Заголовок:

Анализ популярности станций метро с помощью Google Trends

Подготовил: Павел Артюшков

Цель:

Этим летом, в Москве проходил чемпионат Мира по футболу. В Метро было много людей, и тогда я задумался, какие станции самые популярные среди туристов и жителей города.

Факторы популярности:

А что влияет на популярность? Может быть архитектура, дизайн, отдалённость от центра, достопримечательности или что-то ещё.

Первым этапом мы решили не собирать данные вручную, а воспользовалться открытым источником **Google Trends.** Этот источник выдаёт рейтинг запросов пользователей на любую тему.

Инструменты:

Для работы с этим сервисом была найдена библиотека PyTrends на языке Python. Этот язык позволяет быстрее реализовывать новые алгоритмы.

Для визуализации полученных результатов была выбрана библиотека Qt, т.к. Tkinter не позволял работать с необходимыми нам функциями. С сайта MosMetro мы загрузили карту ввиде картинки, вручную определили кординаты всех станций, и сохранили их в CSV файл.

Популярность стаций:

Для получения данных по всем станциям (257) требуется примерно 5 минут.

Первые результаты показали нам популярность обычных слов. И первое место заняла станция под названием **Сокол**, т.к. это ещё и птица :) С этого момента мы стали использовать дополнительный фильтр для запроса.

Полученные данные мы визуализировали на карте Метро и провели анализ в **Excel**. Дополнительная информация отображается в всплывающей подсказке.

В итоге мы получили список самых популярных станций по версии **Google Trends**.  **+**Первое место (83.1)

Планы на будушее:

В дальнейшм мы планируем:

Сделать просмотр популярности по годам c помощью ползунка и анимации.

Добавить график в всплывающюю подсказку.

Подключить новые источники данных для завершения модели.

Использовать модель для сравнения рассчитанной популярности с посещаемостью.

Так же можно применить машинное обучение для подбора коэффицентов полученной модели (для того чтобы понять что важнее дизайн или отдалённость от центра и тд).