Домашна работа 1

СЕМ

Изготвена от: Станислав Станчев, ИС, група 3, Фн. 71283

Обект на изследването: 25 произволно избрани игри от 4 големи студиа.

Разгледани наблюдения: Игрите ще бъдат характеризирани чрез следните данни:

Rating – средно оценка на играта взета от Metacritic;

Studio–студиото, което е разработило играта;

Genre – нейния жанр;

Описание на данните в R:

> ratings = scan()

1: 92 90 93 96 88 91 86 96 89 90 93 77 83 93 89 88 88 85 87 83 94 90 91 92 65

> studios = c("EA", "EA", "Blizzard", "Valve", "Blizzard", "EA", "EA", "Valve", "Ubisoft", "Activision", "Activision", "EA", "Blizzard", "Blizzard", "Valve", "Valve", "EA", "EA", "EA", "Ubisoft", "Valve", "EA", Activision", "EA", "Ubisoft")

> genres = c("Simulator", "Simulator", "MMORPG", "FPS", "RTS", "FPS", "Simulator", "FPS", "MMORPG", "MMORPG", "FPS", "Simulator", "MMORPG", "RTS", "FPS", "FPS", "FPS", "FPS", "FPS", "RTS", "RTS", "RTS", "FPS", "RTS", "FPS")

Изграждане на Data frame:

> games = data.frame(genres, ratings, sold)

> games

 studios genres ratings

1 EA Simulator 92

2 EA Simulator 90

3 Blizzard MMORPG 93

4 Valve FPS 96

5 Blizzard RTS 88

6 EA FPS 91

7 EA Simulator 86

8 Valve FPS 96

9 Ubisoft MMORPG 89

10 Activision MMORPG 90

11 Activision FPS 93

12 EA Simulator 77

13 Blizzard MMORPG 83

14 Blizzard RTS 93

15 Valve FPS 89

16 Valve FPS 88

17 EA FPS 88

18 EA FPS 85

19 EA FPS 87

20 Ubisoft RTS 83

21 Valve RTS 94

22 EA RTS 90

23 Activision FPS 91

24 EA RTS 92

25 Ubisoft FPS 65

Дескриптивни статистики:

> mean(ratings)

[1] 88.36

> sd(ratings)

[1] 6.499231

> median(ratings)

[1] 90

> summary(ratings)

 Min. 1st Qu. Median Mean 3rd Qu. Max.

 65.00 87.00 90.00 88.36 92.00 96.00

> boxplot(ratings, horizontal = T)

|  |  |
| --- | --- |
|  | От boxplot диаграмата на отделните рейтинги на изследваните игри, отбелязваме, че има 2 силно отличаващи се рейтинга – 65 и 77. Останалите данни са относително симетрични, спрямо своето разпределение. |

> barplot(table(genres, studios), legend.text = T)

|  |  |
| --- | --- |
|  | От тази barplot диаграма, може да видим дали някое от студиата се е специализирало в определен жанр. Както виждаме, всяко студио е разработвало игри в поне 2 жанра, но никое няма игри и от 4те жанра. Относно специализация, виждаме че Valve имам основно FPS игри и една RTS игра, докато останалите студиа са с по-балансиран каталог от игри. |

> barplot(table(genres))

|  |  |
| --- | --- |
|  | Лесно се вижда, че от наблюдаваните игри, основна част са от жанра FPS, което не и изненадващо, тъй като това е най-динамичния и лесен за навлизане жанр в индустрията. |

> barplot(table(studios))

|  |  |
| --- | --- |
|  | Отново е лесно забележимо, кое от студиата е с най-голям брой разработени игри. Това също не е изненадващо поради причината, че EA притежават най-голям брой first-party developers, т.е. по-малки студия, които са част от ЕА. |

> hist(ratings, prob=T)

> lines(density(ratings))

|  |  |
| --- | --- |
|  | От наблюдения можем да заключим, че вероятноста рейтинга на игра да е в интервала от 65 до 85 е значително по-малка от това да е в интервала [85; 95]. В частност това се дължи на факта, че част от наблюдаваните игри са сред най-продаваните заглавия за компютър, което означава, че има пряка връзка между рейтинг и продажби, която не бе цел на сегашното наблюдения, но е любопитно допълнение. |

> boxplot(ratings ~ studios)



Диаграма разглежда разпределението на рейтингите според студиата. Най-голямо впечатление прави Ubisoft със своя размах от 65 до 90 рейтинг на свойте игри. Студиото EA има тейтинги групирани в интервала [85; 95] като има outlier от 76, като те са единствените с подобен тип наблюдение. Нужно е да отбележим, че Valve са студиото с най-висок среден рейтинг и с максимален рейтинг от всички студиа.

boxplot(ratings ~ genres)



На тази диаграма виждаме разпределението на рейтингите според жанровете. От данните се вижда, че FPS жанрът е с максимален рейтинг, но средния рейтинг е 3ти по-големи, а също се наблюдава и наблюдение със силно отличаваща се стойност, което не се открива при останалите жанрове. От тях най-успешни като рейтинг е RTS жанрът със среден рейтинг равен на 90 ( mean(games[genres == "RTS", 3])).

Индекс на наблюдението с максимален рейтинг:

> which.max(games[, 3])

[1] 4

Извеждане на наблюдението с максимален рейтинг:

> games[ratings == max(ratings),]

 studios genres ratings

4 Valve FPS 96

8 Valve FPS 96

Забелязваме, че имаме 2 наблюдения, които са с максималет рейтинг, които също са и от един жанр и от едно студио.

|  |  |
| --- | --- |
|  |  |