

R Notebook: IMDB (predict gender from biography)

Contents

(IMDB) Custom library	2
Load data	3
Data Dictionary	3
Movie Count by Year	4
Lists: movies	4
Lists: actors	6
Lists: headliners	7
Companies	8
Creatives	9
Actors: Info (e.g., Bio)	10
Actor on Movie: Rank & Character Name	12
Movie Info: from Actor Search Page	13
Actor's movie Rank	15
Glue Table: Top-Creatives	15
Extra data	16
Will Smith	21
Movies	21
Rankings: Lots of Ways to Internally Rank	30
Unique Character Names	43
Gender	43
Actor Ranking on Given Film	45
Will's Rankings on Some Films	45
Movie Details	47
Movie Details: Extra	49
Denzel Washington	82
Conclusion	82

```
library(devtools);

## Loading required package: usethis
library(humanVerseWSU);

path.github = "https://raw.githubusercontent.com/MonteShaffer/humanVerseWSU/master/";

include.me = paste0(path.github, "misc/functions-nlp.R");
source_url( include.me );

## SHA-1 hash of file is 704afa69d52215d315cb5f59cdc020b0bbfd0b13

## Warning: package 'tm' was built under R version 4.0.3

## Loading required package: NLP

## Warning: package 'NLP' was built under R version 4.0.3

## Warning: package 'SentimentAnalysis' was built under R version 4.0.3
```

```
##
## Attaching package: 'SentimentAnalysis'

## The following object is masked from 'package:base':
##
##      write

include.me = paste0(path.github, "misc/functions-nlp-str.R");
source_url( include.me );

## SHA-1 hash of file is 6bdb234fa84eea995969dc29d6ff2a78f3982131
include.me = paste0(path.github, "misc/functions-nlp-stack.R");
source_url( include.me );

## SHA-1 hash of file is 034efbce0405954198545f8798e119b77a4809c9
include.me = paste0(path.github, "misc/functions-nlp-pos.R");
source_url( include.me );

## SHA-1 hash of file is d8c8cf01c8ead1b6d4228891aa52bac77084a6e7

## Warning: package 'openNLP' was built under R version 4.0.3
include.me = paste0(path.github, "humanVerseWSU/R/functions-encryption.R");
source_url( include.me );

## SHA-1 hash of file is da71dde620bed33db055778b752eefb476f7bf6b
path.to.nascent = "C:/Users/Alexander Nevsky/Dropbox/WSU-419/Fall 2020/__student_access__/unit_02_confidential";
folder.nlp = "nlp/";
path.to.nlp = paste0(path.to.nascent, folder.nlp);

##### UPDATES TO dataframe subset function #####
# inflation adjustments for NA ... and improvements on subsetting
include.me = paste0(path.github, "humanVerseWSU/R/functions-dataframe.R");
source_url( include.me );

## SHA-1 hash of file is 1149cbf3e865f692b50d4d1983e6364dc56ce62d
include.me = paste0(path.github, "humanVerseWSU/R/functions-inflation.R");
source_url( include.me );

## SHA-1 hash of file is b6d29327e3fe030ca132b135f4a89b6fc6a61a66
```

(IMDB) Custom library

This is a large dataset I harvested in September. It will allow us to explore more comprehensively the relationships of various features of the movie database. It is large (about 50MB), so installing may take some time if you are on a slow internet connection.

This dataset will be the source you will use on your final exam to answer the question posed earlier in the semester about Will Smith and Denzel Washington. You now have more analytics skills and with the new dataset there are more features you can extract.

```
# library(devtools);
# install_github("MonteShaffer/imdb/imdb"); # choose #3 to humanVerseWSU
# detach(package:imdb);
```

```
library(imdb);  
packageVersion("imdb"); # '0.1.1'
```

```
## [1] '0.1.1'
```

```
# ?loadDataIMDB
```

Load data

Once this is run, a lot of memory will be required to read in the 23 compressed files.

```
imdb::loadDataIMDB();  
names(imdb.data);
```

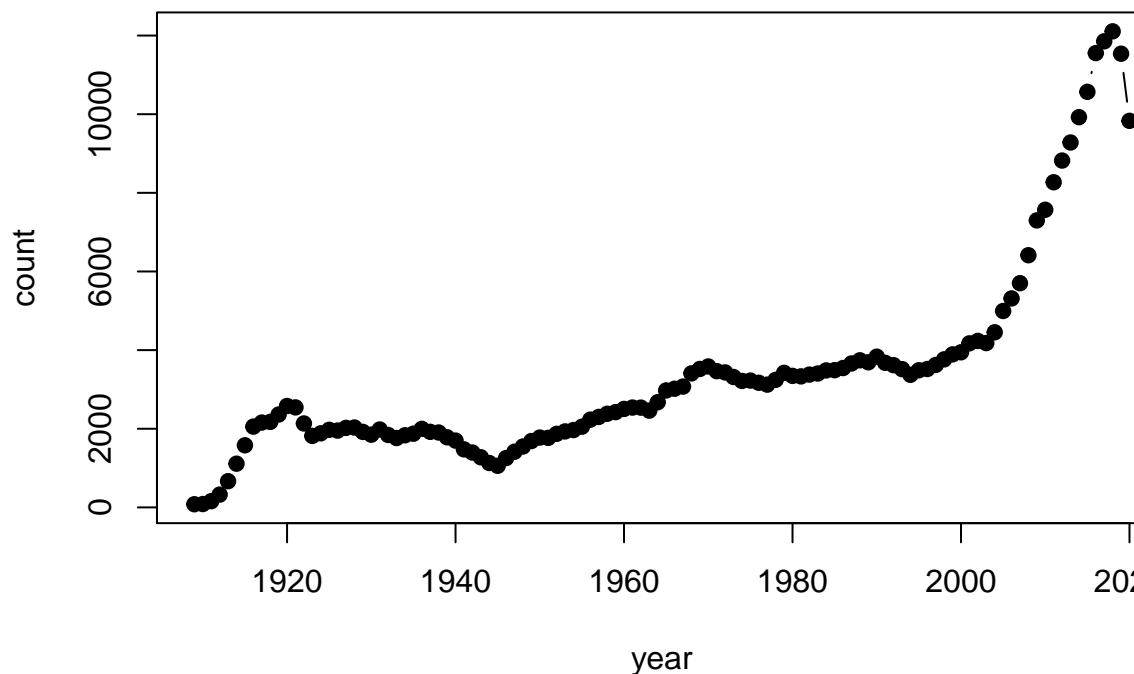
```
## [1] "all.movies.creatives"      "all.movies.companies"  
## [3] "all.movies.actors.characters" "all.actors.rank"  
## [5] "all.actors.movies"        "all.actors.info"  
## [7] "moviecount.byyear"        "actors"  
## [9] "glue"                     "headliners"  
## [11] "movies"                   "movies.df"
```

```
humanVerseWSU::loadInflationData();
```

A global variable is automatically created with several dataframes.

Data Dictionary

```
plot(imdb.data$moviecount.byyear, type="b", pch=19);
```



Movie Count by Year

Even though I harvested this data in September 2020, it seems there was a movie fall-off in 2019, before COVID. Often this can be understood using change-point analyses. Maybe it is linked to a “Trump-factor”; the media suggests that money is moved from block-buster movies into streaming options (Netflix, and so on). It is an interesting question.

Lists: movies

- all :: defined as the unique “ttid”s of all other lists

```
length(imdb.data$movies$all);
```

```
## [1] 74520
```

```
head(imdb.data$movies$all, 20);
```

```
## [1] "tt0000001" "tt0000002" "tt0000003" "tt0000004" "tt0000005" "tt0000006"
## [7] "tt0000007" "tt0000008" "tt0000009" "tt0000010" "tt0000011" "tt0000012"
## [13] "tt0000013" "tt0000014" "tt0000015" "tt0000016" "tt0000017" "tt0000018"
## [19] "tt0000019" "tt0000020"
```

- top250 :: < <https://www.imdb.com/chart/top>> Accessed September 2020.

```
length(imdb.data$movies$top250);
```

```
## [1] 250
```

```
head(imdb.data$movies$top250, 20);
```

```
## [1] "tt0012349" "tt0015324" "tt0015864" "tt0017136" "tt0017925" "tt0018773"
## [7] "tt0019254" "tt0021749" "tt0022100" "tt0025316" "tt0027977" "tt0031381"
```

```
## [13] "tt0031679" "tt0032553" "tt0032976" "tt0033467" "tt0034583" "tt0035446"
## [19] "tt0036775" "tt0038650"
```

- popular50 (dataframe) :: top-50 most popular for each year (by votes)

```
dim(imdb.data$movies$popular50);
```

```
## [1] 5704 3
```

```
names(imdb.data$movies$popular50);
```

```
## [1] "ttid" "year" "rank.popular"
```

```
head(imdb.data$movies$popular50, 20);
```

```
##      ttid year rank.popular
## 1  tt0219574 1898          1
## 2  tt0229287 1898          2
## 3  tt0221384 1898          3
## 4  tt0229399 1898          4
## 5  tt0237889 1898          5
## 6  tt0229764 1898          6
## 7  tt0229389 1898          7
## 8  tt0221654 1898          8
## 9  tt0317653 1898          9
## 10 tt0237861 1898         10
## 11 tt0237775 1898         11
## 12 tt0138538 1898         12
## 13 tt0317652 1898         13
## 14 tt0317654 1898         14
## 15 tt0229690 1898         15
## 16 tt0240795 1898         16
## 17 tt0156186 1899          1
## 18 tt0263933 1899          2
## 19 tt0242588 1899          3
## 20 tt0242282 1899          4
```

- gem50 (dataframe) :: top-50 highest rated for each year (by ratings)

```
dim(imdb.data$movies$gem50);
```

```
## [1] 5704 3
```

```
names(imdb.data$movies$gem50);
```

```
## [1] "ttid" "year" "rank.gem"
```

```
head(imdb.data$movies$gem50, 20);
```

```
##      ttid year rank.gem
## 1  tt0219574 1898          1
## 10 tt0237861 1898          2
## 2  tt0229287 1898          3
## 12 tt0138538 1898          4
## 14 tt0317654 1898          5
## 5  tt0237889 1898          6
## 11 tt0237775 1898          7
## 15 tt0229690 1898          8
## 8  tt0221654 1898          9
```

```
## 3  tt0221384 1898      10
## 6  tt0229764 1898      11
## 13 tt0317652 1898      12
## 7   tt0229389 1898      13
## 9   tt0317653 1898      14
## 16 tt0240795 1898      15
## 4   tt0229399 1898      16
## 20 tt0242282 1899       1
## 18 tt0263933 1899       2
## 19 tt0242588 1899       3
## 17 tt0156186 1899       4
```

- tt5000 :: numeric movie ids from 1:5000 several bad ones ... length is less than that.

```
length(imdb.data$movies$tt5000);
```

```
## [1] 4942
```

```
head(imdb.data$movies$tt5000, 20);
```

```
## [1] "tt0000001" "tt0000002" "tt0000003" "tt0000004" "tt0000005" "tt0000006"
## [7] "tt0000007" "tt0000008" "tt0000009" "tt0000010" "tt0000011" "tt0000012"
## [13] "tt0000013" "tt0000014" "tt0000015" "tt0000016" "tt0000017" "tt0000018"
## [19] "tt0000019" "tt0000020"
```

Lists: actors

- all :: defined as the unique “nmid”s of all other lists

```
length(imdb.data$actors$all);
```

```
## [1] 71341
```

```
head(imdb.data$actors$all, 20);
```

```
## [1] "nm0000001" "nm0000002" "nm0000003" "nm0000004" "nm0000005" "nm0000006"
## [7] "nm0000007" "nm0000008" "nm0000009" "nm0000010" "nm0000011" "nm0000012"
## [13] "nm0000013" "nm0000014" "nm0000015" "nm0000016" "nm0000017" "nm0000018"
## [19] "nm0000019" "nm0000020"
```

- top250 :: < <https://www.imdb.com/chart/top> > Accessed September 2020. These are the actors on those movies

```
length(imdb.data$actors$top250);
```

```
## [1] 3092
```

```
head(imdb.data$actors$top250, 20);
```

```
## [1] "nm0000002" "nm0000006" "nm0000007" "nm0000008" "nm0000012" "nm0000017"
## [7] "nm0000018" "nm0000020" "nm0000021" "nm0000022" "nm0000023" "nm0000024"
## [13] "nm0000026" "nm0000027" "nm0000032" "nm0000034" "nm0000036" "nm0000037"
## [19] "nm0000038" "nm0000044"
```

- popular50 :: top-50 most popular for each year (by votes)

```
length(imdb.data$actors$popular50);
```

```
## [1] 35245
```

```
head(imdb.data$actors$popular50, 20);
```

```
## [1] "nm0000001" "nm0000002" "nm0000003" "nm0000004" "nm0000006" "nm0000007"
## [7] "nm0000008" "nm0000009" "nm0000010" "nm0000011" "nm0000012" "nm0000013"
## [13] "nm0000014" "nm0000015" "nm0000017" "nm0000018" "nm0000020" "nm0000021"
## [19] "nm0000022" "nm0000023"
```

- gem50 :: top-50 highest rated for each year (by ratings)

```
length(imdb.data$actors$gem50);
```

```
## [1] 37997
```

```
head(imdb.data$actors$gem50, 20);
```

```
## [1] "nm0000001" "nm0000002" "nm0000006" "nm0000007" "nm0000008" "nm0000010"
## [7] "nm0000011" "nm0000012" "nm0000014" "nm0000017" "nm0000018" "nm0000019"
## [13] "nm0000020" "nm0000021" "nm0000022" "nm0000023" "nm0000024" "nm0000026"
## [19] "nm0000027" "nm0000028"
```

- tt5000 :: numeric movie ids from 1:5000 several bad ones ... length is less than that.

```
length(imdb.data$actors$tt5000);
```

```
## [1] 5452
```

```
head(imdb.data$actors$tt5000, 20);
```

```
## [1] "nm0000122" "nm0000252" "nm0000406" "nm0000428" "nm0000779" "nm0000856"
## [7] "nm0000858" "nm0000859" "nm0000875" "nm0000891" "nm0001273" "nm0001316"
## [13] "nm0001366" "nm0001908" "nm0002001" "nm0002031" "nm0002034" "nm0002061"
## [19] "nm0002118" "nm0002154"
```

- nm5000 :: numeric actor ids from 1:5000 several bad ones ... length is less than that.

```
length(imdb.data$actors$nm5000);
```

```
## [1] 4968
```

```
head(imdb.data$actors$nm5000, 20);
```

```
## [1] "nm0000001" "nm0000002" "nm0000003" "nm0000004" "nm0000005" "nm0000006"
## [7] "nm0000007" "nm0000008" "nm0000009" "nm0000010" "nm0000011" "nm0000012"
## [13] "nm0000013" "nm0000014" "nm0000015" "nm0000016" "nm0000017" "nm0000018"
## [19] "nm0000019" "nm0000020"
```

- headliners :: actors that were the “headliners” in 15 or more movies (from `imdb.data$glue$movies.creatives`). This is equivalent to `imdb.data$headliners$actors`.

```
length(imdb.data$actors$headliners);
```

```
## [1] 2407
```

```
head(imdb.data$actors$headliners, 20);
```

```
## [1] "nm0000001" "nm0000002" "nm0000003" "nm0000006" "nm0000007" "nm0000008"
## [7] "nm0000009" "nm0000010" "nm0000011" "nm0000012" "nm0000013" "nm0000014"
## [13] "nm0000017" "nm0000018" "nm0000020" "nm0000021" "nm0000022" "nm0000023"
## [19] "nm0000024" "nm0000026"
```

Lists: headliners A headliner has a lot of experience as a top-3 (either writer, director, actor, or company);

- actors :: actors that were the “headliners” in 15 or more movies

```
length(imdb.data$headliners$actors);
```

```
## [1] 2407
```

```
head(imdb.data$headliners$actors, 20);
```

```
## [1] "nm0000001" "nm0000002" "nm0000003" "nm0000006" "nm0000007" "nm0000008"
## [7] "nm0000009" "nm0000010" "nm0000011" "nm0000012" "nm0000013" "nm0000014"
## [13] "nm0000017" "nm0000018" "nm0000020" "nm0000021" "nm0000022" "nm0000023"
## [19] "nm0000024" "nm0000026"
```

- directors :: directors that were the “headliners” in 15 or more movies

```
length(imdb.data$headliners$directors);
```

```
## [1] 606
```

```
head(imdb.data$headliners$directors, 20);
```

```
## [1] "nm0000005" "nm0000033" "nm0000036" "nm0000041" "nm0000080" "nm0000095"
## [7] "nm0000108" "nm0000110" "nm0000118" "nm0000122" "nm0000127" "nm0000142"
## [13] "nm0000165" "nm0000180" "nm0000217" "nm0000229" "nm0000231" "nm0000265"
## [19] "nm0000318" "nm0000338"
```

- writers :: writers that were the “headliners” in 15 or more movies

```
length(imdb.data$headliners$writers);
```

```
## [1] 670
```

```
head(imdb.data$headliners$writers, 20);
```

```
## [1] "nm0000005" "nm0000019" "nm0000041" "nm0000076" "nm0000080" "nm0000095"
## [7] "nm0000108" "nm0000116" "nm0000118" "nm0000122" "nm0000175" "nm0000184"
## [13] "nm0000230" "nm0000231" "nm0000233" "nm0000252" "nm0000338" "nm0000341"
## [19] "nm0000361" "nm0000390"
```

- companies :: companies that were the “headliners” in 15 or more movies

```
length(imdb.data$headliners$companies);
```

```
## [1] 770
```

```
head(imdb.data$headliners$companies, 20);
```

```
## [1] "co0002663" "co0007143" "co0023400" "co0050868" "co0005073" "co0000756"
## [7] "co0002219" "co0053157" "co0007001" "co0041421" "co0022792" "co0028557"
## [13] "co0020540" "co0028775" "co0021358" "co0024010" "co0067278" "co0046718"
## [19] "co0021661" "co0008970"
```

Companies These are the (at most) top-3 company credits plus the (at most) top-3 actors found in the “Details” section of a movie page.

```
dim(imdb.data$all.movies.companies);
```

```
## [1] 120376      4
```

```
names(imdb.data$all.movies.companies);
```

```
## [1] "ttid"      "co.rank" "coid"      "name"
```



```
head(imdb.data$all.movies.companies, 20);
```

```
##          ttid co.rank      coid          name
## 121552 tt0000001      1 co0037035 Edison Manufacturing Company
## 121553 tt0000002      NA NO-INFO
## 121555 tt0000003      NA NO-INFO
## 121557 tt0000004      NA NO-INFO
## 121559 tt0000005      1 co0037035 Edison Manufacturing Company
## 121560 tt0000006      NA NO-INFO
## 121562 tt0000007      1 co0037035 Edison Manufacturing Company
## 121563 tt0000008      1 co0037035 Edison Manufacturing Company
## 121564 tt0000009      1 co0014292 Alexander Black Photoplays
## 121565 tt0000010      1 co0014436 Lumière
## 121566 tt0000011      NA NO-INFO
## 121568 tt0000012      1 co0014436 Lumière
## 121569 tt0000013      1 co0014436 Lumière
## 121570 tt0000014      1 co0014436 Lumière
## 121571 tt0000014      2 co0059061 Société des Etablissements L. Gaumont
## 121572 tt0000015      NA NO-INFO
## 121574 tt0000016      1 co0014436 Lumière
## 121575 tt0000017      NA NO-INFO
## 121577 tt0000018      NA NO-INFO
## 121579 tt0000019      1 co0103007 Williamson Kinematograph Company
```

Creatives These are the (at most) top-3 creatives plus the (at most) top-3 actors found at the top of a movie page.

```
dim(imdb.data$all.movies.creatives);
```

```
## [1] 407481      4
```

```
names(imdb.data$all.movies.creatives);
```

```
## [1] "ttid"          "creative.rank" "nmid"          "role"
```

```
head(imdb.data$all.movies.creatives, 20);
```

```
##          ttid creative.rank      nmid      role
## 389665 tt0000001      1 nm0005690 director
## 389666 tt0000001      2 nm1588970      star
## 389667 tt0000002      1 nm0721526 director
## 389668 tt0000003      1 nm0721526 director
## 389669 tt0000004      1 nm0721526 director
## 389670 tt0000005      1 nm0005690 director
## 389671 tt0000005      2 nm0443482      star
## 389672 tt0000005      3 nm0653042      star
## 389673 tt0000006      1 nm0005690 director
## 389674 tt0000007      1 nm0005690 director
## 389675 tt0000007      2 nm0374658 director
## 389676 tt0000007      3 nm0179163      star
## 389677 tt0000007      4 nm0183947      star
## 389678 tt0000008      1 nm0005690 director
## 389679 tt0000008      2 nm0653028      star
## 389680 tt0000009      1 nm0085156 director
## 389681 tt0000009      2 nm0085156 writer
## 389682 tt0000009      3 nm0063086      star
```

```
## 389683 tt0000009          4 nm0183823    star
## 389684 tt0000009          5 nm1309758    star
```

Actors: Info (e.g., Bio) This is data from the top of an actor's (any creative really) page: <https://www.imdb.com/name/nm0000115/>

```
dim(imdb.data$all.actors.info);
```

```
## [1] 71332    10
```

```
names(imdb.data$all.actors.info);
```

```
## [1] "nmid"          "name"          "roles"         "born.when"
## [5] "born.where"    "died.when"     "died.where"    "starmeter.rank"
## [9] "starmeter.delta" "bio"
```

```
head(imdb.data$all.actors.info, 20);
```

```
##          nmid          name
## 1 nm0000001    Fred Astaire
## 2 nm0000002    Lauren Bacall
## 3 nm0000003    Brigitte Bardot
## 4 nm0000004    John Belushi
## 5 nm0000005    Ingmar Bergman
## 6 nm0000006    Ingrid Bergman
## 7 nm0000007    Humphrey Bogart
## 8 nm0000008    Marlon Brando
## 9 nm0000009    Richard Burton
## 10 nm0000010    James Cagney
## 11 nm0000011    Gary Cooper
## 12 nm0000012    Bette Davis
## 13 nm0000013    Doris Day
## 14 nm0000014    Olivia de Havilland
## 15 nm0000015    James Dean
## 16 nm0000016    Georges Delerue
## 17 nm0000017    Marlene Dietrich
## 18 nm0000018    Kirk Douglas
## 19 nm0000019    Federico Fellini
## 20 nm0000020    Henry Fonda
##
##                                roles  born.when
## 1                                Soundtrack,Actor,Miscellaneous Crew 1899-5-10
## 2                                Actress,Soundtrack 1924-9-16
## 3                                Actress,Soundtrack,Music Department 1934-9-28
## 4                                Actor,Writer,Soundtrack 1949-1-24
## 5                                Writer,Director,Actor 1918-7-14
## 6                                Actress,Soundtrack,Producer 1915-8-29
## 7                                Actor,Soundtrack,Producer 1899-12-25
## 8                                Actor,Soundtrack,Director 1924-4-3
## 9                                Actor,Soundtrack,Producer 1925-11-10
## 10                               Actor,Soundtrack,Director 1899-7-17
## 11                               Actor,Soundtrack,Stunts 1901-5-7
## 12                               Actress,Soundtrack,Make-Up Department 1908-4-5
## 13                               Soundtrack,Actress,Producer 1922-4-3
## 14                               Actress,Soundtrack 1916-7-1
## 15                               Actor,Miscellaneous Crew 1931-2-8
## 16                               Composer,Soundtrack,Music Department 1925-3-12
```

## 17	Soundtrack, Actress, Music Department	1901-12-27
## 18	Actor, Producer, Soundtrack	1916-12-9
## 19	Writer, Director, Second Unit Director or Assistant Director	1920-1-20
## 20	Actor, Producer, Soundtrack	1905-5-16

##	born.where	died.when
----	------------	-----------

## 1	Omaha, Nebraska, USA	1987-6-22
## 2	The Bronx, New York City, New York, USA	2014-8-12
## 3	Paris, France	<NA>
## 4	Chicago, Illinois, USA	1982-3-5
## 5	Uppsala, Uppsala län, Sweden	2007-7-30
## 6	Stockholm, Sweden	1982-8-29
## 7	New York City, New York, USA	1957-1-14
## 8	Omaha, Nebraska, USA	2004-7-1
## 9	Pontrhydyfen, Wales, UK	1984-8-5
## 10	New York City, New York, USA	1986-3-30
## 11	Helena, Montana, USA	1961-5-13
## 12	Lowell, Massachusetts, USA	1989-10-6
## 13	Cincinnati, Ohio, USA	2019-5-13
## 14	Tokyo, Japan	2020-7-26
## 15	Marion, Indiana, USA	1955-9-30
## 16	Roubaix, Nord, France	1992-3-20
## 17	Schöneberg, Berlin, Germany	1992-5-6
## 18	Amsterdam, New York, USA	2020-2-5
## 19	Rimini, Emilia-Romagna, Italy	1993-10-31
## 20	Grand Island, Nebraska, USA	1982-8-12

##	died.where	starmeter.rank	starmeter.delta
----	------------	----------------	-----------------

## 1	Los Angeles, California, USA	NA	990
## 2	New York City, New York, USA	5000	-32
## 3	<NA>	5000	201
## 4	Hollywood, Los Angeles, California, USA	NA	149
## 5	Fårö, Gotlands län, Sweden	5000	-64
## 6	Chelsea, London, England, UK	5000	1065
## 7	Los Angeles, California, USA	5000	-257
## 8	Westwood, Los Angeles, California, USA	5000	5
## 9	Céligny, Geneva, Switzerland	NA	103
## 10	Stanfordville, New York, USA	NA	-56
## 11	Beverly Hills, Los Angeles, California, USA	5000	1325
## 12	Neuilly-sur-Seine, Hauts-de-Seine, France	5000	-1011
## 13	Carmel Valley, California, USA	NA	723
## 14	Paris, France	NA	-289
## 15	Cholame, California, USA	5000	78
## 16	Los Angeles, California, USA	NA	-8260
## 17	Paris, France	5000	2982
## 18	Los Angeles, California, USA	5000	280
## 19	Rome, Lazio, Italy	NA	-156
## 20	Los Angeles, California, USA	5000	289

##

## 1	Fred Astaire was born in Omaha, Nebraska, to Johanna (Geilus) and Fritz Austerlitz, a brewer. I
## 2	Lauren Bacall was born Betty Joan Perske on September 16, 1924, in New York City. She was the
## 3	Brigitte Bardot was born on September 28, 1934 in Paris, France. Her father had an engineering d
## 4	John Belushi was born in Chicago, Illinois, USA, on January 24, 1949, to Agnes Demetri (Samara
## 5	Ernst Ingmar Bergman was born July 14, 1918, the son of a priest. The film and T.V. series,
## 6	Ingrid Bergman was one of the greatest actresses from Hollywood's lamented Golden Era. Her m
## 7	Humphrey DeForest Bogart was born in New York City, New York, to Maud Humphrey, a famed maga

```
## 8   Marlon Brando is widely considered the greatest movie actor of all time, rivaled only by the mor
## 9   Probably best-remembered for his turbulent personal life with Elizabeth Taylor (whom he married
## 10  One of Hollywood's preeminent male stars of all time, James Cagney was also an accomplished dan
## 11  Born to Alice Cooper and Charles Cooper. Gary attended school at Dunstable school England, Hel
## 12 Ruth Elizabeth Davis was born April 5, 1908, in Lowell, Massachusetts, to Ruth Augusta (Favor) an
## 13 One of America's most prolific actresses was born Doris Mary Ann Kappelhoff on April 3, 1922, in C
## 14   Olivia Mary de Havilland was born July 1, 1916, in Tokyo, Japan, to British parents, Lilian A
## 15 James Byron Dean was born February 8, 1931 in Marion, Indiana, to Mildred Marie (Wilson) and Win
## 16                                     Georges Delerue was born on March 12, 19
## 17 Her father was a police lieutenant and imbued in her a military attitude to life. Marlene was kn
## 18   Cleft-chinned, steely-eyed and virile star of international cinema who rose from being
## 19 The women who both attracted and frightened him and an Italy dominated in his youth by Mussolini
## 20   Henry Jaynes Fonda was born in Grand Island, Nebraska, to Elma Herberta (Jaynes) and William B
```

Actor on Movie: Rank & Character Name These are the (at most) top-3 creatives plus the (at most) top-3 actors found at the top of a movie page.

```
dim(imdb.data$all.movies.actors.characters);
```

```
## [1] 901324      4
```

```
names(imdb.data$all.movies.actors.characters);
```

```
## [1] "ttid"      "actor.rank" "nmid"      "character"
```

```
head(imdb.data$all.movies.actors.characters, 20);
```

##	ttid	actor.rank	nmid	character
## 879347	tt0000001	1	nm1588970	Self
## 879348	tt0000005	1	nm0443482	Blacksmith (uncredited)
## 879349	tt0000005	2	nm0653042	Assistant (uncredited)
## 879350	tt0000007	1	nm0179163	<NA>
## 879351	tt0000007	2	nm0183947	<NA>
## 879352	tt0000008	1	nm0653028	Sneezing Man
## 879353	tt0000009	1	nm0063086	Miss Geraldine Holbrook (Miss Jerry)
## 879354	tt0000009	2	nm0183823	
## 879355	tt0000009	3	nm1309758	
## 879356	tt0000011	1	nm3692297	
## 879357	tt0000012	1	nm2880396	
## 879358	tt0000012	2	nm9735580	
## 879359	tt0000012	3	nm0525900	
## 879360	tt0000012	4	nm9735581	
## 879361	tt0000012	5	nm9735579	
## 879362	tt0000012	6	nm9653419	
## 879363	tt0000013	1	nm0525908	
## 879364	tt0000014	1	nm0166380	
## 879365	tt0000014	2	nm0244989	
## 879366	tt0000016	1	nm0525900	

```
## 879354 Mr. Hamilton
## 879355 Chauncey Depew - the Director of the New York Central Railroad
## 879356 Acrobats (as Familie Grunato)
## 879357 Self
## 879358 Self
## 879359 Self
## 879360 Self
## 879361 Self
## 879362 Self
## 879363 Self
## 879364 The Gardener
## 879365 The Boy
## 879366 Self (on the pier) (uncredited)
```

Movie Info: from Actor Search Page This contains ratings, dollars, votes, and other features we used in the earlier homework.

```
dim(imdb.data$all.actors.movies);
```

```
## [1] 282062 11
```

```
names(imdb.data$all.actors.movies);
```

```
## [1] "ttid" "year" "title" "genre" "rated"
## [6] "minutes" "ratings" "metacritic" "votes" "millions"
## [11] "paragraph"
```

```
head(imdb.data$all.actors.movies, 20);
```

```
##          ttid year          title
## 1823544 tt0000009 1894          Miss Jerry
## 494917  tt0000147 1897          The Corbett-Fitzsimmons Fight
## 146344  tt0229676 1897  Reproduction of the Corbett and Fitzsimmons Fight
## 146348  tt0219574 1898          Can Can
## 1840210 tt0221032 1898          Branding Cattle
## 1840211 tt0221040 1898          Buck Dance, Ute Indians
## 1840209 tt0221384 1898          Masked Procession
## 146342  tt0229287 1898          Cake Walk
## 1840208 tt0229690 1898          S.S. 'Coptic'
## 1822438 tt0230366 1899          Jeffries-Sharkey Contest
## 146351  tt0230765 1899          Sharkey-McCoy Fight Reproduced in 10 Rounds
## 146350  tt0242588 1899          Lady Contortionist
## 146343  tt0263933 1899  Reproduction of the Corbett and Jeffries Fight
## 1228276 tt0000335 1900          Soldiers of the Cross
## 146375  tt0266332 1900          Birdseye View of Galveston, Showing Wreckage
## 1822445 tt0327448 1900          At Break-Neck Speed
## 1718321 tt0156019 1902          Sr. Wrangler Mr. R.P. Paranjpe
## 146387  tt0232530 1902  Reproduction of the Jeffries-Fitzsimmons Fight
## 146345  tt0376198 1902          Snow White
## 582790  tt0259649 1903          The Three Musketeers
##          genre rated minutes ratings metacritic votes
## 1823544          Romance <NA>      45      5.9      NA      NA
## 494917  Documentary, News, Sport <NA>      20      5.2      NA      NA
## 146344  Documentary, News, Sport <NA>      NA      4.4      NA      NA
## 146348          <NA> <NA>      NA      5.7      NA      NA
## 1840210  Documentary, Sci-Fi <NA>      1      4.8      NA      NA
```

## 1840211	Documentary	<NA>	NA	NA	NA	NA
## 1840209		<NA>	<NA>	NA	3.6	NA
## 146342		<NA>	<NA>	NA	4.0	NA
## 1840208		<NA>	<NA>	NA	3.6	NA
## 1822438	Documentary, News, Sport	<NA>	135	3.8	NA	NA
## 146351	Documentary, News, Sport	<NA>	NA	3.5	NA	NA
## 146350		<NA>	<NA>	NA	4.1	NA
## 146343	News, Sport	<NA>	NA	4.2	NA	NA
## 1228276	Biography, Drama	<NA>	NA	6.1	NA	NA
## 146375	Documentary	<NA>	NA	NA	NA	NA
## 1822445	Documentary	<NA>	NA	NA	NA	NA
## 1718321		<NA>	<NA>	NA	3.9	NA
## 146387	Documentary, Sport	<NA>	NA	NA	NA	NA
## 146345		<NA>	<NA>	NA	4.7	NA
## 582790		<NA>	<NA>	NA	4.1	NA

millions

## 1823544	NA
## 494917	NA
## 146344	NA
## 146348	NA
## 1840210	NA
## 1840211	NA
## 1840209	NA
## 146342	NA
## 1840208	NA
## 1822438	NA
## 146351	NA
## 146350	NA
## 146343	NA
## 1228276	NA
## 146375	NA
## 1822445	NA
## 1718321	NA
## 146387	NA
## 146345	NA
## 582790	NA

##

1823544

494917

146344

146348 "You are not often afforded the privilege of viewing a real French can-can and this film is "

1840210

1840211

1840209

146342 "Here is a film, one of the finest of its kind, portraying a number of colored professional.

1840208

1822438

146351

146350 "Here is shown Mlle. De La Roy, the well-known dancer who has a national reputation. The la

146343

1228276 The plot outlined the story of the early Christian martyrs with a compendium of horrors gu

146375

1822445

1718321

```
## 146387
## 146345
## 582790
```

Actor's movie Rank An actor has a ranking of their movies based on popularity (by votes), see https://www.imdb.com/filmosearch/?explore=title_type&role=nm0000243&ref_=filmo_ref_typ&sort=num_votes,desc&mode=detail&page=1&title_type=movie

Will Smith's most popular film is I am Legend, Rank 1. Can also be achieved by analyzing and sorting "votes" or "stars" or "metacritic"...

```
dim(imdb.data$all.actors.rank);
```

```
## [1] 1842305      3
```

```
names(imdb.data$all.actors.rank);
```

```
## [1] "nmid" "ttid" "rank"
```

```
head(imdb.data$all.actors.rank, 20);
```

```
##      nmid      ttid rank
## 1 nm0000001 tt0120689   1
## 2 nm0000001 tt7286456   2
## 3 nm0000001 tt0095953   3
## 4 nm0000001 tt0212720   4
## 5 nm0000001 tt0116209   5
## 6 nm0000001 tt0249462   6
## 7 nm0000001 tt0309987   7
## 8 nm0000001 tt3874544   8
## 9 nm0000001 tt1193631   9
## 10 nm0000001 tt0089853  10
## 11 nm0000001 tt4595882  11
## 12 nm0000001 tt4385888  12
## 13 nm0000001 tt0072308  13
## 14 nm0000001 tt3442006  14
## 15 nm0000001 tt0308383  15
## 16 nm0000001 tt0050419  16
## 17 nm0000001 tt1767372  17
## 18 nm0000001 tt0027125  18
## 19 nm0000001 tt0034862  19
## 20 nm0000001 tt0028333  20
```

Glue Table: Top-Creatives For example "tt2671706" has Denzel Washington ("nm0000243") as "director" and "actor" ...

```
dim(imdb.data$glue$movies.creatives);
```

```
## [1] 361077      3
```

```
names(imdb.data$glue$movies.creatives);
```

```
## [1] "ttid" "nmid" "role"
```

```
head(imdb.data$glue$movies.creatives, 20);
```

```
##      ttid      nmid      role
## 1 tt0423409 nm0935863 director
## 2 tt0423409 nm1038177  writer
```

```
## 3 tt0423409 nm0101639 writer
## 4 tt0423409 nm0176869 star
## 5 tt0423409 nm0000562 star
## 6 tt0423409 nm0117339 star
## 7 tt0062108 nm0892650 director
## 8 tt0062108 nm0053168 writer
## 9 tt0062108 nm0297391 writer
## 10 tt0062108 nm0001260 star
## 11 tt0062108 nm0529964 star
## 12 tt0062108 nm0671185 star
## 13 tt6014370 nm1312641 director
## 14 tt6014370 nm1312641 writer
## 15 tt6014370 nm0230649 star
## 16 tt6014370 nm0477810 star
## 17 tt6014370 nm0295429 star
## 18 tt0039075 nm0329574 director
## 19 tt0039075 nm0730905 writer
## 20 tt0039075 nm0743232 writer
```

Extra data

I added three dataframes to the dataset (November 27, 2020), so we have all of the Will/Denzel movies and their information.

- `movies.df$cast`

```
dim(imdb.data$movies.df$cast);
```

```
## [1] 120678      5
```

```
names(imdb.data$movies.df$cast);
```

```
## [1] "ttid"      "actor.rank" "nmid"      "name"      "character"
```

```
head(imdb.data$movies.df$cast, 20);
```

##	ttid	actor.rank	nmid	name	character
## 1	tt0219574	1	nm5005049	Mlle. Fraidora	Self
## 2	tt0242588	1	nm5178266	Mlle. De La Roy	
## 3	tt0000335	1	nm1010955	Beatrice Day	
## 4	tt0000335	2	nm1012612	Harold Graham	
## 5	tt0000335	3	nm1011210	Mr. Graham	
## 6	tt0000335	4	nm1012621	John Jones	
## 7	tt0000335	5	nm0675239	Orrie Perry	Lion
## 8	tt0000335	6	nm0675260	Reg Perry	Lion
## 9	tt0156019	1	nm1532630	R.P. Paranjpye	Self
## 10	tt0000502	1	nm0215752	Antonio del Pozo	
## 11	tt0000502	2	nm0252720	El Mochuelo	
## 12	tt0295963	1	nm1180997	Frida Cotrelli	Frau Schultze
## 13	tt0295963	2	nm0804433	Eugen Skladanowsky	Der Componist
## 14	tt1325665	1	nm0237070	Dranem	Self
## 15	tt0000574	1	nm0846887	Elizabeth Tait	Kate Kelly
## 16	tt0000574	2	nm0846894	John Tait	School Master
## 17	tt0000574	3	nm3002376	Norman Campbell	Steve Hart
## 18	tt0000574	4	nm0170118	Bella Cola	
## 19	tt0000574	5	nm3001683	Will Coyne	Joe Byrne
## 20	tt0000574	6	nm0187697	Sam Crewes	Dan Kelly


```

• movies.df$text
dim(imdb.data$movies.df$text);

## [1] 10523      5
names(imdb.data$movies.df$text);

## [1] "summary"  "storyline" "keywords"  "taglines"  "ttid"
head(imdb.data$movies.df$text, 20);

##
## my.row      "You are not often afforded the privilege of viewing a real French can-can and this film i
## my.row.1     "Here is a film, one of the finest of its kind, portraying a number of colored profession
## my.row.2                                           "Masqueraders dressed in most grotesque costumes
## my.row.3
## my.row.4
## my.row.5
## my.row.6
## my.row.7     "A stunning view of one of the fastest boats in Uncle Sam's torpedo fleet. She is runn
## my.row.8
## my.row.9                                           "This turns out to be as much of a bathing
## my.row.10
## my.row.11
## my.row.12
## my.row.13
## my.row.14                                           "Continuation of the above picture [S.S. COPTI
## my.row.15
## my.row.16
## my.row.17
## my.row.18    "Here is shown Mlle. De La Roy, the well-known dancer who has a national reputation. The
## my.row.19
##
## my.row      "You are not often afforded the privilege of viewing a real French can-can and this film i
## my.row.1
## my.row.2
## my.row.3
## my.row.4
## my.row.5
## my.row.6
## my.row.7
## my.row.8
## my.row.9
## my.row.10
## my.row.11
## my.row.12
## my.row.13
## my.row.14
## my.row.15
## my.row.16
## my.row.17
## my.row.18
## my.row.19
##
##                                     keywords taglines
## my.row      dancer, dancing, dance, cancan dance, repetition in title      <NA>

```

```
## my.row.1      NA      <NA>
## my.row.2      costume  <NA>
## my.row.3      NA      <NA>
## my.row.4      NA      <NA>
## my.row.5      NA      <NA>
## my.row.6      NA      <NA>
## my.row.7      NA      <NA>
## my.row.8      NA      <NA>
## my.row.9      race     <NA>
## my.row.10     NA      <NA>
## my.row.11     NA      <NA>
## my.row.12     NA      <NA>
## my.row.13     NA      <NA>
## my.row.14     picture  <NA>
## my.row.15     NA      <NA>
## my.row.16     NA      <NA>
## my.row.17     boxing   <NA>
## my.row.18     NA      <NA>
## my.row.19     NA      <NA>
```

```
##          ttid
## my.row    tt0219574
## my.row.1  tt0229287
## my.row.2  tt0221384
## my.row.3  tt0229399
## my.row.4  tt0237889
## my.row.5  tt0229764
## my.row.6  tt0229389
## my.row.7  tt0221654
## my.row.8  tt0317653
## my.row.9  tt0237861
## my.row.10 tt0237775
## my.row.11 tt0138538
## my.row.12 tt0317652
## my.row.13 tt0317654
## my.row.14 tt0229690
## my.row.15 tt0240795
## my.row.16 tt0156186
## my.row.17 tt0263933
## my.row.18 tt0242588
## my.row.19 tt0242282
```

```
• movies.df$info
```

```
dim(imdb.data$movies.df$info);
```

```
## [1] 10523    19
```

```
names(imdb.data$movies.df$info);
```

```
## [1] "ttid"          "country"        "language"        "release.date"
## [5] "aka"           "filming.location" "budget"          "usa.opening"
## [9] "usa.gross"     "world.gross"    "production.name" "runtime"
## [13] "sound"        "color"          "aspect"          "release.location"
## [17] "budget.est"   "usa.opening.date" "production.co"
```

```
head(imdb.data$movies.df$info, 20);
```

##	ttid	country	language	release.date						
## 1	tt0219574	USA	None	<NA>						
## 2	tt0229287	USA	None	20 May 1898						
## 3	tt0221384	USA	None	February 1898						
## 4	tt0229399	USA	None	<NA>						
## 5	tt0237889	USA	None	July 1898						
## 6	tt0229764	USA	None	July 1898						
## 7	tt0229389	USA	None	<NA>						
## 8	tt0221654	USA	None	February 1898						
## 9	tt0317653	UK	None	<NA>						
## 10	tt0237861	USA	None	July 1898						
## 11	tt0237775	USA	None	July 1898						
## 12	tt0138538	Portugal	None	<NA>						
## 13	tt0317652	UK	None	<NA>						
## 14	tt0317654	UK	None	<NA>						
## 15	tt0229690	USA	None	20 May 1898						
## 16	tt0240795	USA	None	October 1898						
## 17	tt0156186	India	None	1 November 1899						
## 18	tt0263933	USA	None	November 1899						
## 19	tt0242588	USA	None	<NA>						
## 20	tt0242282	USA	None	March 1899						
##					aka					
## 1					French Can-Can					
## 2					<NA>					
## 3					<NA>					
## 4					Tropas en Cuba					
## 5					<NA>					
## 6					<NA>					
## 7					<NA>					
## 8					Torpedo Boat 'Dupont' on Mobile River Under Full Headway					
## 9					<NA>					
## 10					<NA>					
## 11					<NA>					
## 12					<NA>					
## 13					<NA>					
## 14					<NA>					
## 15					<NA>					
## 16					<NA>					
## 17					<NA>					
## 18					Jeffries-Corbett, 23 Rounds					
## 19					<NA>					
## 20					<NA>					
##						filming.location	budget	usa.opening	usa.gross	world.gross
## 1					<NA>	<NA>	<NA>	<NA>	<NA>	<NA>
## 2					<NA>	<NA>	<NA>	<NA>	<NA>	<NA>
## 3					Denver, Colorado, USA	<NA>	<NA>	<NA>	<NA>	<NA>
## 4					<NA>	<NA>	<NA>	<NA>	<NA>	<NA>
## 5					<NA>	<NA>	<NA>	<NA>	<NA>	<NA>
## 6					<NA>	<NA>	<NA>	<NA>	<NA>	<NA>
## 7					<NA>	<NA>	<NA>	<NA>	<NA>	<NA>
## 8					New Orleans, Louisiana, USA	<NA>	<NA>	<NA>	<NA>	<NA>
## 9					<NA>	<NA>	<NA>	<NA>	<NA>	<NA>

## 10		<NA>	<NA>	<NA>	<NA>	<NA>
## 11		<NA>	<NA>	<NA>	<NA>	<NA>
## 12		<NA>	<NA>	<NA>	<NA>	<NA>
## 13		<NA>	<NA>	<NA>	<NA>	<NA>
## 14		<NA>	<NA>	<NA>	<NA>	<NA>
## 15		<NA>	<NA>	<NA>	<NA>	<NA>
## 16		<NA>	<NA>	<NA>	<NA>	<NA>
## 17	Mumbai, Maharashtra, India	<NA>	<NA>	<NA>	<NA>	<NA>
## 18		<NA>	<NA>	<NA>	<NA>	<NA>
## 19		<NA>	<NA>	<NA>	<NA>	<NA>
## 20	New York City, New York, USA	<NA>	<NA>	<NA>	<NA>	<NA>
##	production.name	runtime	sound	color	aspect	
## 1	S. Lubin	<NA>	Silent	Black and White	<NA>	
## 2	S. Lubin	<NA>	Silent	Black and White	<NA>	
## 3	Edison Manufacturing Company	<NA>	Silent	Black and White	<NA>	
## 4	Amet	<NA>	Silent	Black and White	<NA>	
## 5	American Mutoscope Company	<NA>	Silent	Black and White	1.36 : 1	
## 6	Edison Manufacturing Company	<NA>	Silent	Black and White	<NA>	
## 7	Amet	<NA>	Silent	Black and White	<NA>	
## 8	American Mutoscope Company	<NA>	Silent	Black and White	1.36 : 1	
## 9	Riley Brothers	<NA>	Silent	Black and White	<NA>	
## 10	Edison Manufacturing Company	<NA>	Silent	Black and White	<NA>	
## 11	American Mutoscope Company	<NA>	Silent	Black and White	1.36 : 1	
## 12	<NA>	<NA>	Silent	Black and White	<NA>	
## 13	Riley Brothers	<NA>	Silent	Black and White	<NA>	
## 14	Leviathan Films	<NA>	Silent	Black and White	<NA>	
## 15	Edison Manufacturing Company	<NA>	Silent	Black and White	<NA>	
## 16	American Mutoscope Company	<NA>	Silent	Black and White	1.36 : 1	
## 17	<NA>	<NA>	Silent	Black and White	<NA>	
## 18	S. Lubin	<NA>	Silent	Black and White	<NA>	
## 19	S. Lubin	<NA>	Silent	Black and White	<NA>	
## 20	American Mutoscope Company	<NA>	Silent	Black and White	1.36 : 1	
##	release.location	budget.est	usa.opening.date	production.co		
## 1	<NA>	<NA>	<NA>	co0010327		
## 2	USA	<NA>	<NA>	co0010327		
## 3	USA	<NA>	<NA>	co0037035		
## 4	<NA>	<NA>	<NA>	co0028744		
## 5	USA	<NA>	<NA>	co0032679		
## 6	USA	<NA>	<NA>	co0037035		
## 7	<NA>	<NA>	<NA>	co0028744		
## 8	USA	<NA>	<NA>	co0032679		
## 9	<NA>	<NA>	<NA>	co0051449		
## 10	USA	<NA>	<NA>	co0037035		
## 11	USA	<NA>	<NA>	co0032679		
## 12	<NA>	<NA>	<NA>			
## 13	<NA>	<NA>	<NA>	co0051449		
## 14	<NA>	<NA>	<NA>	co0155594		
## 15	USA	<NA>	<NA>	co0037035		
## 16	USA	<NA>	<NA>	co0032679		
## 17	India	<NA>	<NA>			
## 18	USA	<NA>	<NA>	co0010327		
## 19	<NA>	<NA>	<NA>	co0010327		
## 20	USA	<NA>	<NA>	co0032679		

Will Smith

The function `IMDB.getMoviesForPerson` requires a known “nmid” value. We can search for it:

```
will.search = IMDB.searchPersonName("Will* Smith*");  
  
will.search;
```

```
##          nmid          name          roles  born.when  
## 226   nm0000226   Will Smith Music Department,Actor,Producer 1968-9-25  
## 41738 nm0810342   William Smith Actor,Stunts,Miscellaneous Crew 1933-3-24  
## 41739 nm0810349   William Smith          Actor          <NA>  
## 41744 nm0810461 William Smithers          Actor 1927-7-10  
## 58036 nm2405238   Willow Smith          Soundtrack,Actress,Writer 2000-10-31  
##  
##          born.where died.when died.where starmeter.rank  
## 226   Philadelphia, Pennsylvania, USA          <NA>          <NA>          500  
## 41738          Columbia, Missouri, USA          <NA>          <NA>          5000  
## 41739          <NA>          <NA>          <NA>          NA  
## 41744          Richmond, Virginia, USA          <NA>          <NA>          NA  
## 58036          Los Angeles, California, USA          <NA>          <NA>          NA  
##          starmeter.delta  
## 226          67  
## 41738          1386  
## 41739          -191098  
## 41744          627  
## 58036          -1525  
##  
## 226   Willard Carroll "Will" Smith, Jr. (born September 25, 1968) is an American actor, comedian, p  
## 41738 William Smith is probably best known for his portrayal as "Falconetti" in Rich Man, Poor Man (   
## 41739  
## 41744 Although character actor William (or Bill) Smithers is not recognizable perhaps by name, the f  
## 58036
```

Movies

```
will.nmid = will.search$nmid[1];  
will.nmid = "nm0000226";  
will.movies = IMDB.getMoviesForPerson(will.nmid);  
will.movies;
```

```
##          ttid          nmid rank year          title  
## 39   tt0480249 nm0000226    1 2007          I Am Legend  
## 65   tt1386697 nm0000226    2 2016          Suicide Squad  
## 6    tt0116629 nm0000226    3 1996          Independence Day  
## 7    tt0119654 nm0000226    4 1997          Men in Black  
## 26   tt0343818 nm0000226    5 2004          I, Robot  
## 35   tt0454921 nm0000226    6 2006          The Pursuit of Happyness  
## 34   tt0448157 nm0000226    7 2008          Hancock  
## 12   tt0120912 nm0000226    8 2002          Men in Black II  
## 66   tt1409024 nm0000226    9 2012          Men in Black 3  
## 31   tt0386588 nm0000226   10 2005          Hitch  
## 41   tt0814314 nm0000226   11 2008          Seven Pounds  
## 4    tt0112442 nm0000226   12 1995          Bad Boys  
## 17   tt0172156 nm0000226   13 2003          Bad Boys II  
## 8    tt0120660 nm0000226   14 1998          Enemy of the State
```

## 104	tt6139732	nm0000226	15	2019	Aladdin
## 13	tt0145660	nm0000226	16	1999	Austin Powers: The Spy Who Shagged Me
## 90	tt2381941	nm0000226	17	2015	Focus
## 80	tt1815862	nm0000226	18	2013	After Earth
## 74	tt1596350	nm0000226	19	2012	This Means War
## 57	tt1229340	nm0000226	20	2013	Anchorman 2: The Legend Continues
## 100	tt5519340	nm0000226	21	2017	Bright
## 23	tt0307453	nm0000226	22	2004	Shark Tale
## 54	tt1155076	nm0000226	23	2010	The Karate Kid
## 11	tt0120891	nm0000226	24	1999	Wild Wild West
## 10	tt0120783	nm0000226	25	1998	The Parent Trap
## 70	tt1502397	nm0000226	26	2020	Bad Boys For Life
## 19	tt0248667	nm0000226	27	2001	Ali
## 97	tt4682786	nm0000226	28	2016	Collateral Beauty
## 93	tt3322364	nm0000226	29	2015	Concussion
## 47	tt1025100	nm0000226	30	2019	Gemini Man
## 5	tt0114558	nm0000226	31	1995	Strange Days
## 22	tt0300051	nm0000226	32	2004	Jersey Girl
## 21	tt0284490	nm0000226	33	2002	Showtime
## 14	tt0146984	nm0000226	34	2000	The Legend of Bagger Vance
## 83	tt1837709	nm0000226	35	2014	Winter's Tale
## 25	tt0338466	nm0000226	36	2003	Stuck on You
## 44	tt0947802	nm0000226	37	2008	Lakeview Terrace
## 82	tt1823664	nm0000226	38	2014	Annie
## 20	tt0268397	nm0000226	39	2001	Jimmy Neutron: Boy Genius
## 50	tt1082886	nm0000226	40	2008	The Wackness
## 101	tt5814534	nm0000226	41	2019	Spies in Disguise
## 94	tt3721964	nm0000226	42	2018	Gringo
## 33	tt0416212	nm0000226	43	2008	The Secret Life of Bees
## 3	tt0108149	nm0000226	44	1993	Six Degrees of Separation
## 24	tt0328099	nm0000226	45	2003	Malibu's Most Wanted
## 15	tt0167427	nm0000226	46	1999	Superstar
## 37	tt0466839	nm0000226	47	2007	I Could Never Be Your Woman
## 2	tt0107478	nm0000226	48	1993	Made in America
## 106	tt7255568	nm0000226	49	2019	Student of the Year 2
## 38	tt0466856	nm0000226	50	2006	ATL
## 30	tt0384504	nm0000226	51	2004	Saving Face
## 108	tt7440432	nm0000226	52	2018	Quincy
## 1	tt0105810	nm0000226	53	1992	Where the Day Takes You
## 29	tt0380277	nm0000226	54	2004	The Cookout
## 53	tt1109477	nm0000226	55	2008	The Human Contract
## 64	tt1383607	nm0000226	56	2008	Ramadan Mabrouk Abul-Alamein Hamouda
## 46	tt1015971	nm0000226	57	2009	The Boys: The Sherman Brothers' Story
## 95	tt4209900	nm0000226	58	2015	Fresh Dressed
## 16	tt0169376	nm0000226	59	1998	Welcome to Hollywood
## 9	tt0120707	nm0000226	60	1998	Ride
## 51	tt10883124	nm0000226	61	2019	Dads
## 58	tt1230214	nm0000226	62	2009	Stuntmen
## 28	tt0379487	nm0000226	63	2004	The Seat Filler
## 89	tt2350432	nm0000226	64	2012	Free Angela and All Political Prisoners
## 103	tt5903926	nm0000226	65	2018	Sprinter
## 67	tt1409772	nm0000226	66	2010	A Man's Story
## 27	tt0375663	nm0000226	67	2003	A Closer Walk
## 73	tt1584918	nm0000226	68	2014	Murder101

## 86	tt2074488	nm0000226	69	2005	Will Smith: Live in Concert
## 99	tt5230508	nm0000226	70	2005	AMV Hell 3: The Motion Picture
## 32	tt0414587	nm0000226	71	2005	There's a God on the Mic
## 18	tt0205314	nm0000226	72	1998	Pesel Ha'Zahav
## 71	tt1546668	nm0000226	73	2007	The 100 Best Black Movies (Ever)
## 102	tt5839150	nm0000226	74	2009	The Muhammad Ali Story
## 98	tt4919268	nm0000226	75	NA	Bad Boys 4
## 76	tt1674782	nm0000226	76	NA	The Karate Kid 2
## 111	tt9620288	nm0000226	77	2021	King Richard
## 105	tt6598238	nm0000226	78	2020	Life in a Year
## 110	tt7820302	nm0000226	79	NA	Bright 2
## 52	tt11012318	nm0000226	80	NA	The Council
## 62	tt12915962	nm0000226	81	NA	Planes, Trains & Automobiles
## 69	tt1498780	nm0000226	82	NA	Hancock 2
## 60	tt12530246	nm0000226	83	NA	Emancipation
## 72	tt1578276	nm0000226	84	NA	Uptown Saturday Night
## 91	tt2776710	nm0000226	85	NA	Brilliance
## 55	tt1181790	nm0000226	86	NA	The Billionaire's Vinegar
## 96	tt4463004	nm0000226	87	NA	Bounty
## 61	tt12608018	nm0000226	88	NA	Bounce
## 107	tt7331400	nm0000226	89	NA	1921
## 78	tt1692497	nm0000226	90	NA	The Redemption of Cain
## 40	tt0812268	nm0000226	91	NA	The Last Pharaoh
## 49	tt10344550	nm0000226	92	NA	Untitled Solvan Naim Project
## 68	tt1470831	nm0000226	93	NA	The City That Sailed
## 63	tt1368421	nm0000226	94	NA	Angelology
## 92	tt2846190	nm0000226	95	NA	La escribana de uraba
## 42	tt0911023	nm0000226	96	NA	My Wife Hates Your Wife
## 36	tt0464265	nm0000226	97	NA	Time Share
## 75	tt1609486	nm0000226	98	NA	Harold and the Purple Crayon
## 81	tt1817789	nm0000226	99	NA	Wheeler Dealers
## 77	tt1691439	nm0000226	100	NA	Untitled Dr. S. Allen Counter Project
## 85	tt1879098	nm0000226	101	NA	What Would Kenny Do?
## 88	tt2246761	nm0000226	102	NA	Interpol
## 59	tt1235550	nm0000226	103	NA	Welcome to the Sticks
## 79	tt1714201	nm0000226	104	NA	Monster Witness Relocation Program
## 43	tt0942916	nm0000226	105	NA	Sisters of Mercy
## 48	tt1028582	nm0000226	106	NA	Newton's Law
## 45	tt0979959	nm0000226	107	NA	Cooked
## 109	tt7737796	nm0000226	108	NA	The Prospect
## 56	tt1206884	nm0000226	109	NA	The Long Run
## 87	tt2168954	nm0000226	110	NA	Ghost Graduation
## 84	tt1852816	nm0000226	111	NA	Joe
##		genre	rated	minutes	ratings metacritic votes
## 39	Action, Adventure, Drama	PG-13	101	7.2	65 675160
## 65	Action, Adventure, Fantasy	PG-13	123	6.0	40 588064
## 6	Action, Adventure, Sci-Fi	PG-13	145	7.0	59 520592
## 7	Action, Adventure, Comedy	PG-13	98	7.3	71 507597
## 26	Action, Drama, Sci-Fi	PG-13	115	7.1	59 491474
## 35	Biography, Drama	PG-13	117	8.0	64 438121
## 34	Action, Fantasy	PG-13	92	6.4	49 437887
## 12	Action, Adventure, Comedy	PG-13	88	6.2	49 337458
## 66	Action, Adventure, Comedy	PG-13	106	6.8	58 329948
## 31	Comedy, Romance	PG-13	118	6.6	58 291991

## 41	Drama	PG-13	123	7.6	36	282568
## 4	Action, Comedy, Crime	R	119	6.9	41	235766
## 17	Action, Comedy, Crime	R	147	6.6	38	228329
## 8	Action, Thriller	R	132	7.3	67	224019
## 104	Adventure, Family, Fantasy	PG	128	7.0	53	216257
## 13	Action, Adventure, Comedy	PG-13	95	6.6	59	214064
## 90	Comedy, Crime, Drama	R	105	6.6	56	212789
## 80	Action, Adventure, Sci-Fi	PG-13	100	4.8	33	189337
## 74	Action, Comedy, Romance	PG-13	103	6.3	31	174205
## 57	Comedy	PG-13	119	6.3	61	165909
## 100	Action, Fantasy, Thriller	TV-MA	117	6.3	29	164913
## 23	Animation, Adventure, Comedy	PG	90	6.0	48	164132
## 54	Action, Drama, Family	PG	140	6.2	61	158842
## 11	Action, Comedy, Sci-Fi	PG-13	106	5.0	38	151132
## 10	Adventure, Comedy, Drama	PG	128	6.5	64	117780
## 70	Action, Comedy, Crime	R	124	6.6	59	112826
## 19	Biography, Drama, Sport	R	157	6.8	65	92508
## 97	Drama, Romance	PG-13	97	6.8	23	89138
## 93	Biography, Drama, Sport	PG-13	123	7.1	55	84750
## 47	Action, Drama, Sci-Fi	PG-13	117	5.7	38	76145
## 5	Action, Crime, Drama	R	145	7.2	66	66102
## 22	Comedy, Drama, Romance	PG-13	102	6.2	43	62700
## 21	Action, Comedy, Crime	PG-13	95	5.5	32	59984
## 14	Drama, Fantasy, Sport	PG-13	126	6.7	47	52831
## 83	Drama, Fantasy, Mystery	PG-13	118	6.2	31	51469
## 25	Comedy	PG-13	118	5.7	62	50682
## 44	Crime, Drama, Thriller	PG-13	110	6.1	47	48231
## 82	Comedy, Drama, Family	PG	118	5.3	33	33448
## 20	Animation, Action, Adventure	G	82	6.0	65	30195
## 50	Comedy, Drama, Romance	R	99	7.0	61	29701
## 101	Animation, Action, Adventure	PG	102	6.8	54	28764
## 94	Action, Comedy, Crime	R	111	6.1	46	25794
## 33	Drama	PG-13	114	7.3	57	23962
## 3	Comedy, Drama, Mystery	R	112	6.8	72	19473
## 24	Comedy, Crime	PG-13	86	5.2	43	18356
## 15	Comedy, Romance	PG-13	81	5.2	42	17781
## 37	Comedy, Drama, Romance	PG-13	97	6.0	NA	16855
## 2	Comedy	PG-13	111	5.1	NA	15447
## 106	Drama, Romance, Sport	Not Rated	146	2.3	NA	15200
## 38	Comedy, Crime, Drama	PG-13	105	6.1	63	10580
## 30	Comedy, Drama, Romance	R	91	7.4	65	10016
## 108	Documentary, Biography	TV-MA	124	7.6	60	3359
## 1	Crime, Drama, Thriller	R	105	6.6	NA	3112
## 29	Comedy	PG-13	97	3.7	15	2741
## 53	Drama	R	103	5.1	NA	1970
## 64	Comedy, Family	<NA>	105	6.4	NA	NA
## 46	Documentary	PG	101	7.8	78	853
## 95	Documentary, History, Music	Not Rated	90	6.6	69	840
## 16	Comedy	R	89	4.9	NA	715
## 9	Comedy	R	90	4.6	NA	683
## 51	Documentary, Comedy, Family	TV-14	87	6.3	59	668
## 58	Comedy	R	90	5.1	NA	627
## 28	Comedy, Romance	PG-13	90	5.5	NA	607
## 89	Documentary	Not Rated	102	6.9	73	565

## 103	Drama, Sport	<NA>	114	6.1	NA	NA
## 67	Documentary	Not Rated	98	6.2	49	126
## 27	Documentary	<NA>	NA	8.6	NA	NA
## 73	Thriller	PG-13	90	4.4	NA	41
## 86	Documentary, Music	Not Rated	52	5.7	NA	34
## 99	Comedy	<NA>	68	7.5	NA	NA
## 32	Documentary	<NA>	NA	5.6	NA	NA
## 18	Documentary	<NA>	55	5.2	NA	NA
## 71	Documentary	<NA>	NA	5.0	NA	NA
## 102	Documentary	<NA>	52	5.7	NA	NA
## 98	Action, Adventure, Comedy	<NA>	NA	NA	NA	NA
## 76	Action, Drama	<NA>	NA	NA	NA	NA
## 111	Drama	<NA>	NA	NA	NA	NA
## 105	Drama, Romance	PG-13	NA	NA	NA	NA
## 110	Action, Adventure, Crime	<NA>	NA	NA	NA	NA
## 52	Biography	<NA>	NA	NA	NA	NA
## 62	Comedy	<NA>	NA	NA	NA	NA
## 69	Action, Fantasy	<NA>	NA	NA	NA	NA
## 60	Action, Thriller	<NA>	NA	NA	NA	NA
## 72	Comedy	<NA>	NA	NA	NA	NA
## 91	Sci-Fi, Thriller	<NA>	NA	NA	NA	NA
## 55	Drama	<NA>	NA	NA	NA	NA
## 96	Action	<NA>	NA	NA	NA	NA
## 61	Drama	<NA>	NA	NA	NA	NA
## 107	Drama	<NA>	NA	NA	NA	NA
## 78	Drama	<NA>	NA	NA	NA	NA
## 40	Drama	<NA>	NA	NA	NA	NA
## 49	Musical	<NA>	NA	NA	NA	NA
## 68	Drama, Fantasy	<NA>	NA	NA	NA	NA
## 63	Drama	<NA>	NA	NA	NA	NA
## 92	Drama	<NA>	NA	NA	NA	NA
## 42	Comedy	<NA>	NA	NA	NA	NA
## 36	Comedy	<NA>	NA	NA	NA	NA
## 75	Animation	<NA>	NA	NA	NA	NA
## 81	Drama	<NA>	NA	NA	NA	NA
## 77	Drama	<NA>	NA	NA	NA	NA
## 85	Comedy	<NA>	NA	NA	NA	NA
## 88	Comedy	<NA>	NA	NA	NA	NA
## 59	Comedy	<NA>	NA	NA	NA	NA
## 79	Sci-Fi	<NA>	NA	NA	NA	NA
## 43	Comedy, Drama, Romance	<NA>	NA	NA	NA	NA
## 48	Adventure, Comedy, Family	<NA>	NA	NA	NA	NA
## 45	Drama	<NA>	NA	NA	NA	NA
## 109	<NA>	<NA>	NA	NA	NA	NA
## 56	Drama	<NA>	NA	NA	NA	NA
## 87	Comedy, Fantasy	<NA>	NA	NA	NA	NA
## 84	Drama	<NA>	NA	NA	NA	NA
##	millions					
## 39	256.39					
## 65	325.10					
## 6	306.17					
## 7	250.69					
## 26	144.80					
## 35	163.57					

## 34	227.95
## 12	190.42
## 66	179.02
## 31	179.50
## 41	69.95
## 4	65.81
## 17	138.61
## 8	111.55
## 104	355.56
## 13	206.04
## 90	53.86
## 80	60.52
## 74	54.76
## 57	127.35
## 100	NA
## 23	160.86
## 54	176.59
## 11	113.81
## 10	66.31
## 70	204.42
## 19	58.20
## 97	31.02
## 93	34.54
## 47	20.55
## 5	7.92
## 22	25.27
## 21	38.08
## 14	30.70
## 83	0.02
## 25	33.83
## 44	39.26
## 82	85.91
## 20	80.94
## 50	2.08
## 101	NA
## 94	4.97
## 33	37.77
## 3	6.41
## 24	34.31
## 15	30.63
## 37	NA
## 2	44.94
## 106	0.78
## 38	21.16
## 30	1.19
## 108	NA
## 1	0.39
## 29	11.54
## 53	NA
## 64	NA
## 46	0.05
## 95	NA
## 16	NA
## 9	5.48

## 51	NA
## 58	NA
## 28	NA
## 89	NA
## 103	NA
## 67	NA
## 27	NA
## 73	NA
## 86	NA
## 99	NA
## 32	NA
## 18	NA
## 71	NA
## 102	NA
## 98	NA
## 76	NA
## 111	NA
## 105	NA
## 110	NA
## 52	NA
## 62	NA
## 69	NA
## 60	NA
## 72	NA
## 91	NA
## 55	NA
## 96	NA
## 61	NA
## 107	NA
## 78	NA
## 40	NA
## 49	NA
## 68	NA
## 63	NA
## 92	NA
## 42	NA
## 36	NA
## 75	NA
## 81	NA
## 77	NA
## 85	NA
## 88	NA
## 59	NA
## 79	NA
## 43	NA
## 48	NA
## 45	NA
## 109	NA
## 56	NA
## 87	NA
## 84	NA
##	
## 39	
## 65	

6
7
26
35
34
12
66
31
41
4
17
8
104
13
90
80
74
57
100
23
54
11
10
70
19
97
93
47
5
22
21
14
83
25
44
82
20
50
101
94
33
3
24
15
37
2
106
38
30
108
1
29
53
64

It's the summer of 1994, and the streets of New York are puls

A group of tee

46 Their music is unforg
 ## 95 Fresh Dressed chron
 ## 16
 ## 9
 ## 51 Director Bryce Dallas Howard teams up with her father, <a href="/n
 ## 58
 ## 28
 ## 89
 ## 103
 ## 67
 ## 27 This document
 ## 73
 ## 86 Will Smith: L
 ## 99 An enormou
 ## 32
 ## 18
 ## 71
 ## 102
 ## 98
 ## 76
 ## 111
 ## 105
 ## 110
 ## 52
 ## 62
 ## 69
 ## 60
 ## 72
 ## 91
 ## 55
 ## 96
 ## 61
 ## 107 Green
 ## 78
 ## 40
 ## 49
 ## 68
 ## 63 When God crea
 ## 92
 ## 42
 ## 36
 ## 75
 ## 81
 ## 77
 ## 85
 ## 88
 ## 59
 ## 79
 ## 43
 ## 48
 ## 45
 ## 109
 ## 56
 ## 87

```
## 84
```

```
will.movies.ttids = will.movies$ttids;
```

Rankings: Lots of Ways to Internally Rank

```
# we could rank on other columns ...
```

```
# loadInflationData();
```

```
will.movies = standardizeDollarsInDataFrame(will.movies, 2000, "millions", "year", "millions2000");
```

```
# adjust for inflation ...
```

```
will.movies = sortDataFrameByNumericColumns(will.movies, "millions2000", "DESC"); ## NA's go to end ..
```

```
will.movies$rank.money = 1:nrow(will.movies);
```

```
will.movies = sortDataFrameByNumericColumns(will.movies, "minutes", "DESC");
```

```
will.movies$rank.minutes = 1:nrow(will.movies);
```

```
will.movies = sortDataFrameByNumericColumns(will.movies, "ratings", "DESC");
```

```
will.movies$rank.ratings = 1:nrow(will.movies);
```

```
will.movies = sortDataFrameByNumericColumns(will.movies, "metacritic", "DESC");
```

```
will.movies$rank.metacritic = 1:nrow(will.movies);
```

```
will.movies = sortDataFrameByNumericColumns(will.movies, "votes", "DESC");
```

```
will.movies$rank.votes = 1:nrow(will.movies);
```

```
will.movies = sortDataFrameByNumericColumns(will.movies, "year", "ASC");
```

```
will.movies$rank.year = 1:nrow(will.movies);
```

```
will.movies;
```

##	ttid	nmid	rank	year	title
## 1	tt0105810	nm0000226	53	1992	Where the Day Takes You
## 3	tt0108149	nm0000226	44	1993	Six Degrees of Separation
## 2	tt0107478	nm0000226	48	1993	Made in America
## 4	tt0112442	nm0000226	12	1995	Bad Boys
## 5	tt0114558	nm0000226	31	1995	Strange Days
## 6	tt0116629	nm0000226	3	1996	Independence Day
## 7	tt0119654	nm0000226	4	1997	Men in Black
## 8	tt0120660	nm0000226	14	1998	Enemy of the State
## 10	tt0120783	nm0000226	25	1998	The Parent Trap
## 16	tt0169376	nm0000226	59	1998	Welcome to Hollywood
## 9	tt0120707	nm0000226	60	1998	Ride
## 18	tt0205314	nm0000226	72	1998	Pesel Ha'Zahav
## 13	tt0145660	nm0000226	16	1999	Austin Powers: The Spy Who Shagged Me
## 11	tt0120891	nm0000226	24	1999	Wild Wild West
## 15	tt0167427	nm0000226	46	1999	Superstar
## 14	tt0146984	nm0000226	34	2000	The Legend of Bagger Vance
## 19	tt0248667	nm0000226	27	2001	Ali
## 20	tt0268397	nm0000226	39	2001	Jimmy Neutron: Boy Genius

## 12	tt0120912	nm0000226	8	2002	Men in Black II
## 21	tt0284490	nm0000226	33	2002	Showtime
## 17	tt0172156	nm0000226	13	2003	Bad Boys II
## 25	tt0338466	nm0000226	36	2003	Stuck on You
## 24	tt0328099	nm0000226	45	2003	Malibu's Most Wanted
## 27	tt0375663	nm0000226	67	2003	A Closer Walk
## 26	tt0343818	nm0000226	5	2004	I, Robot
## 23	tt0307453	nm0000226	22	2004	Shark Tale
## 22	tt0300051	nm0000226	32	2004	Jersey Girl
## 30	tt0384504	nm0000226	51	2004	Saving Face
## 29	tt0380277	nm0000226	54	2004	The Cookout
## 28	tt0379487	nm0000226	63	2004	The Seat Filler
## 31	tt0386588	nm0000226	10	2005	Hitch
## 86	tt2074488	nm0000226	69	2005	Will Smith: Live in Concert
## 99	tt5230508	nm0000226	70	2005	AMV Hell 3: The Motion Picture
## 32	tt0414587	nm0000226	71	2005	There's a God on the Mic
## 35	tt0454921	nm0000226	6	2006	The Pursuit of Happyness
## 38	tt0466856	nm0000226	50	2006	ATL
## 39	tt0480249	nm0000226	1	2007	I Am Legend
## 37	tt0466839	nm0000226	47	2007	I Could Never Be Your Woman
## 71	tt1546668	nm0000226	73	2007	The 100 Best Black Movies (Ever)
## 34	tt0448157	nm0000226	7	2008	Hancock
## 41	tt0814314	nm0000226	11	2008	Seven Pounds
## 44	tt0947802	nm0000226	37	2008	Lakeview Terrace
## 50	tt1082886	nm0000226	40	2008	The Wackness
## 33	tt0416212	nm0000226	43	2008	The Secret Life of Bees
## 53	tt1109477	nm0000226	55	2008	The Human Contract
## 64	tt1383607	nm0000226	56	2008	Ramadan Mabrouk Abul-Alamein Hamouda
## 46	tt1015971	nm0000226	57	2009	The Boys: The Sherman Brothers' Story
## 58	tt1230214	nm0000226	62	2009	Stuntmen
## 102	tt5839150	nm0000226	74	2009	The Muhammad Ali Story
## 54	tt1155076	nm0000226	23	2010	The Karate Kid
## 67	tt1409772	nm0000226	66	2010	A Man's Story
## 66	tt1409024	nm0000226	9	2012	Men in Black 3
## 74	tt1596350	nm0000226	19	2012	This Means War
## 89	tt2350432	nm0000226	64	2012	Free Angela and All Political Prisoners
## 80	tt1815862	nm0000226	18	2013	After Earth
## 57	tt1229340	nm0000226	20	2013	Anchorman 2: The Legend Continues
## 83	tt1837709	nm0000226	35	2014	Winter's Tale
## 82	tt1823664	nm0000226	38	2014	Annie
## 73	tt1584918	nm0000226	68	2014	Murder101
## 90	tt2381941	nm0000226	17	2015	Focus
## 93	tt3322364	nm0000226	29	2015	Concussion
## 95	tt4209900	nm0000226	58	2015	Fresh Dressed
## 65	tt1386697	nm0000226	2	2016	Suicide Squad
## 97	tt4682786	nm0000226	28	2016	Collateral Beauty
## 100	tt5519340	nm0000226	21	2017	Bright
## 94	tt3721964	nm0000226	42	2018	Gringo
## 108	tt7440432	nm0000226	52	2018	Quincy
## 103	tt5903926	nm0000226	65	2018	Sprinter
## 104	tt6139732	nm0000226	15	2019	Aladdin
## 47	tt1025100	nm0000226	30	2019	Gemini Man
## 101	tt5814534	nm0000226	41	2019	Spies in Disguise
## 106	tt7255568	nm0000226	49	2019	Student of the Year 2

## 51	tt10883124	nm0000226	61	2019	Dads		
## 70	tt1502397	nm0000226	26	2020	Bad Boys For Life		
## 105	tt6598238	nm0000226	78	2020	Life in a Year		
## 111	tt9620288	nm0000226	77	2021	King Richard		
## 98	tt4919268	nm0000226	75	NA	Bad Boys 4		
## 76	tt1674782	nm0000226	76	NA	The Karate Kid 2		
## 110	tt7820302	nm0000226	79	NA	Bright 2		
## 52	tt11012318	nm0000226	80	NA	The Council		
## 62	tt12915962	nm0000226	81	NA	Planes, Trains & Automobiles		
## 69	tt1498780	nm0000226	82	NA	Hancock 2		
## 60	tt12530246	nm0000226	83	NA	Emancipation		
## 72	tt1578276	nm0000226	84	NA	Uptown Saturday Night		
## 91	tt2776710	nm0000226	85	NA	Brilliance		
## 55	tt1181790	nm0000226	86	NA	The Billionaire's Vinegar		
## 96	tt4463004	nm0000226	87	NA	Bounty		
## 61	tt12608018	nm0000226	88	NA	Bounce		
## 107	tt7331400	nm0000226	89	NA	1921		
## 78	tt1692497	nm0000226	90	NA	The Redemption of Cain		
## 40	tt0812268	nm0000226	91	NA	The Last Pharaoh		
## 49	tt10344550	nm0000226	92	NA	Untitled Solvan Naim Project		
## 68	tt1470831	nm0000226	93	NA	The City That Sailed		
## 63	tt1368421	nm0000226	94	NA	Angelology		
## 92	tt2846190	nm0000226	95	NA	La escribana de uraba		
## 42	tt0911023	nm0000226	96	NA	My Wife Hates Your Wife		
## 36	tt0464265	nm0000226	97	NA	Time Share		
## 75	tt1609486	nm0000226	98	NA	Harold and the Purple Crayon		
## 81	tt1817789	nm0000226	99	NA	Wheeler Dealers		
## 77	tt1691439	nm0000226	100	NA	Untitled Dr. S. Allen Counter Project		
## 85	tt1879098	nm0000226	101	NA	What Would Kenny Do?		
## 88	tt2246761	nm0000226	102	NA	Interpol		
## 59	tt1235550	nm0000226	103	NA	Welcome to the Sticks		
## 79	tt1714201	nm0000226	104	NA	Monster Witness Relocation Program		
## 43	tt0942916	nm0000226	105	NA	Sisters of Mercy		
## 48	tt1028582	nm0000226	106	NA	Newton's Law		
## 45	tt0979959	nm0000226	107	NA	Cooked		
## 109	tt7737796	nm0000226	108	NA	The Prospect		
## 56	tt1206884	nm0000226	109	NA	The Long Run		
## 87	tt2168954	nm0000226	110	NA	Ghost Graduation		
## 84	tt1852816	nm0000226	111	NA	Joe		
##		genre	rated	minutes	ratings	metacritic	votes
## 1	Crime, Drama, Thriller	R	105	6.6	NA	3112	
## 3	Comedy, Drama, Mystery	R	112	6.8	72	19473	
## 2	Comedy	PG-13	111	5.1	NA	15447	
## 4	Action, Comedy, Crime	R	119	6.9	41	235766	
## 5	Action, Crime, Drama	R	145	7.2	66	66102	
## 6	Action, Adventure, Sci-Fi	PG-13	145	7.0	59	520592	
## 7	Action, Adventure, Comedy	PG-13	98	7.3	71	507597	
## 8	Action, Thriller	R	132	7.3	67	224019	
## 10	Adventure, Comedy, Drama	PG	128	6.5	64	117780	
## 16	Comedy	R	89	4.9	NA	715	
## 9	Comedy	R	90	4.6	NA	683	
## 18	Documentary	<NA>	55	5.2	NA	NA	
## 13	Action, Adventure, Comedy	PG-13	95	6.6	59	214064	
## 11	Action, Comedy, Sci-Fi	PG-13	106	5.0	38	151132	

## 15	Comedy, Romance	PG-13	81	5.2	42	17781
## 14	Drama, Fantasy, Sport	PG-13	126	6.7	47	52831
## 19	Biography, Drama, Sport	R	157	6.8	65	92508
## 20	Animation, Action, Adventure	G	82	6.0	65	30195
## 12	Action, Adventure, Comedy	PG-13	88	6.2	49	337458
## 21	Action, Comedy, Crime	PG-13	95	5.5	32	59984
## 17	Action, Comedy, Crime	R	147	6.6	38	228329
## 25	Comedy	PG-13	118	5.7	62	50682
## 24	Comedy, Crime	PG-13	86	5.2	43	18356
## 27	Documentary	<NA>	NA	8.6	NA	NA
## 26	Action, Drama, Sci-Fi	PG-13	115	7.1	59	491474
## 23	Animation, Adventure, Comedy	PG	90	6.0	48	164132
## 22	Comedy, Drama, Romance	PG-13	102	6.2	43	62700
## 30	Comedy, Drama, Romance	R	91	7.4	65	10016
## 29	Comedy	PG-13	97	3.7	15	2741
## 28	Comedy, Romance	PG-13	90	5.5	NA	607
## 31	Comedy, Romance	PG-13	118	6.6	58	291991
## 86	Documentary, Music	Not Rated	52	5.7	NA	34
## 99	Comedy	<NA>	68	7.5	NA	NA
## 32	Documentary	<NA>	NA	5.6	NA	NA
## 35	Biography, Drama	PG-13	117	8.0	64	438121
## 38	Comedy, Crime, Drama	PG-13	105	6.1	63	10580
## 39	Action, Adventure, Drama	PG-13	101	7.2	65	675160
## 37	Comedy, Drama, Romance	PG-13	97	6.0	NA	16855
## 71	Documentary	<NA>	NA	5.0	NA	NA
## 34	Action, Fantasy	PG-13	92	6.4	49	437887
## 41	Drama	PG-13	123	7.6	36	282568
## 44	Crime, Drama, Thriller	PG-13	110	6.1	47	48231
## 50	Comedy, Drama, Romance	R	99	7.0	61	29701
## 33	Drama	PG-13	114	7.3	57	23962
## 53	Drama	R	103	5.1	NA	1970
## 64	Comedy, Family	<NA>	105	6.4	NA	NA
## 46	Documentary	PG	101	7.8	78	853
## 58	Comedy	R	90	5.1	NA	627
## 102	Documentary	<NA>	52	5.7	NA	NA
## 54	Action, Drama, Family	PG	140	6.2	61	158842
## 67	Documentary	Not Rated	98	6.2	49	126
## 66	Action, Adventure, Comedy	PG-13	106	6.8	58	329948
## 74	Action, Comedy, Romance	PG-13	103	6.3	31	174205
## 89	Documentary	Not Rated	102	6.9	73	565
## 80	Action, Adventure, Sci-Fi	PG-13	100	4.8	33	189337
## 57	Comedy	PG-13	119	6.3	61	165909
## 83	Drama, Fantasy, Mystery	PG-13	118	6.2	31	51469
## 82	Comedy, Drama, Family	PG	118	5.3	33	33448
## 73	Thriller	PG-13	90	4.4	NA	41
## 90	Comedy, Crime, Drama	R	105	6.6	56	212789
## 93	Biography, Drama, Sport	PG-13	123	7.1	55	84750
## 95	Documentary, History, Music	Not Rated	90	6.6	69	840
## 65	Action, Adventure, Fantasy	PG-13	123	6.0	40	588064
## 97	Drama, Romance	PG-13	97	6.8	23	89138
## 100	Action, Fantasy, Thriller	TV-MA	117	6.3	29	164913
## 94	Action, Comedy, Crime	R	111	6.1	46	25794
## 108	Documentary, Biography	TV-MA	124	7.6	60	3359
## 103	Drama, Sport	<NA>	114	6.1	NA	NA

## 104	Adventure, Family, Fantasy	PG	128	7.0	53	216257
## 47	Action, Drama, Sci-Fi	PG-13	117	5.7	38	76145
## 101	Animation, Action, Adventure	PG	102	6.8	54	28764
## 106	Drama, Romance, Sport	Not Rated	146	2.3	NA	15200
## 51	Documentary, Comedy, Family	TV-14	87	6.3	59	668
## 70	Action, Comedy, Crime	R	124	6.6	59	112826
## 105	Drama, Romance	PG-13	NA	NA	NA	NA
## 111	Drama	<NA>	NA	NA	NA	NA
## 98	Action, Adventure, Comedy	<NA>	NA	NA	NA	NA
## 76	Action, Drama	<NA>	NA	NA	NA	NA
## 110	Action, Adventure, Crime	<NA>	NA	NA	NA	NA
## 52	Biography	<NA>	NA	NA	NA	NA
## 62	Comedy	<NA>	NA	NA	NA	NA
## 69	Action, Fantasy	<NA>	NA	NA	NA	NA
## 60	Action, Thriller	<NA>	NA	NA	NA	NA
## 72	Comedy	<NA>	NA	NA	NA	NA
## 91	Sci-Fi, Thriller	<NA>	NA	NA	NA	NA
## 55	Drama	<NA>	NA	NA	NA	NA
## 96	Action	<NA>	NA	NA	NA	NA
## 61	Drama	<NA>	NA	NA	NA	NA
## 107	Drama	<NA>	NA	NA	NA	NA
## 78	Drama	<NA>	NA	NA	NA	NA
## 40	Drama	<NA>	NA	NA	NA	NA
## 49	Musical	<NA>	NA	NA	NA	NA
## 68	Drama, Fantasy	<NA>	NA	NA	NA	NA
## 63	Drama	<NA>	NA	NA	NA	NA
## 92	Drama	<NA>	NA	NA	NA	NA
## 42	Comedy	<NA>	NA	NA	NA	NA
## 36	Comedy	<NA>	NA	NA	NA	NA
## 75	Animation	<NA>	NA	NA	NA	NA
## 81	Drama	<NA>	NA	NA	NA	NA
## 77	Drama	<NA>	NA	NA	NA	NA
## 85	Comedy	<NA>	NA	NA	NA	NA
## 88	Comedy	<NA>	NA	NA	NA	NA
## 59	Comedy	<NA>	NA	NA	NA	NA
## 79	Sci-Fi	<NA>	NA	NA	NA	NA
## 43	Comedy, Drama, Romance	<NA>	NA	NA	NA	NA
## 48	Adventure, Comedy, Family	<NA>	NA	NA	NA	NA
## 45	Drama	<NA>	NA	NA	NA	NA
## 109	<NA>	<NA>	NA	NA	NA	NA
## 56	Drama	<NA>	NA	NA	NA	NA
## 87	Comedy, Fantasy	<NA>	NA	NA	NA	NA
## 84	Drama	<NA>	NA	NA	NA	NA
##	millions					
## 1	0.39					
## 3	6.41					
## 2	44.94					
## 4	65.81					
## 5	7.92					
## 6	306.17					
## 7	250.69					
## 8	111.55					
## 10	66.31					
## 16	NA					

## 9	5.48
## 18	NA
## 13	206.04
## 11	113.81
## 15	30.63
## 14	30.70
## 19	58.20
## 20	80.94
## 12	190.42
## 21	38.08
## 17	138.61
## 25	33.83
## 24	34.31
## 27	NA
## 26	144.80
## 23	160.86
## 22	25.27
## 30	1.19
## 29	11.54
## 28	NA
## 31	179.50
## 86	NA
## 99	NA
## 32	NA
## 35	163.57
## 38	21.16
## 39	256.39
## 37	NA
## 71	NA
## 34	227.95
## 41	69.95
## 44	39.26
## 50	2.08
## 33	37.77
## 53	NA
## 64	NA
## 46	0.05
## 58	NA
## 102	NA
## 54	176.59
## 67	NA
## 66	179.02
## 74	54.76
## 89	NA
## 80	60.52
## 57	127.35
## 83	0.02
## 82	85.91
## 73	NA
## 90	53.86
## 93	34.54
## 95	NA
## 65	325.10
## 97	31.02

```

## 100      NA
## 94      4.97
## 108      NA
## 103      NA
## 104    355.56
## 47     20.55
## 101      NA
## 106      0.78
## 51      NA
## 70    204.42
## 105      NA
## 111      NA
## 98      NA
## 76      NA
## 110      NA
## 52      NA
## 62      NA
## 69      NA
## 60      NA
## 72      NA
## 91      NA
## 55      NA
## 96      NA
## 61      NA
## 107      NA
## 78      NA
## 40      NA
## 49      NA
## 68      NA
## 63      NA
## 92      NA
## 42      NA
## 36      NA
## 75      NA
## 81      NA
## 77      NA
## 85      NA
## 88      NA
## 59      NA
## 79      NA
## 43      NA
## 48      NA
## 45      NA
## 109      NA
## 56      NA
## 87      NA
## 84      NA
##
## 1
## 3
## 2
## 4
## 5
## 6

```

A group of te

7
8
10
16
9
18
13
11
15
14
19
20
12
21
17
25
24
27
26
23
22
30
29
28
31
86
99
32
35
38
39
37
71
34
41
44
50
33
53
64
46
58
102
54
67
66
74
89
80
57
83
82
73
90

This document

Will Smith: L.

An enormously

It's the summer of 1994, and the streets of New York are puls

Their music is unforg

```

## 93
## 95
## 65
## 97
## 100
## 94
## 108
## 103
## 104
## 47
## 101
## 106
## 51 Director <a href="/name/nm0397171">Bryce Dallas Howard</a> teams up with her father, <a href="/n
## 70
## 105
## 111
## 98
## 76
## 110
## 52
## 62
## 69
## 60
## 72
## 91
## 55
## 96
## 61
## 107
## 78
## 40
## 49
## 68
## 63
## 92
## 42
## 36
## 75
## 81
## 77
## 85
## 88
## 59
## 79
## 43
## 48
## 45
## 109
## 56
## 87
## 84
##      millions2000 rank.money rank.minutes rank.ratings rank.metacritic
## 1      0.47867427      50      36      30      58
## 3      7.63876817      44      28      21      3

```

Fresh Dressed chron

Green

When God crea

## 2	53.55479585	26	29	64	67
## 4	74.36011811	21	17	18	42
## 5	8.94897638	43	5	11	7
## 6	336.02596558	1	4	15	21
## 7	268.96459813	2	47	10	4
## 8	117.84607362	17	7	8	6
## 10	70.05265031	22	9	33	13
## 16	NA	60	62	69	71
## 9	5.78930061	45	57	71	72
## 18	NA	72	69	63	66
## 13	212.96571429	5	52	31	23
## 11	117.63554622	18	33	67	46
## 15	31.65957983	33	67	62	41
## 14	30.70000000	35	10	25	36
## 19	56.58972332	24	1	20	10
## 20	78.70055336	20	66	52	11
## 12	182.26972763	8	63	44	34
## 21	36.45011673	30	53	58	50
## 17	129.72088043	16	2	26	44
## 25	31.66046739	32	20	53	15
## 24	32.10968478	31	65	61	40
## 27	NA	67	72	1	56
## 26	131.99872949	15	25	14	20
## 23	146.63892006	10	56	51	35
## 22	23.03596612	38	40	42	39
## 30	1.08479619	48	55	7	8
## 29	10.51978825	42	50	73	55
## 28	NA	63	60	59	65
## 31	158.26881720	9	18	28	26
## 86	NA	69	70	55	62
## 99	NA	70	68	6	57
## 32	NA	71	73	57	64
## 35	139.71604167	11	22	2	12
## 38	18.07416667	40	35	48	14
## 39	212.93494806	6	43	12	9
## 37	NA	55	51	50	61
## 71	NA	73	74	68	70
## 34	182.31510940	7	54	35	32
## 41	55.94622462	25	14	5	47
## 44	31.40026846	34	31	47	37
## 50	1.66359038	47	46	17	16
## 33	30.20856189	36	26	9	27
## 53	NA	57	39	65	68
## 64	NA	58	37	34	59
## 46	0.04013294	51	44	3	1
## 58	NA	62	59	66	69
## 102	NA	74	71	56	63
## 54	139.45407602	12	6	40	18
## 67	NA	66	48	43	33
## 66	134.26850876	14	32	22	25
## 74	41.07107329	28	38	38	51
## 89	NA	64	42	19	2
## 80	44.73591264	27	45	70	49
## 57	94.13612813	19	16	36	17

## 83	0.01454785	52	21	41	52
## 82	62.49029298	23	19	60	48
## 73	NA	68	61	72	73
## 90	39.13091466	29	34	29	28
## 93	25.09435188	37	15	13	29
## 95	NA	59	58	32	5
## 65	233.25244680	4	13	49	43
## 97	22.25620086	39	49	24	54
## 100	NA	53	24	37	53
## 94	3.40657565	46	30	46	38
## 108	NA	56	12	4	19
## 103	NA	65	27	45	60
## 104	239.49006777	3	8	16	31
## 47	13.84160449	41	23	54	45
## 101	NA	54	41	23	30
## 106	0.52537477	49	3	74	74
## 51	NA	61	64	39	24
## 70	135.85869603	13	11	27	22
## 105	NA	78	78	78	78
## 111	NA	77	77	77	77
## 98	NA	75	75	75	75
## 76	NA	76	76	76	76
## 110	NA	79	79	79	79
## 52	NA	80	80	80	80
## 62	NA	81	81	81	81
## 69	NA	82	82	82	82
## 60	NA	83	83	83	83
## 72	NA	84	84	84	84
## 91	NA	85	85	85	85
## 55	NA	86	86	86	86
## 96	NA	87	87	87	87
## 61	NA	88	88	88	88
## 107	NA	89	89	89	89
## 78	NA	90	90	90	90
## 40	NA	91	91	91	91
## 49	NA	92	92	92	92
## 68	NA	93	93	93	93
## 63	NA	94	94	94	94
## 92	NA	95	95	95	95
## 42	NA	96	96	96	96
## 36	NA	97	97	97	97
## 75	NA	98	98	98	98
## 81	NA	99	99	99	99
## 77	NA	100	100	100	100
## 85	NA	101	101	101	101
## 88	NA	102	102	102	102
## 59	NA	103	103	103	103
## 79	NA	104	104	104	104
## 43	NA	105	105	105	105
## 48	NA	106	106	106	106
## 45	NA	107	107	107	107
## 109	NA	108	108	108	108
## 56	NA	109	109	109	109
## 87	NA	110	110	110	110

## 84	NA	111	111	111	111
##	rank.votes	rank.year			
## 1	53	1			
## 3	44	2			
## 2	48	3			
## 4	12	4			
## 5	31	5			
## 6	3	6			
## 7	4	7			
## 8	14	8			
## 10	25	9			
## 16	58	10			
## 9	59	11			
## 18	73	12			
## 13	16	13			
## 11	24	14			
## 15	46	15			
## 14	34	16			
## 19	27	17			
## 20	39	18			
## 12	8	19			
## 21	33	20			
## 17	13	21			
## 25	36	22			
## 24	45	23			
## 27	67	24			
## 26	5	25			
## 23	22	26			
## 22	32	27			
## 30	51	28			
## 29	54	29			
## 28	62	30			
## 31	10	31			
## 86	66	32			
## 99	68	33			
## 32	72	34			
## 35	6	35			
## 38	50	36			
## 39	1	37			
## 37	47	38			
## 71	74	39			
## 34	7	40			
## 41	11	41			
## 44	37	42			
## 50	40	43			
## 33	43	44			
## 53	55	45			
## 64	69	46			
## 46	56	47			
## 58	61	48			
## 102	71	49			
## 54	23	50			
## 67	64	51			
## 66	9	52			

## 74	19	53
## 89	63	54
## 80	18	55
## 57	20	56
## 83	35	57
## 82	38	58
## 73	65	59
## 90	17	60
## 93	29	61
## 95	57	62
## 65	2	63
## 97	28	64
## 100	21	65
## 94	42	66
## 108	52	67
## 103	70	68
## 104	15	69
## 47	30	70
## 101	41	71
## 106	49	72
## 51	60	73
## 70	26	74
## 105	78	75
## 111	77	76
## 98	75	77
## 76	76	78
## 110	79	79
## 52	80	80
## 62	81	81
## 69	82	82
## 60	83	83
## 72	84	84
## 91	85	85
## 55	86	86
## 96	87	87
## 61	88	88
## 107	89	89
## 78	90	90
## 40	91	91
## 49	92	92
## 68	93	93
## 63	94	94
## 92	95	95
## 42	96	96
## 36	97	97
## 75	98	98
## 81	99	99
## 77	100	100
## 85	101	101
## 88	102	102
## 59	103	103
## 79	104	104
## 43	105	105
## 48	106	106

```
## 45      107      107
## 109     108      108
## 56      109      109
## 87      110      110
## 84      111      111
```

We now have all the movies, not just the top 50. This data is similar to what we had in our earlier assignments. We now have a bit more data.

Unique Character Names

```
will.characters = IMDB.getUniqueCharactersForPerson(will.nmid);
will.characters;
```

```
##      count      character
## 30      3      Self
## 8       2      Daryl Ward
## 18      2      Jay
## 1       1      Agent J
## 2       1      Bagger Vance
## 3       1      Ben
## 4       1      Capt. Steven Hiller
## 5       1 Cassius Clay /      Muhammad Ali
## 6       1      Chris Gardner
## 7       1      Cypher Raige
## 9       1      Deadshot
## 10      1      Del Spooner
## 11      1      Detective Mike Lowrey
## 12      1      Dr. Bennet Omalu
## 13      1      Genie /      Mariner
## 14      1      Henry Brogan /      Junior
## 15      1      Hitch
## 16      1      Howard
## 17      1      James West
## 19      1      John Hancock
## 20      1      Lance      (voice)
## 21      1      Manny
## 22      1      Mike
## 23      1      Mike Lowrey
## 24      1      Nicky
## 25      1      Oscar      (voice)
## 26      1      Paul
## 27      1      Richard Williams
## 28      1      Robert Clayton Dean
## 29      1      Robert Neville
## 31      1      Tea Cake Walters
```

Unless he plays another gender in a film (e.g., Adam Sandler playing a woman), we could use this information for a classifier as a dictionary. Or we could just build a classifier from the biography.

Gender

```
will.bio = will.search$bio[1];

# you could get just his record with
```

```
will.info = IMDB.getPersonInfo(will.nmid);
# doesn't rely on the search ...
# imdb$all.actors.info
```

```
will.bio;
```

```
## [1] "Willard Carroll \"Will\" Smith, Jr. (born September 25, 1968) is an American actor, comedian, p
gender = list();
gender[["male"]] = c("he", "his", "him");
gender[["female"]] = c("she", "hers", "her");
```

We could use more features than the above, but the IMDB bio template seems to utilize these gender words. And, if a relative (e.g., mother or brother) is mentioned, it may confuse our classifier.

```
predictGenderFromBio = function(bio, gender = list(
  "male" = c("he", "his", "him"),
  "female" = c("she", "hers", "her"))) {
  {
    bio = gsub("(", "", bio, fixed=TRUE);
    bio = gsub(")", "", bio, fixed=TRUE);
    bio = gsub("'", "", bio, fixed=TRUE);

    words = getRawWords(bio); # from nlp, converts to lower

    mycounts = list("male" = 0, "female" = 0);
    tc = 0;
    gc = 2;
    for(i in 1:gc)
    {
      key = names(gender)[i];
      for(j in 1:length(gender[[key]]))
      {
        mc = sum(words == gender[[key]][j]);
        mycounts[[key]] = mc + mycounts[[key]];
        tc = mc + tc;
      }
    }
    # mycounts;
    # we could return a ratio ... but a decision for this example would likely be best ...
    if(tc == 0) { return (NA); } # no data ...
    result = "male"; # default in hollywood?
    if(mycounts$female > mycounts$male) { result = "female"; }
    result;
  }
}
```

```
will.gender = predictGenderFromBio(will.bio);
will.gender;
```

```
## [1] "male"
```

Actor Ranking on Given Film

```
ttid.legend = "tt0480249";
rank.legend = subsetDataFrame(imdb.data$all.movies.actors.characters, "ttid", "==", ttid.legend);
rank.legend;
```

##	ttid	actor.rank	nmid	character
## 599482	tt0480249	1	nm0000226	
## 599483	tt0480249	2	nm0103797	Robert Neville
## 599484	tt0480249	3	nm2253071	Anna
## 599485	tt0480249	4	nm0724757	Ethan
## 599486	tt0480249	5	nm2405238	(as Salli Richardson)
## 599487	tt0480249	6	nm0287756	Marley Neville
## 599488	tt0480249	7	nm0333318	Mike - Military Escort
## 599489	tt0480249	8	nm0005231	TV Personality
## 599490	tt0480249	9	nm2724906	Alpha Male
## 599491	tt0480249	10	nm8421984	Alpha Female
## 599492	tt0480249	11	nm2869725	Sam
## 599493	tt0480249	12	nm2677021	Sam
## 599494	tt0480249	13	nm0565503	Military Driver - Jay
## 599495	tt0480249	14	nm1677477	Male Evacuee
## 599496	tt0480249	15	nm2857229	Woman Evacuee
##				Sergeant

Will's Rankings on Some Films

This is based on the number of movies in the database from which we grabbed the full movie-info page. If you review how that list was created, it was based on top-50 films each year (either on rating or votes).

```
will.rank = subsetDataFrame(imdb.data$all.movies.actors.characters, "nmid", "==", will.nmid);
will.rank;
```

##	ttid	actor.rank	nmid
## 349353	tt0105810	6	nm0000226
## 353669	tt0107478	3	nm0000226
## 355578	tt0108149	2	nm0000226
## 365501	tt0112442	11	nm0000226
## 376889	tt0116629	1	nm0000226
## 387158	tt0119654	2	nm0000226
## 391606	tt0120660	1	nm0000226

##	393444	tt0120891	1	nm0000226	
##	393576	tt0120912	2	nm0000226	
##	413566	tt0146984	1	nm0000226	
##	434016	tt0172156	2	nm0000226	
##	460734	tt0205314	11	nm0000226	
##	486656	tt0248667	1	nm0000226	
##	521274	tt0307453	1	nm0000226	
##	540127	tt0343818	1	nm0000226	
##	562065	tt0386588	1	nm0000226	
##	586225	tt0448157	1	nm0000226	
##	588916	tt0454921	1	nm0000226	
##	599482	tt0480249	1	nm0000226	
##	613408	tt0814314	1	nm0000226	
##	633149	tt1025100	1	nm0000226	
##	680397	tt1386697	1	nm0000226	
##	682251	tt1409024	1	nm0000226	
##	689408	tt1502397	1	nm0000226	
##	693711	tt1546668	9	nm0000226	
##	715818	tt1815862	2	nm0000226	
##	750095	tt2381941	1	nm0000226	
##	778362	tt3322364	1	nm0000226	
##	809470	tt4682786	1	nm0000226	
##	825451	tt5519340	1	nm0000226	
##	831014	tt5814534	6	nm0000226	
##	837383	tt6139732	1	nm0000226	
##	859453	tt7820302	1	nm0000226	
##	876732	tt9620288	1	nm0000226	
##	642982	tt10883124	13	nm0000226	
##					character
##	349353				Manny
##	353669				Tea Cake Walters
##	355578				Paul
##	365501				Mike Lowrey
##	376889				Capt. Steven Hiller
##	387158				Jay
##	391606				Robert Clayton Dean
##	393444				James West
##	393576				Jay
##	413566				Bagger Vance
##	434016				Detective Mike Lowrey
##	460734				Self
##	486656	Cassius Clay /			Muhammad Ali
##	521274	Oscar			(voice)
##	540127				Del Spooner
##	562065				Hitch
##	586225				John Hancock
##	588916				Chris Gardner
##	599482				Robert Neville
##	613408				Ben
##	633149	Henry Brogan /			Junior
##	680397				Deadshot
##	682251				Agent J
##	689408				Mike
##	693711				Self

```
## 715818                Cypher Raige
## 750095                Nicky
## 778362                Dr. Bennet Omalu
## 809470                Howard
## 825451                Daryl Ward
## 831014                Lance          (voice)
## 837383                Genie /        Mariner
## 859453                Daryl Ward
## 876732                Richard Williams
## 642982                Self
```

We see that 35 of his 111 films are in this dataset.

Movie Details

See http://md5.mshaffer.com/WSU_STATS419/01_imdb-monte.html.

See http://md5.mshaffer.com/WSU_STATS419/02_imdb-monte.html.

```
ttid.legend = will.movies$ttid[1]; # I am legend
ttid.legend = "tt0480249";
```

```
info.legend = IMDB.getMovieInfoFromActorSearch(ttid.legend);
info.legend; # this is the same dataframe as the search for will above
```

```
##          ttid year      title          genre rated minutes ratings
## 19065 tt0480249 2007 I Am Legend Action, Adventure, Drama PG-13    101    7.2
##          metacritic votes millions
## 19065          65 675160    256.39
```

```
##
## 19065 Years after a plague kills most of humanity and transforms the rest into monsters, the sole survivor
```

now we have nmid and ttid, we can access other data directly ...

```
# names(imdb.data);
```

```
creatives = subsetDataFrame(imdb.data$all.movies.creatives, "ttid", "==", ttid.legend);
creatives; # these are top billed
```

```
##          ttid creative.rank      nmid      role
## 263761 tt0480249           1 nm1349376 director
## 263762 tt0480249           2 nm0698873  writer
## 263763 tt0480249           3 nm0326040  writer
## 263764 tt0480249           4 nm0000226   star
## 263765 tt0480249           5 nm0103797   star
## 263766 tt0480249           6 nm2253071   star
```

```
companies = subsetDataFrame(imdb.data$all.movies.companies, "ttid", "==", ttid.legend);
companies;
```

```
##          ttid co.rank      coid          name
## 75588 tt0480249      1 co0002663      Warner Bros.
## 75589 tt0480249      2 co0108864 Village Roadshow Pictures
## 75590 tt0480249      3 co0093488      Weed Road Pictures
```

this is the TOP-15 with character name ...

```
actors = subsetDataFrame(imdb.data$all.movies.actors.characters, "ttid", "==", ttid.legend);
actors;
```

```
##          ttid actor.rank      nmid
## 599482 tt0480249         1 nm0000226
## 599483 tt0480249         2 nm0103797
## 599484 tt0480249         3 nm2253071
## 599485 tt0480249         4 nm0724757
## 599486 tt0480249         5 nm2405238
## 599487 tt0480249         6 nm0287756
## 599488 tt0480249         7 nm0333318
## 599489 tt0480249         8 nm0005231
## 599490 tt0480249         9 nm2724906
## 599491 tt0480249        10 nm8421984
## 599492 tt0480249        11 nm2869725
## 599493 tt0480249        12 nm2677021
## 599494 tt0480249        13 nm0565503
## 599495 tt0480249        14 nm1677477
## 599496 tt0480249        15 nm2857229
```

```
##                                     character
## 599482                               Robert Neville
## 599483                               Anna
## 599484                               Ethan
## 599485 Zoe Neville                   (as Salli Richardson)
## 599486                               Marley Neville
## 599487                               Mike - Military Escort
## 599488                               TV Personality
## 599489                               Alpha Male
## 599490                               Alpha Female
## 599491                               Sam
## 599492                               Sam
## 599493                               Military Driver - Jay
## 599494                               Male Evacuee
## 599495                               Woman Evacuee
## 599496                               Sergeant
```

```
# imdb.data$all.actors.rank
```

```
# there is another dataframe that will show more than 15 persons, because those "actors" were linked ba
others = subsetDataframe(imdb.data$all.actors.rank, "ttid", "==", ttid.legend);
others; # there are 40 persons in the dataset with various "roles": actors, directors, stunt men ...
```

```
##          nmid      ttid rank
## 19065   nm0000226 tt0480249  1
## 49625   nm0000668 tt0480249  2
## 131513  nm0002184 tt0480249  7
## 133698  nm0002211 tt0480249  7
## 142363  nm0002490 tt0480249  1
## 143179  nm0002514 tt0480249  5
## 143373  nm0002523 tt0480249  5
## 148803  nm0002696 tt0480249  3
## 154866  nm0002885 tt0480249  3
## 159374  nm0003058 tt0480249  2
## 162190  nm0003179 tt0480249  4
## 169050  nm0003388 tt0480249  1
## 209357  nm0004581 tt0480249  2
## 225272  nm0005231 tt0480249  1
## 312883  nm0050924 tt0480249  2
## 384721  nm0103797 tt0480249  2
```



```
## 610251 nm0258523 tt0480249 5
## 633758 nm0273975 tt0480249 5
## 653194 nm0287756 tt0480249 1
## 707032 nm0326040 tt0480249 2
## 716738 nm0333318 tt0480249 3
## 793694 nm0387894 tt0480249 2
## 1049662 nm0558577 tt0480249 2
## 1059504 nm0565503 tt0480249 1
## 1216905 nm0666604 tt0480249 1
## 1249618 nm0691189 tt0480249 7
## 1299786 nm0724757 tt0480249 1
## 1327834 nm0745751 tt0480249 2
## 1659101 nm1122170 tt0480249 1
## 1727478 nm1677477 tt0480249 1
## 1734602 nm1777059 tt0480249 1
## 1738812 nm1842950 tt0480249 2
## 1756999 nm2253071 tt0480249 1
## 1761788 nm2405238 tt0480249 1
## 1764041 nm2496005 tt0480249 2
## 1770648 nm2677021 tt0480249 1
## 1771697 nm2724906 tt0480249 1
## 1775333 nm2857229 tt0480249 1
## 1775558 nm2869725 tt0480249 1
## 1818908 nm8421984 tt0480249 1
```

Movie Details: Extra

```
will2 = merge(will.movies, imdb.data$movies.df$info, by="ttid");
will2;
```

##	ttid	nmid	rank	year	title
## 1	tt0105810	nm0000226	53	1992	Where the Day Takes You
## 2	tt0107478	nm0000226	48	1993	Made in America
## 3	tt0108149	nm0000226	44	1993	Six Degrees of Separation
## 4	tt0112442	nm0000226	12	1995	Bad Boys
## 5	tt0114558	nm0000226	31	1995	Strange Days
## 6	tt0116629	nm0000226	3	1996	Independence Day
## 7	tt0119654	nm0000226	4	1997	Men in Black
## 8	tt0120660	nm0000226	14	1998	Enemy of the State
## 9	tt0120707	nm0000226	60	1998	Ride
## 10	tt0120783	nm0000226	25	1998	The Parent Trap
## 11	tt0120891	nm0000226	24	1999	Wild Wild West
## 12	tt0120912	nm0000226	8	2002	Men in Black II
## 13	tt0145660	nm0000226	16	1999	Austin Powers: The Spy Who Shagged Me
## 14	tt0146984	nm0000226	34	2000	The Legend of Bagger Vance
## 15	tt0167427	nm0000226	46	1999	Superstar
## 16	tt0169376	nm0000226	59	1998	Welcome to Hollywood
## 17	tt0172156	nm0000226	13	2003	Bad Boys II
## 18	tt0205314	nm0000226	72	1998	Peset Ha'Zahav
## 19	tt0248667	nm0000226	27	2001	Ali
## 20	tt0268397	nm0000226	39	2001	Jimmy Neutron: Boy Genius
## 21	tt0284490	nm0000226	33	2002	Showtime
## 22	tt0300051	nm0000226	32	2004	Jersey Girl
## 23	tt0307453	nm0000226	22	2004	Shark Tale

## 24	tt0328099	nm0000226	45	2003	Malibu's Most Wanted
## 25	tt0338466	nm0000226	36	2003	Stuck on You
## 26	tt0343818	nm0000226	5	2004	I, Robot
## 27	tt0375663	nm0000226	67	2003	A Closer Walk
## 28	tt0379487	nm0000226	63	2004	The Seat Filler
## 29	tt0380277	nm0000226	54	2004	The Cookout
## 30	tt0384504	nm0000226	51	2004	Saving Face
## 31	tt0386588	nm0000226	10	2005	Hitch
## 32	tt0414587	nm0000226	71	2005	There's a God on the Mic
## 33	tt0416212	nm0000226	43	2008	The Secret Life of Bees
## 34	tt0448157	nm0000226	7	2008	Hancock
## 35	tt0454921	nm0000226	6	2006	The Pursuit of Happyness
## 36	tt0464265	nm0000226	97	NA	Time Share
## 37	tt0466839	nm0000226	47	2007	I Could Never Be Your Woman
## 38	tt0466856	nm0000226	50	2006	ATL
## 39	tt0480249	nm0000226	1	2007	I Am Legend
## 40	tt0812268	nm0000226	91	NA	The Last Pharaoh
## 41	tt0814314	nm0000226	11	2008	Seven Pounds
## 42	tt0911023	nm0000226	96	NA	My Wife Hates Your Wife
## 43	tt0942916	nm0000226	105	NA	Sisters of Mercy
## 44	tt0947802	nm0000226	37	2008	Lakeview Terrace
## 45	tt0979959	nm0000226	107	NA	Cooked
## 46	tt1015971	nm0000226	57	2009	The Boys: The Sherman Brothers' Story
## 47	tt1025100	nm0000226	30	2019	Gemini Man
## 48	tt1028582	nm0000226	106	NA	Newton's Law
## 49	tt10344550	nm0000226	92	NA	Untitled Solvan Naim Project
## 50	tt1082886	nm0000226	40	2008	The Wackness
## 51	tt10883124	nm0000226	61	2019	Dads
## 52	tt11012318	nm0000226	80	NA	The Council
## 53	tt1109477	nm0000226	55	2008	The Human Contract
## 54	tt1155076	nm0000226	23	2010	The Karate Kid
## 55	tt1181790	nm0000226	86	NA	The Billionaire's Vinegar
## 56	tt1206884	nm0000226	109	NA	The Long Run
## 57	tt1229340	nm0000226	20	2013	Anchorman 2: The Legend Continues
## 58	tt1230214	nm0000226	62	2009	Stuntmen
## 59	tt1235550	nm0000226	103	NA	Welcome to the Sticks
## 60	tt12530246	nm0000226	83	NA	Emancipation
## 61	tt12608018	nm0000226	88	NA	Bounce
## 62	tt12915962	nm0000226	81	NA	Planes, Trains & Automobiles
## 63	tt1368421	nm0000226	94	NA	Angelology
## 64	tt1383607	nm0000226	56	2008	Ramadan Mabrouk Abul-Alamein Hamouda
## 65	tt1386697	nm0000226	2	2016	Suicide Squad
## 66	tt1409024	nm0000226	9	2012	Men in Black 3
## 67	tt1409772	nm0000226	66	2010	A Man's Story
## 68	tt1470831	nm0000226	93	NA	The City That Sailed
## 69	tt1498780	nm0000226	82	NA	Hancock 2
## 70	tt1502397	nm0000226	26	2020	Bad Boys For Life
## 71	tt1546668	nm0000226	73	2007	The 100 Best Black Movies (Ever)
## 72	tt1578276	nm0000226	84	NA	Uptown Saturday Night
## 73	tt1584918	nm0000226	68	2014	Murder101
## 74	tt1596350	nm0000226	19	2012	This Means War
## 75	tt1609486	nm0000226	98	NA	Harold and the Purple Crayon
## 76	tt1674782	nm0000226	76	NA	The Karate Kid 2
## 77	tt1691439	nm0000226	100	NA	Untitled Dr. S. Allen Counter Project

## 78	tt1692497	nm0000226	90	NA	The Redemption of Cain		
## 79	tt1714201	nm0000226	104	NA	Monster Witness Relocation Program		
## 80	tt1815862	nm0000226	18	2013	After Earth		
## 81	tt1817789	nm0000226	99	NA	Wheeler Dealers		
## 82	tt1823664	nm0000226	38	2014	Annie		
## 83	tt1837709	nm0000226	35	2014	Winter's Tale		
## 84	tt1852816	nm0000226	111	NA	Joe		
## 85	tt1879098	nm0000226	101	NA	What Would Kenny Do?		
## 86	tt2074488	nm0000226	69	2005	Will Smith: Live in Concert		
## 87	tt2168954	nm0000226	110	NA	Ghost Graduation		
## 88	tt2246761	nm0000226	102	NA	Interpol		
## 89	tt2350432	nm0000226	64	2012	Free Angela and All Political Prisoners		
## 90	tt2381941	nm0000226	17	2015	Focus		
## 91	tt2776710	nm0000226	85	NA	Brilliance		
## 92	tt2846190	nm0000226	95	NA	La escribana de uraba		
## 93	tt3322364	nm0000226	29	2015	Concussion		
## 94	tt3721964	nm0000226	42	2018	Gringo		
## 95	tt4209900	nm0000226	58	2015	Fresh Dressed		
## 96	tt4463004	nm0000226	87	NA	Bounty		
## 97	tt4682786	nm0000226	28	2016	Collateral Beauty		
## 98	tt4919268	nm0000226	75	NA	Bad Boys 4		
## 99	tt5230508	nm0000226	70	2005	AMV Hell 3: The Motion Picture		
## 100	tt5519340	nm0000226	21	2017	Bright		
## 101	tt5814534	nm0000226	41	2019	Spies in Disguise		
## 102	tt5839150	nm0000226	74	2009	The Muhammad Ali Story		
## 103	tt5903926	nm0000226	65	2018	Sprinter		
## 104	tt6139732	nm0000226	15	2019	Aladdin		
## 105	tt6598238	nm0000226	78	2020	Life in a Year		
## 106	tt7255568	nm0000226	49	2019	Student of the Year 2		
## 107	tt7331400	nm0000226	89	NA	1921		
## 108	tt7440432	nm0000226	52	2018	Quincy		
## 109	tt7737796	nm0000226	108	NA	The Prospect		
## 110	tt7820302	nm0000226	79	NA	Bright 2		
## 111	tt9620288	nm0000226	77	2021	King Richard		
##		genre	rated	minutes	ratings	metacritic	votes
## 1		Crime, Drama, Thriller	R	105	6.6	NA	3112
## 2		Comedy	PG-13	111	5.1	NA	15447
## 3		Comedy, Drama, Mystery	R	112	6.8	72	19473
## 4		Action, Comedy, Crime	R	119	6.9	41	235766
## 5		Action, Crime, Drama	R	145	7.2	66	66102
## 6		Action, Adventure, Sci-Fi	PG-13	145	7.0	59	520592
## 7		Action, Adventure, Comedy	PG-13	98	7.3	71	507597
## 8		Action, Thriller	R	132	7.3	67	224019
## 9		Comedy	R	90	4.6	NA	683
## 10		Adventure, Comedy, Drama	PG	128	6.5	64	117780
## 11		Action, Comedy, Sci-Fi	PG-13	106	5.0	38	151132
## 12		Action, Adventure, Comedy	PG-13	88	6.2	49	337458
## 13		Action, Adventure, Comedy	PG-13	95	6.6	59	214064
## 14		Drama, Fantasy, Sport	PG-13	126	6.7	47	52831
## 15		Comedy, Romance	PG-13	81	5.2	42	17781
## 16		Comedy	R	89	4.9	NA	715
## 17		Action, Comedy, Crime	R	147	6.6	38	228329
## 18		Documentary	<NA>	55	5.2	NA	NA
## 19		Biography, Drama, Sport	R	157	6.8	65	92508

## 20	Animation, Action, Adventure	G	82	6.0	65	30195
## 21	Action, Comedy, Crime	PG-13	95	5.5	32	59984
## 22	Comedy, Drama, Romance	PG-13	102	6.2	43	62700
## 23	Animation, Adventure, Comedy	PG	90	6.0	48	164132
## 24	Comedy, Crime	PG-13	86	5.2	43	18356
## 25	Comedy	PG-13	118	5.7	62	50682
## 26	Action, Drama, Sci-Fi	PG-13	115	7.1	59	491474
## 27	Documentary	<NA>	NA	8.6	NA	NA
## 28	Comedy, Romance	PG-13	90	5.5	NA	607
## 29	Comedy	PG-13	97	3.7	15	2741
## 30	Comedy, Drama, Romance	R	91	7.4	65	10016
## 31	Comedy, Romance	PG-13	118	6.6	58	291991
## 32	Documentary	<NA>	NA	5.6	NA	NA
## 33	Drama	PG-13	114	7.3	57	23962
## 34	Action, Fantasy	PG-13	92	6.4	49	437887
## 35	Biography, Drama	PG-13	117	8.0	64	438121
## 36	Comedy	<NA>	NA	NA	NA	NA
## 37	Comedy, Drama, Romance	PG-13	97	6.0	NA	16855
## 38	Comedy, Crime, Drama	PG-13	105	6.1	63	10580
## 39	Action, Adventure, Drama	PG-13	101	7.2	65	675160
## 40	Drama	<NA>	NA	NA	NA	NA
## 41	Drama	PG-13	123	7.6	36	282568
## 42	Comedy	<NA>	NA	NA	NA	NA
## 43	Comedy, Drama, Romance	<NA>	NA	NA	NA	NA
## 44	Crime, Drama, Thriller	PG-13	110	6.1	47	48231
## 45	Drama	<NA>	NA	NA	NA	NA
## 46	Documentary	PG	101	7.8	78	853
## 47	Action, Drama, Sci-Fi	PG-13	117	5.7	38	76145
## 48	Adventure, Comedy, Family	<NA>	NA	NA	NA	NA
## 49	Musical	<NA>	NA	NA	NA	NA
## 50	Comedy, Drama, Romance	R	99	7.0	61	29701
## 51	Documentary, Comedy, Family	TV-14	87	6.3	59	668
## 52	Biography	<NA>	NA	NA	NA	NA
## 53	Drama	R	103	5.1	NA	1970
## 54	Action, Drama, Family	PG	140	6.2	61	158842
## 55	Drama	<NA>	NA	NA	NA	NA
## 56	Drama	<NA>	NA	NA	NA	NA
## 57	Comedy	PG-13	119	6.3	61	165909
## 58	Comedy	R	90	5.1	NA	627
## 59	Comedy	<NA>	NA	NA	NA	NA
## 60	Action, Thriller	<NA>	NA	NA	NA	NA
## 61	Drama	<NA>	NA	NA	NA	NA
## 62	Comedy	<NA>	NA	NA	NA	NA
## 63	Drama	<NA>	NA	NA	NA	NA
## 64	Comedy, Family	<NA>	105	6.4	NA	NA
## 65	Action, Adventure, Fantasy	PG-13	123	6.0	40	588064
## 66	Action, Adventure, Comedy	PG-13	106	6.8	58	329948
## 67	Documentary	Not Rated	98	6.2	49	126
## 68	Drama, Fantasy	<NA>	NA	NA	NA	NA
## 69	Action, Fantasy	<NA>	NA	NA	NA	NA
## 70	Action, Comedy, Crime	R	124	6.6	59	112826
## 71	Documentary	<NA>	NA	5.0	NA	NA
## 72	Comedy	<NA>	NA	NA	NA	NA
## 73	Thriller	PG-13	90	4.4	NA	41

## 74	Action, Comedy, Romance	PG-13	103	6.3	31	174205
## 75	Animation	<NA>	NA	NA	NA	NA
## 76	Action, Drama	<NA>	NA	NA	NA	NA
## 77	Drama	<NA>	NA	NA	NA	NA
## 78	Drama	<NA>	NA	NA	NA	NA
## 79	Sci-Fi	<NA>	NA	NA	NA	NA
## 80	Action, Adventure, Sci-Fi	PG-13	100	4.8	33	189337
## 81	Drama	<NA>	NA	NA	NA	NA
## 82	Comedy, Drama, Family	PG	118	5.3	33	33448
## 83	Drama, Fantasy, Mystery	PG-13	118	6.2	31	51469
## 84	Drama	<NA>	NA	NA	NA	NA
## 85	Comedy	<NA>	NA	NA	NA	NA
## 86	Documentary, Music	Not Rated	52	5.7	NA	34
## 87	Comedy, Fantasy	<NA>	NA	NA	NA	NA
## 88	Comedy	<NA>	NA	NA	NA	NA
## 89	Documentary	Not Rated	102	6.9	73	565
## 90	Comedy, Crime, Drama	R	105	6.6	56	212789
## 91	Sci-Fi, Thriller	<NA>	NA	NA	NA	NA
## 92	Drama	<NA>	NA	NA	NA	NA
## 93	Biography, Drama, Sport	PG-13	123	7.1	55	84750
## 94	Action, Comedy, Crime	R	111	6.1	46	25794
## 95	Documentary, History, Music	Not Rated	90	6.6	69	840
## 96	Action	<NA>	NA	NA	NA	NA
## 97	Drama, Romance	PG-13	97	6.8	23	89138
## 98	Action, Adventure, Comedy	<NA>	NA	NA	NA	NA
## 99	Comedy	<NA>	68	7.5	NA	NA
## 100	Action, Fantasy, Thriller	TV-MA	117	6.3	29	164913
## 101	Animation, Action, Adventure	PG	102	6.8	54	28764
## 102	Documentary	<NA>	52	5.7	NA	NA
## 103	Drama, Sport	<NA>	114	6.1	NA	NA
## 104	Adventure, Family, Fantasy	PG	128	7.0	53	216257
## 105	Drama, Romance	PG-13	NA	NA	NA	NA
## 106	Drama, Romance, Sport	Not Rated	146	2.3	NA	15200
## 107	Drama	<NA>	NA	NA	NA	NA
## 108	Documentary, Biography	TV-MA	124	7.6	60	3359
## 109	<NA>	<NA>	NA	NA	NA	NA
## 110	Action, Adventure, Crime	<NA>	NA	NA	NA	NA
## 111	Drama	<NA>	NA	NA	NA	NA
##	millions					
## 1	0.39					
## 2	44.94					
## 3	6.41					
## 4	65.81					
## 5	7.92					
## 6	306.17					
## 7	250.69					
## 8	111.55					
## 9	5.48					
## 10	66.31					
## 11	113.81					
## 12	190.42					
## 13	206.04					
## 14	30.70					
## 15	30.63					

## 16	NA
## 17	138.61
## 18	NA
## 19	58.20
## 20	80.94
## 21	38.08
## 22	25.27
## 23	160.86
## 24	34.31
## 25	33.83
## 26	144.80
## 27	NA
## 28	NA
## 29	11.54
## 30	1.19
## 31	179.50
## 32	NA
## 33	37.77
## 34	227.95
## 35	163.57
## 36	NA
## 37	NA
## 38	21.16
## 39	256.39
## 40	NA
## 41	69.95
## 42	NA
## 43	NA
## 44	39.26
## 45	NA
## 46	0.05
## 47	20.55
## 48	NA
## 49	NA
## 50	2.08
## 51	NA
## 52	NA
## 53	NA
## 54	176.59
## 55	NA
## 56	NA
## 57	127.35
## 58	NA
## 59	NA
## 60	NA
## 61	NA
## 62	NA
## 63	NA
## 64	NA
## 65	325.10
## 66	179.02
## 67	NA
## 68	NA
## 69	NA

## 70	204.42
## 71	NA
## 72	NA
## 73	NA
## 74	54.76
## 75	NA
## 76	NA
## 77	NA
## 78	NA
## 79	NA
## 80	60.52
## 81	NA
## 82	85.91
## 83	0.02
## 84	NA
## 85	NA
## 86	NA
## 87	NA
## 88	NA
## 89	NA
## 90	53.86
## 91	NA
## 92	NA
## 93	34.54
## 94	4.97
## 95	NA
## 96	NA
## 97	31.02
## 98	NA
## 99	NA
## 100	NA
## 101	NA
## 102	NA
## 103	NA
## 104	355.56
## 105	NA
## 106	0.78
## 107	NA
## 108	NA
## 109	NA
## 110	NA
## 111	NA
##	
## 1	
## 2	
## 3	
## 4	
## 5	
## 6	
## 7	
## 8	
## 9	
## 10	
## 11	

A group of tea

12
13
14
15
16
17
18
19
20
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

This document

Their music is unforg

It's the summer of 1994, and the streets of New York are pulsing.
Director [Bryce Dallas Howard](/name/nm0397171) teams up with her father, [Bryce Dallas Howard](/name/nm0397171)

When God crea

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

Will Smith: L

Fresh Dressed chron

An enormousl

Green

##	millions2000	rank.money	rank.minutes	rank.ratings	rank.metacritic
## 1	0.47867427	50	36	30	58
## 2	53.55479585	26	29	64	67
## 3	7.63876817	44	28	21	3
## 4	74.36011811	21	17	18	42
## 5	8.94897638	43	5	11	7
## 6	336.02596558	1	4	15	21
## 7	268.96459813	2	47	10	4

## 8	117.84607362	17	7	8	6
## 9	5.78930061	45	57	71	72
## 10	70.05265031	22	9	33	13
## 11	117.63554622	18	33	67	46
## 12	182.26972763	8	63	44	34
## 13	212.96571429	5	52	31	23
## 14	30.70000000	35	10	25	36
## 15	31.65957983	33	67	62	41
## 16	NA	60	62	69	71
## 17	129.72088043	16	2	26	44
## 18	NA	72	69	63	66
## 19	56.58972332	24	1	20	10
## 20	78.70055336	20	66	52	11
## 21	36.45011673	30	53	58	50
## 22	23.03596612	38	40	42	39
## 23	146.63892006	10	56	51	35
## 24	32.10968478	31	65	61	40
## 25	31.66046739	32	20	53	15
## 26	131.99872949	15	25	14	20
## 27	NA	67	72	1	56
## 28	NA	63	60	59	65
## 29	10.51978825	42	50	73	55
## 30	1.08479619	48	55	7	8
## 31	158.26881720	9	18	28	26
## 32	NA	71	73	57	64
## 33	30.20856189	36	26	9	27
## 34	182.31510940	7	54	35	32
## 35	139.71604167	11	22	2	12
## 36	NA	97	97	97	97
## 37	NA	55	51	50	61
## 38	18.07416667	40	35	48	14
## 39	212.93494806	6	43	12	9
## 40	NA	91	91	91	91
## 41	55.94622462	25	14	5	47
## 42	NA	96	96	96	96
## 43	NA	105	105	105	105
## 44	31.40026846	34	31	47	37
## 45	NA	107	107	107	107
## 46	0.04013294	51	44	3	1
## 47	13.84160449	41	23	54	45
## 48	NA	106	106	106	106
## 49	NA	92	92	92	92
## 50	1.66359038	47	46	17	16
## 51	NA	61	64	39	24
## 52	NA	80	80	80	80
## 53	NA	57	39	65	68
## 54	139.45407602	12	6	40	18
## 55	NA	86	86	86	86
## 56	NA	109	109	109	109
## 57	94.13612813	19	16	36	17
## 58	NA	62	59	66	69
## 59	NA	103	103	103	103
## 60	NA	83	83	83	83
## 61	NA	88	88	88	88

## 62	NA	81	81	81	81
## 63	NA	94	94	94	94
## 64	NA	58	37	34	59
## 65	233.25244680	4	13	49	43
## 66	134.26850876	14	32	22	25
## 67	NA	66	48	43	33
## 68	NA	93	93	93	93
## 69	NA	82	82	82	82
## 70	135.85869603	13	11	27	22
## 71	NA	73	74	68	70
## 72	NA	84	84	84	84
## 73	NA	68	61	72	73
## 74	41.07107329	28	38	38	51
## 75	NA	98	98	98	98
## 76	NA	76	76	76	76
## 77	NA	100	100	100	100
## 78	NA	90	90	90	90
## 79	NA	104	104	104	104
## 80	44.73591264	27	45	70	49
## 81	NA	99	99	99	99
## 82	62.49029298	23	19	60	48
## 83	0.01454785	52	21	41	52
## 84	NA	111	111	111	111
## 85	NA	101	101	101	101
## 86	NA	69	70	55	62
## 87	NA	110	110	110	110
## 88	NA	102	102	102	102
## 89	NA	64	42	19	2
## 90	39.13091466	29	34	29	28
## 91	NA	85	85	85	85
## 92	NA	95	95	95	95
## 93	25.09435188	37	15	13	29
## 94	3.40657565	46	30	46	38
## 95	NA	59	58	32	5
## 96	NA	87	87	87	87
## 97	22.25620086	39	49	24	54
## 98	NA	75	75	75	75
## 99	NA	70	68	6	57
## 100	NA	53	24	37	53
## 101	NA	54	41	23	30
## 102	NA	74	71	56	63
## 103	NA	65	27	45	60
## 104	239.49006777	3	8	16	31
## 105	NA	78	78	78	78
## 106	0.52537477	49	3	74	74
## 107	NA	89	89	89	89
## 108	NA	56	12	4	19
## 109	NA	108	108	108	108
## 110	NA	79	79	79	79
## 111	NA	77	77	77	77
##	rank.votes	rank.year		country	
## 1	53	1		USA	
## 2	48	3		USA , France	
## 3	44	2		USA	

## 4	12	4	USA
## 5	31	5	USA
## 6	3	6	USA
## 7	4	7	USA
## 8	14	8	USA
## 9	59	11	USA
## 10	25	9	USA , UK
## 11	24	14	USA
## 12	8	19	USA
## 13	16	13	USA
## 14	34	16	USA
## 15	46	15	USA
## 16	58	10	USA
## 17	13	21	USA
## 18	73	12	Israel
## 19	27	17	USA
## 20	39	18	USA
## 21	33	20	USA , Australia
## 22	32	27	USA
## 23	22	26	USA
## 24	45	23	USA
## 25	36	22	USA
## 26	5	25	USA , Germany
## 27	67	24	USA
## 28	62	30	USA
## 29	54	29	USA
## 30	51	28	USA
## 31	10	31	USA
## 32	72	34	USA
## 33	43	44	USA
## 34	7	40	USA
## 35	6	35	USA
## 36	97	97	<NA>
## 37	47	38	USA
## 38	50	36	USA
## 39	1	37	USA
## 40	91	91	<NA>
## 41	11	41	USA
## 42	96	96	<NA>
## 43	105	105	<NA>
## 44	37	42	USA
## 45	107	107	<NA>
## 46	56	47	USA
## 47	30	70	China , USA , India
## 48	106	106	<NA>
## 49	92	92	USA
## 50	40	43	UK , USA
## 51	60	73	USA
## 52	80	80	USA
## 53	55	45	USA
## 54	23	50	USA , China
## 55	86	86	<NA>
## 56	109	109	<NA>
## 57	20	56	USA

## 58	61	48	USA
## 59	103	103	<NA>
## 60	83	83	<NA>
## 61	88	88	<NA>
## 62	81	81	<NA>
## 63	94	94	<NA>
## 64	69	46	Egypt
## 65	2	63	USA
## 66	9	52	USA , United Arab Emirates
## 67	64	51	UK
## 68	93	93	<NA>
## 69	82	82	USA
## 70	26	74	USA , Mexico
## 71	74	39	USA
## 72	84	84	USA
## 73	65	59	USA
## 74	19	53	USA
## 75	98	98	<NA>
## 76	76	78	USA
## 77	100	100	<NA>
## 78	90	90	<NA>
## 79	104	104	<NA>
## 80	18	55	USA
## 81	99	99	<NA>
## 82	38	58	USA
## 83	35	57	USA
## 84	111	111	<NA>
## 85	101	101	<NA>
## 86	66	32	USA
## 87	110	110	<NA>
## 88	102	102	<NA>
## 89	63	54	USA , France
## 90	17	60	USA , Argentina
## 91	85	85	<NA>
## 92	95	95	<NA>
## 93	29	61	UK , Australia , USA
## 94	42	66	Australia , USA , Mexico
## 95	57	62	France , USA
## 96	87	87	<NA>
## 97	28	64	USA
## 98	75	77	USA
## 99	68	33	USA
## 100	21	65	USA
## 101	41	71	USA
## 102	71	49	USA
## 103	70	68	Jamaica , USA
## 104	15	69	UK , USA , Australia
## 105	78	75	USA
## 106	49	72	India
## 107	89	89	<NA>
## 108	52	67	USA
## 109	108	108	<NA>
## 110	79	79	USA
## 111	77	76	<NA>

##		language	release.date
## 1		English	11 September 1992
## 2		English , Japanese	28 May 1993
## 3		English , Italian	8 December 1993
## 4		English	7 April 1995
## 5		English	13 October 1995
## 6		English	3 July 1996
## 7		English , Spanish	2 July 1997
## 8		English	20 November 1998
## 9		English	27 March 1998
## 10		English , French	29 July 1998
## 11		English	30 June 1999
## 12		English	3 July 2002
## 13		English , German	11 June 1999
## 14		English	3 November 2000
## 15		English	8 October 1999
## 16		English	October 1998
## 17		English , Spanish , Russian	18 July 2003
## 18		English	1998
## 19		English , French , Swahili	25 December 2001
## 20		English	21 December 2001
## 21		English	15 March 2002
## 22		English	26 March 2004
## 23		English	1 October 2004
## 24		English	18 April 2003
## 25		English	12 December 2003
## 26		English	16 July 2004
## 27		English	5 February 2003
## 28		English	20 July 2005
## 29		English	3 September 2004
## 30		English , Mandarin , Shanghainese	24 June 2005
## 31		English	11 February 2005
## 32		English	15 January 2005
## 33		English	17 October 2008
## 34		English , Japanese , Vietnamese	2 July 2008
## 35		English , Cantonese	15 December 2006
## 36		<NA>	<NA>
## 37		English	11 May 2007
## 38		English	31 March 2006
## 39		English	14 December 2007
## 40		<NA>	<NA>
## 41		English , Spanish	19 December 2008
## 42		<NA>	<NA>
## 43		<NA>	<NA>
## 44		English	19 September 2008
## 45		<NA>	<NA>
## 46		English	24 April 2009
## 47		English , French	11 October 2019
## 48		<NA>	<NA>
## 49		English	<NA>
## 50		English	1 August 2008
## 51		English , Japanese , Portuguese	19 June 2020
## 52		<NA>	<NA>
## 53		English , French	18 March 2010

## 54	English , Mandarin	11 June 2010
## 55	<NA>	<NA>
## 56	<NA>	<NA>
## 57	English	18 December 2013
## 58	English	18 August 2009
## 59	<NA>	<NA>
## 60	<NA>	<NA>
## 61	<NA>	<NA>
## 62	<NA>	<NA>
## 63	<NA>	<NA>
## 64	Arabic	4 December 2008
## 65	English , Japanese , Spanish	5 August 2016
## 66	English , Ukrainian	25 May 2012
## 67	English	9 March 2012
## 68	<NA>	<NA>
## 69	English	<NA>
## 70	English , Spanish	17 January 2020
## 71	English	5 February 2007
## 72	<NA>	<NA>
## 73	English	10 June 2014
## 74	English , German , Japanese , Serbian	17 February 2012
## 75	<NA>	<NA>
## 76	English	<NA>
## 77	<NA>	<NA>
## 78	<NA>	<NA>
## 79	<NA>	<NA>
## 80	English	31 May 2013
## 81	<NA>	<NA>
## 82	English	19 December 2014
## 83	English	14 February 2014
## 84	<NA>	<NA>
## 85	<NA>	<NA>
## 86	English	14 June 2005
## 87	<NA>	<NA>
## 88	<NA>	<NA>
## 89	English	5 April 2013
## 90	English , Spanish	27 February 2015
## 91	<NA>	<NA>
## 92	<NA>	<NA>
## 93	English	25 December 2015
## 94	English , Spanish , French , Yoruba	9 March 2018
## 95	English	26 June 2015
## 96	<NA>	<NA>
## 97	English	16 December 2016
## 98	English	<NA>
## 99	English	23 September 2005
## 100	English , Spanish	22 December 2017
## 101	English , Japanese	25 December 2019
## 102	English	2009
## 103	English	6 September 2019
## 104	English , Arabic	24 May 2019
## 105	English	2020
## 106	Hindi	10 May 2019
## 107	<NA>	<NA>

## 108	English 21 September 2018
## 109	<NA> <NA>
## 110	English <NA>
## 111	<NA> 19 November 2021
##	aka
## 1	Where the Day Takes You
## 2	Genesis
## 3	Six Degrees of Separation
## 4	Bulletproof Hearts
## 5	Strange Days 1999
## 6	ID4
## 7	MIB
## 8	Enemy of the State
## 9	I-95
## 10	Disney's The Parent Trap
## 11	Wild Wild West
## 12	MIIB
## 13	Austin Powers 2: The Spy Who Shagged Me
## 14	The Legend of Baggar Vance
## 15	Superstar
## 16	Welcome to Hollywood
## 17	Bad Boys 2
## 18	Inside the Golden Statue
## 19	Ali
## 20	Jimmy Neutron: Boy Genius
## 21	Showtime
## 22	Untitled Kevin Smith Project
## 23	Sharkslayer
## 24	Suckaz
## 25	Stuck on You
## 26	Hardwired
## 27	<NA>
## 28	The Seat Filler
## 29	The Cookout
## 30	Saving Face
## 31	The Last First Kiss
## 32	<NA>
## 33	The Secret Life of Bees
## 34	Tonight, He Comes
## 35	Untitled Chris Gardner Project
## 36	<NA>
## 37	I Could Be Your Mother
## 38	Jellybeans
## 39	I Am Legend: The IMAX Experience
## 40	<NA>
## 41	Seven Pounds
## 42	<NA>
## 43	<NA>
## 44	Lakeview Terrace
## 45	<NA>
## 46	The Boys: The Sherman Brothers' Story
## 47	Gemini Man
## 48	<NA>
## 49	<NA>

## 50	The Wackness
## 51	Dads
## 52	<NA>
## 53	The Human Contract
## 54	Untitled Karate Kid Remake
## 55	<NA>
## 56	<NA>
## 57	Anchorman: The Legend Continues
## 58	<NA>
## 59	<NA>
## 60	<NA>
## 61	<NA>
## 62	<NA>
## 63	<NA>
## 64	<NA>
## 65	Task Force X
## 66	Men in Black 3D
## 67	Meska historia
## 68	<NA>
## 69	<NA>
## 70	Bad Boys 3
## 71	<NA>
## 72	Uptown Saturday Night
## 73	<NA>
## 74	This Means War
## 75	<NA>
## 76	Karate Kid 2
## 77	<NA>
## 78	<NA>
## 79	<NA>
## 80	1000 A.E.
## 81	<NA>
## 82	Annie
## 83	Winter's Tale
## 84	<NA>
## 85	<NA>
## 86	<NA>
## 87	<NA>
## 88	<NA>
## 89	Free Angela and All Political Prisoners
## 90	Focus
## 91	<NA>
## 92	<NA>
## 93	Game Brain
## 94	Gringo
## 95	Fresh Dressed
## 96	<NA>
## 97	Collateral Beauty
## 98	Bad Boys 4
## 99	<NA>
## 100	Bright
## 101	Spies in Disguise
## 102	Muhammad Ali: Fighting Spirit
## 103	Sprinter

## 104	Aladdin	
## 105	Year in a Life	
## 106	Student of the Year 2	
## 107	<NA>	
## 108	<NA>	
## 109	<NA>	
## 110	Bright 2	
## 111	King Richard	
##		
## 1		
## 2		
## 3		
## 4		
## 5		
## 6		
## 7		
## 8		
## 9		
## 10		
## 11		
## 12		Brooklyn Battery Tunnel Ventilator
## 13		
## 14		
## 15		
## 16		
## 17		
## 18		
## 19		
## 20		DNA Pro
## 21		
## 22		
## 23		
## 24		
## 25		Barney's Beanitos
## 26	British Columbia Institute of Technology - Aerospace Technology Campus, Vancouver International Airport	
## 27		
## 28		
## 29		
## 30		TWA Hotel, John F. Kennedy International Airport
## 31		Washington Square Park
## 32		
## 33		
## 34		
## 35		
## 36		
## 37		Ava
## 38		
## 39		11 Washington Square North
## 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
##      budget usa.opening usa.gross world.gross
## 1      3000000      190961      390152      390152
## 2     22000000     11821326     44942695     104942695
## 3     12000000      53058      6405918      6405918
## 4     19000000     15523358     65807024     141407024
## 5     42000000      31062      7959291      7959291
## 6     75000000     50228264     306169268     817400891
## 7     90000000     51068455     250690539     589390539
## 8     90000000     20038573     111549836     250849789
## 9          <NA>      2513334      5485295      5485295
## 10         <NA>     11148497     66308518      92108518
## 11    170000000     27687484     113804681     222104681
## 12    140000000     52148751     193735288     445135288
## 13     33000000     17981900     206040086     312016928
## 14     80000000     11516712      30919168      39459427
## 15     34000000      8912743      30636478      30636478
## 16         <NA>         <NA>         <NA>         <NA>
## 17    130000000     46522560     138608444     273339556
## 18         <NA>         <NA>         <NA>         <NA>
## 19    107000000     14710892     58203105      87812729
## 20     30000000     13832786     80936232     102992536
## 21     85000000     15011430     38082712      77885672
## 22     35000000      8300000     25268157      35492921
## 23     75000000     47604606     160861908     374583879
## 24     15000000     12618159      34432201      34622504
## 25     55000000      9411055      33832741      65784503
## 26    120000000     52179887     144801023     353133898
## 27         <NA>         <NA>         <NA>         <NA>
## 28     80000000         <NA>         <NA>         <NA>
## 29         <NA>      5600000     11814019     12009070
## 30         <NA>      75104      1187266      1236518
## 31     70000000     43142214     179495555     371594210
## 32     1000000         <NA>         <NA>         <NA>
## 33     11000000     10527799      37770162      39952437
## 34    150000000     62603879     227946274     629443428
## 35     55000000     26541709     163566459     307127625
## 36         <NA>         <NA>         <NA>         <NA>
## 37     24000000         <NA>         <NA>      9576495
## 38     20000000     11554404     21170563      21170563
## 39    150000000     77211321     256393010     585410052
## 40         <NA>         <NA>         <NA>         <NA>
## 41     55000000     14851136     69951824     169748929

```

## 42	<NA>	<NA>	<NA>	<NA>
## 43	<NA>	<NA>	<NA>	<NA>
## 44	20000000	15004672	39263506	44655002
## 45	<NA>	<NA>	<NA>	<NA>
## 46	<NA>	14632	55513	55513
## 47	138000000	20552372	48546770	173469516
## 48	<NA>	<NA>	<NA>	<NA>
## 49	<NA>	<NA>	<NA>	<NA>
## 50	6000000	138401	2077116	3175469
## 51	<NA>	<NA>	<NA>	<NA>
## 52	<NA>	<NA>	<NA>	<NA>
## 53	<NA>	<NA>	<NA>	154527
## 54	40000000	55665805	176591618	359126022
## 55	<NA>	<NA>	<NA>	<NA>
## 56	<NA>	<NA>	<NA>	<NA>
## 57	50000000	26232425	127352707	173649931
## 58	<NA>	<NA>	<NA>	<NA>
## 59	<NA>	<NA>	<NA>	<NA>
## 60	<NA>	<NA>	<NA>	<NA>
## 61	<NA>	<NA>	<NA>	<NA>
## 62	<NA>	<NA>	<NA>	<NA>
## 63	<NA>	<NA>	<NA>	<NA>
## 64	<NA>	<NA>	<NA>	1012809
## 65	175000000	133682248	325100054	746846894
## 66	225000000	54592779	179020854	624026776
## 67	<NA>	<NA>	<NA>	14612
## 68	<NA>	<NA>	<NA>	<NA>
## 69	<NA>	<NA>	<NA>	<NA>
## 70	90000000	62504105	204417855	424617855
## 71	10000	<NA>	<NA>	<NA>
## 72	<NA>	<NA>	<NA>	<NA>
## 73	1500000	<NA>	<NA>	<NA>
## 74	65000000	17405930	54760791	156491279
## 75	<NA>	<NA>	<NA>	<NA>
## 76	<NA>	<NA>	<NA>	<NA>
## 77	<NA>	<NA>	<NA>	<NA>
## 78	<NA>	<NA>	<NA>	<NA>
## 79	<NA>	<NA>	<NA>	<NA>
## 80	130000000	27520040	60522097	243611982
## 81	<NA>	<NA>	<NA>	<NA>
## 82	65000000	15861939	85911262	136853506
## 83	60000000	7297694	12600231	30800231
## 84	<NA>	<NA>	<NA>	<NA>
## 85	<NA>	<NA>	<NA>	<NA>
## 86	<NA>	<NA>	<NA>	<NA>
## 87	<NA>	<NA>	<NA>	<NA>
## 88	<NA>	<NA>	<NA>	<NA>
## 89	<NA>	69306	194102	195005
## 90	50100000	18685137	53862963	158762963
## 91	<NA>	<NA>	<NA>	<NA>
## 92	<NA>	<NA>	<NA>	<NA>
## 93	35000000	10513749	34542474	48623572
## 94	<NA>	2722420	4969853	11001627
## 95	<NA>	<NA>	<NA>	<NA>

## 96	<NA>	<NA>	<NA>	<NA>
## 97	36000000	7102085	31016021	88528280
## 98	<NA>	<NA>	<NA>	<NA>
## 99	<NA>	<NA>	<NA>	<NA>
## 100	90000000	<NA>	<NA>	<NA>
## 101	100000000	13354798	66757013	171616764
## 102	<NA>	<NA>	<NA>	<NA>
## 103	<NA>	<NA>	<NA>	11622
## 104	183000000	91500929	355559216	1050693953
## 105	<NA>	<NA>	<NA>	<NA>
## 106	INR800000000	462108	778566	1635907
## 107	<NA>	<NA>	<NA>	<NA>
## 108	<NA>	<NA>	<NA>	<NA>
## 109	<NA>	<NA>	<NA>	<NA>
## 110	<NA>	<NA>	<NA>	<NA>
## 111	100000000	<NA>	<NA>	<NA>

##				product.
## 1				Cinetel Films, John Cornett & Doc Pro
## 2				Canal+, New Regency Productions, Alce
## 3				Metro-Goldwyn-Mayer (MGM), Maiden Movies, New Regency Pro
## 4				Don Simpson/Jerry Bruckheim
## 5				Lightstorm Enter
## 6				Twentieth Century Fox, Centropolis Enter
## 7				Columbia Pictures, Amblin Entertainment, Parkes/MacDonald Imag
## 8				Touchstone Pictures, Jerry Bruckheimer Films, Don Simpson/Jerry Bruckheim
## 9				Dimension Films, The Hudlin I
## 10				Cinema Vehicle Services, The Meyers/Shyer Company, Walt Disney I
## 11	Peters Entertainment, Sonnenfeld Josephson Worldwide Entertainment, Todman, Simon, LeMasters Pro			
## 12				Columbia Pictures, Amblin Entertainment, Parkes/MacDonald Imag
## 13				New Line Cinema, Gratitude, Moving Pict
## 14				Twentieth Century Fox, DreamWorks, Epsilon Motion I
## 15				Broadway Pictures, SNL
## 16	Crystal Spring Productions, Stone Canyon Entertainment, Filmsmith Production & Mar			
## 17				Columbia Pictures, Don Simpson/Jerry Bruckheimer Films, Lightstorm Enter
## 18				Tamuz Pro
## 19				Columbia Pictures, Forward Pass, Initial Entertainment Gro
## 20				Paramount Pictures, Nickelodeon Movies, DNA Pro
## 21				Material, NPV Entertainment, Overbrook Enter
## 22				Miramax, View Askew Productions, Beverly
## 23				DreamWorks Animation, Dre
## 24				Warner Bros., Karz Entertainment, Big Ticket Pro
## 25				Twentieth Century Fox, Conundrum Enter
## 26	Twentieth Century Fox, Mediastream Vierte Film GmbH & Co. Vermarktungs KG, Davis Enter			
## 27				
## 28				Seat Filler Productions, Shake Martin Films, Strange Fru
## 29				Cookout Productions, Flavor Unit Enter
## 30				Destination Films, Forensic Films, GreeneStre
## 31				Columbia Pictures, Overbrook Enter
## 32				Creanspeak Productions, Velvet Steamroller Enter
## 33				Fox Searchlight Pictures, Overbrook Entertainment, Donners'
## 34				Columbia Pictures, Relativity Media, Bl
## 35				Columbia Pictures, Relativity Media, Overbrook Enter
## 36				
## 37				Bauer Martinez Studios, Templar Films Unlimited, Formu

38 Sweet Tea Pictures LLC, Overbrook Entertainment, Syndrom
 ## 39 Warner Bros., Village Roadshow Pictures, Weed Road
 ## 40
 ## 41 Columbia Pictures, Relativity Media, Overbrook Enter
 ## 42
 ## 43
 ## 44 Screen Gems, Overbrook Enter
 ## 45
 ## 46 Walt Disney Pictures, Crescendo Productions, Traveling Light
 ## 47 Skydance Media, Jerry Bruckheimer Films, Fosun Group Forever
 ## 48
 ## 49 25/7 Productions, Flavor Unit Entertainment, Overbrook Enter
 ## 50 Occupant Entertainment, SBK
 ## 51 Imagine Documentaries, Nine Muses Entertainment, Dove
 ## 52 Anonymous Nobodies, Jackson Pictures, Overbrook Enter
 ## 53 Overbrook Entertainment, Tycoon Entertainment, 100
 ## 54 Columbia Pictures, Overbrook Entertainment, Sony Pictures Re
 ## 55
 ## 56
 ## 57 Paramount Pictures, Apatow Productions, Digital Image Ass
 ## 58 Final Cut Co
 ## 59
 ## 60
 ## 61
 ## 62
 ## 63
 ## 64
 ## 65 Atlas Entertainment, DC Comics, DC Enter
 ## 66 Columbia Pictures, Hemisphere Media Capital, Amblin Enter
 ## 67 Wellington Films, Almega
 ## 68
 ## 69 Columbia
 ## 70 Columbia Pictures, 2.0 Entertainment, Don Simpson/Jerry Bruckheim
 ## 71 Nikki Love Pro
 ## 72 Overbrook Entertainment, Warn
 ## 73 Murder 101, Upper Lav, Upper Laventile Pro
 ## 74 Overbrook Entertainment, Robert Simonds Productions, Dune Enter
 ## 75
 ## 76 Columbia Pictures, Overbrook Enter
 ## 77
 ## 78
 ## 79
 ## 80 Columbia Pictures, Overbrook Entertainment, Blinding Edge
 ## 81
 ## 82 Marcy Media, Olive Bridge Entertainment, Overbrook Enter
 ## 83 Village Roadshow Pictures, Warner Bros., Weed Road
 ## 84
 ## 85
 ## 86 Sony Pictures Home Enter
 ## 87
 ## 88
 ## 89 Realside Productions, De Films En Aiguille, Direc
 ## 90 Kramer & Sigman Films, RatPac-Dune Entertainment, Zaft
 ## 91

## 92	
## 93	LStar Capital, Scott Free Productions, The Cantillon
## 94	Amazon Studios, Blue-Tongue Films, Denver and Delilah Pro
## 95	Mass
## 96	
## 97	New Line Cinema, Village Roadshow Pictures, Anonymous
## 98	Columbia Pictures, Jerry Bruckheimer Films, Sony Pictures Entertainme
## 99	
## 100	Clubhouse Pictures (II), Overbrook Entertainment,
## 101	Blue Sky Studios, Chernin Entertainment, Twentieth Century Fox An
## 102	Jans
## 103	Mental Telepathy Pictures, MBG Entertainment, Flipped Out Enter
## 104	Walt Disney Pictures, Lin Pictures, I
## 105	Overbrook Enter
## 106	Fox STAR Studios, Dharma Productions, SIF 309 Fi
## 107	
## 108	Le Train Train, Tribeca Pro
## 109	
## 110	Clubhouse Pictures (II), Netflix, Screen
## 111	Overbrook Entertainment, Star Thrower Entertainment, Warn
##	runtime
## 1	105 min
## 2	111 min
## 3	112 min
## 4	119 min
## 5	145 min
## 6	145 min , 154 min (extended cut)
## 7	98 min
## 8	132 min , 140 min (extended)
## 9	90 min , 81 min (DVD)
## 10	128 min
## 11	106 min
## 12	88 min
## 13	95 min
## 14	126 min
## 15	81 min
## 16	89 min
## 17	147 min
## 18	55 min
## 19	157 min , 165 min (director's cut)
## 20	82 min
## 21	95 min
## 22	102 min
## 23	90 min
## 24	<NA>
## 25	118 min
## 26	115 min
## 27	<NA>
## 28	90 min
## 29	<NA>
## 30	91 min , 96 min
## 31	118 min
## 32	<NA>
## 33	114 min (director's cut) , 110 min

## 34	92 min , 102 min (unrated)
## 35	117 min
## 36	<NA>
## 37	97 min
## 38	105 min
## 39	101 min , 104 min (Alternate ending)
## 40	<NA>
## 41	123 min
## 42	<NA>
## 43	<NA>
## 44	110 min
## 45	<NA>
## 46	<NA>
## 47	117 min
## 48	<NA>
## 49	<NA>
## 50	99 min
## 51	81 min
## 52	<NA>
## 53	103 min
## 54	140 min
## 55	<NA>
## 56	<NA>
## 57	119 min , 143 min (supersize edition)
## 58	96 min
## 59	<NA>
## 60	<NA>
## 61	<NA>
## 62	<NA>
## 63	<NA>
## 64	<NA>
## 65	123 min , 134 min (Extended Cut)
## 66	106 min
## 67	98 min
## 68	<NA>
## 69	<NA>
## 70	124 min
## 71	<NA>
## 72	<NA>
## 73	<NA>
## 74	97 min (edited) , 103 min (director's cut)
## 75	<NA>
## 76	<NA>
## 77	<NA>
## 78	<NA>
## 79	<NA>
## 80	100 min
## 81	<NA>
## 82	118 min
## 83	118 min
## 84	<NA>
## 85	<NA>
## 86	52 min
## 87	<NA>

## 88	<NA>
## 89	102 min
## 90	105 min
## 91	<NA>
## 92	<NA>
## 93	123 min
## 94	111 min
## 95	90 min
## 96	<NA>
## 97	97 min
## 98	<NA>
## 99	68 min
## 100	117 min
## 101	102 min
## 102	52 min
## 103	114 min
## 104	128 min
## 105	<NA>
## 106	146 min
## 107	<NA>
## 108	124 min
## 109	<NA>
## 110	<NA>
## 111	<NA>
##	
## 1	
## 2	
## 3	
## 4	
## 5	
## 6	
## 7	
## 8	
## 9	
## 10	
## 11	
## 12	
## 13	
## 14	
## 15	
## 16	
## 17	
## 18	
## 19	
## 20	
## 21	
## 22	
## 23	
## 24	
## 25	
## 26	
## 27	
## 28	
## 29	

DTS 70 mm (70 mm print), DTS

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

Dolby Digit

```

## 84
## 85
## 86
## 87
## 88
## 89
## 90
## 91
## 92
## 93
## 94
## 95
## 96
## 97
## 98
## 99
## 100
## 101 Dolby Atmos (Dolby Atmos+Vision), DTS (DTS: X), IMAX 6-Track (IMAX version), 12-Track Digital So
## 102
## 103
## 104 IMAX 6-Track (IMAX version), DTS (DTS: X), Dolby Dig
## 105
## 106
## 107
## 108
## 109
## 110
## 111
##
##          color    aspect release.location
## 1          Color 1.85 : 1          USA
## 2          Color 1.85 : 1          USA
## 3          Color 2.35 : 1          USA
## 4          Color 1.85 : 1          USA
## 5          Color (CFI) 2.39 : 1      USA
## 6          Color 2.39 : 1          USA
## 7          Color (Technicolor) 1.85 : 1 USA
## 8          Color 2.39 : 1          USA
## 9          Color 1.85 : 1          USA
## 10         Color (Technicolor) 1.85 : 1 USA
## 11         Color (Technicolor) 1.85 : 1 USA
## 12         Color 1.37 : 1          USA
## 13         Color 2.39 : 1          USA
## 14         Color (Technicolor) 1.85 : 1 USA
## 15         Color 1.85 : 1          USA
## 16         Color    <NA>          USA
## 17         Color 2.39 : 1          USA
## 18         Color    <NA>          Israel
## 19 Black and White (archive footage), Color 2.39 : 1 USA
## 20         Color 1.66 : 1          USA
## 21         Color 2.35 : 1          USA
## 22         Color 2.39 : 1          USA
## 23         Color 1.85 : 1          USA
## 24         Color 2.35 : 1          USA
## 25         Color 2.39 : 1          USA

```

## 26	Color (Alphacine) 2.35 : 1	USA
## 27	Color <NA>	USA
## 28	Color 1.85 : 1	USA
## 29	Color 1.85 : 1	USA
## 30	Color 1.85 : 1	USA
## 31	Color (Technicolor) 2.39 : 1	USA
## 32	Color <NA>	USA
## 33	Color 2.35 : 1	USA
## 34	Color 2.39 : 1	USA
## 35	Color 2.35 : 1	USA
## 36	<NA> <NA>	<NA>
## 37	Color 1.85 : 1	Spain
## 38	Color 2.35 : 1	USA
## 39	Color (Technicolor) 2.35 : 1	USA
## 40	<NA> <NA>	<NA>
## 41	Color 2.35 : 1	USA
## 42	<NA> <NA>	<NA>
## 43	<NA> <NA>	<NA>
## 44	Color 2.35 : 1	USA
## 45	<NA> <NA>	<NA>
## 46	Color 1.78 : 1	USA
## 47	Color 1.85 : 1	USA
## 48	<NA> <NA>	<NA>
## 49	<NA> <NA>	<NA>
## 50	Color 2.35 : 1	USA
## 51	Color <NA>	USA
## 52	<NA> <NA>	<NA>
## 53	Color 1.85 : 1	Israel
## 54	Color 2.39 : 1	USA
## 55	<NA> <NA>	<NA>
## 56	<NA> <NA>	<NA>
## 57	Color 2.35 : 1	USA
## 58	Color <NA>	USA
## 59	<NA> <NA>	<NA>
## 60	<NA> <NA>	<NA>
## 61	<NA> <NA>	<NA>
## 62	<NA> <NA>	<NA>
## 63	<NA> <NA>	<NA>
## 64	Color <NA>	Egypt
## 65	Color 2.39 : 1	USA
## 66	Color 1.85 : 1	USA
## 67	Color <NA>	UK
## 68	<NA> <NA>	<NA>
## 69	<NA> <NA>	<NA>
## 70	Color 2.39 : 1	USA
## 71	Color <NA>	USA
## 72	<NA> <NA>	<NA>
## 73	Color <NA>	USA
## 74	Color 2.35 : 1	USA
## 75	<NA> <NA>	<NA>
## 76	Color <NA>	<NA>
## 77	<NA> <NA>	<NA>
## 78	<NA> <NA>	<NA>
## 79	<NA> <NA>	<NA>

## 80		Color (ACES) 2.35 : 1	USA
## 81		<NA> <NA>	<NA>
## 82		Color 2.39 : 1	USA
## 83		Color 2.35 : 1	USA
## 84		<NA> <NA>	<NA>
## 85		<NA> <NA>	<NA>
## 86		Color 1.85 : 1	USA
## 87		<NA> <NA>	<NA>
## 88		<NA> <NA>	<NA>
## 89		Color <NA>	USA
## 90		Color 1.85 : 1	USA
## 91		<NA> <NA>	<NA>
## 92		<NA> <NA>	<NA>
## 93		Color 2.39 : 1	USA
## 94		Color 2.39 : 1	USA
## 95		Color <NA>	USA
## 96		<NA> <NA>	<NA>
## 97		Color 2.35 : 1	USA
## 98		Color <NA>	<NA>
## 99		Color 1.33 : 1	USA
## 100		Color 2.39 : 1	USA
## 101		Color 2.35 : 1	USA
## 102		Color <NA>	USA
## 103		Color <NA>	UK
## 104		Color 2.39 : 1	USA
## 105		Color 2.35 : 1	USA
## 106		Color , Color (ACES) <NA>	USA
## 107		<NA> <NA>	<NA>
## 108		Color 1.78 : 1	USA
## 109		<NA> <NA>	<NA>
## 110		Color <NA>	<NA>
## 111		<NA> <NA>	USA
##	budget.est	usa.opening.date	production.co
## 1	TRUE	13 September 1992	co0017447, co0178734
## 2	TRUE	31 May 1993	co0028557, co0007127, co0021159
## 3	TRUE	12 December 1993	co0007143, co0050210, co0007127
## 4	TRUE	9 April 1995	co0102770
## 5	TRUE	8 October 1995	co0038663
## 6	TRUE	7 July 1996	co0000756, co0050111
## 7	TRUE	6 July 1997	co0050868, co0009119, co0108321
## 8	TRUE	22 November 1998	co0049348, co0026281, co0102770
## 9	<NA>	29 March 1998	co0019626, co0055790
## 10	<NA>	2 August 1998	co0020336, co0057284, co0008970
## 11	TRUE	4 July 1999	co0004624, co0046771, co0196312
## 12	TRUE	7 July 2002	co0050868, co0009119, co0108321
## 13	TRUE	13 June 1999	co0046718, co0048779, co0037668
## 14	TRUE	5 November 2000	co0000756, co0040938, co0166202
## 15	TRUE	10 October 1999	co0113592, co0098530
## 16	<NA>	<NA>	co0086763, co0027546, co0017423
## 17	TRUE	20 July 2003	co0050868, co0102770, co0038663
## 18	<NA>	<NA>	co0079754
## 19	TRUE	30 December 2001	co0050868, co0035930, co0011877
## 20	TRUE	23 December 2001	co0023400, co0004265, co0073465
## 21	TRUE	17 March 2002	co0049929, co0024939, co0072775

## 22	TRUE	28 March	2004	co0022594,	co0007545,	co0058058
## 23	TRUE	3 October	2004		co0129164,	co0040938
## 24	TRUE	20 April	2003	co0002663,	co0033868,	co0037296
## 25	TRUE	14 December	2003		co0000756,	co0030016
## 26	TRUE	18 July	2004	co0000756,	co0133294,	co0022730
## 27	<NA>			<NA>		
## 28	TRUE			<NA>	co0110544,	co0158077,
## 29	<NA>	5 September	2004		co0111298,	co0050873
## 30	<NA>	29 May	2005	co0017834,	co0024292,	co0025423
## 31	TRUE	13 February	2005		co0050868,	co0072775
## 32	TRUE			<NA>	co0097231,	co0108388
## 33	TRUE	19 October	2008	co0028932,	co0072775,	co0001946
## 34	TRUE	6 July	2008	co0050868,	co0142678,	co0237209
## 35	TRUE	17 December	2006	co0050868,	co0142678,	co0072775
## 36	<NA>			<NA>		
## 37	TRUE			<NA>	co0025891,	co0264916,
## 38	TRUE	2 April	2006	co0165993,	co0072775,	co0388505
## 39	TRUE	16 December	2007	co0002663,	co0108864,	co0093488
## 40	<NA>			<NA>		
## 41	TRUE	21 December	2008	co0050868,	co0142678,	co0072775
## 42	<NA>			<NA>		
## 43	<NA>			<NA>		
## 44	TRUE	21 September	2008		co0010568,	co0072775
## 45	<NA>			<NA>		
## 46	<NA>	24 May	2009	co0008970,	co0025116,	co0288376
## 47	TRUE	13 October	2019	co0152219,	co0026281,	co0630951
## 48	<NA>			<NA>		
## 49	<NA>			<NA>	co0200336,	co0050873,
## 50	TRUE	6 July	2008		co0166660,	co0236322
## 51	<NA>			<NA>	co0732934,	co0677683,
## 52	<NA>			<NA>	co0766067,	co0766124,
## 53	<NA>			<NA>	co0072775,	co0015206,
## 54	TRUE	13 June	2010	co0050868,	co0072775,	co0110101
## 55	<NA>			<NA>		
## 56	<NA>			<NA>		
## 57	TRUE	22 December	2013	co0023400,	co0073081,	co0437457
## 58	<NA>			<NA>		co0215421
## 59	<NA>			<NA>		
## 60	<NA>			<NA>		
## 61	<NA>			<NA>		
## 62	<NA>			<NA>		
## 63	<NA>			<NA>		
## 64	<NA>			<NA>		co0134421
## 65	TRUE	7 August	2016	co0028338,	co0038332,	co0283444
## 66	TRUE	27 May	2012	co0050868,	co0316575,	co0009119
## 67	<NA>			<NA>	co0166504,	co0243623
## 68	<NA>			<NA>		
## 69	<NA>			<NA>		co0050868
## 70	TRUE	19 January	2020	co0050868,	co0605949,	co0102770
## 71	TRUE			<NA>		co0289471
## 72	<NA>			<NA>	co0072775,	co0002663
## 73	TRUE			<NA>	co0339181,	co0610944,
## 74	TRUE	19 February	2012	co0072775,	co0033486,	co0174373
## 75	<NA>			<NA>		

```

## 76      <NA>      <NA>      co0050868, co0072775
## 77      <NA>      <NA>
## 78      <NA>      <NA>
## 79      <NA>      <NA>
## 80      TRUE      2 June 2013 co0050868, co0072775, co0054054
## 81      <NA>      <NA>
## 82      TRUE      21 December 2014 co0448488, co0219609, co0072775
## 83      TRUE      16 February 2014 co0108864, co0002663, co0093488
## 84      <NA>      <NA>
## 85      <NA>      <NA>
## 86      <NA>      <NA>      co0137851
## 87      <NA>      <NA>
## 88      <NA>      <NA>
## 89      <NA>      7 April 2013 co0151922, co0239430, co0307871
## 90      TRUE      1 March 2015 co0393274, co0449458, co0330850
## 91      <NA>      <NA>
## 92      <NA>      <NA>
## 93      TRUE      27 December 2015 co0479739, co0259281, co0562694
## 94      <NA>      11 March 2018 co0319272, co0062793, co0116500
## 95      <NA>      <NA>      co0541614
## 96      <NA>      <NA>
## 97      TRUE      18 December 2016 co0046718, co0108864, co0017525
## 98      <NA>      <NA> co0050868, co0026281, co0026545
## 99      <NA>      <NA>
## 100     TRUE      <NA> co0637285, co0072775, co0144901
## 101     TRUE      29 December 2019 co0047265, co0286257, co0179259
## 102     <NA>      <NA>      co0159509
## 103     <NA>      <NA> co0668501, co0668502, co0694052
## 104     TRUE      26 May 2019 co0008970, co0226345, co0699871
## 105     <NA>      <NA>      co0072775
## 106     TRUE      12 May 2019 co0266060, co0078768, co0133843
## 107     <NA>      <NA>
## 108     <NA>      <NA>      co0441292, co0063961
## 109     <NA>      <NA>
## 110     <NA>      <NA> co0637285, co0144901, co0688564
## 111     TRUE      <NA> co0072775, co0306981, co0002663

```

We now have data for all of his movies

```

ttid.legend = "tt0480249";

cast = subsetDataFrame(imdb.data$movies.df$cast, "ttid", "==", ttid.legend);
cast;

```

```

##      ttid actor.rank      nmid      name
## 62104 tt0480249      1 nm0000226      Will Smith
## 62105 tt0480249      2 nm0103797      Alice Braga
## 62106 tt0480249      3 nm2253071      Charlie Tahan
## 62107 tt0480249      4 nm0724757 Salli Richardson-Whitfield
## 62108 tt0480249      5 nm2405238      Willow Smith
## 62109 tt0480249      6 nm0287756      Darrell Foster
## 62110 tt0480249      7 nm0333318      April Grace
## 62111 tt0480249      8 nm0005231      Dash Mihok
## 62112 tt0480249      9 nm2724906      Joanna Numata
## 62113 tt0480249     10 nm8421984      Abbey

```



```
## 62114 tt0480249      11 nm2869725      Kona
## 62115 tt0480249      12 nm2677021      Samuel Glen
## 62116 tt0480249      13 nm0565503      James McCauley
## 62117 tt0480249      14 nm1677477      Marin Ireland
## 62118 tt0480249      15 nm2857229      Pedro Mojica
##
##                character
## 62104                Robert Neville
## 62105                Anna
## 62106                Ethan
## 62107 Zoe Neville (as Salli Richardson)
## 62108                Marley Neville
## 62109                Mike - Military Escort
## 62110                TV Personality
## 62111                Alpha Male
## 62112                Alpha Female
## 62113                Sam
## 62114                Sam
## 62115                Military Driver - Jay
## 62116                Male Evacuee
## 62117                Woman Evacuee
## 62118                Sergeant
```

```
info = subsetDataFrame(imdb.data$movies.df$info, "ttid", "==", ttid.legend);
info;
```

```
##                ttid country language      release.date
## 5020 tt0480249      USA   English 14 December 2007
##
##                aka
## 5020 I Am Legend: The IMAX Experience
##
##                filming.location
## 5020 11 Washington Square North, Greenwich Village, Manhattan, New York City, New York, USA
##                budget usa.opening usa.gross world.gross
## 5020 150000000      77211321 256393010   585410052
##
##                production.name
## 5020 Warner Bros., Village Roadshow Pictures, Weed Road Pictures
##
##                runtime                sound
## 5020 101 min , 104 min (Alternate ending) SDDS , Dolby Digital , DTS
##                color  aspect release.location budget.est usa.opening.date
## 5020 Color (Technicolor) 2.35 : 1      USA      TRUE 16 December 2007
##
##                production.co
## 5020 co0002663, co0108864, co0093488
```

```
# notice the money is not yet numeric, some cleansing may be required
# some elements in the column are in "other currencies"
```

```
text = subsetDataFrame(imdb.data$movies.df$text, "ttid", "==", ttid.legend);
text;
```

```
##
## my.row.5019 Years after a plague kills most of humanity and transforms the rest into monsters, the s
##
## my.row.5019 Robert Neville (Will Smith) is a scientist who was unable to stop the spread of the terr
##
##                keywords
## my.row.5019 post apocalypse, new york city, virus, loneliness, survivalist
##
##                taglines      ttid
## my.row.5019 Welcome to earth. Population 1. tt0480249
```

Denzel Washington

[I would suggest reviewing Will and seeing how many features you can extract, then copy those same features here, as you do Denzel.

The more features for comparison, the better. Even if the features don't align with your final answer, you can use language like: although Denzel was in more top-50 movies by year (as a proportion), ... THAT would be an example if your final answer is TEAM WILL

]

– TODO –

Conclusion

We now have all of the movie information for each of these two actors. And you can possibly predict gender. Think of some gender-specific features you could consider: how many headliners with Will/Denzel are the same-sex? Opposite sex? What about the cast? What about the extended cast (others = 40)?