RWorksheet#5_group(1&9)

GROUP 1&9 BSIT-2B

2023-11-30

##Extracting TV Shows Reviews

1. Each group needs to extract the top 50 tv shows in Imdb.com. It will include the rank, the title of the tv show, tv rating, the number of people who voted, the number of episodes, the year it was released.

```
library(rvest)
```

```
## Warning: package 'rvest' was built under R version 4.3.2
```

```
library(httr)
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 4.3.2
```

```
##
## 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(polite)
```

```
## Warning: package 'polite' was built under R version 4.3.2
```

```
library(stringr)
```

```
## Warning: package 'stringr' was built under R version 4.3.2
```

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

```
## v Setting active project to 'C:/Users/missy/OneDrive/Documents/Worksheet#5'
```

```r
url <- 'https://m.imdb.com/chart/toptv/?ref_=nv_tvv_250'

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

```
## <polite session> https://m.imdb.com/chart/toptv/?ref_=nv_tvv_250
##     User-agent: Educational
##     robots.txt: 23 rules are defined for 2 bots
##    Crawl delay: 5 sec
##   The path is scrapable for this user-agent
```

```r
title_show <- character(0)
list_year_ep <-character(0)

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

title_show_only <- as.data.frame(title_show[2:51])
title_show_only
```

```
##                           title_show[2:51]
## 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. The World at War
## 15                        15. Rick and Morty
## 16                                16. Bluey
## 17     17. Fullmetal Alchemist Brotherhood
## 18                         18. The Last Dance
## 19                                 19. Life
## 20                    20. The Twilight Zone
## 21                             21. Sherlock
## 22                       22. The Vietnam War
## 23          23. Batman: The Animated Series
## 24                         24. Attack on Titan
## 25  25. Scam 1992: The Harshad Mehta Story
## 26                             26. The Office
## 27                                 27. Arcane
## 28                       28. The Blue Planet
## 29                      29. Better Call Saul
## 30                          30. Human Planet
## 31                               31. Firefly
```

```
## 32                       32. Frozen Planet
## 33                     33. Clarkson's Farm
## 34                          34. Death Note
## 35           35. Only Fools and Horses....
## 36                     36. Hunter x Hunter
## 37                       37. The Civil War
## 38                      38. True Detective
## 39                           39. Seinfeld
## 40               40. The Beatles: Get Back
## 41                              41. Dekalog
## 42                             42. Sahsiyet
## 43                                43. Fargo
## 44                         44. Cowboy Bebop
## 45                        45. Gravity Falls
## 46                       46. Nathan for You
## 47  47. Last Week Tonight with John Oliver
## 48                     48. When They See Us
## 49                          49. Succession
## 50 50. Apocalypse: La 2ème guerre mondiale
```

```
colnames(title_show_only) <- "Rank"

show_df <- strsplit(as.character(title_show_only$Rank),".",fixed = TRUE)
show_df <- data.frame(do.call(rbind,show_df))
```

```
## Warning in (function (..., deparse.level = 1) : number of columns of result is
## not a multiple of vector length (arg 1)
```

```
show_df <- show_df[-c(3:5)]

#renaming column 1 and 2
colnames(show_df) <- c("Rank","Title")


list_year_ep <- scrape(session) %>%
  html_nodes('span.sc-479faa3c-8.bNrEFi.cli-title-metadata-item
') %>%
  html_text
years_only <- c()

for (i in seq(1, length(list_year_ep), by = 3)) {
  years_only <- c(years_only, list_year_ep[i])
}
Year <- years_only[1:50]

ep_only <- c()
for (i in seq(2, length(list_year_ep), by = 3)) {
  ep_only <- c(ep_only, list_year_ep[i])
}
Episode <- ep_only[1:50]

df_title_ep <- data.frame(Episode, Year)
```

```
df_title_ep$Year <- gsub("^(\\d{4}).*", "\\1", df_title_ep$Year)

colnames(df_title_ep) <- c("Number Of Episodes", "Year Released")

list_rating <- scrape(session) %>%
  html_nodes('span.ipc-rating-star.ipc-rating-star--base.ipc-rating-star--imdb.ratingGroup--imdb-rating
  html_text()

wholeRATING <- as.data.frame(list_rating[1:50])
colnames(wholeRATING) <- "Rating"

wholeRATING$Rating <- gsub("\\s*\\([^)]+\\)\\s*", "", wholeRATING$Rating)
wholeRATING$Vote_Count <- gsub(".*\\(([^)]+)\\)", "\\1", list_rating[1:50])


df_rating_vote <- wholeRATING
colnames(df_rating_vote) <- c("Rating","Vote Count")


final_df <- cbind(show_df, df_rating_vote,df_title_ep )
final_df
```

```
##    Rank                              Title Rating Vote Count
## 1     1                       Breaking Bad   9.5        2.1M
## 2     2                    Planet Earth II   9.5        155K
## 3     3                       Planet Earth   9.4        217K
## 4     4                    Band of Brothers   9.4        507K
## 5     5                          Chernobyl   9.3        834K
## 6     6                           The Wire   9.3        366K
## 7     7       Avatar: The Last Airbender   9.3        349K
## 8     8                     Blue Planet II   9.3         45K
## 9     9                       The Sopranos   9.2        449K
## 10   10     Cosmos: A Spacetime Odyssey   9.3        127K
## 11   11                            Cosmos   9.3         43K
## 12   12                        Our Planet   9.3         49K
## 13   13                    Game of Thrones   9.2        2.2M
## 14   14                   The World at War   9.2         28K
## 15   15                     Rick and Morty   9.1        580K
## 16   16                              Bluey   9.4         22K
## 17   17  Fullmetal Alchemist Brotherhood   9.1        191K
## 18   18                     The Last Dance   9.1        145K
## 19   19                               Life   9.1         42K
## 20   20                The Twilight Zone   9.1         90K
## 21   21                           Sherlock   9.1        976K
## 22   22                    The Vietnam War   9.1         27K
## 23   23      Batman: The Animated Series   9.0        113K
## 24   24                    Attack on Titan   9.1        475K
## 25   25 Scam 1992: The Harshad Mehta Story   9.3        153K
## 26   26                        The Office   9.0        680K
## 27   27                            Arcane   9.0        252K
## 28   28                    The Blue Planet   9.0         42K
## 29   29                   Better Call Saul   9.0        619K
## 30   30                       Human Planet   9.0         27K
```

```
## 31   31                            Firefly   9.0        277K
## 32   32                       Frozen Planet   9.0         32K
## 33   33                     Clarkson's Farm   9.0         52K
## 34   34                          Death Note   8.9        364K
## 35   35                Only Fools and Horses   9.0         55K
## 36   36                      Hunter x Hunter   9.0        125K
## 37   37                        The Civil War   9.0         18K
## 38   38                      True Detective   8.9        617K
## 39   39                             Seinfeld   8.9        342K
## 40   40                The Beatles: Get Back   9.0         27K
## 41   41                              Dekalog   9.0         27K
## 42   42                              Sahsiyet   9.0         45K
## 43   43                                 Fargo   8.9        396K
## 44   44                        Cowboy Bebop   8.9        134K
## 45   45                        Gravity Falls   8.9        129K
## 46   46                      Nathan for You   8.9         37K
## 47   47   Last Week Tonight with John Oliver   8.9         94K
## 48   48                     When They See Us   8.9        134K
## 49   49                            Succession   8.9        246K
## 50   50  Apocalypse: La 2ème guerre mondiale   9.0         14K
##    Number Of Episodes Year Released
## 1              62 eps          2008
## 2               6 eps          2016
## 3              11 eps          2006
## 4              10 eps          2001
## 5               5 eps          2019
## 6              60 eps          2002
## 7              62 eps          2005
## 8               7 eps          2017
## 9              86 eps          1999
## 10             13 eps          2014
## 11             13 eps          1980
## 12             12 eps          2019
## 13             73 eps          2011
## 14             26 eps          1973
## 15             74 eps          2013
## 16            171 eps          2018
## 17             68 eps          2009
## 18             10 eps          2020
## 19             11 eps          2009
## 20            156 eps          1959
## 21             15 eps          2010
## 22             10 eps          2017
## 23             85 eps          1992
## 24             98 eps          2013
## 25             10 eps          2020
## 26            188 eps          2005
## 27             10 eps          2021
## 28              8 eps          2001
## 29             63 eps          2015
## 30              8 eps          2011
## 31             14 eps          2002
## 32             10 eps          2011
## 33             18 eps          2021
```

```
## 34             37 eps          2006
## 35             64 eps          1981
## 36            148 eps          2011
## 37              9 eps          1990
## 38             30 eps          2014
## 39            173 eps          1989
## 40              3 eps          2021
## 41             10 eps          1989
## 42             17 eps          2018
## 43             51 eps          2014
## 44             26 eps          1998
## 45             41 eps          2012
## 46             32 eps          2013
## 47            339 eps          2014
## 48              4 eps          2019
## 49             39 eps          2018
## 50              6 eps          2009
```

From the 50 tv shows, select at least 5 tv shows to scrape the user reviews that will include the reviewer's name, date of reviewed, user rating, title of the review, and text reviews.

```
tv_show_urls <- c(
  "https://www.imdb.com/title/tt0081846/reviews",  # COSMOS
  "https://www.imdb.com/title/tt0903747/reviews",  # BREAKING BAD
  "https://www.imdb.com/title/tt0185906/reviews",  # BAND OF BROTHERS
  "https://www.imdb.com/title/tt7366338/reviews",  # CHERNOBYL
  "https://www.imdb.com/title/tt0417299/reviews"   # Avatar: The Last Airbender
)

all_shows<-list()

for (url in tv_show_urls) {
  page <- read_html(url)

  reviewers_name <- page %>% html_nodes(".display-name-link") %>% html_text()
  dates <- page %>% html_nodes("span.review-date") %>% html_text()
  user_ratings <- page %>% html_nodes("span.rating-other-user-rating") %>% html_text()
  text_reviews <- page %>% html_nodes("div.text") %>% html_text()

  tvshow_reviews_df <- data.frame(
    Reviewer_Name = reviewers_name[1:25],
    Date = dates[1:25],
    User_Rating = user_ratings[1:25],
    Text_Review = text_reviews[1:25]
  )
all_shows[[url]] <- tvshow_reviews_df

}
final_reviews_df <- do.call(rbind, all_shows)
colnames(final_reviews_df) <- c("Name", "Date of Review", "User Rating", "Text Reviews")
rownames(final_reviews_df) <- NULL


final_reviews_df
```

```
##                         Name    Date of Review
## 1                  khatcher-2  22 February 2004
## 2                      Cari-8      22 July 1999
## 3                 Steve_Nyland    26 August 2007
## 4                     phynigan 19 September 2005
## 5                   Cheese-18     27 March 2001
## 6                       Matum      8 March 1999
## 7                     dlevine      28 July 2004
## 8                       sfear 30 September 1998
## 9                  sent2null    15 August 2005
## 10             stephen_bounds  6 September 2004
## 11                  tcooke-2   24 January 2000
## 12                  Tweekums        6 May 2015
## 13                porphyrous   11 January 2005
## 14                    graphi   4 November 2006
## 15              timdalton007   8 November 2010
## 16             lee_eisenberg   20 January 2013
## 17               christian94  19 December 2007
## 18               dunmore_ego      21 July 2011
## 19                     Stoh80    8 October 2011
## 20             nickenchuggets       9 July 2021
## 21              coasterdude44   25 January 2021
## 22                 rmax304823   1 December 2009
## 23                      gring0  29 December 2005
## 24                        FKDZ      29 June 2022
## 25                   Blueghost     22 March 2009
## 26                     FiRE010       4 July 2021
## 27            Supermanfan-13   9 November 2021
## 28          TheLittleSongbird  13 November 2017
## 29               KinoKoopaKid      30 July 2021
## 30                jehuschultz  19 February 2020
## 31                gogoschka-1   12 January 2014
## 32                 bruhperson      7 March 2019
## 33             Leofwine_draca        5 May 2021
## 34                FishDrowned   9 November 2021
## 35         dhanushreddy-14919      18 July 2021
## 36              DiCaprioFan13   9 December 2022
## 37                    EVON1TY      3 March 2023
## 38                napierslogs       17 May 2010
## 39                  valis1949  26 February 2009
## 40             Leofwine_draca   7 December 2020
## 41                 Quinoa1984   8 February 2015
## 42              otnememento-2  21 February 2020
## 43          manishsingh-03299       31 May 2019
## 44                   FeastMode    18 August 2019
## 45             alanbenfieldjr      27 June 2017
## 46                     lerrom   11 October 2022
## 47                   xpinerhd  16 November 2019
## 48                fatcat-73450   3 February 2023
## 49       anthonymichael-93585   29 October 2022
## 50                       d_sa   13 October 2019
## 51                  rbverhoef  14 February 2003
## 52         philip_vanderveken 17 September 2004
## 53                  bsmith5552   6 November 2001
```

```
## 54              planktonrules       31 May 2015
## 55              DiCaprioFan13 28 September 2022
## 56                 theshape79   5 November 2001
## 57                 yodaschoda   24 January 2005
## 58                SnoopyStyle   5 December 2015
## 59                  grahamsj3  26 November 2002
## 60                   Libretio        4 May 2005
## 61             erwan_ticheler   8 December 2002
## 62                mickman91-1        4 May 2022
## 63                wildcatt268   19 January 2002
## 64                    arjay24     19 April 2004
## 65       faded_english_monkey    26 August 2004
## 66             Supermanfan-13        7 May 2022
## 67                  TusharViv     2 August 2022
## 68                  mhorg2018       6 June 2019
## 69                kipmcmillan   25 October 2018
## 70                bob the moo   7 February 2016
## 71            MovieCriticDave     13 April 2005
## 72                    jazmodo       4 June 2019
## 73                 paulimiles      25 July 2010
## 74               aucoinandrew       7 June 2020
## 75                       <NA>              <NA>
## 76             Leofwine_draca  28 November 2019
## 77                   jfirebug       21 May 2019
## 78                ahmetkozan       8 June 2019
## 79              deepfrieddodo  6 September 2022
## 80                    EVON1TY     13 March 2023
## 81                 emholberg       27 May 2019
## 82             Sleepin_Dragon       2 June 2019
## 83                SnoopyStyle       5 June 2019
## 84                    thegldt        7 May 2019
## 85                   Lladerat        8 May 2019
## 86  DimitrisPassas-TapTheLine      24 June 2019
## 87             classicsoncall   6 February 2020
## 88                     Vivkon        7 May 2019
## 89           justahunch-70549   10 October 2022
## 90                wmeduardowm        7 May 2019
## 91                      jazz1       30 May 2019
## 92            gregoryblanch88       31 May 2019
## 93              DiCaprioFan13 28 September 2022
## 94                   anyakiss        7 May 2019
## 95                 m-porpaczi       15 May 2019
## 96                    dannehh       4 June 2019
## 97                raggingbull       14 May 2019
## 98                   manga-th       12 May 2019
## 99           claudio_carvalho   6 February 2020
## 100              natashapekar       10 May 2019
## 101                mjplysaght   2 February 2017
## 102       A_Different_Drummer 18 September 2016
## 103                 cuzzinman 21 September 2018
## 104                Quinoa1984    16 August 2009
## 105               ericjcant-1   30 October 2005
## 106              KineticSeoul   11 October 2010
## 107            theblackscythe    15 August 2014
```

```
## 108           Bluereviews4you      29 July 2020
## 109           BadgerStorm4345       23 May 2020
## 110         TheLittleSongbird     8 August 2010
## 111            Trixie_reviews    3 October 2018
## 112                 salinaqvi       13 May 2020
## 113                 poseyfan      11 June 2022
## 114           Alex_Hodgkinson     13 April 2014
## 115             joshkej-84077  10 February 2018
## 116            BoxwoodExpress  23 February 2017
## 117   Rectangular_businessman      16 June 2010
## 118               haris_manda      17 June 2011
## 119             MissSimonetta     7 August 2014
## 120                  dee.reid     9 August 2006
## 121                   EVON1TY     13 March 2023
## 122                purplepoka       25 May 2020
## 123                doomedmac  21 December 2019
## 124            Styrophoamicus     6 August 2006
## 125                 antonwaas   29 January 2018
##                         User Rating
## 1    \n            9/10\n
## 2   \n            10/10\n
## 3   \n            10/10\n
## 4   \n            10/10\n
## 5   \n            10/10\n
## 6   \n            10/10\n
## 7   \n            10/10\n
## 8   \n            10/10\n
## 9   \n            10/10\n
## 10  \n            10/10\n
## 11  \n            10/10\n
## 12  \n            10/10\n
## 13  \n            10/10\n
## 14  \n            10/10\n
## 15   \n            9/10\n
## 16  \n            10/10\n
## 17  \n            10/10\n
## 18  \n            10/10\n
## 19   \n            9/10\n
## 20   \n            9/10\n
## 21   \n            9/10\n
## 22                                 <NA>
## 23                                 <NA>
## 24                                 <NA>
## 25                                 <NA>
## 26  \n            10/10\n
## 27  \n            10/10\n
## 28  \n            10/10\n
## 29  \n            10/10\n
## 30  \n            10/10\n
## 31  \n            10/10\n
## 32  \n            10/10\n
## 33  \n            10/10\n
## 34  \n            10/10\n
## 35  \n            10/10\n
```

```
## 36  \n            10/10\n
## 37  \n            10/10\n
## 38  \n            10/10\n
## 39  \n            10/10\n
## 40  \n            10/10\n
## 41  \n            10/10\n
## 42  \n            10/10\n
## 43  \n            10/10\n
## 44  \n            10/10\n
## 45  \n            10/10\n
## 46  \n            10/10\n
## 47  \n            10/10\n
## 48  \n            10/10\n
## 49  \n            10/10\n
## 50  \n            10/10\n
## 51  \n            10/10\n
## 52  \n            10/10\n
## 53  \n            10/10\n
## 54  \n            10/10\n
## 55  \n            10/10\n
## 56  \n            10/10\n
## 57  \n            10/10\n
## 58  \n            10/10\n
## 59  \n            10/10\n
## 60   \n            9/10\n
## 61  \n            10/10\n
## 62  \n            10/10\n
## 63  \n            10/10\n
## 64  \n            10/10\n
## 65  \n            10/10\n
## 66  \n            10/10\n
## 67  \n            10/10\n
## 68  \n            10/10\n
## 69  \n            10/10\n
## 70  \n            10/10\n
## 71  \n            10/10\n
## 72                              <NA>
## 73                              <NA>
## 74                              <NA>
## 75                              <NA>
## 76  \n            10/10\n
## 77  \n            10/10\n
## 78  \n            10/10\n
## 79  \n            10/10\n
## 80  \n            10/10\n
## 81   \n            9/10\n
## 82  \n            10/10\n
## 83   \n            9/10\n
## 84  \n            10/10\n
## 85  \n            10/10\n
## 86  \n            10/10\n
## 87  \n            10/10\n
## 88  \n            10/10\n
## 89  \n            10/10\n
```

```
## 90  \n            10/10\n
## 91  \n            10/10\n
## 92  \n            10/10\n
## 93   \n            9/10\n
## 94   \n            9/10\n
## 95  \n            10/10\n
## 96  \n            10/10\n
## 97  \n            10/10\n
## 98  \n            10/10\n
## 99   \n            9/10\n
## 100 \n            10/10\n
## 101 \n            10/10\n
## 102 \n            10/10\n
## 103 \n            10/10\n
## 104 \n            10/10\n
## 105 \n            10/10\n
## 106 \n            10/10\n
## 107 \n            10/10\n
## 108 \n            10/10\n
## 109 \n            10/10\n
## 110 \n            10/10\n
## 111 \n            10/10\n
## 112 \n            10/10\n
## 113 \n            10/10\n
## 114 \n            10/10\n
## 115 \n            10/10\n
## 116 \n            10/10\n
## 117 \n            10/10\n
## 118 \n            10/10\n
## 119 \n            10/10\n
## 120 \n            10/10\n
## 121  \n            9/10\n
## 122 \n            10/10\n
## 123 \n            10/10\n
## 124 \n            10/10\n
## 125 \n            10/10\n
##
## 1
## 2
## 3
## 4
## 5
## 6
## 7
## 8
## 9
## 10
## 11
## 12
## 13
## 14
## 15
## 16
## 17
```

```
## 18
## 19
## 20  Imagine if you will a tv series that manages to seamlessly combine elements of humor, fear, sadn
## 21
## 22
## 23
## 24
## 25
## 26
## 27
## 28
## 29
## 30
## 31
## 32
## 33
## 34
## 35
## 36
## 37
## 38
## 39
## 40
## 41
## 42
## 43
## 44
## 45
## 46
## 47
## 48
## 49
## 50
## 51
## 52
## 53
## 54
## 55
## 56
## 57
## 58
## 59
## 60
## 61
## 62
## 63
## 64
## 65
## 66
## 67
## 68
## 69
## 70
## 71
```

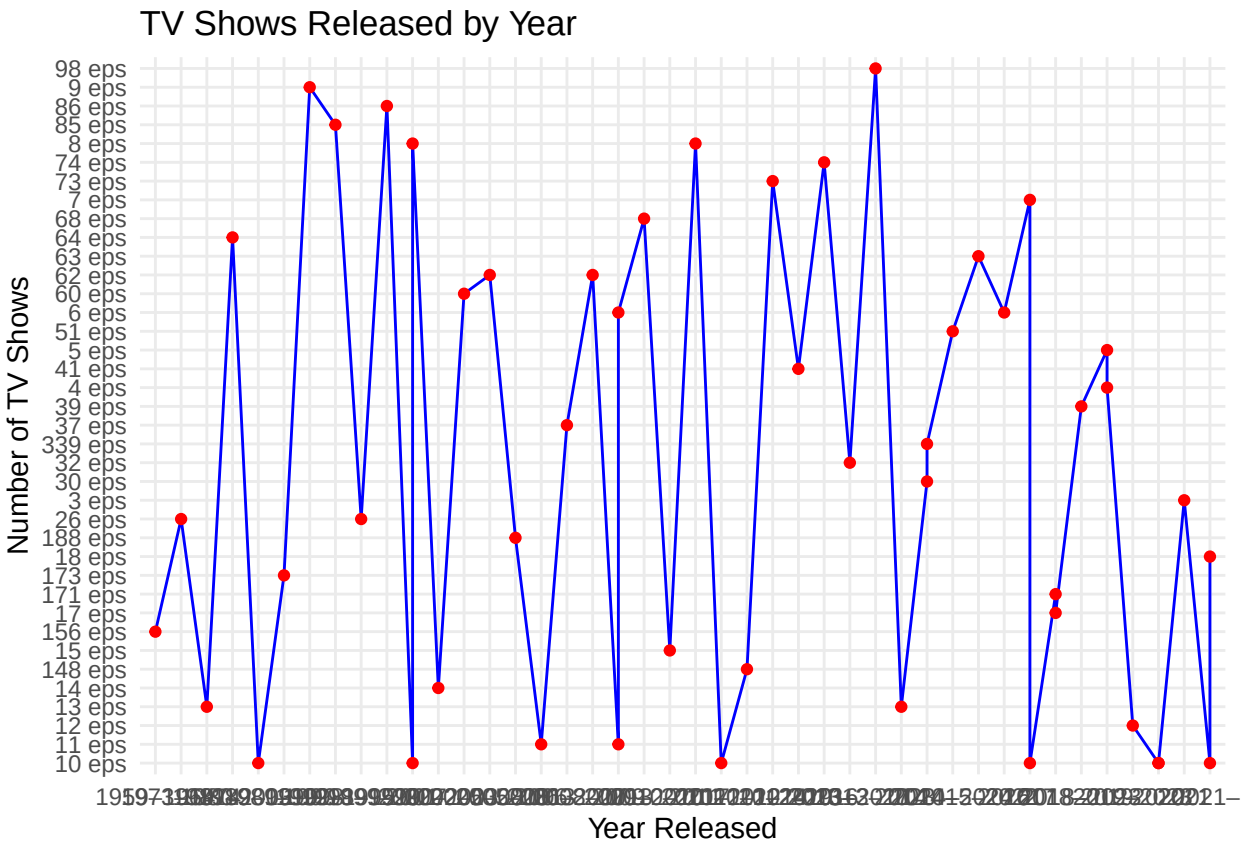
```
## 72
## 73
## 74
## 75
## 76
## 77
## 78
## 79
## 80
## 81
## 82
## 83
## 84
## 85
## 86
## 87
## 88
## 89
## 90
## 91
## 92
## 93
## 94
## 95
## 96
## 97
## 98
## 99
## 100
## 101
## 102
## 103
## 104
## 105
## 106
## 107
## 108
## 109
## 110
## 111
## 112
## 113
## 114
## 115
## 116
## 117
## 118
## 119
## 120
## 121
## 122
## 123
## 124
## 125
```

3. Create a time series graph for the tv shows released by year. Which year has the most number of tv shows released?

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.3.2
```

```
ggplot(final_df, aes(x = Year, y = Episode, group = 1)) +
  geom_line(color = "blue") +
  geom_point(color = "red") +
  labs(title = "TV Shows Released by Year",
       x = "Year Released",
       y = "Number of TV Shows") +
  theme_minimal()
```



4. Select 3 products from Amazon of the same category. Extract the price, description, ratings and reviews of each product.

```
url<-"https://www.amazon.com/s?k=tee+back+tank+tops&crid=1OB1A9SM5CVRK&sprefix=%2Caps%2C456&ref=nb_sb_s

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

```
## <polite session> https://www.amazon.com/s?k=tee+back+tank+tops&crid=1OB1A9SM5CVRK&sprefix=%2Caps%2C45
```

```
##     User-agent: Educational
##     robots.txt: 154 rules are defined for 4 bots
##    Crawl delay: 5 sec
##   The path is scrapable for this user-agent
```

```
prices<- character(0)
ratings<- character(0)

prices<-scrape(session)%>%
  html_nodes('span.a-price-whole')%>%
  html_text

ratings<-scrape(session)%>%
  html_nodes('span.a-icon-alt')%>%
  html_text

prices_ratings<-data.frame(prices[2:4],
                           ratings[2:4]

                           )


colnames(prices_ratings) <- c("Price", "Overall Rating");

teeback_urls<-c(
  "https://www.amazon.com/TELALEO-Athletic-Racerback-Activewear-Sleeveless/dp/B08T1Z7FP1/ref=sr_1_5?cri
  "https://www.amazon.com/Aeuui-Workout-Racerback-Shirts-Clothes/dp/B088CYZNKS/ref=sr_1_6?crid=1OB1A9SM
  "https://www.amazon.com/SOLY-HUX-Womens-Sleeveless-Shirts/dp/B0BYD8H1DZ/ref=sr_1_8?crid=1OB1A9SM5CVRK
)

all_reviews<-list()

for(url in teeback_urls){
  page<-read_html(url)

user_comment<-page%>%html_nodes('span.a-size-base.review-text')%>%
    html_text()


user_descr<-page%>%html_nodes('span.a-list-item.a-size-base.a-color-base')%>%
  html_text()

    user_comment_cleaned <- str_trim(user_comment[2:4])

  reviews.df<-data.frame(user_comment_cleaned,
                         user_descr[2:4])
all_reviews[[url]]<-reviews.df



}
colnames(reviews.df)<-c("Product Review", "Product Description")

do.call(rbind, reviews.df)
```

```
##                      [,1]
## Product Review      "This is really cute on. The pink is pretty. Soft fabric. Wore a cross back spor
## Product Description "Fabric: Made of soft and lightweight fabric, slight stretch, comfortable to wea
##                      [,2]
## Product Review      "This shirt appears well made, and it's pretty. But I'd size up twice. I bought a
## Product Description "Feature: Sleeveless, round neck, criss cross, solid color, fashion tank top for
##                      [,3]
## Product Review      "It's like a bathing suit tried to have a baby with an XXS tank top on top but ar
## Product Description "Size recommendation: Please refer to size chart which we provide in our photos
```

```
View(reviews.df)


final_amazon_scrape<-cbind(prices_ratings,reviews.df)
final_amazon_scrape
```

```
##   Price     Overall Rating
## 1   40. 4.4 out of 5 stars
## 2   17. 4.5 out of 5 stars
## 3   28. 4.2 out of 5 stars
##
## 1
## 2 This shirt appears well made, and it's pretty. But I'd size up twice. I bought a large, and it's si
## 3
##
## 1                                                                  Fabric: Made of soft and lightweight
## 2             Feature: Sleeveless, round neck, criss cross, solid color, fashion tank top for women,
## 3 Size recommendation: Please refer to size chart which we provide in our photos (The product measure
```