

# SAMUEL PUCEK

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

## ABOUT ME

---

I'm friendly and open-minded person, with a passion for technology. With strong mathematical background (MSc in STEM) I focus on statistics and programming. In my spare time I love running and hiking.

My daily routine consists of developing ML models and pipelines, creating ad-hoc analysis, and supporting other colleagues. I'm fluent in Python, SQL and Git. In my tasks I am systematic and persistent.

## WORK EXPERIENCE

---

**Gen (formerly Avast and NortonLifeLock)**  
*Principal Data Scientist*

*Prague, Czech Republic*  
*Feb 2023 - Present*

- Experimentation and A/B testing

I improve experimentation, specifically implement new metrics tailored to the business, model customer LTV and support business with various pre-experiment and post-experiment analyses (Jupyter Notebooks). I prototype variance reduction methods, i.e., stratification and CUPED.

**Avast (now Gen)**  
*Data Scientist*

*Prague, Czech Republic*  
*Jan 2022 - Jan 2023*

- Architecture and development of ML pipelines

I worked on architecture and development of Machine Learning pipelines which supported retention strategy and personalisation in Avast (Dynamic Discounting). I was responsible for end-to-end pipelines architecture and development, the quality of code delivered by third parties, communication with other internal and external teams, and support of other colleagues, including onboarding.

In the Dynamic Discounting project were used XGBoost models followed by Reinforcement Learning model. The pipeline was scheduled daily, predictions were passed to other internal systems.

**Avast (now Gen)**  
*Junior Data Scientist*

*Prague, Czech Republic*  
*Sep 2019 - Dec 2021*

- Avast's Experimentation Platform

In Avast, we developed internally the Experimentation Platform, self-service tool for A/B testing. My main contribution to Experimentation Platform was to design a statistical engine for evaluating experiments, which consisted of sequential evaluation, multiple comparison correction, delta method etc. The solution was open-sourced in fall 2020, check GitHub repo **ep-stats**. The corresponding Python package 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. The project met all object-oriented programming standards. Next, I provided various deep-dive analyses (Jupyter Notebooks) to answer business-related questions, maintained and updated a few legacy data pipelines.

*After a year since the launch of the Experimentation Platform we served over 100 experiments per month. The experimentation mindset grew enormously; all ideas in the sales and marketing were A/B tested before being full scaled. EP provided evaluation of the experiments using a unified set of business*

*metrics. The goal was to further evangelise experimentation mindset into engineering and product development.*

## **O2 Czech Republic**

*Internship and QA Tester (part-time)*

*Prague, Czech Republic*

*Aug 2016 - Aug 2019*

- IT reporting in SAP Solution Manager
- Integration and regression testing of the global CRM system (SQL, HP Quality Center, SAP)

## **TECHNICAL STRENGTHS**

---

<b>Programming</b>	Python (Object-oriented programming, Jupyter Notebooks)
<b>Libraries</b>	Pandas, NumPy, Plotly, scikit-learn
<b>Software &amp; Tools</b>	SQL, Airflow, Git, L <sup>A</sup> T <sub>E</sub> X
<b>Data</b>	GCP (cloud), Hadoop (on prem), Hive, ClickHouse

## **EDUCATION**

---

**Charles University in Prague, Faculty of Mathematics and Physics**

*2014 - 2019*

Master's degree in Probability, Mathematical Statistics and Econometrics

## **LANGUAGES**

---

<b>English</b>	Professional working proficiency
<b>Slovak</b>	Native

## **PROJECTS**

---

### **Blackjack with RL agent**

*Dec 2022*

I coded Blackjack game in Python with standard actions, i.e. hit, stand, double down, split, surrender and insurance. The goal is to add Reinforcement Learning agent on top of the game. The agent should learn how to play Blackjack on its own, and wisely use strategy when the game is "hot".

See the `blackjack` repo in my personal GitHub.

## **INDIVIDUAL INTERESTS**

---

Running (marathon 3:00:03, half-marathon 1:23:51), hiking, cycling.