

# 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**

*Data Scientist*

*Prague, Czech Republic*

*Jan 2022 - Present*

- Retention & Loyalty

I work on development and productionalisation of Machine Learning pipelines and models which are part of the retention strategy.

### **Avast Software**

*Junior Data Scientist*

*Prague, Czech Republic*

*Jan 2021 - Dec 2021*

- Experimentation Platform (A/B Testing Platform)
- Data Quality

I focused exclusively on further development of Experimentation Platform, mainly on automation of Data Quality Checks. Next I provided deep-dive analyses to answer questions related to personalization and further evangelized experimentation mindset across the organization.

### **Avast Software**

*Junior Data Analyst*

*Prague, Czech Republic*

*Sep 2019 - Dec 2020*

- Experimentation Platform (A/B Testing Platform)
- 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**

*QA Tester (part-time)*

*Prague, Czech Republic*

*Nov 2017 - Aug 2019*

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

### **O2 Czech Republic**

*Internship program (part-time)*

*Prague, Czech Republic*

*Aug 2016 - Oct 2017*

- IT reporting in SAP Solution Manager

## TECHNICAL STRENGTHS

---

<b>Statistics</b>	R, MATLAB, GAMS
<b>Programming</b>	Python, C# (Object-oriented programming)
<b>Software &amp; Tools</b>	SQL, Git, L <sup>A</sup> T <sub>E</sub> X

## EDUCATION

---

<b>Charles University in Prague, Faculty of Mathematics and Physics</b> Master's degree in Probability, Mathematical Statistics and Econometrics	<i>Sep 2017 - Sep 2019</i>
<b>Charles University in Prague, Faculty of Mathematics and Physics</b> Bachelor's degree in General Mathematics	<i>Sep 2014 - Sep 2017</i>

## THESES

---

Master thesis: <b>Risk aversion in portfolio efficiency</b>	<i>Sep 2019</i>
Bachelor thesis: <b>Scheduling optimization problems in education</b>	<i>Sep 2017</i>

## LANGUAGES

---

<b>English</b>	fluent (spoken and written)
<b>Slovak</b>	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 1:26:02), hiking, cycling