



РОССИЙСКИЙ УНИВЕРСИТЕТ  
ТРАНСПОРТА  
РУТ (МИИТ)

Высшая  
инженерная  
школа

# **Аналитика пассажиропотоков Московского метрополитена:** оценка распределения пассажиров по сети Московского метрополитена, заходящих на станции Киевская в утренний час-пик

Состав проектной команды:

Глава проекта: Васильев Даниель - написание кода

Кузов Иван - сопоставление файлов, редактирование

Землянский Павел - работа в Excel, редактирование

Бокова Мария - аналитик данных

Итулин Максим - помощь в написании кода

# Постановка задачи

## ❑ Проблематика вопроса



МЦД

Оценить потенциальный спрос на маршруты Московских центральных диаметров среди пассажиров, которые сегодня совершают поездки на метро со станции Киевская

## ❑ Цель исследования



Определить направления следования пассажиров по станциям метро со станции Киевская за утренний час-пик

8:00-9:00 по валидациям билетов на вход.

## ❑ Функциональный заказчик проекта



Центр цифровых высокоскоростных  
транспортных систем РУТ (МИИТ)

## ❑ Исходные данные



Массив BigData (выгрузка валидаций билетов ГУП «Московский метрополитен»). Данные по маршрутам движения и конечным станциям назначения пассажиров в массиве отсутствуют.

**Формирование матрицы корреспонденции на основе следующих исходных данных:**

- станции, вестибюль отправления
- время валидации пассажиров
- уникальные номера билетов

# Алгоритм решения задачи

1

Чтение  
датафрейма

Выборка необходимых  
полей исходной  
таблицы для анализа

2

Определение  
станций  
назначения

Допускаем, что  
станцией назначения  
является следующая по  
времени входа в  
метрополитен

3

Фильтр  
датафрейма

Фильтрация по станции  
и времени отправления  
(с 8:00 до 9:00)

4

Формирование  
матрицы  
корреспонденции

Вывод данных в  
удобном для заказчика  
формате

5

Визуализация  
данных на карте

Отображение данных  
на интерактивной карте

# Фрагменты разработанного кода

```
name: STATION, dtype: object

In [48]: result_02 = result_01.loc[(result_01.STATION == 'Петр.-Разумовск.')] | (result_01.STATION == 'Петровско-Разум.')]

In [49]: result_02

Out[49]:
```

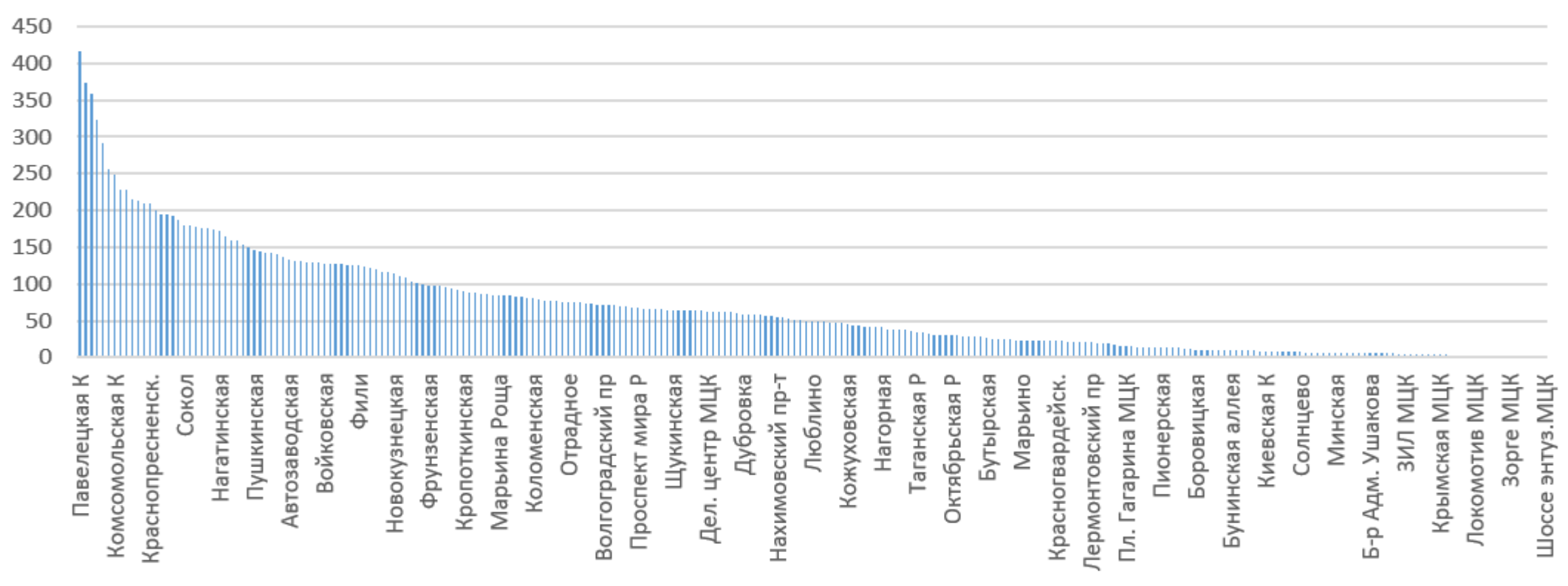
	dttm	ticket_id	STATION	STATION_out
20	2021-05-01 11:01:46	26536053	Петровско-Разум.	Б.Рокоссовского
20	2021-05-03 15:14:55	33637941	Петровско-Разум.	Б.Рокоссовского
20	2021-05-03 15:15:10	8730103	Петровско-Разум.	Б.Рокоссовского
20	2021-05-03 15:28:41	22866482	Петровско-Разум.	Б.Рокоссовского
20	2021-05-04 06:31:56	17254392	Петровско-Разум.	Б.Рокоссовского
...	...	...	...	...
2420	2021-05-23 19:00:57	1360041719	Петровско-Разум.	Мнёвники
2420	2021-05-26 07:14:35	1360041719	Петровско-Разум.	Мнёвники
2420	2021-05-27 19:04:40	1360041719	Петровско-Разум.	Мнёвники
2420	2021-05-30 07:14:28	1360041719	Петровско-Разум.	Мнёвники
2420	2021-05-31 18:42:25	1360041719	Петровско-Разум.	Мнёвники

	dttm	vestibule_id	ticket_id
67929139	2021-05-15 21:57:37	310	7
69797482	2021-05-16 13:37:02	2183	7
71344086	2021-05-16 19:39:17	366	7
72158202	2021-05-17 00:21:42	2183	7
73318909	2021-05-17 08:21:35	2230	7
...	...	...	...
13504565	2021-05-04 22:08:06	2238	4294965448
13615638	2021-05-04 22:59:07	47	4294965448
13571064	2021-05-04 23:33:43	47	4294965448
63120447	2021-05-14 20:08:46	2238	4294965448
66030185	2021-05-15 14:54:48	2043	4294965448

# Результаты анализа Big Data:

потенциальный объем переключаемого пасс. потока со станции Киевская – 19 300 пасс./час пик

❑ Гистограмма распределения пассажиров метрополитена со ст. Петровско-Разумовская в утренний пик (с 8:00 до 9:00), пасс./час



- ❑ Потенциальный объем переключаемого пасс. потока со станции Петровско-Разумовская – **19 300 пасс./час пик**
- ❑ Обработанный объем Больших данных составляет **25 gb**, вмещает в себя **более 355 млн строк**.
- ❑ **Для решения задачи** был написан **код** и визуализирован в виде **тепловой карты**

