

# Web Scraping in R

```
listpage <- read_html("https://www.metacritic.com/browse/games/score/metascore/90day/ps4/filtered") #Ti
titlelist<- html_nodes(listpage, "#main .product_title a")#just titles
titles<-html_text(titlelist)
titles# as html
```

```
## [1] "\n      NieR: Automata - Game of the YoRHa Edition\n## [2] "\n      Sekiro: Shadows Die Twice\n## [3] "\n      Devil May Cry 5\n## [4] "\n      Steins;Gate Elite\n## [5] "\n      MLB The Show 19\n## [6] "\n      DiRT Rally 2.0\n## [7] "\n      Mortal Kombat 11\n## [8] "\n      Tom Clancy's The Division 2\n## [9] "\n      The Legend of Heroes: Trails of Cold Steel\n## [10] "\n      The Messenger\n## [11] "\n      Ghost Giant\n## [12] "\n      Metro Exodus\n## [13] "\n      Phoenix Wright: Ace Attorney Trilogy\n## [14] "\n      Trials Rising\n## [15] "\n      Pode\n## [16] "\n      Stellaris: Console Edition\n## [17] "\n      Borderlands: Game of the Year Edition\n## [18] "\n      Fate/Extella Link\n## [19] "\n      My Time At Portia\n## [20] "\n      The Walking Dead: The Telltale Series - The Final Season Episode\n## [21] "\n      God's Trigger\n## [22] "\n      Zanki Zero: Last Beginning\n## [23] "\n      SNK 40th Anniversary Collection\n## [24] "\n      Assassin's Creed III Remastered\n## [25] "\n      Dead or Alive 6\n## [26] "\n      Vaporum\n## [27] "\n      Death end re;Quest\n## [28] "\n      Days Gone\n## [29] "\n      Falcon Age\n## [30] "\n      Far Cry New Dawn\n## [31] "\n      God Eater 3\n## [32] "\n      Heaven's Vault\n## [33] "\n      Chocobo's Mystery Dungeon: Every Buddy!\n## [34] "\n      The Liar Princess and the Blind Prince\n## [35] "\n      Earth Defense Force: Iron Rain\n## [36] "\n      The Wizards: Enhanced Edition\n## [37] "\n      Degrees of Separation\n## [38] "\n      Nelke & the Legendary Alchemists: Ateliers of the New World\n## [39] "\n      World War Z\n## [40] "\n      ToeJam & Earl: Back in the Groove\n## [41] "\n      Xenon Racer\n## [42] "\n      Conarium\n## [43] "\n      Arcade Classics Anniversary Collection\n## [44] "\n      The Caligula Effect: Overdose\n## [45] "\n      ChromaGun VR\n## [46] "\n      Space Junkies
```

```
## [47] "\n                The Occupation\n
## [48] "\n                Intruders: Hide and Seek\n
## [49] "\n                Dangerous Driving\n
## [50] "\n                Outward\n
## [51] "\n                Massira\n
## [52] "\n                Jupiter & Mars\n
## [53] "\n                One Piece: World Seeker\n
## [54] "\n                The LEGO Movie 2 Videogame\n
## [55] "\n                Jump Force\n
## [56] "\n                Anthem\n
## [57] "\n                Fade to Silence\n
## [58] "\n                Fimbul\n
## [59] "\n                Eden-Tomorrow\n
## [60] "\n                Generation Zero\n
## [61] "\n                Left Alive\n
## [62] "\n                Eternity: The Last Unicorn
```

```
for(i in 1:62){
  titles[i]= str_sub(titles[i], 30, -54)#format of titles
}
```

```
Full.df<-data.frame("Title"=1:62,"MetaScore"=1:62,"User"=1:62,"Rating"=1:62)
for(i in 1:62){
```

```
  urltitle<-titles[i]
  urltitle<-str_replace_all(urltitle,"&","")
  urltitle<-str_replace_all(urltitle,":","")#changes colons to delete
  urltitle<-str_replace_all(urltitle," "," ")#changes double space to space
  urltitle<-str_replace_all(urltitle," ","-")#changes spaces to dashes
  urltitle<-str_replace_all(urltitle,";","")#deletes semicolons
  urltitle<-str_replace_all(urltitle,"\\.","")
  urltitle<-str_replace_all(urltitle,"'", "")
  urltitle<-str_replace_all(urltitle,"/","")
```

```
  url<-str_c("https://www.metacritic.com/game/playstation-4/",urltitle)#combines strings
  url<-tolower(url)#puts into lower case
```

```
  url<-read_html(url)#so this is the url needed for the individual titles
```

```
  data<- html_nodes(url,".large , h1 , .product_rating .data , .game span")#green yellow and red have d
```

```
  textdata<-html_text(data)
```

```
  Full.df[i,1]=textdata[1]#math to make sure 1,5,11 go to section 1
```

```
  Full.df[i,2]=textdata[2]
```

```
  Full.df[i,3]=textdata[3]
```

```
  Full.df[i,4]=textdata[5]#some games seem to not have ratings
```

```
  print(textdata)
```

```
}
```

```
## [1] "NieR: Automata - Game of the YoRHa Edition"
```

```
## [2] "91"
```

```
## [3] "8.4"
```

```
## [4] "tbd"
```

```
## [5] "M"
```

```
## [1] "Sekiro: Shadows Die Twice" "90"
```

```
## [3] "7.9" "7.9"
```

```
## [5] "M"
```

```
## [1] "Devil May Cry 5" "88" "8.4" "tbd"
```

```

## [5] "M"
## [1] "Steins;Gate Elite" "86" "7.4"
## [4] "7.4" "M"
## [1] "MLB The Show 19" "86" "8.0" "8.0"
## [5] "E"
## [1] "DiRT Rally 2.0" "84" "5.7" "5.7"
## [5] "E"
## [1] "Mortal Kombat 11" "83" "2.8"
## [4] "tbd" "M"
## [1] "Tom Clancy's The Division 2" "82"
## [3] "7.2" "7.2"
## [5] "M"
## [1] "The Legend of Heroes: Trails of Cold Steel"
## [2] "82"
## [3] "9.0"
## [4] "tbd"
## [5] "T"
## [1] "The Messenger" "81" "7.8" "7.8"
## [5] "E10+"
## [1] "Ghost Giant" "80" "7.6" "tbd" "E"
## [1] "Metro Exodus" "80" "6.7" "6.7"
## [5] "M"
## [1] "Phoenix Wright: Ace Attorney Trilogy"
## [2] "80"
## [3] "7.3"
## [4] "tbd"
## [5] "T"
## [1] "Trials Rising" "79" "6.3" "6.3"
## [5] "T"
## [1] "Pode" "77" "4.4" "4.4"
## [1] "Stellaris: Console Edition" "77"
## [3] "6.5" "tbd"
## [5] "E10+"
## [1] "Borderlands: Game of the Year Edition"
## [2] "76"
## [3] "6.0"
## [4] "6.0"
## [5] "M"
## [1] "Fate/Extella Link" "75" "6.9"
## [4] "6.9" "T"
## [1] "My Time At Portia" "75" "5.4"
## [4] "tbd" "E10+"
## [1] "The Walking Dead: The Telltale Series - The Final Season Episode 4: Take Us Back"
## [2] "75"
## [3] "7.5"
## [4] "tbd"
## [1] "God's Trigger" "73" "7.7" "7.7"
## [1] "Zanki Zero: Last Beginning" "73"
## [3] "6.4" "6.4"
## [5] "M"
## [1] "SNK 40th Anniversary Collection" "73"
## [3] "5.0" "5.0"
## [5] "T"
## [1] "Assassin's Creed III Remastered" "72"

```

```

## [3] "4.3"                                "tbd"
## [5] "M"
## [1] "Dead or Alive 6" "72"                "5.4"                "5.4"
## [5] "M"
## [1] "Vaporum" "72"                "0.4"                "0.4"
## [1] "Death end re;Quest" "72"                "7.5"
## [4] "tbd"                "M"
## [1] "Days Gone" "72"                "8.0"                "tbd"                "M"
## [1] "Falcon Age" "72"                "6.5"                "6.5"                "T"
## [1] "Far Cry New Dawn" "71"                "3.6"
## [4] "3.6"                "M"
## [1] "God Eater 3" "71"                "7.2"                "tbd"                "T"
## [1] "Heaven's Vault" "71"                "4.3"                "4.3"
## [1] "Chocobo's Mystery Dungeon: Every Buddy!"
## [2] "69"
## [3] "7.5"
## [4] "7.5"
## [5] "E10+"
## [1] "The Liar Princess and the Blind Prince"
## [2] "69"
## [3] "6.4"
## [4] "6.4"
## [5] "E10+"
## [1] "Earth Defense Force: Iron Rain" "69"
## [3] "6.3"                "tbd"
## [1] "The Wizards: Enhanced Edition" "69"
## [3] "6.6"                "6.6"
## [5] "E10+"
## [1] "Degrees of Separation" "68"                "4.8"
## [4] "tbd"
## [1] "Nelke & the Legendary Alchemists: Ateliers of the New World"
## [2] "68"
## [3] "3.6"
## [4] "tbd"
## [5] "E10+"
## [1] "World War Z" "66"                "6.2"                "tbd"                "M"
## [1] "ToeJam & Earl: Back in the Groove" "66"
## [3] "6.2"                "tbd"
## [5] "E10+"
## [1] "Xenon Racer" "65"                "5.8"                "5.8"                "E"
## [1] "Conarium" "65"                "2.7"                "2.7"                "T"
## [1] "Arcade Classics Anniversary Collection"
## [2] "64"
## [3] "4.0"
## [4] "tbd"
## [5] "T"
## [1] "The Caligula Effect: Overdose" "64"
## [3] "5.6"                "5.6"
## [1] "ChromaGun VR" "64"                "5.9"                "5.9"
## [1] "Space Junkies" "63"                "tbd"                "tbd"
## [5] "T"
## [1] "The Occupation" "63"                "3.5"                "3.5"
## [5] "T"
## [1] "Intruders: Hide and Seek" "62"

```

```

## [3] "6.6" "tbd"
## [1] "Dangerous Driving" "62" "4.1"
## [4] "tbd" "E10+"
## [1] "Outward" "62" "6.9" "tbd" "M"
## [1] "Massira" "60" "6.7" "6.7"
## [1] "Jupiter & Mars" "60" "5.5" "5.5"
## [5] "E"
## [1] "One Piece: World Seeker" "58"
## [3] "3.5" "3.5"
## [5] "T"
## [1] "The LEGO Movie 2 Videogame" "57"
## [3] "1.5" "tbd"
## [5] "E10+"
## [1] "Jump Force" "56" "4.4" "tbd" "T"
## [1] "Anthem" "54" "3.5" "3.5" "T"
## [1] "Fade to Silence" "54" "3.4" "tbd"
## [5] "M"
## [1] "Fimbul" "52" "1.8" "1.8"
## [1] "Eden-Tomorrow" "52" "6.1" "6.1"
## [5] "T"
## [1] "Generation Zero" "46" "5.4" "tbd"
## [5] "T"
## [1] "Left Alive" "37" "8.4" "tbd" "M"
## [1] "Eternity: The Last Unicorn" "36"
## [3] "5.3" "5.3"

```