## 2       <NA>
## 3       <NA>
## 4       <NA>
## 5       <NA>
## 6       <NA>
## 7       <NA>
## 8       <NA>
## 9       <NA>
## 10      <NA>
## 11      <NA>
## 12      <NA>
## 13      <NA>
## 14      <NA>
## 15      <NA>
## 16      <NA>
## 17      <NA>
```

```
## 18      <NA>
## 19      <NA>
## 20      <NA>
```

```
df[[5]]
```

```
##            user_Name Date_Reviewed user_Rating
## 1    curiosityonmars  May 23, 2019          10
## 2          Permalink  May 10, 2019          10
## 3           stelmakh   May 9, 2019          10
## 4          Permalink  May 14, 2019          10
## 5       natashapekar   May 7, 2019          10
## 6          Permalink  May 20, 2019          10
## 7         m-porpaczi   May 6, 2019          10
## 8          Permalink  May 13, 2019          10
## 9           Lladerat   May 6, 2019          10
## 10         Permalink  Nov 27, 2019          10
## 11          jfirebug  May 23, 2019           5
## 12         Permalink  Jun 15, 2019           8
## 13           thegldt  May 20, 2019          10
## 14         Permalink  May 30, 2019          10
## 15 alexander-phoenix   Jun 7, 2019           9
## 16         Permalink  Sep 27, 2022          10
## 17       wmeduardowm   May 6, 2019           9
## 18         Permalink  Jul 10, 2022           9
## 19    Leofwine_draca  May 26, 2019          10
## 20         Permalink  May 15, 2019           7
##                                                                  Review_Title Helpful
## 1                                                          They got it right    <NA>
## 2                                                      Goosebumps and tears     <NA>
## 3                                                I highly recommend this film!    <NA>
## 4                                             No hero wakes up wanting to die     <NA>
## 5                                                     So far looks excellent     <NA>
## 6                                                               Incredible       <NA>
## 7                                 Bleak, Unsettling, Haunting All Throughout     <NA>
## 8                                                           Unbelievable       <NA>
## 9                                                       HBO did it again!      <NA>
## 10                                                             Exemplary       <NA>
## 11                                                              Amazing!       <NA>
## 12 The movie is far from thuth. A lot of fake info to create a drama...        <NA>
## 13                                                    Emotionally drained...    <NA>
## 14                                                         Just watch it (!)    <NA>
## 15                                  Now you look like the minister of coal!     <NA>
## 16                                                              Must Watch!     <NA>
## 17                                                                Cracking.     <NA>
## 18                                                               Brilliant!     <NA>
## 19                 It is hard to overestimate the importance of this show.     <NA>
## 20                                                                Amazing      <NA>
##    Unhelpful
## 1       <NA>
## 2       <NA>
## 3       <NA>
## 4       <NA>
## 5       <NA>
```

```
## 6       <NA>
## 7       <NA>
## 8       <NA>
## 9       <NA>
## 10      <NA>
## 11      <NA>
## 12      <NA>
## 13      <NA>
## 14      <NA>
## 15      <NA>
## 16      <NA>
## 17      <NA>
## 18      <NA>
## 19      <NA>
## 20      <NA>
```

```
#just for space
```

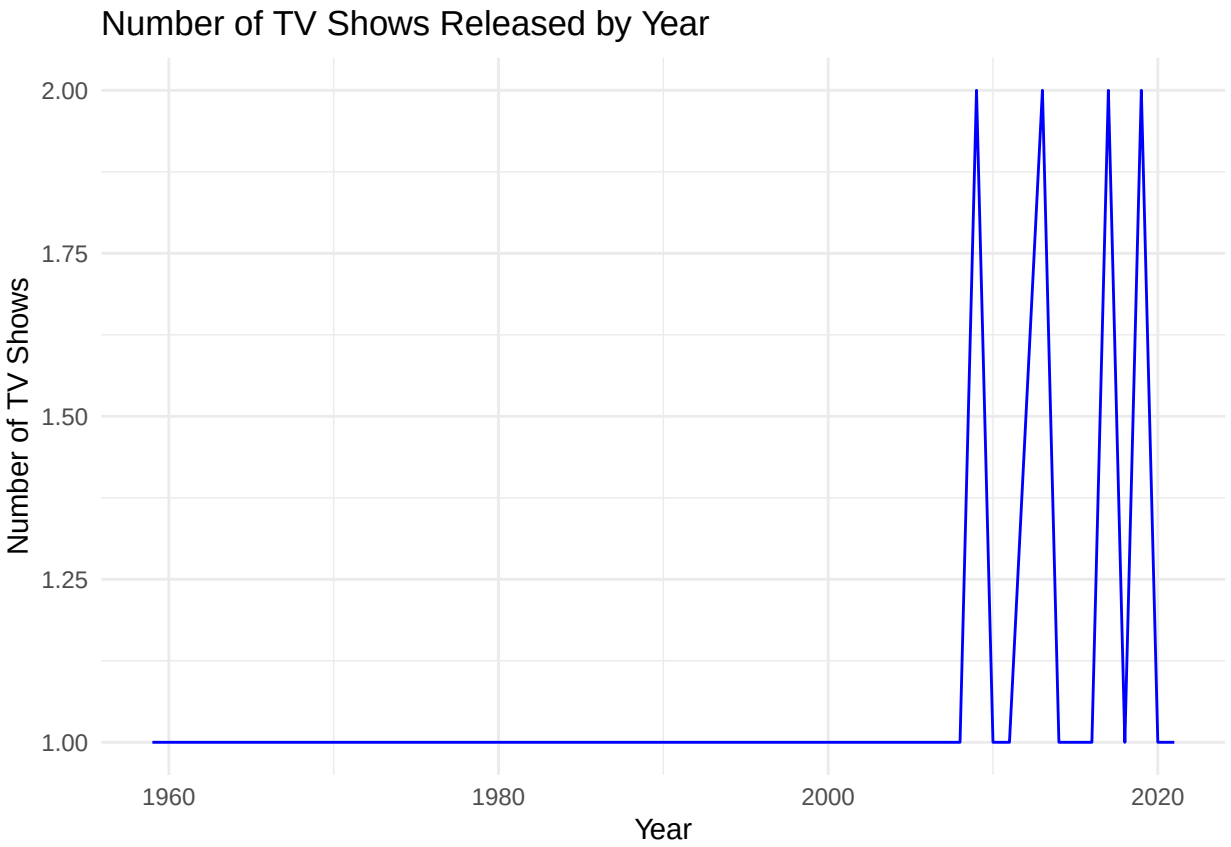
3.

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.4.2
```

```
years <- substr(year_list, 1,4)
years <- as.numeric(years)

ggplot(data.frame(Year = years), aes(x = Year)) +
  geom_line(stat = "count", fill = "skyblue", color = "blue") +
  labs(title = "Number of TV Shows Released by Year",
       x = "Year",
       y = "Number of TV Shows") +
  theme_minimal()
```

```
## Warning in geom_line(stat = "count", fill = "skyblue", color = "blue"):
## Ignoring unknown parameters: `fill`
```




```
most_shows_year <- as.data.frame(table(years))
most_shows_year <- most_shows_year[which.max(most_shows_year$Freq), ]
print(most_shows_year)
```

```
##    years Freq
## 11  2009    2
```

```
#just for space
```

4-5. [**EXTRACTING AMAZON PRODUCT REVIEWS**]

Graphics Card

```
  gc_url <- c("https://www.amazon.com/s?k=graphics+card&crid=6809YKENO5VZ&sprefix=graphics+%2Caps%2C516
              "https://www.amazon.com/s?k=graphics+card&page=2&crid=NY6LX0H811DR&qid=1732062550&sprefix=%

graphics <- list()

for(i in seq_along(gc_url)){
  Sys.sleep(2)
  gSession <- bow(gc_url[i], user_agent = "Educational")

  # Scrape data
```

```
  product_titles <- scrape(gSession) %>%
    html_nodes("span.a-text-normal") %>%
    html_text() %>%
    head(30)
  product_titles <- product_titles[!grepl("Check each product page for other buying options", product_ti

  price <- scrape(gSession) %>%
    html_nodes('.a-price .a-offscreen') %>%
    html_text() %>%
    head(30)
  price <- as.numeric(str_extract(price, "\\d+\\.\\d"))

  ratings <- scrape(gSession) %>%
    html_nodes('span.a-icon-alt') %>%
    html_text() %>%
    head(30)
 ratings <- as.numeric(str_extract(ratings, "\\d+\\.\\d"))

  reviews <- scrape(gSession) %>%
    html_nodes('.s-link-style .s-underline-text') %>%
    html_text() %>%
    head(30)

  descriptions <- scrape(gSession) %>%
    html_nodes('.a-row.a-size-base.a-color-secondary') %>%
    html_text() %>%
    head(30)

  # Find the maximum length
  max_length <- max(
    length(product_titles), length(price), length(ratings),
    length(reviews), length(descriptions)
  )

  # Standardize all vectors to the same length
  length(product_titles) <- max_length
  length(price) <- max_length
  length(ratings) <- max_length
  length(reviews) <- max_length
  length(descriptions) <- max_length

  # Create a data frame
  card <- data.frame(
    ProductTitle = product_titles[1:15],
    Price = price[1:15],
    Ratings = ratings[1:15],
    Reviews = reviews[1:15],
    Description = descriptions[1:15],
    stringsAsFactors = FALSE
  )

graphics[[i]] <- card
}
```

```
Gcard <- rbind(graphics[[1]], graphics[[2]])

kable(Gcard, caption = "Graphics Card Category")
```

Table 1: Graphics Card Category

ProductTitle	Price	Ratings	Reviews	Description
NA	509.9	5.0	2	FREE delivery Mon, Jan 6 to Philippines
NA	119.9	4.7	3,527	Only 2 left in stock - order soon.
NA	279.9	4.5	96	FREE delivery to Philippines
NA	839.9	4.7	381	FREE delivery Mon, Jan 6 to Philippines
NA	89.9	4.4	39	More Buying Choices$274.50(19+ used & new offers)
NA	105.9	4.0	38	FREE delivery Mon, Jan 6 to Philippines
NA	78.9	4.3	15	No featured offers available$293.03(8 used & new offers)
NA	84.9	4.4	290	$10.00 off coupon appliedSave $10.00 with coupon
NA	102.9	4.2	121	5% off coupon appliedSave 5% with coupon
NA	90.9	4.5	172	FREE delivery to Philippines
NA	314.9	4.5	130	FREE delivery to Philippines
NA	329.9	4.5	12,229	FREE delivery Mon, Jan 6 to Philippines
NA	169.9	4.7	1,901	More Buying Choices$208.34(7 used & new offers)
NA	179.9	4.7	2,370	FREE delivery to Philippines
NA	210.0	4.7	2,665	FREE delivery to Philippines
NA	59.9	4.4	39	FREE delivery Mon, Jan 6 to Philippines
NA	99.9	5.0	2	FREE delivery to Philippines
NA	129.9	4.5	8,660	FREE delivery Thu, Jan 2 to Philippines
NA	199.9	4.5	4,588	More Buying Choices$79.99(2+ used & new offers)
NA	379.9	4.5	710	FREE delivery to Philippines
NA	489.9	4.6	174	More Buying Choices$197.99(11 used & new offers)
NA	299.9	4.6	207	FREE delivery to Philippines
NA	42.9	4.5	1,138	FREE delivery to Philippines
NA	49.9	4.5	92	FREE delivery to Philippines
NA	298.9	4.7	2,490	FREE delivery Mon, Jan 6 to Philippines on $49 of eligible items
NA	94.9	4.0	323	More Buying Choices$37.61(4 used & new offers)
NA	108.9	5.0	2	No featured offers available$2,299.00(8 used & new offers)
NA	119.9	4.4	89	FREE delivery Mon, Jan 6 to Philippines
NA	92.9	5.0	7	FREE delivery Thu, Jan 2 to Philippines
NA	564.9	4.7	414	FREE delivery to Philippines

```
write.csv(Gcard, "Gcard.csv", row.names = F)
```

Laptop

```
lap_url <- c("https://www.amazon.com/s?k=laptop&crid=15FICF43OPKJ2&sprefix=laptop%2Caps%2C1359&ref=nb_sb
             "https://www.amazon.com/s?k=laptop&page=2&crid=15FICF43OPKJ2&qid=1732065167&sprefix=laptop%2

laptop <- list()

for(i in seq_along(lap_url)){

  gSession <- bow(lap_url[i], user_agent = "Educational")
```

```
  # Scrape data
  product_titles <- scrape(gSession) %>%
    html_nodes("span.a-text-normal") %>%
    html_text() %>%
    head(30)
  product_titles <- product_titles[!grepl("Check each product page for other buying options", product_ti

  price <- scrape(gSession) %>%
    html_nodes('.a-price .a-offscreen') %>%
    html_text() %>%
    head(30)
  price <- as.numeric(str_extract(price, "\\d+\\.\\d"))

  ratings <- scrape(gSession) %>%
    html_nodes('span.a-icon-alt') %>%
    html_text() %>%
    head(30)
  ratings <- as.numeric(str_extract(ratings, "\\d+\\.\\d"))

  reviews <- scrape(gSession) %>%
    html_nodes('.s-link-style .s-underline-text') %>%
    html_text() %>%
    head(30)

  descriptions <- scrape(gSession) %>%
    html_nodes('.a-row.a-size-base.a-color-secondary') %>%
    html_text() %>%
    head(30)

  # Find the maximum length
  max_length <- max(
    length(product_titles), length(price), length(ratings),
    length(reviews), length(descriptions)
  )

  # Standardize all vectors to the same length
  length(product_titles) <- max_length
  length(price) <- max_length
  length(ratings) <- max_length
  length(reviews) <- max_length
  length(descriptions) <- max_length

  # Create a data frame
  lappy <- data.frame(
    ProductTitle = product_titles[1:15],
    Price = price[1:15],
    Ratings = ratings[1:15],
    Reviews = reviews[1:15],
    Description = descriptions[1:15],
    stringsAsFactors = FALSE
  )

laptop[[i]] <- lappy
```

```
}
```

```
laptops <- rbind(laptop[[1]], laptop[[2]])

kable(laptops, caption = "Laptop Category")
```

Table 2: Laptop Category

ProductTitle	Price	Ratings	Reviews	Description
NA	212.0	4.4	10,815	FREE delivery Thu, Jan 2 to Philippines
NA	249.9	4.4	18,017	FREE delivery to Philippines
NA	176.0	4.1	2,081	More Buying Choices$149.99(45+ used & new offers)
NA	209.9	4.0	1,964	FREE delivery to Philippines
NA	134.0	4.0	694	Only 14 left in stock - order soon.
NA	148.0	4.4	38	More Buying Choices$133.43(29+ used & new offers)
NA	269.0	4.2	41	FREE delivery Thu, Jan 2 to Philippines
NA	649.9	4.2	1,426	Only 7 left in stock - order soon.
NA	769.9	4.1	18	FREE delivery Mon, Jan 6 to Philippines
NA	449.9	4.3	2,053	Only 1 left in stock - order soon.
NA	529.9	4.6	1,632	More Buying Choices$439.99(10+ used & new offers)
NA	299.9	4.1	92	FREE delivery Thu, Jan 2 to Philippines
NA	349.9	4.3	19	FREE delivery to Philippines
NA	309.0	4.4	14	FREE delivery to Philippines
NA	399.0	4.0	1,812	More Buying Choices$193.00(12 new offers)
NA	289.9	5.0	12	FREE delivery Thu, Jan 2 to Philippines
NA	791.1	4.5	32	FREE delivery Mon, Jan 6 to Philippines
NA	879.0	4.6	11	Only 4 left in stock - order soon.
NA	179.9	4.2	733	FREE delivery Mon, Jan 6 to Philippines
NA	219.9	4.2	323	More Buying Choices$178.99(9 used & new offers)
NA	320.0	4.3	169	FREE delivery Mon, Jan 6 to Philippines
NA	399.0	4.1	23	Only 4 left in stock - order soon.
NA	348.0	4.7	4	More Buying Choices$210.00(10 used & new offers)
NA	249.9	4.3	551	FREE delivery Mon, Jan 6 to Philippines
NA	329.9	3.9	99	Only 2 left in stock - order soon.
NA	199.9	4.4	38	More Buying Choices$269.09(10+ used & new offers)
NA	623.9	5.0	17	FREE delivery to Philippines
NA	671.9	4.2	63	Only 4 left in stock (more on the way).
NA	204.9	5.0	6	More Buying Choices$185.99(4+ used & new offers)
NA	299.9	4.1	733	FREE delivery Mon, Jan 6 to Philippines

```
write.csv(laptops, "Laptop.csv", row.names = F)
```

PC Case

```
pc_url <- c("https://www.amazon.com/s?k=pc+cases&crid=3232B9XX8J30L&sprefix=pccase%2Caps%2C674&ref=nb_sb
            "https://www.amazon.com/s?k=pc+cases&page=2&crid=3232B9XX8J30L&qid=1732066185&sprefix=pccase

case <- list()

for(i in seq_along(pc_url)){
  Sys.sleep(2)
```

```r
gSession <- bow(pc_url[i], user_agent = "Educational")

# Scrape data
product_titles <- scrape(gSession) %>%
  html_nodes("span.a-text-normal") %>%
  html_text() %>%
  head(30)
product_titles <- product_titles[!grepl("Check each product page for other buying options", product_ti

price <- scrape(gSession) %>%
  html_nodes('.a-price .a-offscreen') %>%
  html_text() %>%
  head(30)
price <- as.numeric(str_extract(price, "\\d+\\.\\d"))

ratings <- scrape(gSession) %>%
  html_nodes('span.a-icon-alt') %>%
  html_text() %>%
  head(30)
ratings <- as.numeric(str_extract(ratings, "\\d+\\.\\d"))

reviews <- scrape(gSession) %>%
  html_nodes('.s-link-style .s-underline-text') %>%
  html_text() %>%
  head(30)

descriptions <- scrape(gSession) %>%
  html_nodes('.a-row.a-size-base.a-color-secondary') %>%
  html_text() %>%
  head(30)

# Find the maximum length
max_length <- max(
  length(product_titles), length(price), length(ratings),
  length(reviews), length(descriptions)
)

# Standardize all vectors to the same length
length(product_titles) <- max_length
length(price) <- max_length
length(ratings) <- max_length
length(reviews) <- max_length
length(descriptions) <- max_length

# Create a data frame
pCase <- data.frame(
  ProductTitle = product_titles[1:15],
  Price = price[1:15],
  Ratings = ratings[1:15],
  Reviews = reviews[1:15],
  Description = descriptions[1:15],
  stringsAsFactors = FALSE
)
```

```
case[[i]] <- pCase
}
```

```
pcCase <- rbind(case[[1]], case[[2]])

kable(pcCase, caption = "PC Case Category")
```

Table 3: PC Case Category

ProductTitle	Price	Ratings	Reviews	Description
NA	103.9	4.1	149	Delivery Mon, Jan 6 to Philippines
NA	129.9	4.8	4,275	Delivery Mon, Jan 6 to Philippines
NA	94.9	4.7	18,603	Delivery Thu, Jan 2 to Philippines
NA	104.9	4.6	197	More Buying Choices$77.17(3+ used & new offers)
NA	106.9	4.8	4,275	Delivery Thu, Jan 2 to Philippines
NA	69.9	4.3	66	More Buying Choices$99.06(2+ used & new offers)
NA	94.9	4.3	71	More Buying Choices$69.29(3+ used & new offers)
NA	89.9	4.6	43	5% off coupon appliedSave 5% with coupon
NA	68.0	4.1	91	Delivery Mon, Jan 6 to Philippines
NA	58.1	4.8	5,452	More Buying Choices$126.71(3+ used & new offers)
NA	63.9	4.0	2	FREE delivery Thu, Jan 2 to Philippines on $49 of eligible items
NA	99.9	4.5	12,924	Delivery Mon, Jan 6 to Philippines
NA	128.9	4.5	177	More Buying Choices$37.06(3+ used & new offers)
NA	174.9	4.5	386	Delivery Mon, Jan 6 to Philippines
NA	34.9	4.5	197	Only 9 left in stock - order soon.
NA	149.9	4.6	82	Delivery Fri, Jan 10 to Philippines
NA	34.9	4.0	2	Only 6 left in stock - order soon.
NA	59.9	4.4	388	FREE delivery Thu, Jan 2 to Philippines on $49 of eligible items
NA	159.9	4.7	247	More Buying Choices$56.99(2+ used & new offers)
NA	106.9	3.9	8	Only 2 left in stock - order soon.
NA	64.9	4.4	692	More Buying Choices$147.00(4+ used & new offers)
NA	129.9	4.7	1,042	Delivery Mon, Jan 6 to Philippines
NA	104.9	4.6	127	Only 10 left in stock - order soon.
NA	139.9	4.5	76	Delivery Mon, Jan 6 to Philippines
NA	149.9	4.4	250	More Buying Choices$51.34(2+ used & new offers)
NA	59.9	4.5	2	More Buying Choices$102.00(3 new offers)
NA	65.9	4.3	71	Delivery Mon, Jan 6 to Philippines
NA	68.0	4.5	251	Delivery Mon, Jan 6 to Philippines
NA	99.9	4.4	45	More Buying Choices$122.87(2+ used & new offers)
NA	89.9	4.8	1,378	Delivery Mon, Jan 6 to Philippines

```
write.csv(pcCase, "PC_Case.csv", row.names = F)
```

Monitors

```
monitor_url <- c("https://www.amazon.com/s?k=monitors&crid=UGL8HSJM01RM&qid=1732067471&refresh=1&sprefi
                 "https://www.amazon.com/s?k=monitors&page=2&crid=UGL8HSJM01RM&qid=1732067581&refresh=1&s
```

```r
mon <- list()

for(i in seq_along(monitor_url)){
  Sys.sleep(2)
  gSession <- bow(monitor_url[i], user_agent = "Educational")

  # Scrape data
  product_titles <- scrape(gSession) %>%
    html_nodes("span.a-text-normal") %>%
    html_text() %>%
    head(30)
  product_titles <- product_titles[!grepl("Check each product page for other buying options", product_ti

  price <- scrape(gSession) %>%
    html_nodes('.a-price .a-offscreen') %>%
    html_text() %>%
    head(30)
  price <- as.numeric(str_extract(price, "\\d+\\.\\d"))

  ratings <- scrape(gSession) %>%
    html_nodes('span.a-icon-alt') %>%
    html_text() %>%
    head(30)
  ratings <- as.numeric(str_extract(ratings, "\\d+\\.\\d"))

  reviews <- scrape(gSession) %>%
    html_nodes('.s-link-style .s-underline-text') %>%
    html_text() %>%
    head(30)

  descriptions <- scrape(gSession) %>%
    html_nodes('.a-row.a-size-base.a-color-secondary') %>%
    html_text() %>%
    head(30)

  # Find the maximum length
  max_length <- max(
    length(product_titles), length(price), length(ratings),
    length(reviews), length(descriptions)
  )

  # Standardize all vectors to the same length
  length(product_titles) <- max_length
  length(price) <- max_length
  length(ratings) <- max_length
  length(reviews) <- max_length
  length(descriptions) <- max_length

  # Create a data frame
  moni <- data.frame(
    ProductTitle = product_titles[1:15],
    Price = price[1:15],
    Ratings = ratings[1:15],
```

```
    Reviews = reviews[1:15],
    Description = descriptions,
    stringsAsFactors = FALSE
  )

mon[[i]] <- moni
}
```

```
mOnitor <- rbind(mon[[1]], mon[[2]])

kable(mOnitor, caption = "Monitor Category")
```

Table 4: Monitor Category

ProductTitle	Price	Ratings	Reviews	Description
NA	129.9	4.3	51	$78.47 delivery Wed, Jan 8
NA	144.9	4.3	944	FREE delivery Wed, Jan 8 to Philippines
NA	249.9	4.5	33,803	FREE delivery Mon, Jan 6 to Philippines
NA	98.1	4.6	5,254	More Buying Choices$96.99(3+ used & new offers)
NA	119.9	4.6	1,455	No featured offers available$57.51(5 used & new offers)
NA	89.9	4.3	2,847	$100.00 off coupon appliedSave $100.00 with coupon
NA	149.9	4.6	3,493	FREE delivery Mon, Jan 6 to Philippines
NA	399.9	4.5	2,552	More Buying Choices$79.17(3+ used & new offers)
NA	799.9	4.4	1,369	FREE delivery Thu, Jan 2 to Philippines on $49 of eligible items
NA	229.9	4.3	503	Delivery Mon, Jan 6 to Philippines
NA	99.9	4.6	23,287	More Buying Choices$82.32(6+ used & new offers)
NA	129.9	4.4	9,005	Delivery Wed, Jan 8 to Philippines
NA	178.9	4.5	5,384	More Buying Choices$174.37(4+ used & new offers)
NA	289.9	4.3	2,847	Delivery Mon, Jan 6 to Philippines
NA	949.9	4.6	725	More Buying Choices$107.10(3+ used & new offers)
NA	129.9	4.3	51	FREE delivery Mon, Jan 6 to Philippines
NA	144.9	4.3	944	More Buying Choices$52.99(7+ used & new offers)
NA	249.9	4.5	33,803	FREE delivery Thu, Jan 2 to Philippines on $49 of eligible items
NA	98.1	4.6	5,254	Delivery Mon, Jan 6 to Philippines
NA	119.9	4.6	1,455	More Buying Choices$95.99(2+ used & new offers)
NA	89.9	4.3	2,847	Delivery Mon, Jan 6 to Philippines
NA	149.9	4.6	3,493	FREE delivery Mon, Jan 6 to Philippines
NA	399.9	4.5	2,552	More Buying Choices$163.17(4+ used & new offers)
NA	799.9	4.4	1,369	FREE delivery Mon, Jan 6 to Philippines
NA	229.9	4.3	503	FREE delivery Thu, Jan 2 to Philippines
NA	99.9	4.6	23,287	FREE delivery Thu, Jan 2 to Philippines
NA	129.9	4.4	9,005	Delivery Mon, Jan 6 to Philippines
NA	178.9	4.5	5,384	Delivery Thu, Jan 2 to Philippines
NA	289.9	4.3	2,847	More Buying Choices$129.59(2+ used & new offers)
NA	949.9	4.6	725	FREE delivery Mon, Jan 6 to Philippines
NA	119.9	5.0	10	Delivery Thu, Jan 2 to Philippines
NA	49.9	4.7	5	FREE delivery Thu, Jan 2 to Philippines
NA	199.9	4.3	2,448	Usually ships within 1 to 4 weeks
NA	329.9	4.4	1,369	More Buying Choices$197.99(2+ used & new offers)
NA	178.9	4.6	1,055	Delivery Mon, Jan 6 to Philippines

ProductTitle	Price	Ratings	Reviews	Description
NA	289.9	4.3	790	More Buying Choices$177.15(9+ used & new offers)
NA	139.9	4.4	1,580	No featured offers available$236.54(15 used & new offers)
NA	89.9	4.4	406	10% off coupon appliedSave 10% with coupon
NA	172.9	4.5	2,552	$20.00 off coupon appliedSave $20.00 with coupon
NA	99.9	4.3	229	More Buying Choices$169.99(25+ used & new offers)
NA	129.9	4.6	12,118	FREE delivery to Philippines
NA	249.9	4.4	397	Only 7 left in stock (more on the way).
NA	399.9	4.5	1,450	$142.61 delivery Wed, Jan 8
NA	257.2	4.4	7,019	Only 8 left in stock - order soon.
NA	299.9	4.5	2,314	More Buying Choices$274.31(16+ used & new offers)
NA	119.9	5.0	10	FREE delivery Mon, Jan 6 to Philippines
NA	49.9	4.7	5	More Buying Choices$70.58(6+ used & new offers)
NA	199.9	4.3	2,448	FREE delivery Thu, Jan 2 to Philippines
NA	329.9	4.4	1,369	More Buying Choices$227.83(15+ used & new offers)
NA	178.9	4.6	1,055	FREE delivery Mon, Jan 6 to Philippines on $49 of eligible items
NA	289.9	4.3	790	FREE delivery Wed, Jan 8 to Philippines
NA	139.9	4.4	1,580	FREE delivery Wed, Jan 8 to Philippines
NA	89.9	4.4	406	More Buying Choices$118.79(2 used & new offers)
NA	172.9	4.5	2,552	FREE delivery Thu, Jan 2 to Philippines
NA	99.9	4.3	229	More Buying Choices$566.87(6+ used & new offers)
NA	129.9	4.6	12,118	Delivery Mon, Jan 6 to Philippines
NA	249.9	4.4	397	More Buying Choices$65.49(4+ used & new offers)
NA	399.9	4.5	1,450	$143.58 delivery Wed, Jan 8
NA	257.2	4.4	7,019	Only 1 left in stock - order soon.
NA	299.9	4.5	2,314	More Buying Choices$260.94(18+ used & new offers)

```
write.csv(mOnitor, "Monitor.csv", row.names = F)
```

CPU

```
cpu_url <- c("https://www.amazon.com/s?k=cpu&crid=3DBICI2YNNKHJ&sprefix=cpu%2Caps%2C768&ref=nb_sb_noss_
             "https://www.amazon.com/s?k=cpu&page=2&crid=3DBICI2YNNKHJ&qid=1732067707&sprefix=cpu%2Ca

processor <- list()

for(i in seq_along(cpu_url)){
  Sys.sleep(2)
  gSession <- bow(cpu_url[i], user_agent = "Educational")

  # Scrape data
  product_titles <- scrape(gSession) %>%
    html_nodes("span.a-text-normal") %>%
    html_text() %>%
    head(30)
  product_titles <- product_titles[!grepl("Check each product page for other buying options", product_t

  price <- scrape(gSession) %>%
    html_nodes('.a-price .a-offscreen') %>%
    html_text() %>%
    head(30)
```

```
  price <- as.numeric(str_extract(price, "\\d+\\.\\d"))

  ratings <- scrape(gSession) %>%
    html_nodes('span.a-icon-alt') %>%
    html_text() %>%
    head(30)
  ratings <- as.numeric(str_extract(ratings, "\\d+\\.\\d"))

  reviews <- scrape(gSession) %>%
    html_nodes('.s-link-style .s-underline-text') %>%
    html_text() %>%
    head(30)

  descriptions <- scrape(gSession) %>%
    html_nodes('.a-row.a-size-base.a-color-secondary') %>%
    html_text() %>%
    head(30)

  # Find the maximum length
  max_length <- max(
    length(product_titles), length(price), length(ratings),
    length(reviews), length(descriptions)
  )

  # Standardize all vectors to the same length
  length(product_titles) <- max_length
  length(price) <- max_length
  length(ratings) <- max_length
  length(reviews) <- max_length
  length(descriptions) <- max_length

  # Create a data frame
  process <- data.frame(
    ProductTitle = product_titles[1:15],
    Price = price[1:15],
    Ratings = ratings[1:15],
    Reviews = reviews[1:15],
    Description = descriptions,
    stringsAsFactors = FALSE
  )

processor[[i]] <- process
}
```

```
proc <- rbind(processor[[1]], processor[[2]])

kable(proc, caption = "CPU Category")
```

Table 5: CPU Category

ProductTitle	Price	Ratings	Reviews	Description
NA	167.9	4.2	243	FREE delivery Mon, Jan 6 to Philippines

ProductTitle	Price	Ratings	Reviews	Description
NA	64.9	4.7	25,192	Only 1 left in stock - order soon.
NA	79.9	4.8	7,241	by Kingston
NA	149.9	4.7	5,076	FREE delivery Mon, Jan 6 to Philippines
NA	319.0	4.8	1,353	FREE delivery to Philippines
NA	79.5	4.8	1,339	More Buying Choices$139.64(34+ used & new offers)
NA	159.0	4.6	2,486	FREE delivery Thu, Jan 2 to Philippines
NA	219.9	4.5	23	FREE delivery Thu, Jan 2 to Philippines
NA	249.0	4.4	1,623	More Buying Choices$199.99(18+ used & new offers)
NA	149.0	4.8	26,936	FREE delivery to Philippines
NA	342.5	4.7	1,911	More Buying Choices$112.09(21+ used & new offers)
NA	449.9	4.8	2,664	FREE delivery to Philippines
NA	629.9	4.7	1,575	Usually ships within 2 to 3 days
NA	123.9	4.7	3,048	More Buying Choices$255.00(23+ used & new offers)
NA	268.9	4.7	12,118	FREE delivery to Philippines
NA	167.9	4.2	243	FREE delivery Mon, Jan 6 to Philippines
NA	64.9	4.7	25,192	FREE delivery Thu, Jan 2 to Philippines
NA	79.9	4.8	7,241	More Buying Choices$180.49(41+ used & new offers)
NA	149.9	4.7	5,076	FREE delivery to Philippines
NA	319.0	4.8	1,353	More Buying Choices$268.00(11+ used & new offers)
NA	79.5	4.8	1,339	Delivery Thu, Jan 2 to Philippines
NA	159.0	4.6	2,486	FREE delivery to Philippines
NA	219.9	4.5	23	More Buying Choices$152.99(19+ used & new offers)
NA	249.0	4.4	1,623	FREE delivery to Philippines
NA	149.0	4.8	26,936	Usually ships within 2 to 3 days
NA	342.5	4.7	1,911	FREE delivery Thu, Jan 2 to Philippines
NA	449.9	4.8	2,664	No featured offers available$334.34(18 used & new offers)
NA	629.9	4.7	1,575	FREE delivery Fri, Jan 10 to Philippines
NA	123.9	4.7	3,048	FREE delivery Thu, Jan 2 to Philippines
NA	268.9	4.7	12,118	More Buying Choices$65.00(12+ used & new offers)
NA	167.9	4.2	243	FREE delivery Mon, Jan 6 to Philippines
NA	65.9	3.9	86	Only 1 left in stock - order soon.
NA	89.9	4.6	2,486	FREE delivery Thu, Jan 2 to Philippines
NA	149.0	4.8	6,251	Delivery Thu, Jan 2 to Philippines
NA	342.5	4.8	1,551	FREE delivery Thu, Jan 2 to Philippines
NA	116.5	4.0	3,777	More Buying Choices$95.00(15+ used & new offers)
NA	199.0	4.8	4,092	FREE delivery Mon, Jan 6 to Philippines
NA	197.9	4.8	7,946	FREE delivery Mon, Jan 6 to Philippines
NA	229.0	4.8	1,774	Only 4 left in stock - order soon.
NA	199.9	4.4	77	More Buying Choices$197.99(9 new offers)
NA	249.9	4.6	43	FREE delivery to Philippines
NA	476.0	4.8	843	FREE delivery to Philippines
NA	159.0	4.4	1,667	FREE delivery Mon, Jan 6 to Philippines
NA	359.0	4.6	103	FREE delivery to Philippines
NA	111.3	4.8	805	Delivery Mon, Jan 6 to Philippines
NA	167.9	4.2	243	FREE delivery Mon, Jan 6 to Philippines
NA	65.9	3.9	86	Only 4 left in stock (more on the way).
NA	89.9	4.6	2,486	Delivery Thu, Jan 2 to Philippines
NA	149.0	4.8	6,251	FREE delivery to Philippines
NA	342.5	4.8	1,551	Usually ships within 2 to 3 days
NA	116.5	4.0	3,777	FREE delivery Mon, Jan 6 to Philippines
NA	199.0	4.8	4,092	Only 8 left in stock - order soon.
NA	197.9	4.8	7,946	FREE delivery to Philippines

ProductTitle	Price	Ratings	Reviews	Description
NA	229.0	4.8	1,774	More Buying Choices$359.99(7 used & new offers)
NA	199.9	4.4	77	FREE delivery to Philippines
NA	249.9	4.6	43	Only 3 left in stock - order soon.
NA	476.0	4.8	843	More Buying Choices$244.59(18 new offers)
NA	159.0	4.4	1,667	FREE delivery to Philippines
NA	359.0	4.6	103	More Buying Choices$374.99(15 used & new offers)
NA	111.3	4.8	805	FREE delivery to Philippines

```
write.csv(proc, "CPU.csv", row.names = F)
```

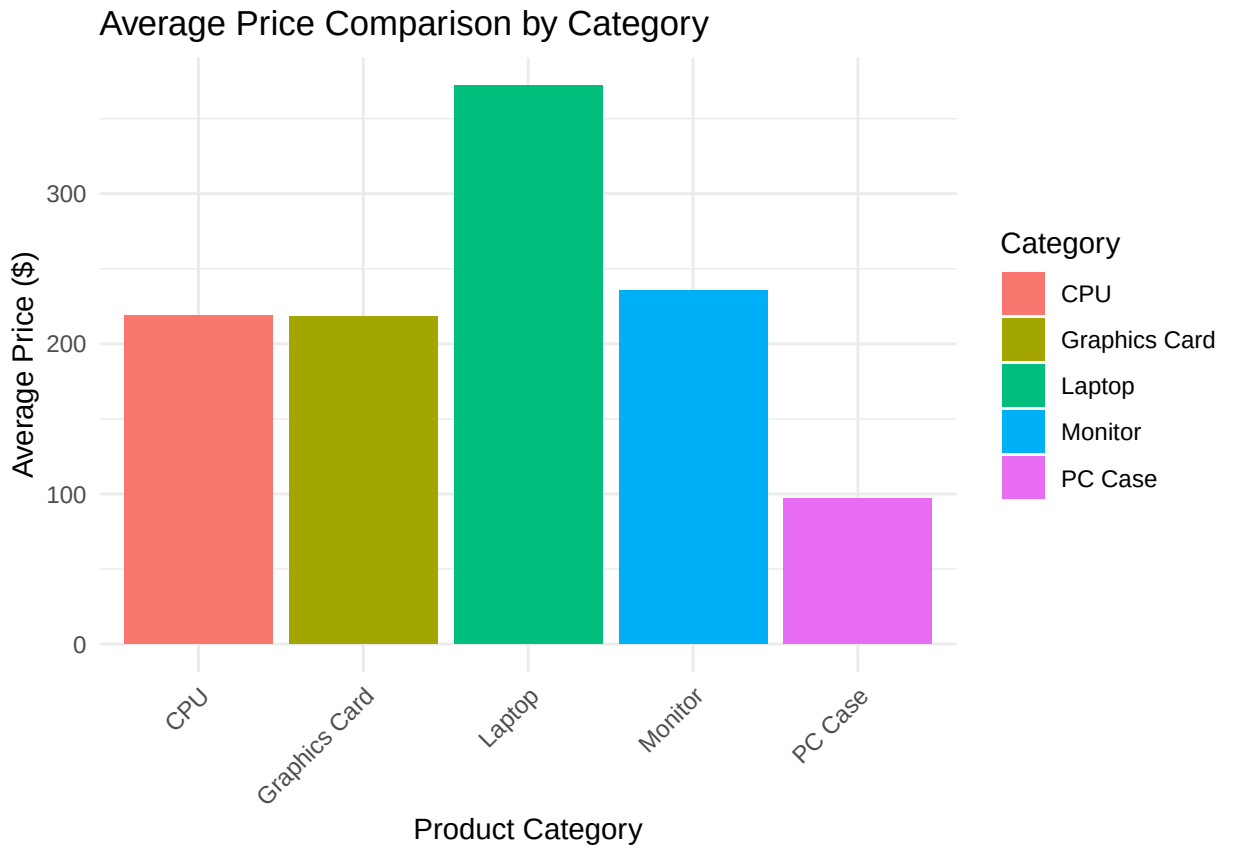
6. The extracted dataset contains information on 150 products across five categories: Laptops, CPU, Graphics Card, Monitors, and PC Cases, with 30 products selected from each category. For each product, the dataset includes the price, description, ratings, and reviews.The product title contains a string which shows the name of the product. The price is represented as a numeric value indicating the cost of the product. The ratings are numerical, ranging from 0 to 5, and reflect the average customer satisfaction score.The reviews field contains an integer representing the total number of customer reviews. Finally, the description is a text field providing a brief overview of the product's features

7. Competitive Benchmarking and Pricing Trends Analysis: The dataset enables businesses to perform competitive benchmarking by comparing product prices within and across categories (e.g., laptops, CPUs, graphics cards, monitors, and PC cases). By analyzing pricing trends, businesses can gain critical insights into market positioning and consumer preferences.

8.

```
library(ggplot2)

# Read the CSV files and calculate the mean price for each category
cpu_price <- read.csv("CPU.csv")
gcard_price <- read.csv("Gcard.csv")
laptop_price <- read.csv("Laptop.csv")
monitor_price <- read.csv("Monitor.csv")
PCcase_price <- read.csv("PC_Case.csv")

# Calculate the mean price for each category
mean_cpu <- mean(cpu_price$Price)
mean_gcard <- mean(gcard_price$Price)
mean_laptop <- mean(laptop_price$Price)
mean_monitor <- mean(monitor_price$Price)
mean_pc_case <- mean(PCcase_price$Price)

# Create a data frame to store the average prices
avg_prices <- data.frame(
  Category = c("CPU", "Graphics Card", "Laptop", "Monitor", "PC Case"),
  Average_Price = c(mean_cpu, mean_gcard, mean_laptop, mean_monitor, mean_pc_case)
)

# Create the bar plot
ggplot(avg_prices, aes(x = Category, y = Average_Price, fill = Category)) +
  geom_bar(stat = "identity") +
  labs(title = "Average Price Comparison by Category",
```

```
      x = "Product Category",
      y = "Average Price ($)") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



The graph provides valuable insights into competitive benchmarking and pricing trends across different product categories in the computing market. Laptops stand out as the most expensive category, with an average price exceeding $300, indicating their premium nature and potential for higher profit margins. CPUs and graphics cards fall into a similar pricing tier, both averaging slightly above $200, which could reflect their high demand among gamers and professionals building custom systems. Monitors, priced around $150 on average, are moderately positioned, suggesting they cater to a broad range of users with varying budgets. Finally, PC cases are the least expensive category, averaging under $100, highlighting their relatively low-cost nature compared to other components. These insights are crucial for businesses looking to align their pricing strategies with market trends and remain competitive in these product segments.

9.

```
# Load the necessary library
library(ggplot2)

# Read the data
cpu_price <- read.csv("CPU.csv")
gcard_price <- read.csv("Gcard.csv")
laptop_price <- read.csv("Laptop.csv")
monitor_price <- read.csv("Monitor.csv")
PCcase_price <- read.csv("PC_Case.csv")
```

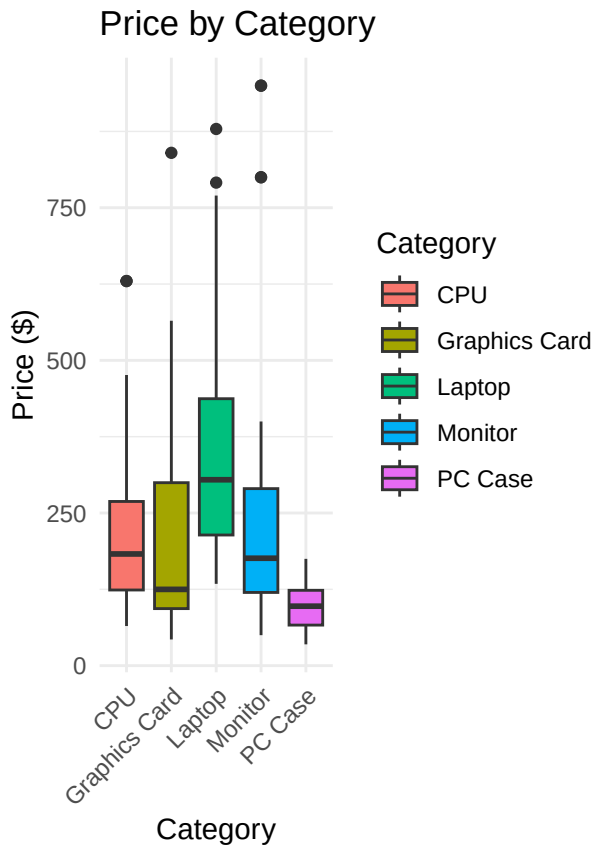
```r
# Combine the data for price and ratings into one data frame
price_data <- data.frame(
  Category = rep(c("CPU", "Graphics Card", "Laptop", "Monitor", "PC Case"),
                 times = c(nrow(cpu_price), nrow(gcard_price), nrow(laptop_price), nrow(monitor_price),
  Price = c(cpu_price$Price, gcard_price$Price, laptop_price$Price, monitor_price$Price, PCcase_price$Pr
  Ratings = c(cpu_price$Ratings, gcard_price$Ratings, laptop_price$Ratings, monitor_price$Ratings, PCcas
)

# Price Plot using ggplot2
price_plot <- ggplot(price_data, aes(x = Category, y = Price, fill = Category)) +
  geom_boxplot() +
  labs(title = "Price by Category", y = "Price ($)", x = "Category") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))  # Rotate x-axis labels for readability

# Ratings Plot using ggplot2
ratings_plot <- ggplot(price_data, aes(x = Category, y = Ratings, fill = Category)) +
  geom_boxplot() +
  labs(title = "Ratings by Category", y = "Ratings", x = "Category") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))  # Rotate x-axis labels for readability

# Print the plots side by side
library(gridExtra)
```

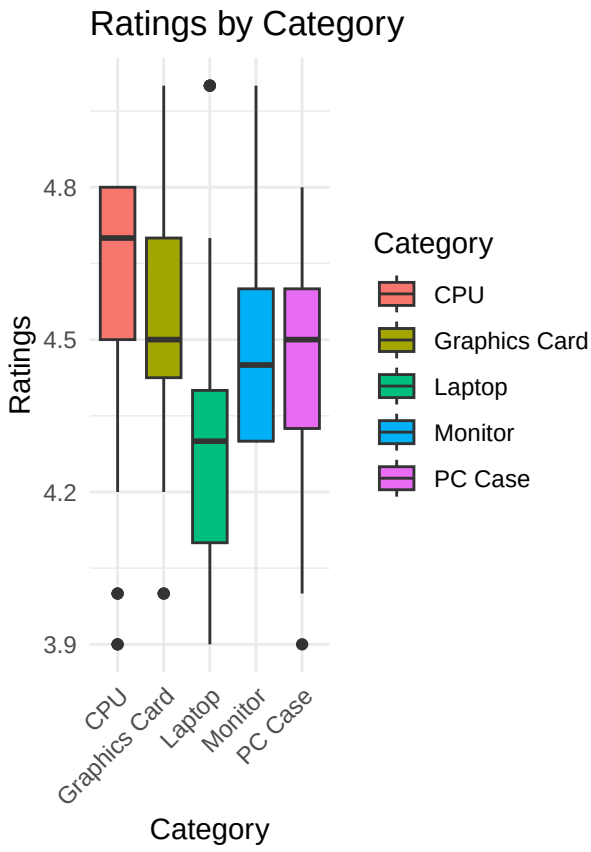
```
## Warning: package 'gridExtra' was built under R version 4.4.2
```

```
##
## Attaching package: 'gridExtra'
```

```
## The following object is masked from 'package:dplyr':
##
##     combine
```

```r
grid.arrange(price_plot, ratings_plot, ncol = 2)
```





10.

```
# Rank products by price and ratings for each category

# Rank by price (highest to lowest)
cpu_price$Price_Rank <- rank(-cpu_price$Price)  # Negative sign for descending order
gcard_price$Price_Rank <- rank(-gcard_price$Price)
laptop_price$Price_Rank <- rank(-laptop_price$Price)
monitor_price$Price_Rank <- rank(-monitor_price$Price)
PCcase_price$Price_Rank <- rank(-PCcase_price$Price)

# Rank by ratings (highest to lowest)
cpu_price$Rating_Rank <- rank(-cpu_price$Ratings)
gcard_price$Rating_Rank <- rank(-gcard_price$Ratings)
laptop_price$Rating_Rank <- rank(-laptop_price$Ratings)
monitor_price$Rating_Rank <- rank(-monitor_price$Ratings)
PCcase_price$Rating_Rank <- rank(-PCcase_price$Ratings)

# View the top ranked products for each category
head(cpu_price)
```

```
##   ProductTitle Price Ratings Reviews
## 1           NA 167.9     4.2     243
## 2           NA  64.9     4.7  25,192
## 3           NA  79.9     4.8   7,241
## 4           NA 149.9     4.7   5,076
## 5           NA 319.0     4.8   1,353
```

```
## 6           NA  79.5     4.8   1,339
##                                       Description Price_Rank Rating_Rank
## 1          FREE delivery Mon, Jan 6 to Philippines        32.5        54.5
## 2               Only 1 left in stock - order soon.        59.5        30.5
## 3                                       by Kingston        53.5        12.5
## 4          FREE delivery Mon, Jan 6 to Philippines        39.5        30.5
## 5                     FREE delivery to Philippines        13.5        12.5
## 6 More Buying Choices$139.64(34+ used & new offers)        55.5        12.5
```

```
head(gcard_price)
```

```
##   ProductTitle Price Ratings Reviews
## 1           NA 509.9     5.0       2
## 2           NA 119.9     4.7   3,527
## 3           NA 279.9     4.5      96
## 4           NA 839.9     4.7     381
## 5           NA  89.9     4.4      39
## 6           NA 105.9     4.0      38
##                                       Description Price_Rank Rating_Rank
## 1          FREE delivery Mon, Jan 6 to Philippines         3.0         2.5
## 2               Only 2 left in stock - order soon.        16.5         8.0
## 3                     FREE delivery to Philippines        10.0        18.0
## 4          FREE delivery Mon, Jan 6 to Philippines         1.0         8.0
## 5 More Buying Choices$274.50(19+ used & new offers)        25.0        24.5
## 6          FREE delivery Mon, Jan 6 to Philippines        19.0        29.5
```

```
head(laptop_price)
```

```
##   ProductTitle Price Ratings Reviews
## 1           NA 212.0     4.4  10,815
## 2           NA 249.9     4.4  18,017
## 3           NA 176.0     4.1   2,081
## 4           NA 209.9     4.0   1,964
## 5           NA 134.0     4.0     694
## 6           NA 148.0     4.4      38
##                                       Description Price_Rank Rating_Rank
## 1          FREE delivery Thu, Jan 2 to Philippines        23.0          10
## 2                     FREE delivery to Philippines        20.5          10
## 3 More Buying Choices$149.99(45+ used & new offers)        28.0          24
## 4                     FREE delivery to Philippines        24.0          28
## 5               Only 14 left in stock - order soon.        30.0          28
## 6 More Buying Choices$133.43(29+ used & new offers)        29.0          10
```

```
head(monitor_price)
```

```
##   ProductTitle Price Ratings Reviews
## 1           NA 129.9     4.3      51
## 2           NA 144.9     4.3     944
## 3           NA 249.9     4.5  33,803
## 4           NA  98.1     4.6   5,254
## 5           NA 119.9     4.6   1,455
## 6           NA  89.9     4.3   2,847
```

```
##                                                Description Price_Rank
## 1                              $78.47 delivery Wed, Jan 8        41.5
## 2                  FREE delivery Wed, Jan 8 to Philippines        35.5
## 3                  FREE delivery Mon, Jan 6 to Philippines        20.5
## 4           More Buying Choices$96.99(3+ used & new offers)       53.5
## 5 No featured offers available$57.51(5 used & new offers)       46.5
## 6    $100.00 off coupon appliedSave $100.00  with coupon        56.5
##   Rating_Rank
## 1        52.5
## 2        52.5
## 3        24.5
## 4        11.5
## 5        11.5
## 6        52.5
```

```
head(PCcase_price)
```

```
##   ProductTitle Price Ratings Reviews
## 1           NA 103.9     4.1     149
## 2           NA 129.9     4.8   4,275
## 3           NA  94.9     4.7  18,603
## 4           NA 104.9     4.6     197
## 5           NA 106.9     4.8   4,275
## 6           NA  69.9     4.3      66
##                                     Description Price_Rank Rating_Rank
## 1             Delivery Mon, Jan 6 to Philippines        13.0        26.5
## 2             Delivery Mon, Jan 6 to Philippines         6.5         2.5
## 3             Delivery Thu, Jan 2 to Philippines        16.5         6.0
## 4 More Buying Choices$77.17(3+ used & new offers)       11.5         9.5
## 5             Delivery Thu, Jan 2 to Philippines         9.5         2.5
## 6 More Buying Choices$99.06(2+ used & new offers)       20.0        24.0
```

Ranking products by price and ratings provides insights into how cost and customer satisfaction relate. The price rank identifies the most expensive products, often premium options, but doesn't necessarily reflect customer satisfaction. In contrast, the rating rank shows which products are most highly regarded by consumers, regardless of their price. Comparing these rankings helps highlight whether consumers prioritize cost or quality. For example, in Laptops, mid-range models may have higher ratings despite lower prices, while in PC Cases, budget-friendly products could be highly rated for their value and functionality. This comparison aids in understanding the balance between price and quality in each category.