# TEODORA MIHAYLOVA

tmind@protonmail.ch ✉
www.linkedin.com/in/tmihaylova 🔗
+356-9935-7070 📞
https://tonymih.github.io 🐙
EU + CANADA 🌐

## DATA SCIENTIST | CERTIFIED TECH ETHICIST

### MOTIVATION

*I am eager to leverage my background in economics, data science and tech ethics to translate business problems into data science solutions. I am looking to apply my skill set in creative ways to bring value to the team, the business and the end-user.*

### SKILLS & TOOLS

**Programming:** SQL, Python (Base, Pandas, Numpy, Matplotlib, Scikit-Learn, Keras)

**Tools:** Excel, PowerBI, Tableau, Github, QGIS, AWS (S3, Lambda, IAM, EC2, SageMaker, RDS, DynamoDB)

**Math:** Linear Algebra, Statistics (Hypothesis Testing, AB Testing, Central Limit Theorem, Distributions)

**Machine Learning:** Linear Regression, Logistic Regression, Decision Trees, Random Forest, KNN, K-means, PCA, Association Rule Learning, Causal Impact Analysis, Neural Networks

### PROJECTS

**"Segmented city" - k-means borough segmentation**

- Used k-means clustering to split boroughs in accordance with crime data. This allows for tailoring crime prevention programs and optimizing budget allocation.

**"All that glitters isn't gold" - A/B testing to determine campaign efficiency**

- Conducted a Chi-square test to determine if a more costly marketing mailer was associated with significantly higher customer sign up rates in a "Delivery club".
- Higher-cost mailer's increased sign up rate was not statistically significant, hence, a recommendation could be made to save costs on marketing material.

**Quantifying sales uplift with Causal Impact Analysis**

- Analysed if there is a sales uplift in customer's spending after joining a "Delivery club", through causal impact analysis.
- Customers that joined the club spent 41.1 % more, than they would have spent, for the same period, if the club had not been in existence.

**More projects**
**https://tonymih.github.io**

### EXPERIENCE

**Data Scientist - Financial inclusion DataKit, DataKind** (volunteer)

APRIL 2025

- Led a data-driven analysis on a 298-variable financial inclusion factors dataset to identify key predictors of a target metric.
- Built a ML model (ridge regression with cross-validation) to select 14 variables explaining 92% of target variance.
- Validated results with Principal Component Analysis, revealing factor groupings.

# TEODORA MIHAYLOVA

tmind@protonmail.ch
www.linkedin.com/in/tmihaylova
+359-895-061200
https://tonymih.github.io
EU + CANADA

## DATA SCIENTIST | CERTIFIED TECH ETHICIST

---

### EXPERIENCE

**Data Analyst - International Center for the Prevention of Crime**
NOVEMBER 2018 - APRIL 2023

In charge of establishing a data analysis modus operandi
(delegated authority from Ministry of Public Safety, QC, Canada):
- Developed data operations protocols (SOP) and policies
- Procured, cleaned and analyzed data from external sources, including geospatial data
- Identified trends, patterns, and insights
- Prepared reports, dashboards and data visualizations
- Performed data validation and quality assurance in view of reliability of analysis results
- Ensured compliance and security in accordance data provider's requirements
- Implemented ethics and privacy safeguards within data projects
- Served as a jury of a Data Jam for tackling Human trafficking (mandate by the UNODC)

---

### EDUCATION

**Microprogram in data mining**
Université de Montréal / HÉC, Montreal, Canada

**Master's in Public Policy and Public Administration**
Concordia University, Montreal, Canada

**Bachelor's in International Economic Relations**
University of Economics, Varna, Bulgaria

---

### COURSES & CERTS

**Data Science Professional Certification (Data Science Infinity), 2025**
Actionable Learnings: Mining data through SQL. Application of statistical concepts such as hypothesis tests for measuring the effect of AB Tests. Utilising Github for version control, and collaboration. Using Python for data analysis, manipulation & visualisation. Applying data preparation steps for ML including missing values, categorical variable encoding, outliers, feature scaling, feature selection & model validation. Applying Machine Learning algorithms for regression, classification, clustering, association rule learning, and causal impact analysis for measuring the impact of an event over time. ML pipelines to streamline the ML pre-processing & modelling phase. Deployment of a ML pipeline onto a live website using Streamlit. Using Tableau to create powerful Data Visualizations.

**Certified Ethical Emerging Technologist (CertNexus), 2023**
Actionable Learnings: the specialization focused on the skills necessary to lead with ethical integrity in any organization focused on the development, maintenance, and governance of data-driven technologies.

**Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans, Course on Research Ethics (TCPS 2: CORE), Government of Canada, 2020**
Actionable Learnings: steps to undertake to ensure ethical treatment of human participants in research projects, regardless of discipline or methodology.