MARIYA PETUKHOVA

DATA SCIENTIST

Naperville, IL | mariyapetukhova13@gmail.com | 630-890-2766 | [/in/mariya-petukhova](#) | [mpetuk](#)

Summary

Senior analytics consultant with nearly 20 years translating business problems into analytical solutions for healthcare and other industries. Strong technical skills in SQL, Python, R, and dashboard development (R Shiny, Tableau) combined with healthcare domain expertise including HEDIS/NCQA certification. Experienced supporting clinical and business stakeholders through the full analytics lifecycle: understanding requirements, querying diverse data sources, performing statistical analysis, and presenting insights through compelling visualizations. Track record of leading analytic initiatives independently and collaboratively, delivering high-quality output under tight timelines. Currently at HCSC providing analytics support and consultancy across Medicare, Commercial, and RFP domains.

Employment

Health Care Service Corporation (HCSC)

Sr Analytics & Reporting Consultant | *Chicago, IL (Remote)* | *March 2024 – Present*

- Built R Shiny dashboard for health disparities analysis providing stakeholders easy access to 42 lines of business × 29 care gap measures; recognized by clinical and business partners for enabling data-driven decision making on quality initiatives
- Provide NCQA analytical support and consultancy to internal stakeholders across Medicare, Commercial HMO/PPO, and Marketplace lines of business; completed HEDIS 2023 certification to strengthen healthcare quality measurement expertise
- Query diverse internal and external healthcare data sources using SQL, R, and Python to extract meaningful insights; present findings through Excel reports and interactive visualizations that enable senior leadership decision-making
- Independently delivered 2x report volume year-over-year while maintaining <2 week turnaround and 100% accuracy; earned multiple Stellar Awards from business partners for responsiveness, innovation, and quality
- Lead technical initiatives including Git migration project for production codebase modernization; champion collaborative development practices and version control adoption across analytics team
- Work independently and collaboratively across cross-functional teams; engage stakeholders to understand business needs, scope analytic initiatives, and deliver solutions under tight timelines with senior leadership visibility
- Automated report generation using OfficeR package (70% manual effort reduction); continuously expand technical capabilities through training (Posit::conf R Shiny workshop) and open-source contributions
- Tech Stack: R, R Shiny, Python, SQL, Excel, Teradata, Azure, Git/GitHub, Confluence, Tableau-ready visualizations, HEDIS measures, healthcare claims data, Agile methodologies

The MirrorStream NFP

President | *Naperville, IL* | *June 2022 – Present*

- Incorporated 501(c)(3) nonprofit to streamline medical supply delivery to Ukraine, providing tax benefits and simplified customs clearance for donors
- Focus on fundraising, sourcing, procurement, and delivery logistics for medical supplies
- Manage website, social media content, and community engagement at local charitable events
- Coordinate vendor relationships and donor communications

Butler/Till via Dataly LLC

Data Science Consultant | *Rochester, NY (Remote)* | *Nov 2022 – Mar 2023*

- Trained and upskilled marketing analytics team in Python and ML model development—translating technical concepts for non-technical audience and enabling team independence
- Built simulation tool for marketing budget optimization using channel spend models, providing leadership with data-driven planning capabilities
- Established quality standards and documentation practices for modeling projects, improving maintainability and knowledge transfer

ShopRunner via SPR Consulting

Data Scientist/ML Engineer | Chicago, IL | Apr 2021 – Jan 2022

- Worked on taxonomy model expansion project
- Troubleshooted infrastructure issues including EC2 launches via Terraform
- Participated in code reviews, PageDuty on-call rotation, and journal club
- Tech Stack: AWS EC2, Databricks, Python, Snowflake, Terraform, image scraping, shell scripts, Git

Aetna, a CVS Health Company

Sr Data Scientist | Hartford, CT (Remote) | May 2020 – Apr 2021

- Designed and executed end-to-end project plan for healthcare data imputation initiative—evaluated multiple data sources (internal, external, licensed), formulated competing methodologies, assessed resource requirements, and presented recommendations to leadership
- Pioneered PySpark implementation for large-scale healthcare data processing, including SMOTE techniques for handling imbalanced datasets
- Tech Stack: Hadoop/Hive/HDFS, DB2, MS SQL Server, Python, PySpark, shell scripts, Git

Vivid Seats

Data Scientist II | Chicago, IL | Apr 2019 – Mar 2020

- Supported pricing forecasting using FB Prophet and H2O AutoML
- Studied customer shopping behaviors and lifetime value relationships
- Created data pipeline from BigQuery to S3 via Google Cloud Storage
- Automated H2O model development and scoring in Python
- Tech Stack: AWS EC2, Python, Snowflake, BigQuery, GCS, shell scripts, Git

Additional Experience:

- **Sogeti**, Westchester, IL – Sr Consultant, Data Science (Jan-Mar 2019): Refactored Python code for 9 data science models (XGBoost, linear regression) in Azure ML
- **DRIVIN**, Chicago, IL – Data Scientist (Oct 2017-Dec 2018): Built dealer segmentation models using agglomerative clustering to drive product and contact strategy; implemented fuzzy string matching for master data consolidation
- **Metis**, Chicago, IL – Data Scientist (Apr-Jun 2017): Completed 12-week intensive data science bootcamp; built portfolio including NLP semantic clustering (15 segments via k-means), financial services customer classification (12 types), and regression models
- **Dataly LLC**, Naperville, IL – Owner (Mar 2017-Present): Data analytics consulting; completed donor segmentation for Hong Kong charity and response/revenue models for transportation company
- **VSA Partners**, Chicago, IL – Director, Research and Analytics (May 2016-Apr 2017): Built market segmentation using unsupervised learning (BayesiaLab) for financial products; customer segmentation for B2B account-based marketing; presented to Fortune 500 board of directors
- **Acxiom Impact**, Conway, AR (Remote) – Expert Analyst/Senior Analyst (Aug 2014-May 2016): Built customer segmentation and response models for retail, insurance, and transportation clients; developed look-alike modeling for prospect targeting
- **Cogensia**, Schaumburg, IL – Manager, Client Solutions/Senior Consultant (Jul 2010-Mar 2014): Built customer segmentation for direct marketing campaigns; developed response/clone models for targeted outreach

Education

Indiana University Purdue University Indianapolis | M.A. Economics | 2002

Kharkiv National University of Economics (Ukraine) | M.S. Economics | 1998

Additional Training: HEDIS 2023 (NCQA), Posit::conf 2024 Advanced R Shiny Workshop, Metis Data Science Bootcamp (2017)

Skills

Programming: Python, R, SQL, SAS, PL/SQL, Shell and DOS scripting, Apache Spark, HTML

Analytics Tools: scikit-learn, NumPy, SciPy, Pandas, NLTK, Statsmodels, BayesiaLab, Aginity, Qubole

Machine Learning: Logistic Regression, Decision Trees, Random Forest, Multilayer Perceptron, k-nearest neighbors, Naive Bayes, DBSCAN, NLP, Neural Networks, k-means and agglomerative clustering, Factorization Machines

DBMS: Oracle, Netezza, PostgreSQL, MongoDB, Snowflake, Teradata, BigQuery

Statistical Techniques: Customer/Market/Behavioral Segmentation (clustering, unsupervised learning), Regression (with or without Regularization), MaxDiff Analysis, Fractional Factorial Design, Uplift Models, Principal Component Analysis, Market Basket Analysis

Data Visualization: Matplotlib, Seaborn, Tableau, Plotly, QlikView, Looker, R Shiny, Leaflet

Cloud/Infrastructure: AWS (EC2, S3), Azure ML, Databricks, Terraform, GCS

Healthcare Domain: HEDIS measures, Healthcare claims data, Medicare/Medicaid analytics

Market Research: Qualtrics, SurveyGizmo, Simmons

Volunteering

Balakun – English Tutor | Remote | 2025 – Present

Volunteer English conversation tutor helping Ukrainians practice spoken English through weekly one-on-one sessions

Project C.U.R.E. – Volunteer | Woodridge, IL | Mar 2022 – Present

Sort, pack, load hospital and other medical supplies