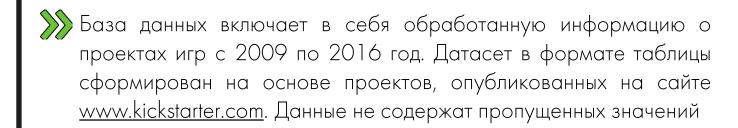


Проекты игр с Кикстартера



Описание базы данных



Данные были собраны с сайта www.kickstarter.com с помощью web-scrapping методов и обработаны для преобразования в машиночитаемый формат. Каждая строчка в таблице соответствует уникальному проекту (игре). Всего в датасете 8421 проект и для каждого проекта в таблице включено 37 характеристик (столбцов)





Поля таблицы 1



Финансовые переменные (указаны в долларах):

- goal сколько проект планировал собрать средств
- pled сколько средств проект собрал в итоге .
- min_pled_t, mean_pled_t, max_pled_t информация о величине пожертвований от пользователей
- n_pled_t, step_pled_t количество опций пожертвований и средняя разница между ними



Переменные, связанные с оформлением страницы проекта на сайте:

- text_am количество слов в на странице
- n_img, n_vid количество картинок и видео на странице
- mgp суммарное количество пикселей у фотографий на странице (в миллионах)
- cont средняя контрастность изображений на странице
 (EBImage, среднее среди стандартных отклонений по трем каналам)



Бинарные переменные, показывающие наличие тега в описании игры, символизирующие жанровую принадлежность (прокси для жанра игры):

- rgp, platformer, shooter, fighting, survival, horror, strategy, arcade, simulator, mmo, indie, action, quest, adventure,
- Уточнение: некоторые проекты имеют несколько тегов (до 10), а некоторые игры не имеют ни одного тега





Поля таблицы 2



Другие переменные

- URL ссылка на страницу проекта на Kickstarter
- date дата опубликования страницы проекта на Kickstarter
- cr_time время со дня регистрации аккаунта автора проекта (прокси опыта автора)
- backed количество проектов профинансированных автором проекта на текущий период
- created количество проектов профинансированных автором проекта до опубликования этого проекта
- site ссылка на сайт проекта (если у проекта есть сайт, если сайта нет, то значение пустое)
- is_site бинарная переменная, показывающая наличие сайта у проекта
- curr валюта в которой производится сбор средств
- period продолжительность онлайн компании (сбора средств), в днях
- success бинарная переменная, показывающая, собрала ли компания нужное количество средств (или больше) или нет

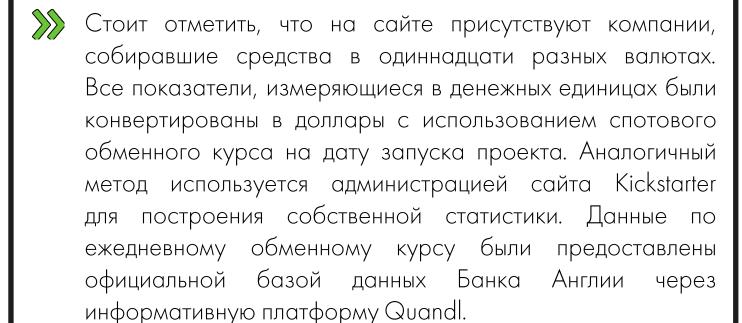




Дополнительные детали по базе



При прогнозировании успешности проекта стоит учитывать только информацию, доступную в момент опубликования проекта. Иначе прогнозирование теряет смысл.







Примеры исследований



Сколько слов в описании проекта необходимо для максимальной успешности проекта?



Есть ли то количество видео, после которого публикация новых видеороликов не влияет на успешность проекта?



Проявление сезонности успешности проекта: влияет ли дата публикации проекта на его успешность?



Влияние тегов в описании на успешность тегов (влияние каждого тега по отдельности и/или влияние количества тегов)



Влияние успешности проекта на активность автора проекта в дальнейшем

