

Исследование данных сервиса “Яндекс.Музыка” — сравнение пользователей двух городов

Статус проекта: *завершен*

Использованные библиотеки: pandas

Описание проекта:

Сравнение Москвы и Санкт-Петербурга окружено мифами.

Например:

- Москва — мегаполис, подчинённый жёсткому ритму рабочей недели;
- Санкт-Петербург — культурная столица, со своими вкусами.

На реальных данных Яндекс.Музыки с помощью библиотеки Pandas и её возможностей проверим данные и сравним поведение и предпочтения пользователей двух столиц — Москвы и Санкт-Петербурга.

Цель исследования — проверить три гипотезы:

1. Активность пользователей зависит от дня недели. Причём в Москве и Санкт-Петербурге это проявляется по-разному.
2. В понедельник утром в Москве преобладают одни жанры, а в Санкт-Петербурге — другие. Так же и вечером пятницы преобладают разные жанры — в зависимости от города.
3. Москва и Санкт-Петербург предпочитают разные жанры музыки. В Москве чаще слушают поп-музыку, в Санкт-Петербурге — русский рэп.

Ход исследования

Данные о поведении пользователей мы получили из файла `yandex_music_project.csv`. О качестве данных ничего не известно.

Поэтому перед проверкой гипотез понадобится обзор данных.

Мы проверим данные на ошибки и оценим их влияние на исследование. Затем, на этапе предобработки поищем возможность исправить самые критичные ошибки данных.

Таким образом, исследование пройдёт в три этапа:

1. Обзор данных.
2. Предобработка данных.
3. Проверка гипотез.

Вывод по проекту:

День недели по-разному влияет на активность пользователей в Москве и Санкт-Петербурге.

Музыкальные предпочтения не сильно меняются в течение недели — будь то Москва или Санкт-Петербург.

Во вкусах пользователей Москвы и Санкт-Петербурга больше общего чем различий.