

Data analysis of Yandex.Music service — comparison of users of two cities

Project status: *completed*

Libraries used: pandas

Description of the project:

Comparison of Moscow and St. Petersburg is surrounded by myths.

For example:

- Moscow is a metropolis subject to the rigid rhythm of the working week;
- St. Petersburg is a cultural capital, with its own tastes.

Using real Yandex.Music data, using the Pandas library and its capabilities, we will check the data and compare the behavior and preferences of users in the two capitals - Moscow and St. Petersburg.

The purpose of the study is to test three hypotheses:

1. User activity depends on the day of the week. Moreover, in Moscow and St. Petersburg, this manifests itself in different ways.
2. On a Monday morning, some genres dominate in Moscow, while others dominate in St. Petersburg. Similarly, Friday evenings are dominated by different genres, depending on the city.
3. Moscow and St. Petersburg prefer different genres of music. In Moscow, they listen to pop music more often, in St. Petersburg - Russian rap.

Research Progress

We got data on user behavior from the `yandex_music_project.csv` file. Nothing is known about the quality of the data. Therefore, before testing hypotheses, a review of the data is needed.

We will check the data for errors and assess their impact on the study. Then, at the preprocessing stage, we will look for an opportunity to correct the most critical data errors.

Thus, the study will take place in three stages:

1. Data review.
2. Data preprocessing.
3. Hypothesis testing.

Conclusion on the project:

The day of the week affects the activity of users in Moscow and St. Petersburg differently.

Musical preferences do not change much during the week - be it Moscow or St. Petersburg.

The tastes of Moscow and St. Petersburg users have more in common than differences.