

+374 55 813358/+7 999 2011679
Yerevan, Armenia
georgemaschmidt@gmail.com

Egor Shmidt

Data Analyst

<https://t.me/idcardforfood>
<https://github.com/Ramrion>
www.linkedin.com/in/ramrion

SKILLS

Programming Languages	Python, R
Technologies & Frameworks	Pandas, Numpy, Scikit-learn, Catboost, ggplot, Shiny dashboards, Data parsing, Markdown reports, SQL, MS SQL, PostgreSQL, Git, Docker, Bash
Knowledge	Data Processing, Machine Learning (Regression, Classification, Clustering, NLP), Data visualization, A/B Testing, Qualitative methods (Interview, Polls)

TECHNICAL EXPERIENCE

Data Analyst **Sep 2021 — Present**
Aroma Market - 600+ stores retail company *Remote*

- **Developed ETL pipelines.** I developed pipelines for data updates and data enrichment for the marketing department, which reduced manual work on daily reports and helped to coordinate work with email notifications.
- **Automated the matching of goods with different names.** We receive data about competitors who record the same goods we have in different ways. I made an algorithm for goods matching, taking into account difference in language and typos. Due to this, we were able to more accurately track the turnover, as well as enrich the data with the characteristics of the goods.
- **Developed a model for optimal store location.** I designed a catboost model to predict revenue for a specified location based on geo features like car traffic and distance to closest grocery store. The model reduced error on 30% relative to previous contractor's model.

Data Analyst / Project: Mapping Russian media discourse **Oct 2020 — Mar 2021**
National Research University Higher School of Economics *Saint-Petersburg, Russia*

- Clustered 200.000 news texts from 62 Russian media outlets to identify the agenda profiles of the outlets.
- Using permutation tests, identified vocabulary characteristic of the contrasting groups of media.
- Developed a method to measure short-term fluctuations of the news discourse.

Programmer / Project: Science studies based on bibliometrics data **Nov 2019 — Jun 2020**
National Research University Higher School of Economics *Saint-Petersburg, Russia*

- Combined data from multiple APIs and XML pages to improve inference of gender based on a person's name.

Programmer / Internship **Feb 2019 — Mar 2019**
Institute of Sociology, Russian Academy of Science *Saint-Petersburg, Russia*

- Statistically analysed temporal and geographical patterns in the topics of the USSR scholars' theses.

EDUCATION

Bachelor's degree in Sociology and Social Informatics **2017 — 2021**
National Research University Higher School of Economics
Thesis: Regional and socioeconomic factors of unemployment exit in Russia

ACTIVITIES

Hackathon: Pandemic Data Hack **2020**
Analyzed who finds job in public employment services using MCA and multilevel regression model

Summer School: Data Analysis for Social Sciences and Humanities **2019**
Student in the advanced track (Bayesian statistics & Mixed effect models)

PUBLICATIONS

Moroz, A., Loseva, A., **Shmidt, E.**, and Alexandrov, D. (2022). Mapping the Russian Media Field through Audience Networks and Agenda Choice. Manuscript revised for the resubmission in the *International Journal of Communication* (SJR Q1 Communication)