

# Szymon Socha

[linkedIn/szymon-socha](https://www.linkedin.com/in/szymon-socha) | [github/szymonsocha](https://github.com/szymonsocha) | [szymonsocha.github.io](https://szymonsocha.github.io)

## SUMMARY

Data Scientist with experience in the gaming industry and an economic background. Currently pursuing an MSc in Data Science and Business Analytics from the University of Warsaw. Uses machine learning and statistical modelling to find new patterns and relationships in data to drive business value. Passionate about using data to solve real-world problems. Ardent traveler.

## EDUCATION

### MSc. Data Science and Business Analytics

Warsaw, PL | 2021 - present

UNIVERSITY OF WARSAW

Awards: Dean's List 2022 (Top 7%)

Coursework: Machine Learning I, Machine Learning II, Unsupervised Learning, Advanced Econometrics, Applied Finance, Statistics, Python, Advanced Programming in R, Big Data Analytics, Algorithms for Data Science

### BSc. Computer Science and Econometrics

Warsaw, PL | 2018 - 2021

UNIVERSITY OF WARSAW

GPA: 4.5/5

Thesis: "The influence of COVID-19 pandemic on the music streaming consumption" Econometric model constructed and evaluated using R.

Python utilized for data manipulation purposes

Coursework: Machine Learning, Econometrics, Linear Algebra, Statistics, Time Series, Probability Calculus, Finance, Python, R, SQL, C++

## WORK EXPERIENCE

### WESUB | DATA SCIENTIST

Warsaw, PL | Apr 2023 – present

### FLYING WILD HOG | DATA SCIENTIST (JUNIOR)

Warsaw (Remote), PL | Jan 2022 – present

- Analyzing video games telemetry data using **R** and **SQL**
- Developing dashboards and reports in **RMarkdown** and **Shiny**
- Developing **ETL** processes with **BigQuery** and **MongoDB**
- Collaborating with UX Research team members in combining quantitative and qualitative research findings
- Working closely with different game development teams and editorial teams around the world in order to understand their requirements and questions related to a project

### PUBLICIS GROUPE | DATA SCIENCE ASSISTANT

Warsaw, PL | Sep 2021 - Nov 2021

- Developing tool for econometric modeling project
- Data cleaning and processing using **Python** and **R** and building custom R tools for econometric analysis

### TESTRONIC | QA TESTER

Warsaw, PL | Jan 2018 – Oct 2018

- Detecting and reporting bugs, software flaws and design issues (**JIRA**)

## SELECTED PROJECTS

### NEURAL NETWORK PERFORMANCE COMPARISON

PYTHON, NEURAL NETWORK, LIGHTGBM, XGBOOST, RANDOM FOREST

Comparison of neural network performance with other ML algorithms for two different types of data (image classification and regression).

### TRAFFIC PREDICTION

PYTHON, SKLEARN, RANDOM FOREST, LINEAR REGRESSION, KNN, SVM, RMD

Traffic prediction using multiple ML algorithms. Data manipulation and feature engineering.

### OPTION PRICING RCPP PACKAGE

C++, R, RCPP, RMD

Path-dependent European style up-and-out call option pricing with Monte Carlo and Rcpp package.

## SKILLS

**Tech stack:** R (tidyverse, dplyr, ggplot2, Rmd, Shiny), Python (sklearn, pandas, numpy), SQL (BigQuery, KQL), Git, Jenkins

**Expertise:** Machine Learning and Econometric data modelling, clusterization, dimension reduction, statistics, hypothesis testing

**Soft skills:** Problem-solving, Stress management, Decision-making, Time management, Communication, Attention to detail