

Google Play Store Apps Data Analysis

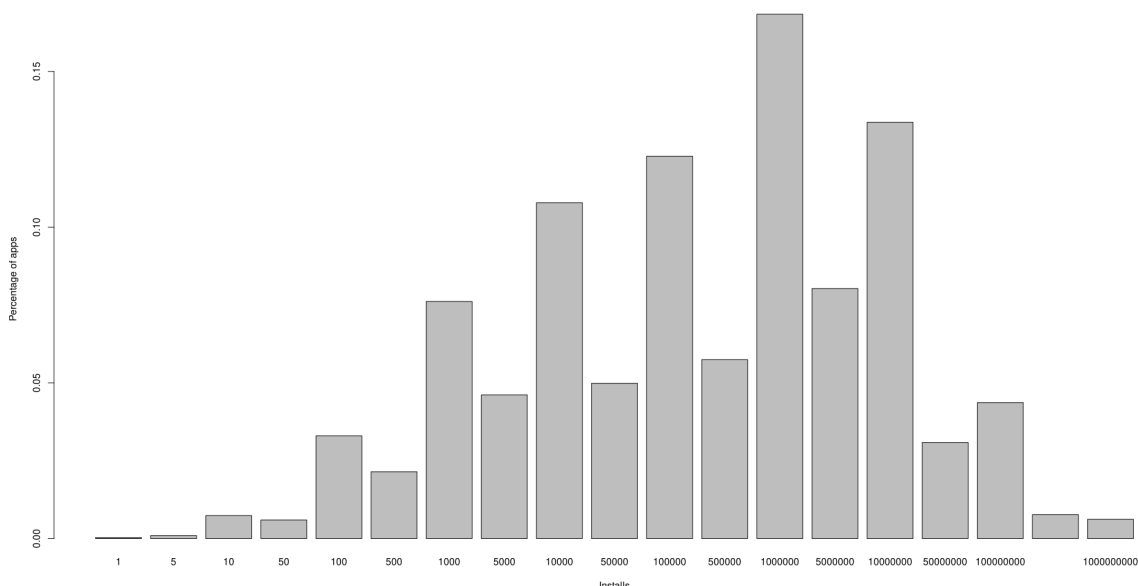
Данните са информация за Андроид приложенията от Google Play Store.

Анализираните данни са взети от Kaggle. Линк към данните: <https://www.kaggle.com/lava18/google-play-store-apps>

Целът на анализа е извличане на информация в коя категория е подходящо да се разработва ново приложение или коя категория има нужда от нови добри приложения.

Анализ

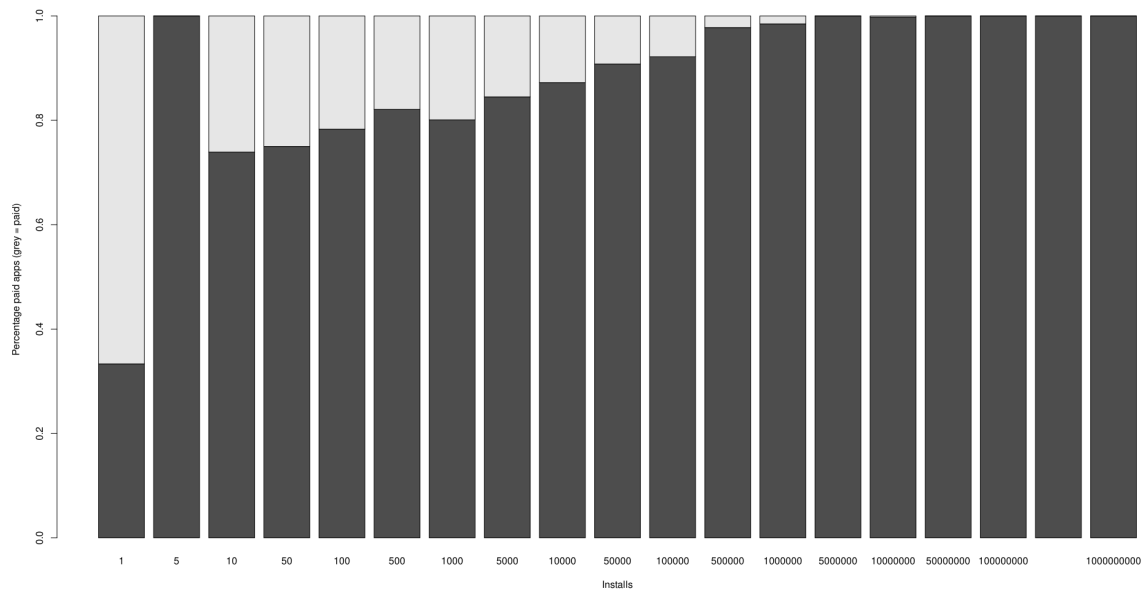
Брой изтегляния на приложение



От горната схема виждаме, че има много малко приложения (< 1%), които са изтеглени 1-5 пъти само, тоест няма опасност приложението ни да остане без изтегляния. Както се очаква има малко приложения (< 1%) с над 1 милиард изтегляния. От друга страна близо 20% от приложенията имат над 1 000 000 изтегляния. Това за нас означава, че има голяма вероятност приложението разработено от нас да има успех. Колко точно може да се види в информацията от summary - средния брой изтегляния е 17 897 444. А само 25% от приложенията са в първия квантил с под 10 000 тегления.

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1	10000	500000	17897444	5000000	1000000000

Брой изтегляния спрямо дали приложението е платено или не

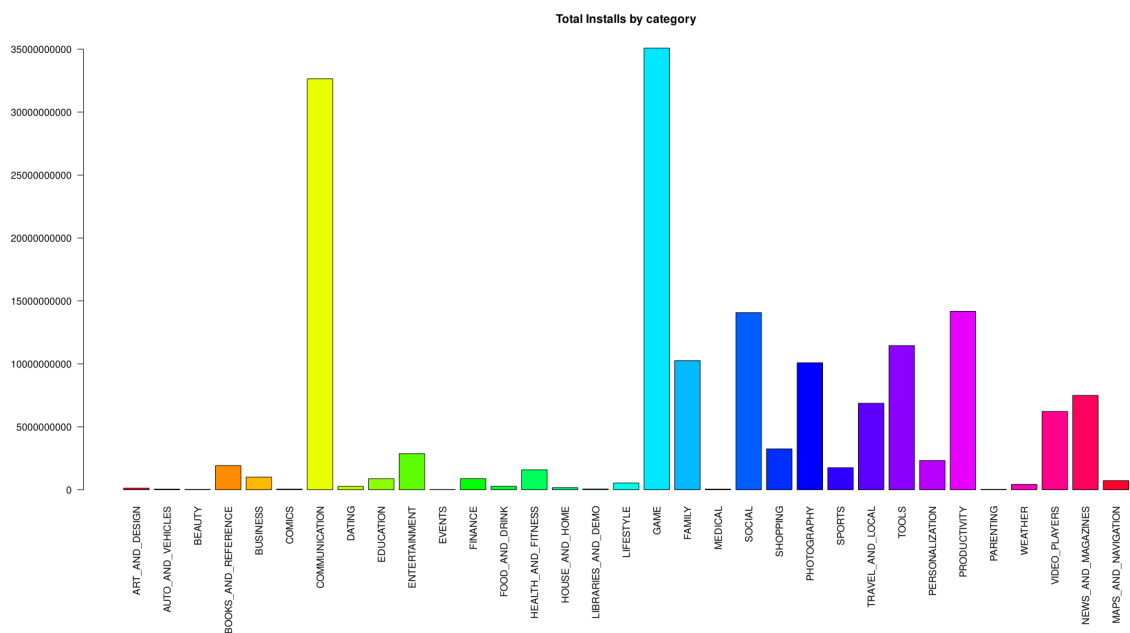


От графиката можем да видим, че има малко платени приложения с много потребители - например има < 1% платени приложения с над 1 милион тегления. Платените приложения имат малък дял от всички теглени приложения и то основно при приложенията теглени < 100 000 пъти. Това означава, че ако искаме приложението ни да има много тегления и съответно много потребители е необходимо да не е платено.

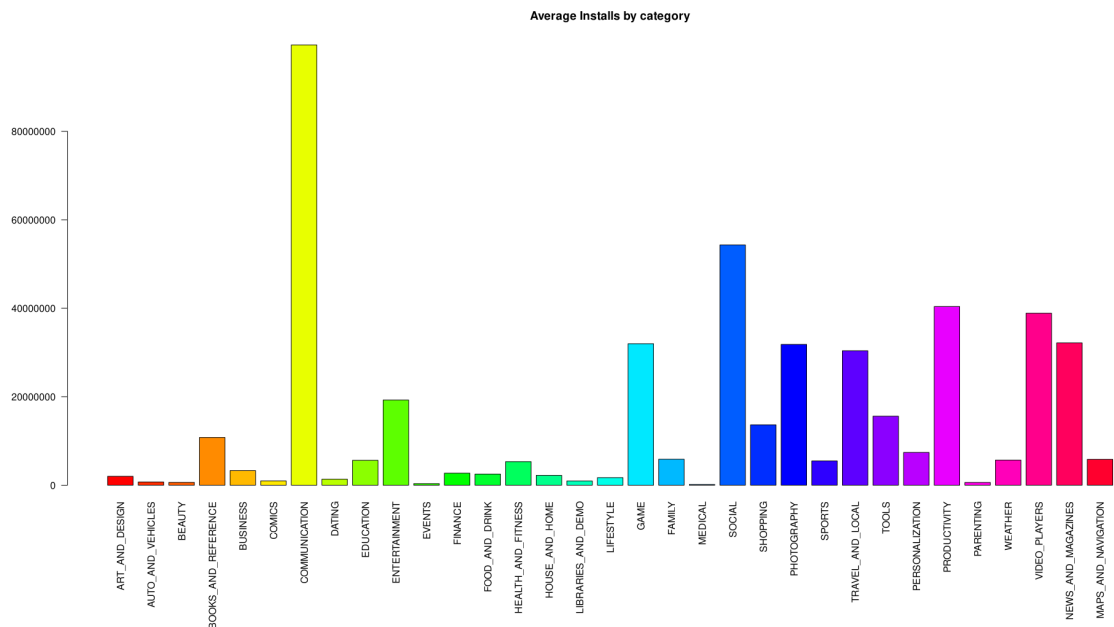
Категории

Нека разгледаме кои категории имат най-много тегления, за да може да изберем от коя категория да бъде приложението което ще разработваме.

Тотален брой тегления по категория



Среден брой тегления по категория



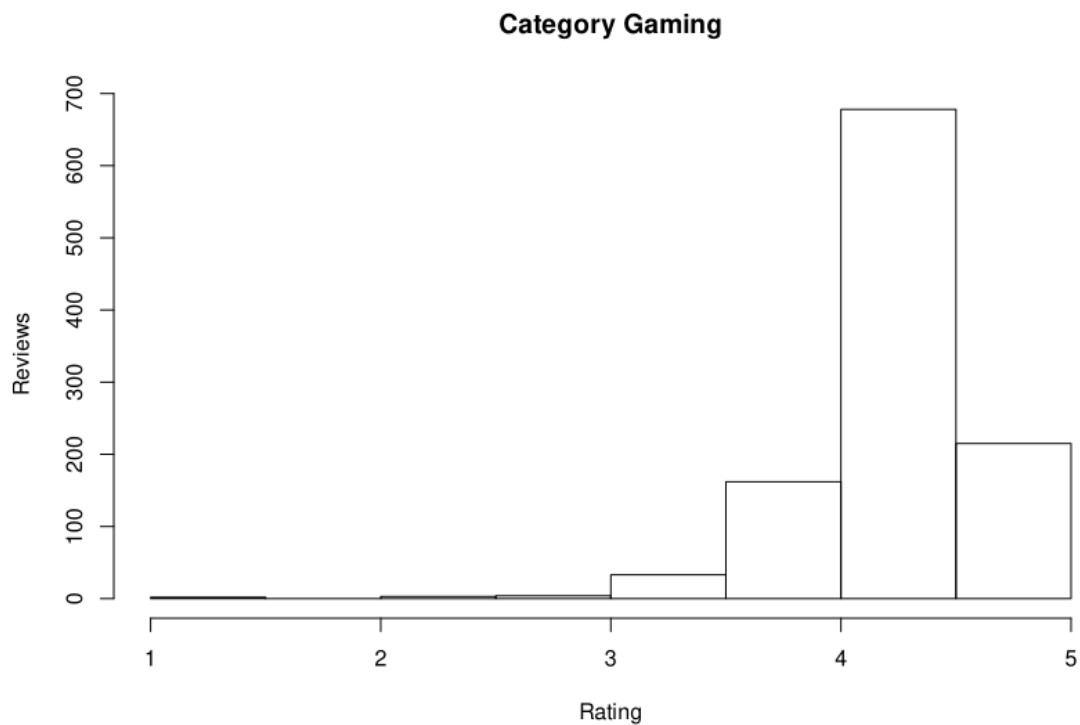
Горната диаграма показва общо колко тегления имат висчки приложения, а долната - средно аритметично на тегленията от дадена категория.

От горната диаграма виждаме, че приложенията в категории Комуникация и Игри са най-теглени от потребителите - общия брой на тегленията им е най-висок. От друга страна приложения от категория Изкуство, Медицински, Готвене и тн. не са популярни сред потребителите, защото имат малък брой тегления общо.

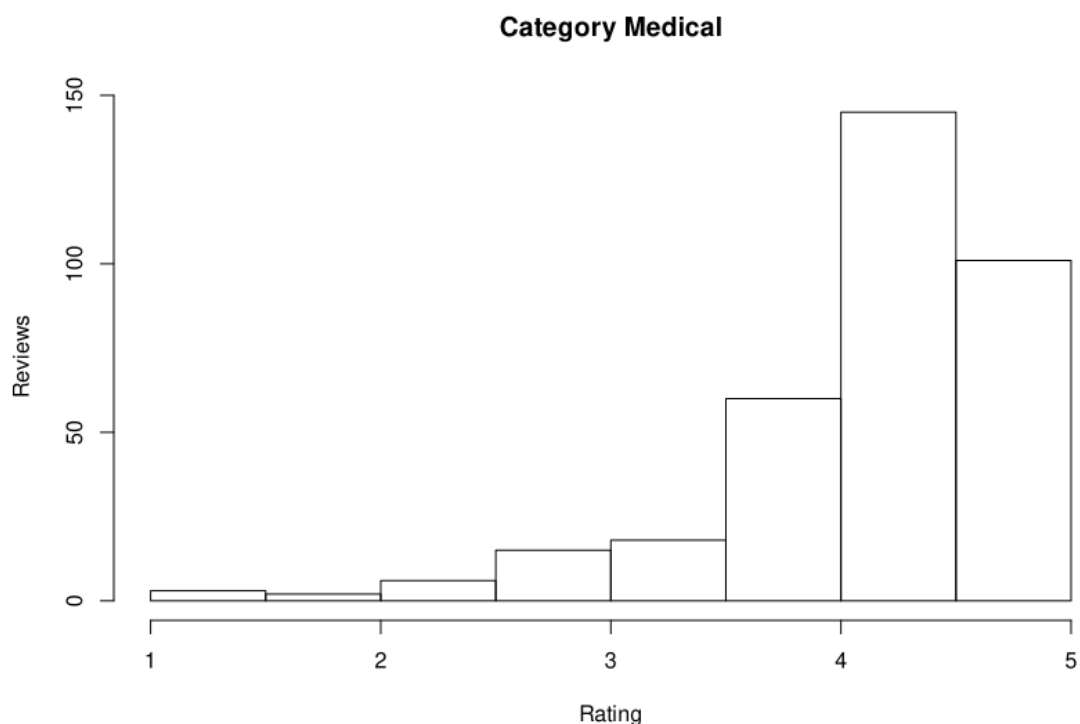
Има 2 насоки, които може да поемем - или да разработваме приложения от популярна категория или от непопулярна. Въпросът е как да изберем кое е по-доброто решение. От информацията виждаме категорията Игри е много популярна и най-свляяните приложения са от категории игри, когато говорим общо, но когато гледаме средните сваления категорията игри не е във водещите. Тоест има голям риск ако разработим игра да не е популярна. От друга страна приложенията от категория Комуникация имат както много общ брой тегления така и среден брой. Тоест ако разработваме приложение от тази категория има по-добър шанс то да е успешно. За категориите с по-малка популярност още нищо не може да заключим.

Рейтинг

Ще разгледаме няколко от категориите и какъв рейтинг получават приложенията в тях.

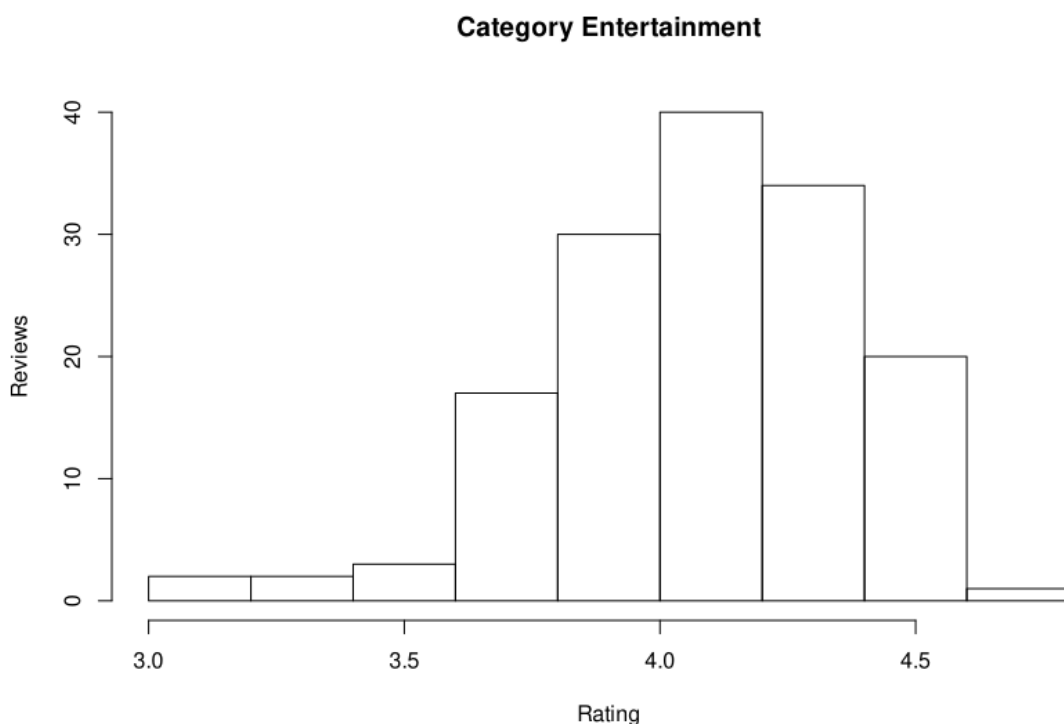
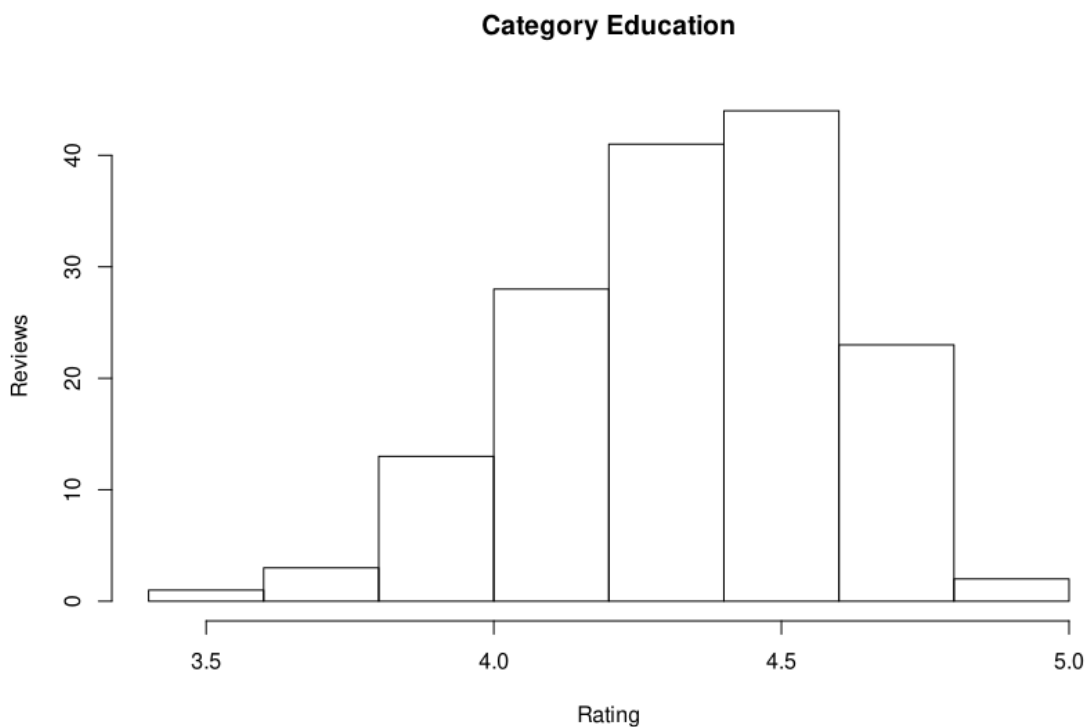


Категорията Игри има голям брой приложения с рейтинг 5 (което е максимума). Това означава, че има много добри игри, от които хората са доволни и най-вероятно нямат причина да ги заменят. За това и в горните графики се получава, че игрите имат много общ брой сваляния, но средният брой е нисък - има много разработчици, които виждат че игрите се свалят много и разработват още игри, но с неуспех. Категорията Игри е рискова и неподходяща.



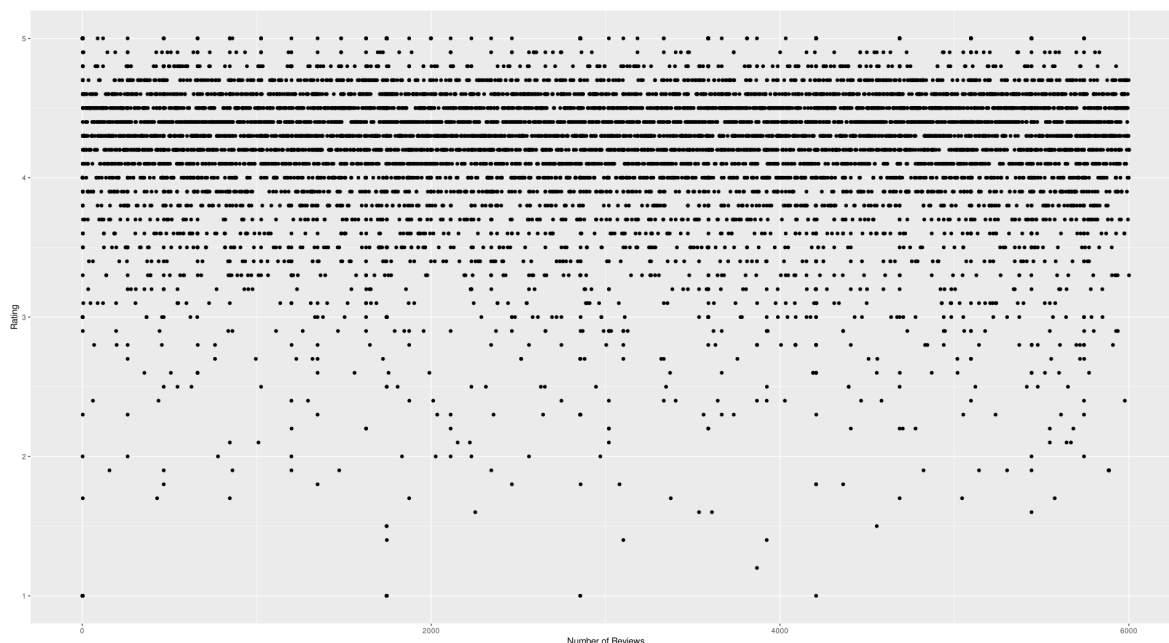
Тъй като категориите с много тегления като Игри са неподходящи, нека разгледаме някоя от категориите с малко тегления - Медицински приложения. Те не са популярни и нямат много потребители, но може би можем да пробием пазара точно там. За жалост се оказва, че малкият брой потребители, които използват такива приложения са напълно задоволени

- приложенията с рейтинг 5 са над 30%. Това означава, че някой добър разработчик вече е направил приложения, които да се ползват в този пазар и няма място за пробив.



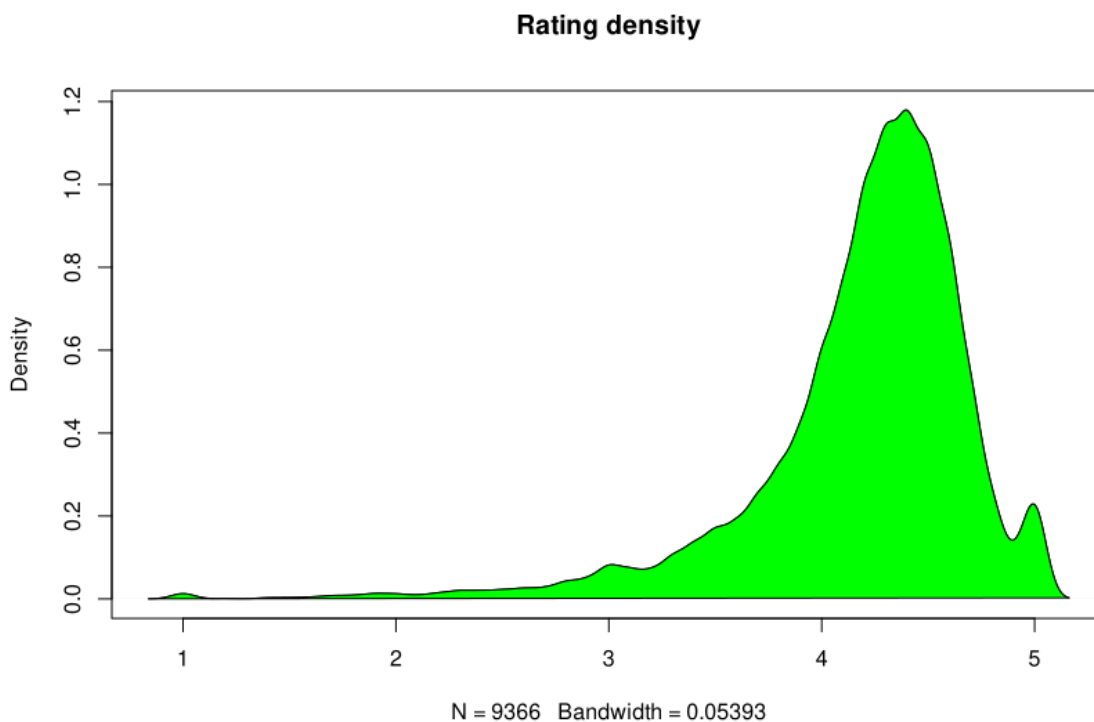
От останалите категории може да разгледаме Образование и Забавление. И в двете категории има по-малко от 1% приложения с рейтинг 5. Това означава, че има потребители (макар и малко), но те не са задоволени от приложенията, които има в момента на пазара. Тоест тези две категории са подходящи за пробив с ново иновативно приложение, което да спечели много потребители.

Рейтингът върши ли ни работа?



До момента почти може да заключим какво приложение е подходящо да разработваме, единствено може да направим проверка дали наистина рейтингът на приложенията има смисъл.

От горната диаграма се вижда ясно, че потребителите дават рейтинг основно на приложения, които са им допаднали поне малко - преобладава рейтинг между 4 и 5. А колко точно преобладава?



Достатъчно много, за да може да заключим, че когато хората харесват едно приложение, те са склонни да оставят рейтинг 4 или 5.

Защо може да се интересуваме от разработване на приложение с висок рейтинг?

