SAMUEL PUCEK

Simunkova 1608/19, Prague 8, Czech Republic
+420 770 189 639 ◇ samuel.pucek@gmail.com ◇ GitHub: samuelpucek

CARRIER OBJECTIVE

I am a friendly guy with a passion for technology. I greatly welcome new challenges and opportunities to learn new things. In my tasks I am systematic and persistent.

WORK EXPERIENCE

Avast Software *Prague, Czech Republic*
Junior Data Scientist *Jan 2021 - Present*

- Experimentation Platform
- Personalization
- Data Quality

I focus exclusively on further development of Experimentation Platform, mainly on automatization of Data Quality Checks. Our next goal is to provide deep-dive analyses to answer questions related to personalization and further evangelization of experimentation mindset across the organization.

Avast Software *Prague, Czech Republic*
Junior Data Analyst *Sep 2019 - Dec 2020*

- A/B testing
- Experimentation Platform
- Ad-hoc analyses
- Data Engineering

My responsibility was to maintain and update data pipelines, create dozens of ad-hoc analyses (Jupyter Notebooks) to answer business-related questions, and to co-work on development of Avast's Experimentation Platform - the cutting-edge tool for automatic evaluation of experiments (A/B tests).

My main contribution to Experimentation Platform was to design a statistical engine for evaluating experiments, which consists of sequential evaluation, multiple comparison problem, delta methods etc. The solution has been open-sourced in fall 2020 - feel free to check Avast GitHub repository `ep-stats`. From `ep-stats` we created an independent Python package which can be found in PyPi.

In `ep-stats` we mainly focused on computational complexity. Using `numpy` we vectorized many math operations. We back-upped all features using unit testing, including statistical features stated above.

O2 Czech Republic *Prague, Czech Republic*
Software tester (part-time) *Nov 2017 - Sep 2019*

- Integration and regression testing of company global CRM system during its development
- SQL, HP Quality Center, SAP, SoapUI, Confluence

O2 Czech Republic *Prague, Czech Republic*
Internship program (part-time) *Nov 2017 - Sep 2019*

- IT reporting in SAP Solution Manager

TECHNICAL STRENGTHS

Statistics	R, Python
Programming	C#, Python (Object-oriented programming)
Software & Tools	SQL, Git, LaTeX
Optimization	GAMS, MATLAB

EDUCATION

Charles University in Prague, Faculty of Mathematics and Physics *Sep 2017 - Sep 2019*
Master's degree in Probability, Mathematical Statistics and Econometrics

Charles University in Prague, Faculty of Mathematics and Physics *Sep 2014 - Sep 2017*
Bachelor's degree in General Mathematics

THESES

Master thesis: **Risk aversion in portfolio efficiency** *Sep 2019*
Bachelor thesis: **Scheduling optimization problems in education** *Sep 2017*

LANGUAGES

English	fluent (spoken and written)
German	basic (passive)
Slovak	mother tongue

PROJECTS

Discrete simulation of music festival *Feb 2019*
The simulation was coded in C#, with focus on user-friendly interface.
User can choose number of visitors, service speed (e.g. beer, hot-dog, toilets), and couple of other parameters. Program itself randomly creates visitors with various preferences and their own music schedule. Then the simulation starts. Randomness is also included in decisions during the simulation. The output is the overall satisfaction of visitors and detailed history of all visitors.

Modeling mortgage rates *Oct 2018 - Jan 2019*
This project aimed at developing a framework for predicting mortgage rates for the purpose of actuarial cash flow models implemented in insurance company. We suggested multiple approaches, i.e. vector autoregression, simple regression. Bootstrap algorithm was used for generating scenarios. In addition, we dealt with client specific data of the insurance company. We used R and EViews for regression. Bootstrap and simulations were made in MATLAB. Client specific data were analysed in R.

Primary school timetable *Aug 2018*
I used the results obtained in my Bachelor thesis and implemented them in real-life problem. Using GAMS I solved large linear optimization problem, which covered all requirements and conditions. The output was optimal timetable for the upcoming academic year. R was used for dealing with data before, and after the optimization.

INDIVIDUAL INTERESTS

Running (marathon 3:15:41, half-marathon 1:28:20), hiking, cycling