

problema3-ck4

Maelson Marques de Lima

21 de maio de 2016

FPCC2-p2-c2

ORGANIZANDO OS DADOS:

Descrição da atividade:

Escolha 2 das perguntas que você listou e construa o relatório final deste problema respondendo-as ut.

Inicialmente a organização dos dados:

```
#Tabela mpvies
movies <- read.csv("~/Rprojetos/Rprojects-fpcc2/bd-movies/movies.csv")
#View(movies)

#Tabela ratings
ratings <- read.csv("~/Rprojetos/Rprojects-fpcc2/bd-movies/ratings.csv")

#Dados Um --Base 01
movies_join_ratings = merge(x = movies, y = ratings, by = "movieId", all.y = TRUE)
#View(movies_join_ratings)

#Dados Dois --Base 02 (Filmes Brasileiros Lançados - 1995 a 2014)
filmes_tratados<-read.csv("~/Rprojetos/Rprojects-fpcc2/bd-movies/filmes_tratados.csv")
#View(filmes_tratados)

knitr::kable(filmes_tratados[1:10, 1:11], caption = "Amostra de Filmes Tratados:")
```

X	AnoLancamento	Titulo	Diretor	Produtor
1	2014	12 de Junho de 1993 - O Dia da Paixão Palmeirense	Jaime Queiroz, Mauro Beting	Canal 1
2	2014	A Balada do Provisório	Felipe David Rodrigues	Maria
3	2014	A Casa Elétrica	Gustavo Fogaça	Panda
4	2014	A Farra do Circo	Roberto Berliner/PEDRO BRONZ	TV ZE
5	2014	A Grande Vitória	Stefano Capuzzi Lapietra	ALFA
6	2014	A História do Homem Henry Sobel	André Bushatsky	Twins
7	2014	A Noite da Virada	Fábio Mendonça	O2 CIN
8	2014	A Oeste do Fim do Mundo	Paulo Nascimento	Accord
9	2014	A Onda da Vida – Uma História de Amor & Surf	José Augusto Costa Henriques	IDÉIA
10	2014	A Pelada	Damien Chemin	W.G. F

```
#Dados Três --Base 03
ratings_filme<- read.csv("~/Rprojetos/Rprojects-fpcc2/bd-movies/ratings-por-filme.csv")
```

```
#View(ratings_filme)
```

```
knitr::kable(ratings_filme[1:10, 1:5], caption = "Ratings por Filmes: ")
```

Table 2: Ratings por Filmes:

movieId	title	genres	rating	popularity
1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	3.894802	60424
2	Jumanji (1995)	Adventure Children Fantasy	3.221086	23950
3	Grumpier Old Men (1995)	Comedy Romance	3.180094	15267
4	Waiting to Exhale (1995)	Comedy Drama Romance	2.879727	2935
5	Father of the Bride Part II (1995)	Comedy	3.080811	14769
6	Heat (1995)	Action Crime Thriller	3.836536	26593
7	Sabrina (1995)	Comedy Romance	3.373367	14846
8	Tom and Huck (1995)	Adventure Children	3.139661	1475
9	Sudden Death (1995)	Action	3.015246	4329
10	GoldenEye (1995)	Action Adventure Thriller	3.436888	31357

Minhas perguntas:

Qual a probabilidade de ocorrência de uma nota menor que 4 em grupos com mais de 1 gênero?

Calcule a probabilidade de ocorrência de uma nota menor que 4 em uma distribuição normal baseados em sua média e seu desvio padrão para grupos com mais de 1 gênero. Calcule seu intervalo de confiança com precisão de 95%. Filmes mais populares tem melhores notas? Calcule a variância e o desvio padrão para estas variáveis. Plote o gráfico com a distribuição de probabilidades do número de notas perfeitas. Qual dos filmes possui maior variância de notas baseados no número de gêneros. Estime com 95% de precisão. Estime a popularidade de filmes com mais de 2 gêneros.