

Praesentation

Márk Reichmann, Simon Keil, Daniel Henke

6 1 2022

Tibble modifizieren (1)

unser Datensatz (die interessanten Variablen):

```
## # A tibble: 15,480 x 4
##   Title                `Release Date` Genre      `Country Availability`
##   <chr>                <chr>      <chr>      <chr>
## 1 Lets Fight Ghost    12 Dec 2008 Crime, Drama~ Thailand
## 2 HOW TO BUILD A GIRL 08 May 2020 Comedy      Canada
## 3 Centigrade          28 Aug 2020 Drama, Thril~ Canada
## 4 ANNE+               01 Oct 2016 Drama        Belgium,Netherlands
## 5 Moxie               22 Sep 2011 Animation, S~ Lithuania,Poland,France,Ice~
## 6 The Con-Heartist    03 Dec 2020 Comedy, Roma~ Thailand
## 7 Gleboka woda        14 Jun 2011 Drama        Poland
## 8 Instynkt            03 Mar 2011 Crime        Poland
## 9 Only a Mother       31 Oct 1949 Drama        Lithuania,Poland,France,Ita~
## 10 Snowroller         04 Oct 1985 Comedy      Lithuania,Poland,France,Ita~
## # ... with 15,470 more rows
```

Tibble modifizieren (2)

Release Date und Title sind nur zusammen eindeutig

```
## # A tibble: 2 x 3
##   Title                                `Release Date` `Series or Movie`
##   <chr>                                <chr>          <chr>
## 1 Cloudy with a Chance of Meatballs 20 Feb 2017   Series
## 2 Cloudy with a Chance of Meatballs 18 Sep 2009   Movie
```

Tibble modifizieren (3)

leider gibt es Duplikate:

```
daten %>%  
  count(Title, `Release Date`)%>%  
  filter(n>1)
```

```
## # A tibble: 148 x 3  
##   Title                                `Release Date`      n  
##   <chr>                                <chr>            <int>  
## 1 A Bride for Rip Van Winkle          26 Mar 2016        2  
## 2 A Love So Beautiful                  09 Nov 2017        2  
## 3 A Perfect Day                        15 Jan 2016        2  
## 4 About Time                           08 Nov 2013        2  
## 5 Aftermath                            07 Apr 2017        2  
## 6 Aliens                               18 Jul 1986        2  
## 7 Alone                                07 Nov 2008        2  
## 8 Anastasia                            21 Nov 1997        2  
## 9 Anohana: The Flower We Saw That Day 14 Apr 2011        2  
## 10 Another                             09 Jan 2012        2  
## # ... with 138 more rows
```

Tibble modifizieren (4)

das sind diejenigen Filme, die mehrmals auf Netflix releast wurden

```
## # A tibble: 2 x 4
##   Title      `Release Date` Genre      `Country Availability`
##   <chr>      <chr>      <chr>      <chr>
## 1 A Perfect Day 15 Jan 2016 Comedy, Drama, War Italy
## 2 A Perfect Day 15 Jan 2016 Comedy, Drama, War Romania,Poland
```

Lösung:

```
daten %>%
  filter(Title == "A Perfect Day") %>%
  select(Title, `Release Date`, Genre, `Country Availability`)%>%
  separate_rows(`Country Availability`, sep = ",")%>%
  distinct(.keep_all = TRUE)
```

```
## # A tibble: 3 x 4
##   Title      `Release Date` Genre      `Country Availability`
##   <chr>      <chr>      <chr>      <chr>
## 1 A Perfect Day 15 Jan 2016 Comedy, Drama, War Italy
## 2 A Perfect Day 15 Jan 2016 Comedy, Drama, War Romania
## 3 A Perfect Day 15 Jan 2016 Comedy, Drama, War Poland
```

Tibble modifizieren (5)

selbiges auch mit Genre

```
## # A tibble: 3 x 3
##   Title      `Release Date` `Country Availability`
##   <chr>      <chr>          <chr>
## 1 A Perfect Day 15 Jan 2016    Italy
## 2 A Perfect Day 15 Jan 2016    Romania
## 3 A Perfect Day 15 Jan 2016    Poland

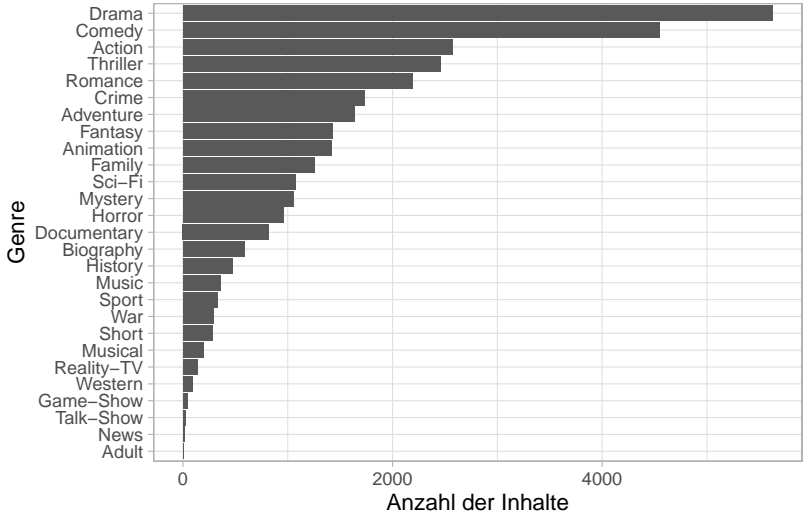
## # A tibble: 3 x 3
##   Title      `Release Date` Genre
##   <chr>      <chr>          <chr>
## 1 A Perfect Day 15 Jan 2016    Comedy
## 2 A Perfect Day 15 Jan 2016    Drama
## 3 A Perfect Day 15 Jan 2016    War
```

diese beiden Tibbles werden jetzt an Title, Release Date gejoint

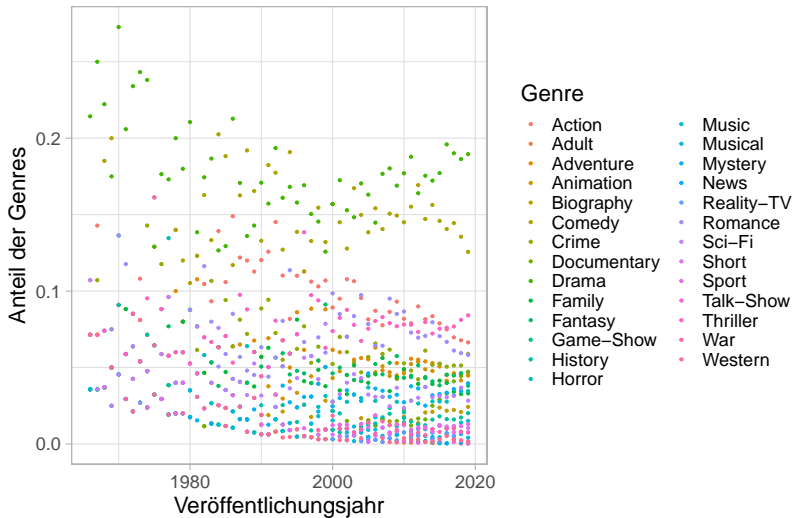
Tibble modifizieren (6)

```
## # A tibble: 9 x 4
##   Title      `Release Date` `Country Availability` Genre
##   <chr>      <chr>          <chr>          <chr>
## 1 A Perfect Day 15 Jan 2016    Italy          Comedy
## 2 A Perfect Day 15 Jan 2016    Italy          Drama
## 3 A Perfect Day 15 Jan 2016    Italy          War
## 4 A Perfect Day 15 Jan 2016    Romania        Comedy
## 5 A Perfect Day 15 Jan 2016    Romania        Drama
## 6 A Perfect Day 15 Jan 2016    Romania        War
## 7 A Perfect Day 15 Jan 2016    Poland         Comedy
## 8 A Perfect Day 15 Jan 2016    Poland         Drama
## 9 A Perfect Day 15 Jan 2016    Poland         War
```

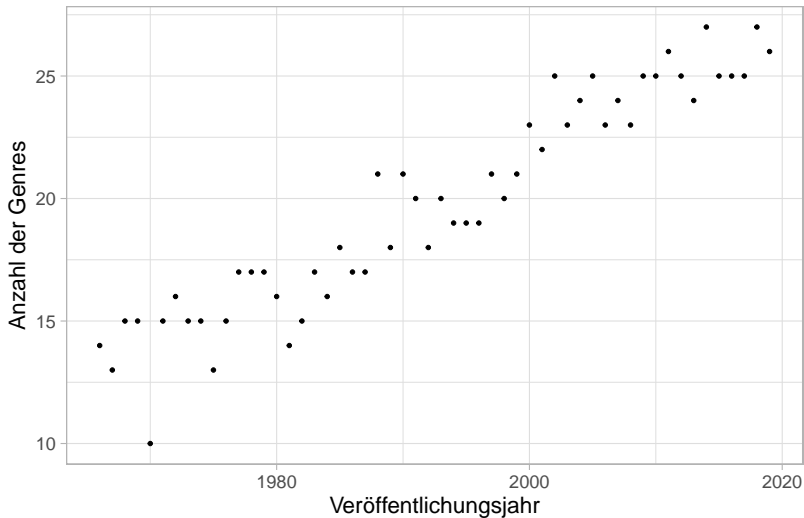
Ranking der Genres



Anteil der verfügbaren Genres nach Veröffentlichungsjahr



Anzahl der Genres nach Veröffentlichungsjahr



Anzahl der Genres nach Veröffentlichungsjahr

