

Проект по “Статистика и емперични методи” (практикум)

Тема: Изследване на футболни звезди

Изготвен от Мартин Илиев, 62197



Съдържание

1. Увод в проблема
2. Представяне на броя на избраните играчи в различните клубове
3. Изследване на възрастта
4. Изследване на представянето на терена
5. Изследване на пазарната цена
6. Изследване на връзката между отбори и възраст
7. Изследване на връзката между отбор и представяне на терена
8. Изследване на връзката между цена и отбор
9. Изследване на връзката между възраст и представяне на терена
10. Изследване на връзката между възраст и цена на пазара
11. Изследване на връзката между представяне на терена и цена
12. Заключение

1. Въведение

В днешно време трансферният прозорец във футбола е нещо много непредвидимо. Трансферните рекорди се подобряват година след година и все повече трансфери се указват изненадващи.

Например, през 2009 трансферът на Кристиано Роналдо от Манчестър Юнайтед в Реал Мадрид за 90 милиона евро подобри почти двойно най-големият трансфер дотогава на стойност 50 милиона евро.

Макар да бяха нужни 4 години за Реал Мадрид да подбери този рекорд с трансфера на Гарет Бейл през 2013 за 95 милиона евро и после още 3 за Манчестър Юнайтед да купят Пол Погба за 100 милиона евро, от тогава насам цари абсолютен хаос на трансферния прозорец.

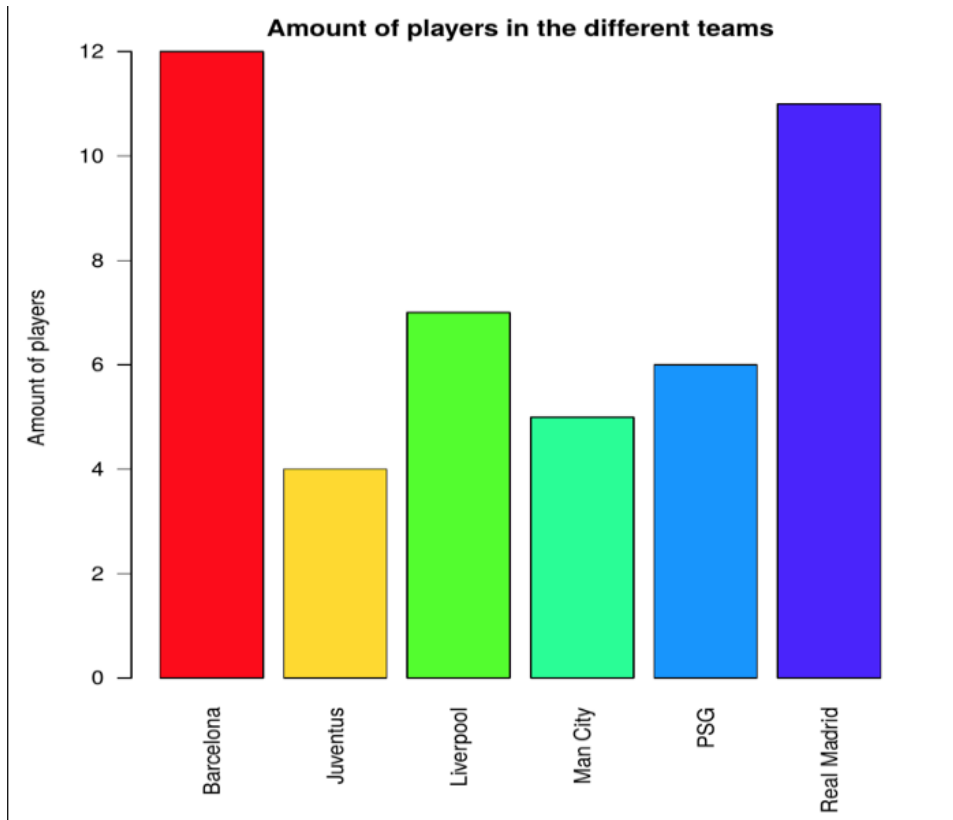
Основен виновник за това се смята да са ПСЖ, които в един трансферен прозорец купиха Неймар да Силва за 222 милиона евро и Килиан Мбапе за 180 милиона евро. Само през последния трансферен прозорец имаше 3 трансфера на стойност над 100 милиона евро.

Цената на един играч би следвало да се определя от фактори като възраст и представяне на терена, защото колкото по-млад е играчът, толкова повече време ще остане в клуба, а колкото по-добре играе, толкова по-вероятно е да играе добре и в новия клуб.

Идеята на този проект е да проверява каква е зависимостта между цена, възраст и представяне на играчите, като за целта съм подбрал титулярите на два от най-големите клубове в Европа, както и някои други звезди на световния футбол.

2. Представяне на броя на избраните играчи в различните клубове

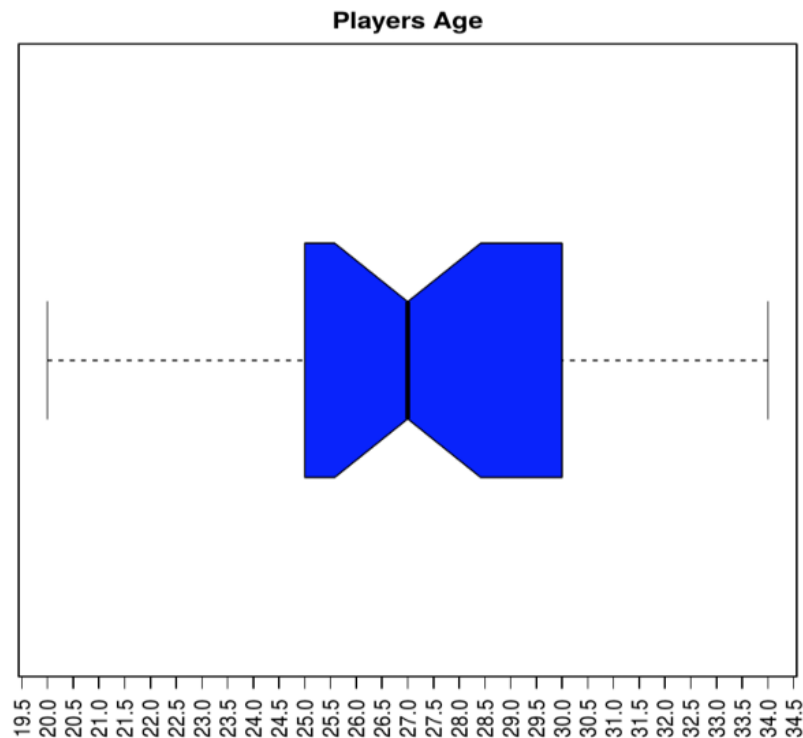
Ето представяне на играчите по отбори представено в barplot



Както може да се види, най-много играчи в извадката имаме от Барселона и Реал Мадрид

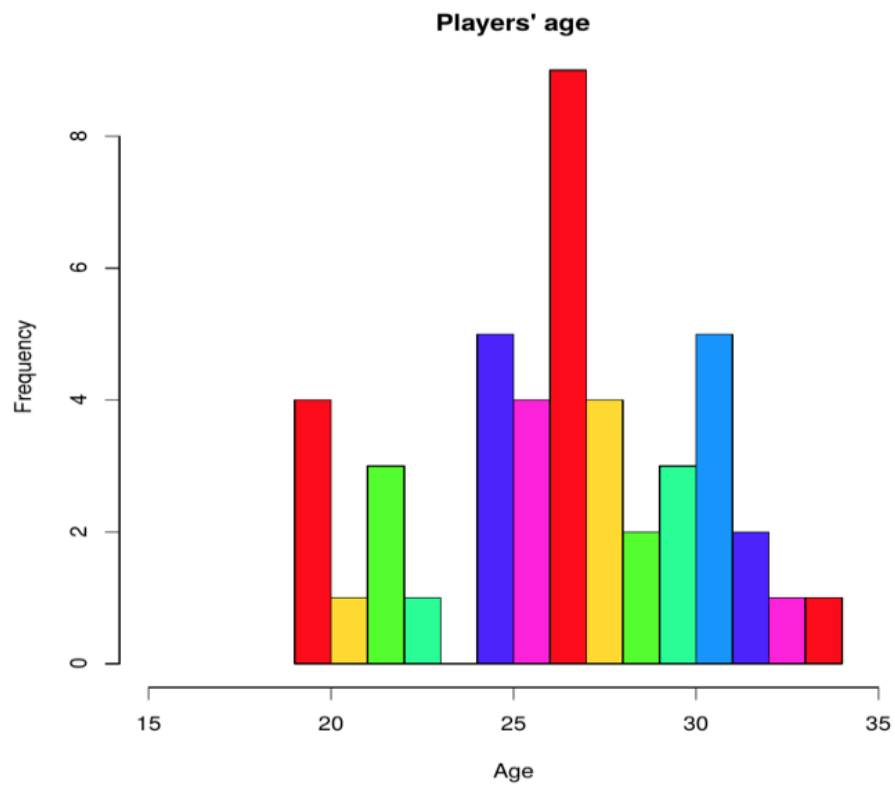
3. Изследване на възрастта

Ето разпределението по години на играчите представено чрез boxplot:



Както може да се види най-много играчи в извадката са на възраст 27 години, като най-младият е на 20, а най-възрастният на 24.

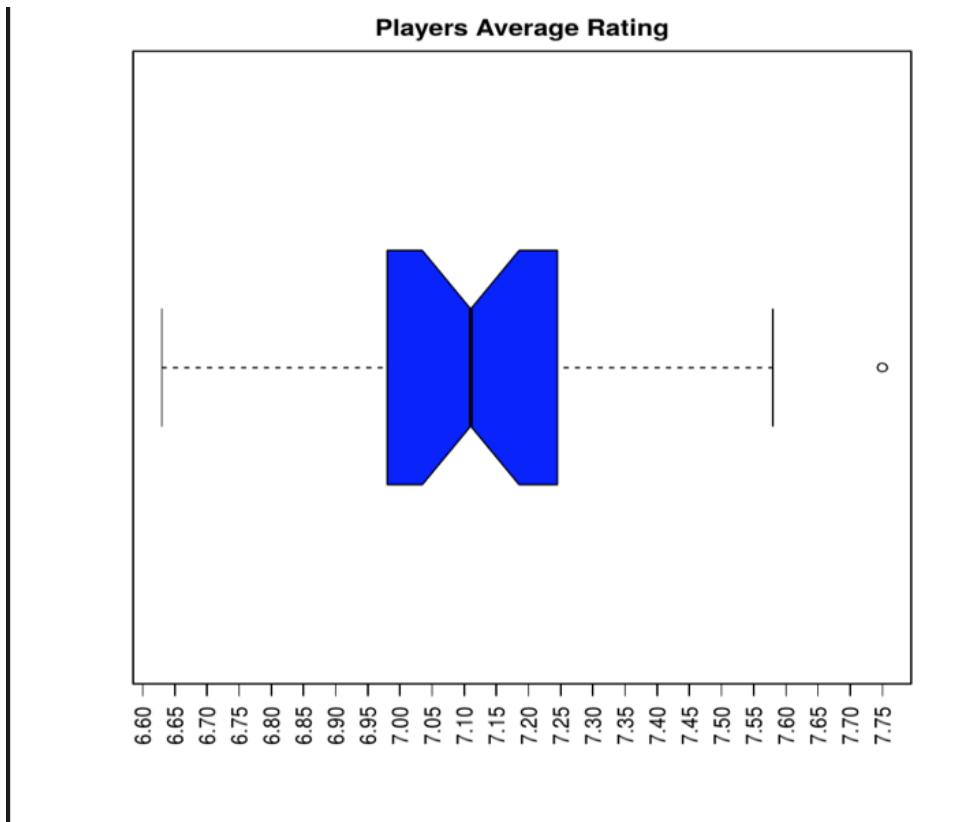
Ето и хистограма на броя играчи на съответна възраст:



Както може да се забележи, най-много играчи има на възраст 27, като в извадката липсват играчи на 23 и под 19 годишна възраст

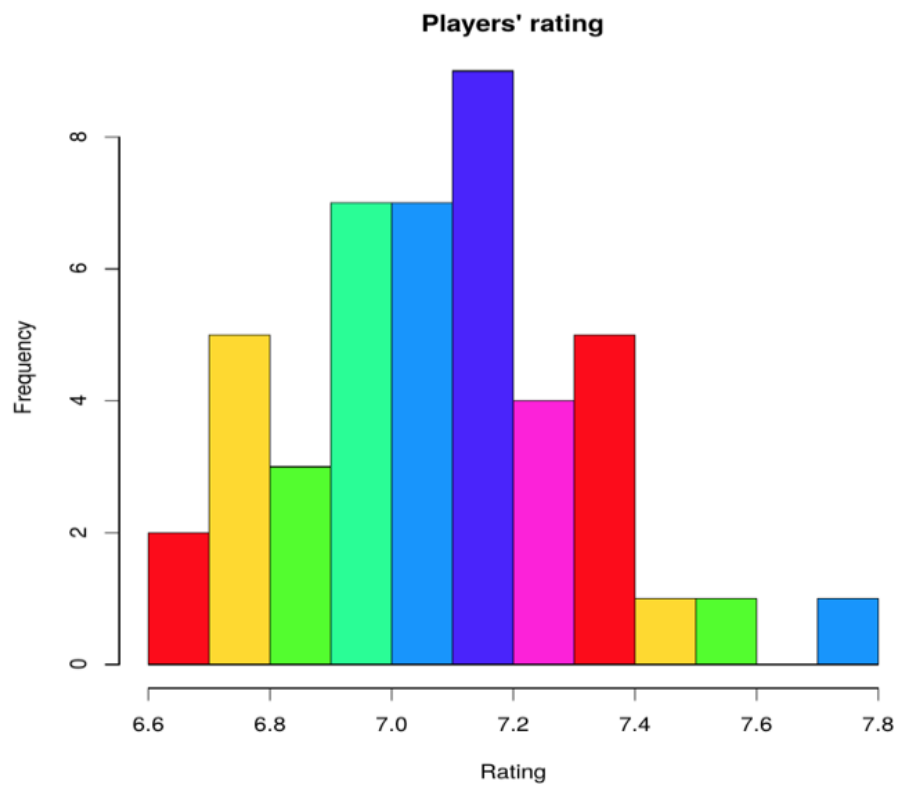
4. Изследване на представянето на терена

Ето разпределението по средно представяне на терена на играчите



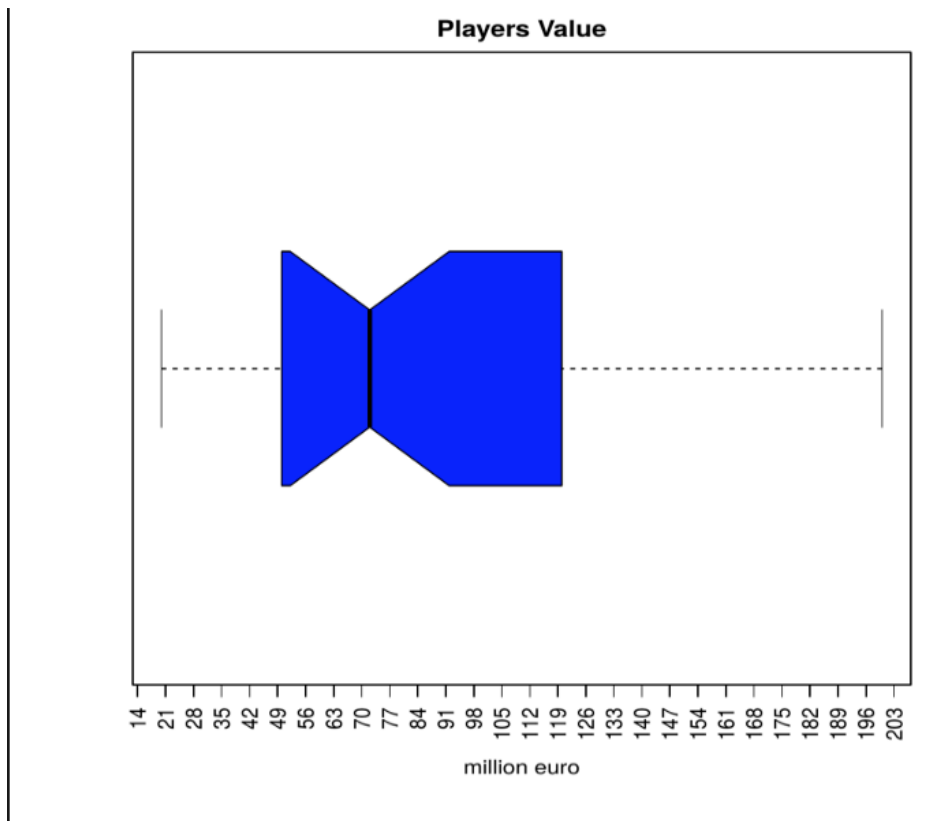
Тъй като говорим за сравнително популярни играчи, очаквано средните им рейтинги са близки, като пак се наблюдават някои outliers.

Ето и хистограма за броя на играчи в различните интервали:



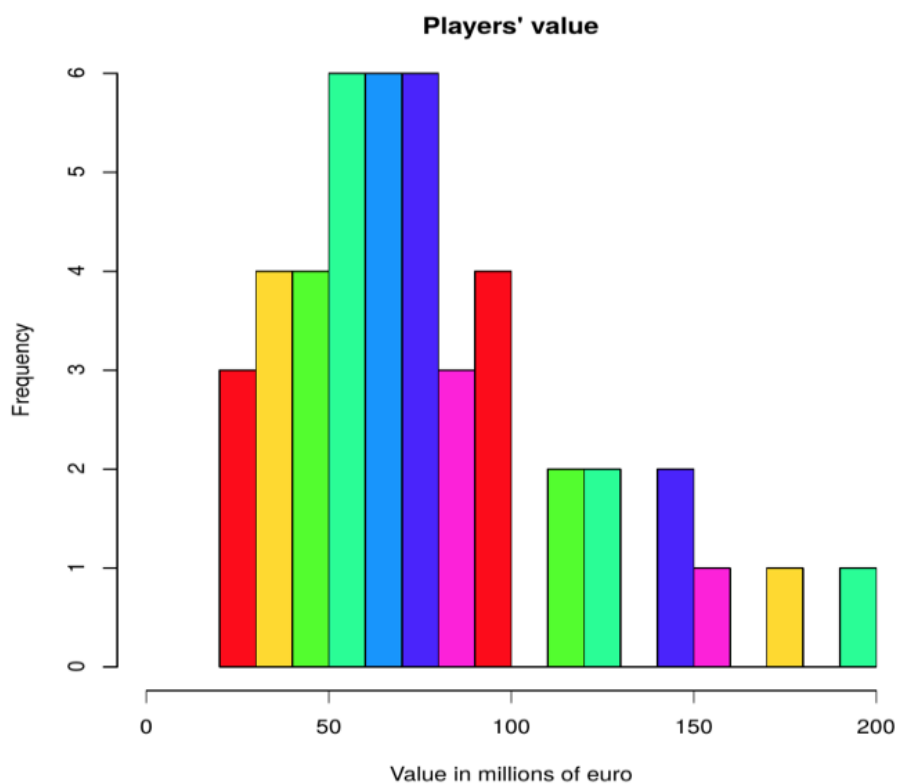
5. Изследване на пазарната цена на играча

Ето разпределението на цената на играчите



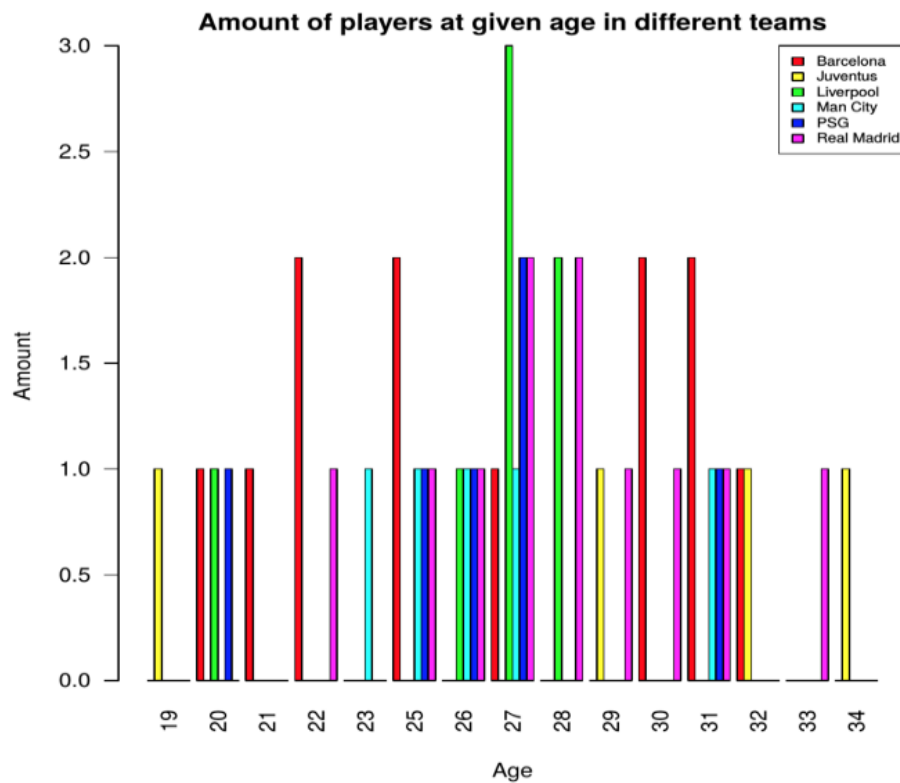
Тук данните са доста по-разхвърляни, но пак има леко струпване около 70 милиона евро.

Ето броя на играчи в различните интервали по цена:



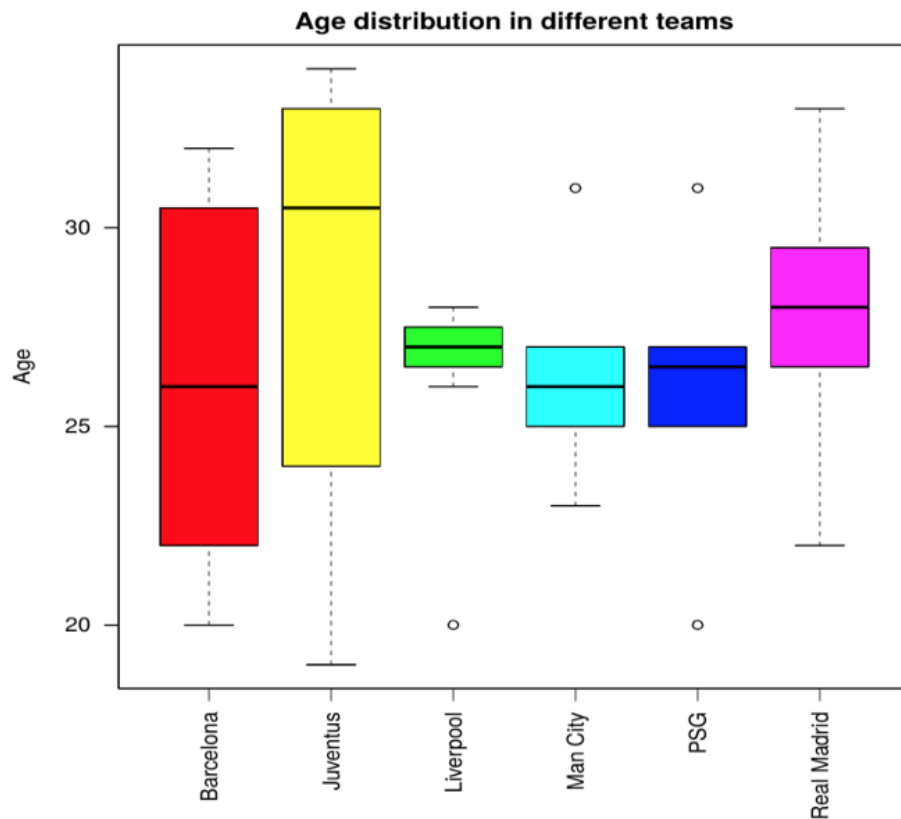
6. Изследване на връзката между отбори и възраст

Ето разпредлението на отбори спрямо възрастта на играчите



Както може да се види всички отбори в извадката имат играчи на 27 години, а Ювентус имат най-младият и най-старият играч в извадката (Матиас де Лихт - 19 години и Кристиано Роналдо - 34 години)

Ето разпределението на години вътрешно за самите отбори:

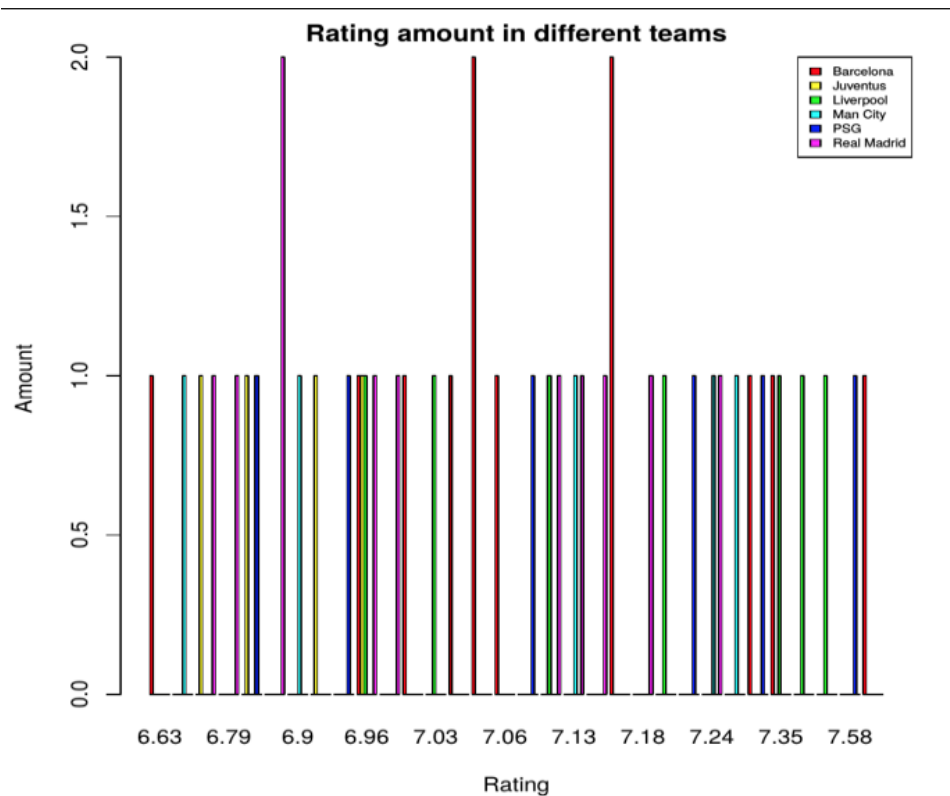


Както може да се види в Ювентус има най-голямо разнообразие от възрасти, което се дължи на факта, че от подбраните данни има много голямо разнообразие.

Барселона също притежава голямо разнообразие от възрасти макар, че при тях има много избрани играчи, въпреки че от Реал Мадрид също има много избрани играчи, там разнообразието от възрасти е силно ограничено.

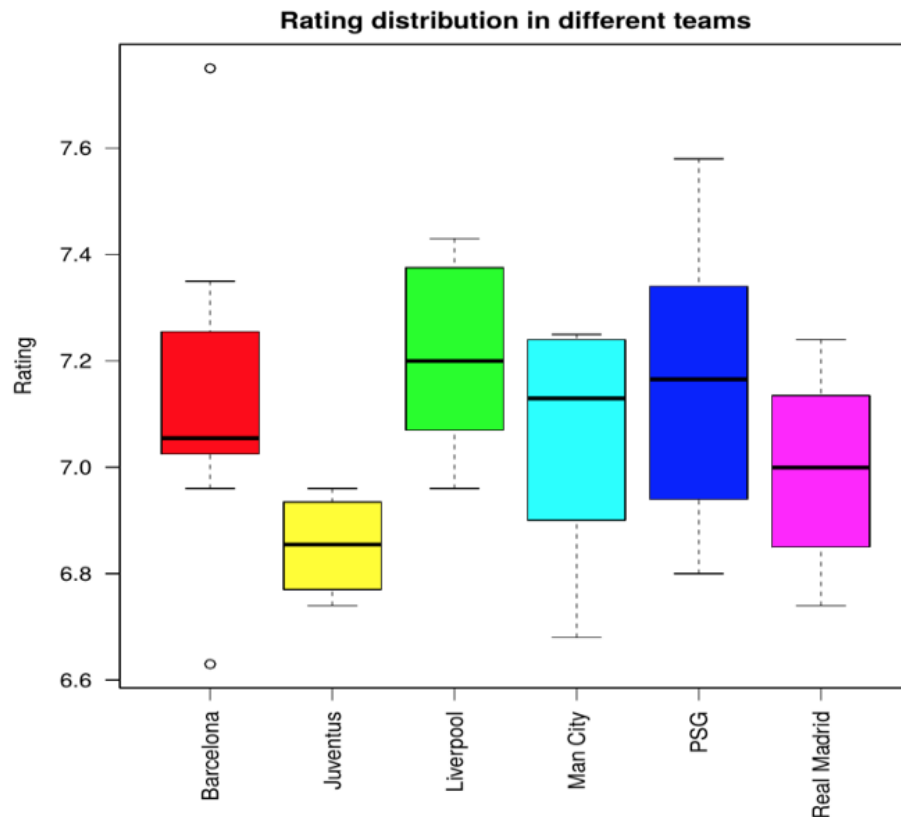
7. Изследване на връзката между отбор и представяне на терена

Ето рейтингите на различните отбори:



Както може да се види, няма особено изпъкващ отбор, който да присъства само в конкретна зона, освен може би Ливърпул, които единствени имат представители в категория 7.43 и 7.50.

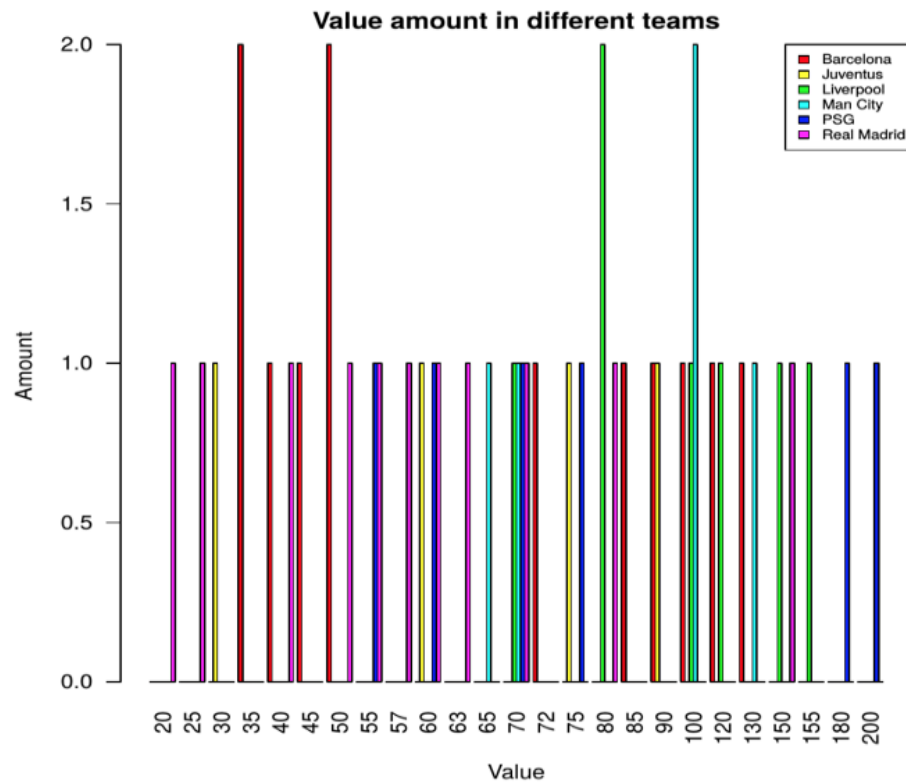
Ето и вътрешното разпределение на възрастта за различните отбори:



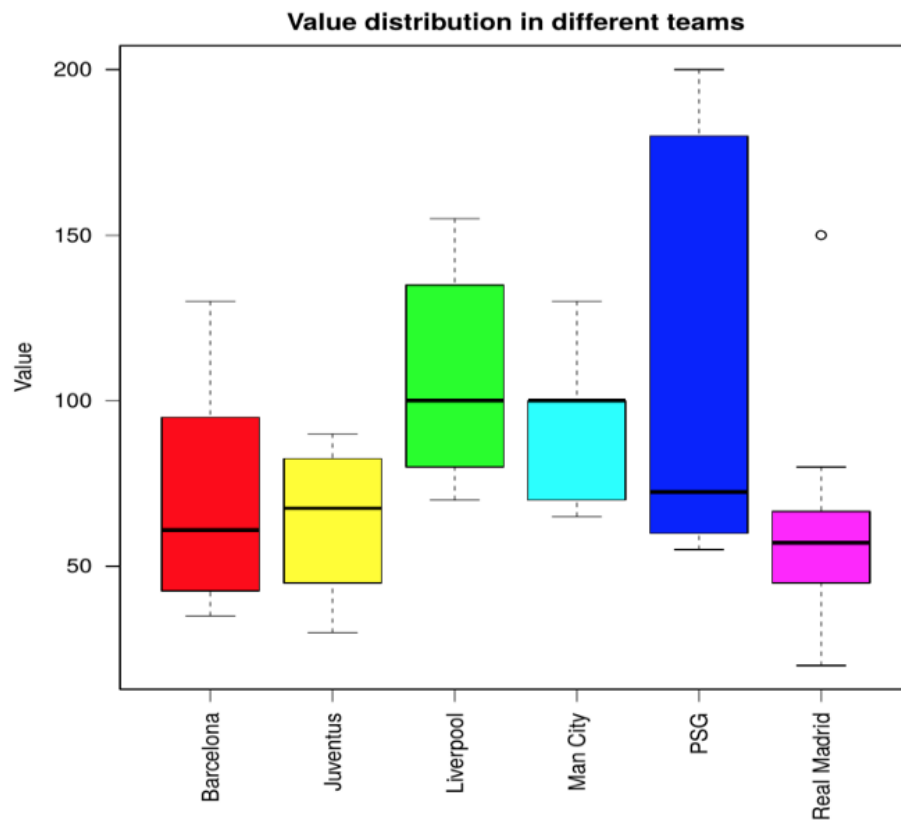
Както може да се види най-сходните като рейтинг са Ювентус, а с най-голямо разнообразие - ПСЖ. Също така в Барселона може да се забележи, че долната половина е много сходна, което означава, че нямат някой играч, който драстично да ги дърпа надолу.

8. Изследване на връзката между цена и отбор

Ето как са разпределени играчите на различните отбори по цени:



Ето разпределнието на играчите по цени вътрешно в различните отбори:

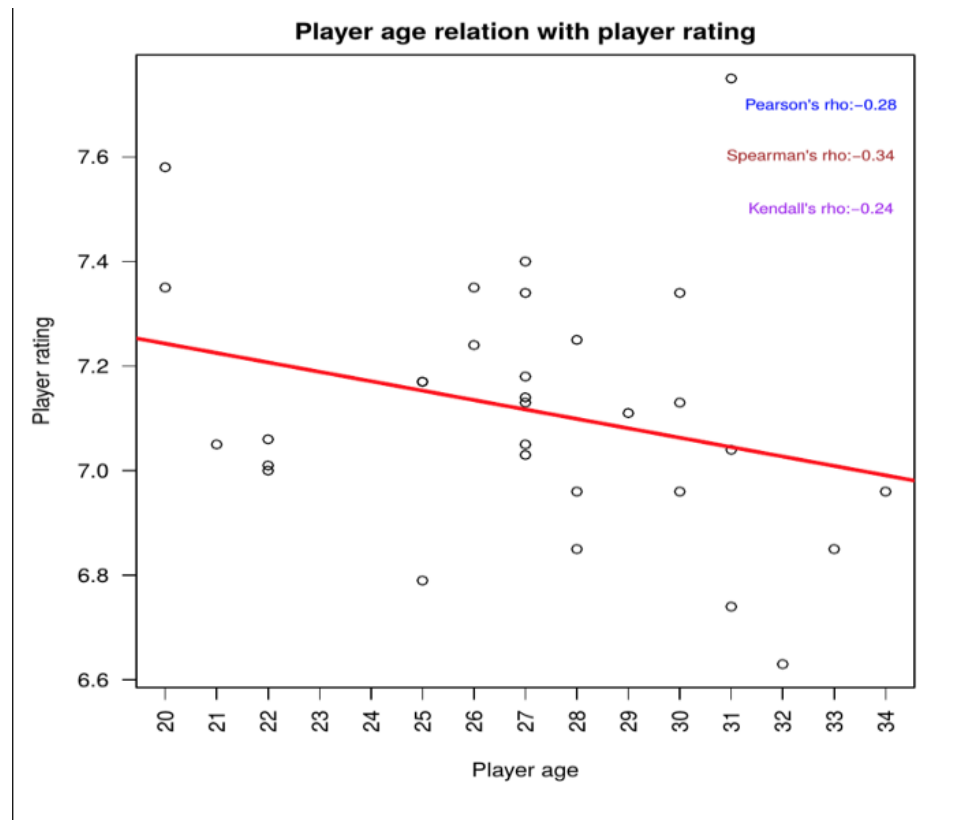


Както може да се забележи, в отбори като Манчестър Сити и ПСЖ, които се смятат за прахосници, има най-голяма разлика в цените на играчите, където горната половина има значително по-висока цена.

В Реал Мадрид отново съществува най-голям баланс между различните играчи.

9. Изследване на връзката между възраст и представяне на терена

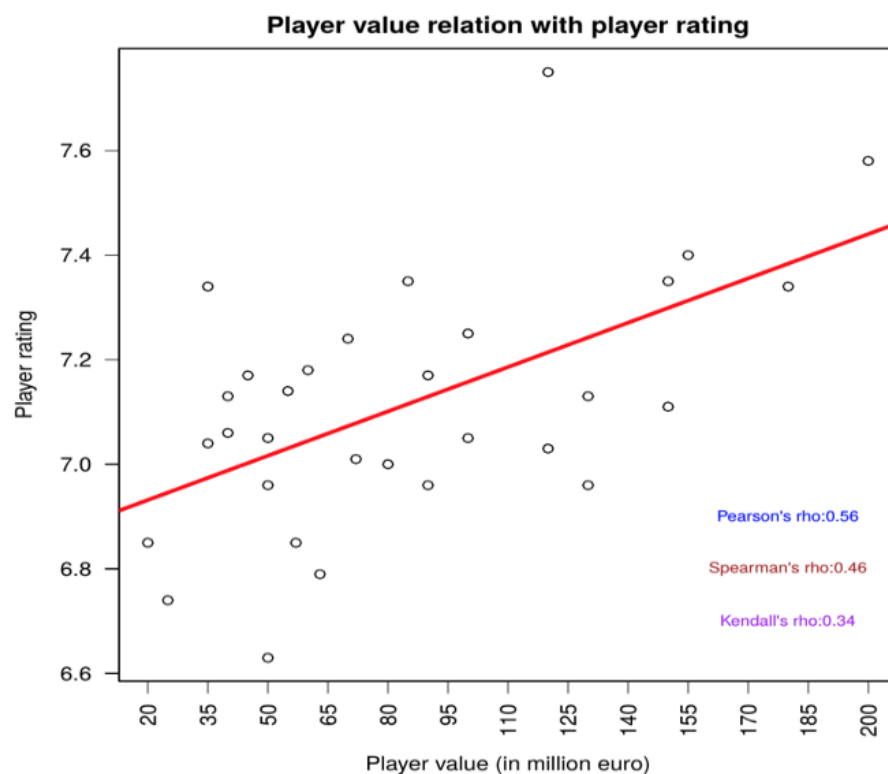
Ето разпределението на възрастта спрямо представянето на терена



Както може да се забележи връзката между двете характеристики е слаба. До известна степен това се дължи на факта, че играчите на възраст около 27 са в разцвета си и се представят най-добре.

10. Изследване на връзката между цена и представяне на терена

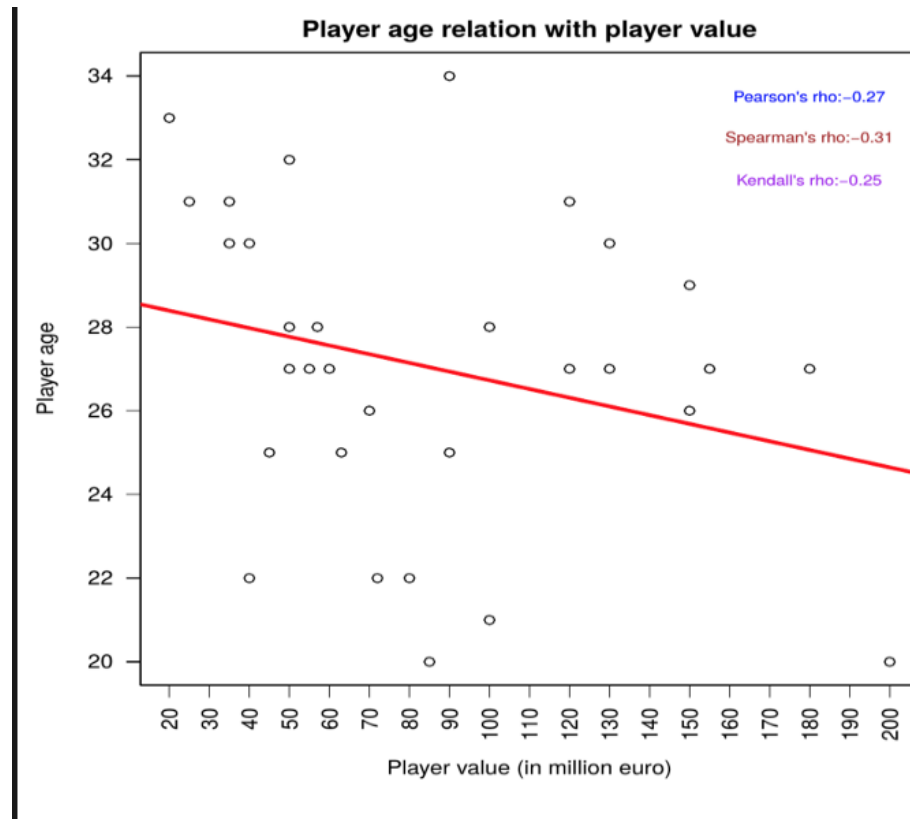
Ето разпределението на цената спрямо представянето на терена



Тук чрез метода на Пиърс излиза, че има връзка между представянето на терена и цената на играча. Тъй като има outliers обаче тези данни са некоректни. Когато се приложат методите на Спийрман и Кендъл си проличава, че реално няма връзка между двете характеристики.

11. Изследване на връзката между цена и възраст

Ето разпределението на цената спрямо представянето на терена:



И тук както в точка 5 се получава липса на линейна корелация, защото най-скъпи са футболистите, които вече са се наложили в съставите си.

12. Заключение

В съвременния футбол има множество фактори, които влияят върху трансферната стойност на един играч, но не може да се определи някаква пряка зависимост.

Това се дължи основно на факта, че се правят трансфери не на играчи, а на имена. Например, трансферът на Кристиано Роналдо в Ювентус беше изцяло провокиран от финансови подбуди. За един сезон клубът изби парите от трансфер чрез продажба на тениски и други фенски артикули, а вече са на голяма печалба от това.

Този трансфер изглежда невероятен погледнат през настоящото изследване, защото Роналдо е стар, а представянето му е под очакваното ниво, но въпреки това е факт и е печеливш.

С това проучване идеята ми, че трансферите във футбола вече не са повлияни от самия футбол се оказва правилно.