

PETR BARANOVSKIY

www.dataenthusiast.ca | gitlab.com/peterbar | Contacts: [please click here](#)

CORE SKILLS

Statistics and Research

- Inferential statistics: linear regression, dependent and independent t-tests, time series, logistic regression, robust statistical methods, etc. Applying non-parametric tests to deal with non-standard situations such as small samples, Likert scale data, and categorical data.
- Conducting surveys: choosing survey methodology, creating questionnaires, gathering data, coding and processing survey data, and writing up the conclusions. Online tools: LimeSurvey, or creating custom-made survey apps with *Shiny* if data privacy is a major issue.
- Qualitative research: focus groups and research interviews, literature reviews. NVivo.

Data Mining, Cleaning and Processing

- Reading all known tabular formats into R, Web scraping through Web APIs, JSONs, and *rvest*; text mining; mining Twitter and Google Search Trends, automated data retrieval from Statistics Canada.
- Data manipulation in R with *tidyverse* tools: converting datasets into the tidy format, subsetting, joining, and splitting datasets, data type conversions, pivoting between wide and long formats, etc.
- Cleaning and processing various data types in R: text data with *stringr* and *stringi*, date-time data with *lubridate*, and categorical data with *forcats*.

Data Visualization and Reporting

- Making fully interactive data dashboards and visualizations with *Shiny*, *plotly*, and *flexdashboard*. Creating reproducible and self-updating reports with *R Markdown*. Deploying *Shiny* apps.
- Making publication-quality plots of any complexity with *ggplot2* and other graphic R packages.
- Creating interactive maps for the web with *Leaflet* and *plotly*. Making thematic maps for reports and publication with *tmap*, *ggmap*, and *OpenStreetMap* R packages.
- Presenting findings to public servants, industry representatives, and other non-specialist audiences.
- Extensive experience writing for academic, government, and business audiences.

SQL and Databases

- PostgreSQL and MariaDB/MySQL: setting up and administrating databases, integrating with R for data processing using *DBI*, *RPostgres*, *RMariaDB*, and *dbplyr*; connection management with *pool*; creating web UIs for databases with *Shiny*. Experience in integrating big databases with R.
- Writing SQL queries to read from and write to databases, alter and update tables. Using parameterized queries to work with databases from inside R (including dynamically updating parameters with *Shiny*).

GIS and Spatial Data

- Processing spatial data with *sf*, *sp*, and *raster*.
- Spatial data analysis: computing centroids, buffers, hulls, areas, bounding boxes, distances between features, spatial feature densities, determining relationships (e.g. intersections) between layers, etc.
- Spatial statistics: Poisson models and quadrat tests with *spatstat*, spatial cluster detection.

Other Skills

- R package development: author and maintainer of [ggwebthemes](#) R package (*ggplot2* themes for blogs).
- Linux, including Ubuntu/Debian server set up and administration. Version control with Git.
- Web development: dynamic websites with WordPress, static web pages with Hugo and *blogdown*.
- MS Excel, MS Word, MS PowerPoint, other MS Office and LibreOffice apps; SPSS.

EXPERIENCE

R Programmer and Data Analyst | Freelance & casual projects | September 2020 – present

- **For VGeo.ca:** Developing a UI app for PostgreSQL database with R *Shiny* in order to enable remote reading from, writing to, and editing of a large-scale relational database using a simple and intuitive web interface. Development server system administration, writing and maintaining system and user-specific documentation, maintaining relevant GitHub repositories.
- **For a private client in Saskatoon:** Developing, testing, and training several alternative time-series price forecasting models for Bitcoin, Ethereum, and Litecoin cryptocurrencies. Data mining through CoinDesk, coinmarketcap.com, and bitbuy.ca APIs.

Research Officer | University of Saskatchewan | January 2018 – December 2019

- Data analysis for a comprehensive overview of the local economic and social indicators of northern Saskatchewan communities, including data mining, cleaning and transforming datasets, building descriptive and predictive models, and data visualization.
- Provided policy advice to local governments based on our findings – as reports and presentations.
- Automated various tasks. For example, we routinely worked with Statistics Canada data, but previously the datasets used to be downloaded manually from the web, and then processed and visualized in MS Excel. I automated data retrieval and analysis using R, so that no relevant data got overlooked, while ensuring time-efficiency and one-click reproducibility.

Research Assistant | University of Saskatchewan | August 2015 – December 2017

- Analyzed survey and focus group data about public attitudes towards different sources of energy in Saskatchewan: I collected and coded survey data, designed statistical methodology, performed statistical analysis, cross-checked the results against qualitative focus-groups data, and visualized the results (originally in SPSS and then replicated the process in R for reproducibility).
- Carried out comparative analysis of carbon pricing policies across Canadian provinces to estimate the potential effects of the federal carbon pricing initiative on the economy of Saskatchewan.

Graduate Service Fellow | University of Saskatchewan | December 2015 – April 2017

- Analyzed current and historic student enrollment data and the impact of international academic cooperation networks on international students recruitment.
- Collected the information for, and coordinated the work on the production and publication of a non-periodic “Discover US” magazine highlighting our university’s most prominent research projects.

Personal Projects and Publications

- Blogging about R and data analytics at www.dataenthusiast.ca. My “[COVID-19 Canada Data Explorer](#)” dashboard app was [endorsed](#) by the Macdonald-Laurier Institute – one of Canada’s leading public policy think tanks – to track the progression of the current pandemic in Canada.
- “[Designing COVID-19 Data Tools](#)” policy brief. Co-authored with Dr. Coates and Dr. Holroyd.
- “Diverse Community Energy Futures in Saskatchewan, Canada”. *Clean Technologies and Environmental Policy* 22(3). 2020. Co-authored with Dr. Rayner et al.

EDUCATION

Master of Public Policy | May 2017 | University of Saskatchewan

Johnson Shoyama Graduate School of Public Policy.

Certifications | Multiple | DataCamp

Some of the key ones include: Importing and Cleaning Data with R Skills Track, Spatial Data with R Skills Track, Shiny Fundamentals with R Skills Track, and many more. Please see [my DataCamp profile](#).