Martin **Čadek**

DATA SCIENTIST

✉ cadekmail@gmail.com | ⌂ www.martincadek.com | in martincadek

Data scientist passionate about healthcare, and former researcher with doctorate degree (PhD) in psychology which focused on health communication, public health, and obesity interventions. User of data analytics software and programming languages suitable for providing analytical and statistical solutions such as R, Python, and SAS. Advocate of reproducible research with the use of version control (git, docker, makefiles), reproducible reports (shiny, markdown).

Skills

DATA SCIENCE

- Exploratory data analysis, machine learning and regression analysis {tidymodels, car, lme4}, data wrangling {tidyverse, data.table}, dashboard development {shiny}, data visualization {ggplot2, plotly}, computerised text mining {quanteda, tidytext}, continuous deployment {netlify}, reproducibility {rmarkdown, renv, venv, miniconda}

APPLICATIONS

- RStudio, VS Code, SAS, Tableau, Excel, Qualtrics

LANGUAGES

- R {ggplot2, Shiny, data.table, tidyverse, tidymodels, quanteda, tidytext}
- Python {pandas, scikit-learn}
- SQL {Postgre}
- Git {GitHub, GitLab}
- SAS

Experience

Roche
Welwyn Garden City, England
DATA SCIENTIST
2023 - ongoing

- I am currently working as part of the Pharmaceutical Research and Early Development organization (pRED) as a Data Scientist:
  - Providing analytical and programming expertise and leadership support for the reporting and analysis of clinical trials and related submissions to Health Authorities worldwide.
  - Writing reusable functions and packages in R.
  - Transforming, organizing and analyzing healthcare data.
  - Contributing to open source software development.
  - Using R, SAS, Git.

Vodafone
London, England
QUANTITATIVE USER RESEARCHER
2022 - 2023

- In my role, I was combining statistical modelling techniques with focus on user behaviour (UX) as a Quantitative Researcher:
  - Developing business-wide metrics to measure user experience across key digital journeys. This is done using my knowledge of psychometric testing, validity, and reliability.
  - Establishing methodologies to support quantitative UX research internally and helping to standardise our approach to user experience measurement. Focusing on developing methods to reproducible survey analysis using Python and R.
  - Collaborating with designers, product owners, data scientists, researchers, and other stakeholders to develop user-centered products.
  - Development of internal R package and Python library to streamline reproducible research, documentation on AzureDevOps repositories, and maintenance of repositories for analytical work in the team.
  - Development of reproducible pipelines using Python digesting data from APIs and plotting the results in Dashboard environment.
  - Using R, Python, Git, Azure DevOps.

Kantar
London, England
DATA SCIENTIST
2021 - 2022

- I was working as part of the Expert Solutions team with clients in the fast-moving consumer goods sector. Here were some of the things I did:
  - developed as a data scientist, gained industry experience and tackled exciting new challenges
  - worked with structured, untidy, and large data
  - developed bespoke solutions for clients, using techniques such as hierarchical clustering algorithms, k-means, or latent class analysis
  - provided solutions such as customer segmentations, market penetration analyses, and market structure
  - worked with tools such as Git, Azure DevOps, SQL, R, and SAS

Department of Health and Social Care
London, England

STATISTICIAN *2020 - 2021*

- I was working as an Statistician in the DHSC Social Care Data and Analysis team since September 2020 until October 2021.
- I was responsible for delivering situational reports and statistical support for briefings in collaboration with policy colleagues and the Test and Trace programme. I contributed to the quality assurance of analyses across the team and promoted data science using R programming language.
- Examples of my work:
  - automated reporting in R for flu immunisation status of residents and staff at care homes in England
  - reproducible analytical pipeline (RAP) to draw and access Care Home data from a JSON feed
  - RAP to extract and process data from the Office for National Statistics (the ONS)
  - advanced visualisations regarding numerous ad-hoc request across topics such as flu immunisation, survey response rates, or covid incidence
  - quality assurance of reports, statistical analyses, and figures

Assessment Systems International a.s.
Remote

SHINY R DEVELOPER *2018 - 2019*

- I worked under contract to develop an R application with a Shiny front-end for machine-learning-driven assessment of candidates based on psychometric data, and I,
  - created a proof-of-concept that was accepted for further development
  - wrote the client-side components
  - designed and implemented a web UI in Shiny that enabled interactive navigation through the ML-model-building process and enabled the user to save and reuse the resulting model
  - coordinated project management remotely with the client using agile methodology
- The application drew on R packages to provide feature selection methods, classification and regression algorithms, and ensemble models, specifically elastic net models (glmnet), tree-based models (rpart), support vector machines (svmRadial), random forest (rf), neural networks (nnet), and k-nearest neighbour (knn)
- The application enabled the client to engage consultants with no programming experience and increased the effectiveness of the assessment process

Leeds Beckett University
Leeds, England

PHD *2017 - 2021*

- My research looked into improving the routine feedback letters with height and weight results delivered to all parents with Reception and Year 6 children as part of the National Child Measurement Programme (the NCMP) in England
- The project engaged multiple stakeholders such as Local Government, Public Health England, and parents, and provided nuanced insights into the NCMP, throughout the project I,
  - developed three complex studies using qualitative and quantitative methods
  - gathered opinions of 92 Local Government Authorities in England about how they use the NCMP and provided the results to the Public Health England as evidence for further improvement
  - prepared cluster randomised controlled trial and stratification procedure with reliance on clustering utilising k-means
  - developed a new version of the feedback letters which resulted in improvement of parental perceptions about the letters
  - used computerised text analysis techniques on a vast corpus of text data. For example, sentiment analysis, topic modelling (LDA), hierarchical clustering
  - used Google Natural Language Processing speech-to-text API to automatically transcribe all the interviews
- I became experienced with content analysis, literature and evidence reviews, and contributed to the writing of reports and journal articles
- I presented my findings at,
  - UKCO, Newcastle upon Tyne, England United Kingdom, September 2019
  - Public Health England Annual Conference 2018, University of Warwick, UK
  - Weight Stigma Conference, Leeds, UK, June 2018
  - UKCO, Newcastle upon Tyne, England United Kingdom, June 2018

Influencer.cz
Prague, Czechia

MARKETING ANALYST *2013 - 2017*

- As a marketing researcher executive, I was primarily responsible for the management of social media pages, content creation, and copywriting
- My clients were OZP (the third largest insurance provider in Czechia), BMW Mini (car manufacturer), and Centropol (energy and gas supplier), for these clients I,
  - prepared engaging content for their pages which led to an increase in sales and engagement of the page visitors
  - developed a transparent marketing strategy and content plans that are still used by the clients
  - provided long-term content plans that engaged their customers and facilitated sales
- Some other clients I have worked with were Lexus (car manufacturer), NKC (a non-profit organisation for the representation of woman in STEM), for them I,
  - provided an analysis of their marketing strategy and content
  - conducted customer segmentation using R and latentclass modelling
  - developed descriptive reports using ggplot2 and RMarkdown proposing a suitable marketing strategy

University of Cambridge
Cambridge, England

VISITING RESEARCHER
2015 - 2016

- I joined the Policy Research Group at the University of Cambridge as a visiting researcher under the supervision of Dr Kai Ruggeri. As a member of the group I,
    - developed infographics, reports, and presentations for the Safer Care Pathways in Mental Health Services project supported by The Health Foundation (https://cli.re/rwmbWR)
    - assisted with data collection and development of a database containing Issue-Based Industry Collectives in the alcohol industry and their characteristics for Dr Jochem Kroezen (https://cli.re/Zejkxx)
    - helped to design the Health@Work (H@W) questionnaire for an organisational setting that aimed to match employees' views on health and wellbeing to those of employers'
    - assisted with data collection of Pro Bono activities of U.S. Law firms for Dr Lionel Paolella
    - attended policy lectures and MBA courses in microeconomics

Junior Research Programme
Europe

JUNIOR DATA ANALYST
2014 - 2015

- I have joined a team of early-career researchers in the position of Junior Data Analyst and collaboratively developed a project focusing on Obesity discrimination in the Workplace supervised by Dr Stuart Flint
- The project was disseminated as a journal article (https://cli.re/JpW7k5)
- Our work was featured in the following media:
    - https://cli.re/Qm54PB
    - https://cli.re/mqrM2P
    - https://cli.re/b1x17e

Courses

Course
www.coursera.org

COURSERA
2023

- Making Data Science Work for Clinical Reporting. Credential: 2PYLWUBACCCF

Course
Welwyn Garden City, England

SAS
2023

- SAS Programming 1: Essentials. Credential: https://www.credly.com/badges/76d7c959-433b-4d91-8911-fc75c0402d4b

Course
www.datacamp.com

DATACAMP
2018

- Data Analyst with R TrackData Analyst with R Track

Course
Leeds, England

UNIVERSITY OF LEEDS
2018

- Summer School of Causal inference with observational data: the challenges and pitfalls (5 days course).

Online
https://cli.re/5Azz8D

AGROCAMPUS OUEST
2017

- Exploratory Multivariate Data Analysis (MOOC) by François Husson.

Course
Cambridge, England

UNIVERSITY OF CAMBRIDGE
2016

- Certified IRT and CAT course using Concerto at Psychometric Centre (3 days course).

Community

Leeds R Meetups
Leeds, England

CO-ORGANISER
2019 - 2021

- As an R enthusiast, I have been involved in organising the R meetups in Leeds
- https://cli.re/E59kPe

Leeds Beckett R Meetups
Leeds, England

ORGANISER
2018

- I held a series of R meetups to provide R training to other PhD students, and supported the statistical analysis that was essential for other research projects.

Qualtrics workshops
Leeds

LECTURER
2018

- Delivering series of free workshops about survey development for research staff at Leeds Beckett University

Education

Leeds Beckett University *Leeds, England*
PhD degree, Psychology *2017 - 2022*

Charles University *Prague, Czechia*
Master's degree, Psychology *2013 - 2016*

Charles University *Prague, Czechia*
Bachelor's degree, Psychology *2010 - 2013*