

# Кранков Семен

@awpxx | +7-925-001-96-69 | krankov21@gmail.com | [github.com/adminov21](https://github.com/adminov21)

## О себе

Студент МИФИ, начинающий data-аналитик. Ищу стажировку в области анализа данных или data-инжиниринга. Имею практический опыт построения ETL пайплайнов, SQL анализа и визуализации данных на реальных датасетах.

## Образование

### НИЯУ МИФИ 2023 - 2028

Институт финансовых технологий и экономической безопасности

*Безопасность информационных технологий в правоохранительной сфере*

## Проекты

### E-commerce analytics (<https://github.com/adminov21/ecommerce-analytics>)

Стек: Python, PostgreSQL, ClickHouse, Apache Airflow, Superset, Docker

- Построил сквозной ETL пайpline для обработки 550 тыс. строк данных бразильского интернет-магазина Olist
- Реализовал RFM сегментацию 96 тыс. покупателей, выявил что 45% клиентов не возвращаются после первой покупки
- Автоматизировал обновление витрин данных через Apache Airflow с DAG для staging и OLAP слоев
- Создал интерактивные дашборды в Apache Superset с метриками выручки, топ категорий и клиентской аналитикой
- Развернул все сервисы в Docker (PostgreSQL, ClickHouse, Airflow, Jupyter, Superset)

### Анализ оттока клиентов банка ([https://github.com/adminov21/Churn\\_Modelling](https://github.com/adminov21/Churn_Modelling))

Стек: Python (Pandas, Matplotlib, Seaborn), SQL (PostgreSQL / DuckDB)

- Провел разведочный анализ датасета 10 тыс. клиентов банка: очистка данных, поиск аномалий, агрегация по географическому и демографическому признакам через SQL внутри Python среды
- Выявил критические сегменты оттока: клиенты с 3 продуктами уходят в 83% случаев, с 4 продуктами - в 100%; уровень оттока в Германии (32%) вдвое выше Франции (16%) и Испании (17%); 62% ушедших клиентов имели средний кредитный рейтинг (500-700)
- Обнаружил что основная группа оттока - клиенты старше 40 лет: в Германии их доля составляет 22% от всех клиентов
- Применил три метода корреляционного анализа - Пирсон выявил слабую связь между количеством продуктов и оттоком, тогда как Спирмен и Кендалл зафиксировали нелинейную обратную зависимость, что указывает на пороговый эффект при использовании 3+ продуктов
- Построил тепловые карты корреляций и стековые диаграммы структуры оттока по странам и возрасту

## Технические навыки

Языки программирования: SQL, Python(Pandas, Numpy, Matplotlib, Seaborn)

Инструменты: PostgreSQL, PyCharm, Jupyter Notebook, Superset, Excel, ClickHouse, Airflow, Docker

Иностранные языки: английский (B2)