

```
1 ---
2 title: "Sophia ibraeva"
3 author: "Sophia"
4 date: "2023-01-30"
5 output:
6   word_document: default
7   pdf_document: default
8   html_document:
9     df_print: paged
10 ---
11
12 ```{r setup, include=FALSE}
13 knitr::opts_chunk$set(echo = TRUE)
14 ```
15
16 ```{r}
17 Movie <- read.csv("~/Downloads/movie_data.csv")
18 Movie
19 ```
```

director_name <chr>	duration <int>	actor_2_name <chr>
James Cameron	178	Joel David Moore
Gore Verbinski	169	Orlando Bloom
Sam Mendes	148	Rory Kinnear
Christopher Nolan	164	Christian Bale
Doug Walker	NA	Rob Walker
Andrew Stanton	132	Samantha Morton
Sam Raimi	156	James Franco
Nathan Greno	100	Donna Murphy
Joss Whedon	141	Robert Downey Jr.
David Yates	153	Daniel Radcliffe

1–10 of 5,043 rows | 1–3 of 14 columns Previous 1 2 3 4 5 6 ... 100 Next

20

21 ````{r}`22 `str(Movie)`23 `````

```

'data.frame':  5043 obs. of  14 variables:
 $ director_name      : chr  "James Cameron" "Gore Verbinski" "Sam Mendes"
"Christopher Nolan" ...
 $ duration           : int   178 169 148 164 NA 132 156 100 141 153 ...
 $ actor_2_name       : chr   "Joel David Moore" "Orlando Bloom" "Rory Kinnear"
"Christian Bale" ...
 $ genres             : chr   "Action|Adventure|Fantasy|Sci-Fi"
"Action|Adventure|Fantasy" "Action|Adventure|Thriller" "Action|Thriller" ...
 $ actor_1_name       : chr   "CCH Pounder" "Johnny Depp" "Christoph Waltz" "Tom
Hardy" ...
 $ movie_title        : chr   "Avatar " "Pirates of the Caribbean: At World's
End " "Spectre " "The Dark Knight Rises " ...
 $ num_voted_users    : int   886204 471220 275868 1144337 8 212204 383056 294810
462669 321795 ...
 $ actor_3_name       : chr   "Wes Studi" "Jack Davenport" "Stephanie Sigman"
"Joseph Gordon-Levitt" ...
 $ movie_imdb_link    : chr   "http://www.imdb.com/title/tt0499549/?
ref_=fn_tt_tt_1" "http://www.imdb.com/title/tt0449088/?ref_=fn_tt_tt_1"
"http://www.imdb.com/title/tt2379713/?ref_=fn_tt_tt_1"
"http://www.imdb.com/title/tt1345836/?ref_=fn_tt_tt_1" ...
 $ num_user_for_reviews: int   3054 1238 994 2701 NA 738 1902 387 1117 973 ...
 $ language           : chr   "English" "English" "English" "English" ...
 $ country            : chr   "USA" "USA" "UK" "USA" ...
 $ title_year         : int   2009 2007 2015 2012 NA 2012 2007 2010 2015 2009 ...
 $ imdb_score         : num   7.9 7.1 6.8 8.5 7.1 6.6 6.2 7.8 7.5 7.5 ...

```

24

25 ````{r}

26 summary(Movie)

27 ````



director_name	duration	actor_2_name	genres
Length:5043	Min. : 7.0	Length:5043	Length:5043
Class :character	1st Qu.: 93.0	Class :character	Class :character
Mode :character	Median :103.0	Mode :character	Mode :character
	Mean :107.2		
	3rd Qu.:118.0		
	Max. :511.0		
	NA's :15		
actor_1_name	movie_title	num_voted_users	actor_3_name
Length:5043	Length:5043	Min. : 5	Length:5043
Class :character	Class :character	1st Qu.: 8594	Class :character
Mode :character	Mode :character	Median : 34359	Mode :character
		Mean : 83668	
		3rd Qu.: 96309	
		Max. :1689764	
movie_imdb_link	num_user_for_reviews	language	
Length:5043	Min. : 1.0	Length:5043	
Class :character	1st Qu.: 65.0	Class :character	
Mode :character	Median : 156.0	Mode :character	
	Mean : 272.8		
	3rd Qu.: 326.0		
	Max. :5060.0		
	NA's :21		
country	title_year	imdb_score	
Length:5043	Min. :1916	Min. :1.600	
Class :character	1st Qu.:1999	1st Qu.:5.800	
Mode :character	Median :2005	Median :6.600	
	Mean :2002	Mean :6.442	
	3rd Qu.:2011	3rd Qu.:7.200	
	Max. :2016	Max. :9.500	
	NA's :108		

28

29

30

31

```\r}

median(Movie\$imdb\_score)

```\r}

[1] 6.6

32

33

34

35

```\r}

summary(Movie\$imdb\_score)

```\r}

| Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. |
|-------|---------|--------|-------|---------|-------|
| 1.600 | 5.800 | 6.600 | 6.442 | 7.200 | 9.500 |

36

37

38

39

```\r}

class(Movie\$director\_name)

```\r}

[1] "character"

40

41

42

43

```\r}

class(Movie\$title\_year)

```\r}

[1] "integer"

44

45

46

47

48

```\r}

genre\_custom=as.factor(Movie\$genres)

class(genre\_custom)

```\r}

[1] "factor"

49

50

51

52

53

```\r}

genre\_custom\_cat\_to\_num<-unclass(genre\_custom)

class(genre\_custom\_cat\_to\_num)

```\r}

```
54
55 ````{r}
56 #plot
57 tab
58 x=c(Movie$num_voted_users)
59 y=c(Movie$num_user_for_reviews)
60 plot(x,y, main = "scatter plot",
61       xlab = "num_voted_users" , ylab = "num_user_for_reviews")
62 abline(lm(y~x), col= "Blue")
63 ````
```

[1,] [1,] [1,]
[2,] "61" "63" "NAHESH"
[2,] "67" "68" "79"

R Console

