Martin Skarzynski

Biostatistics · Data Science · Epidemiology · Genomics · Machine Learning · Pharmacometrics · Public Health · Software Engineering

240-595-3460 | marskar@gmail.com | www.marskar.github.io | www.github.com/marskar | www.linkedin.com/in/mskar | @marskar

Summary

My primary research interest is in understanding health risk factors by combining scientific expertise from diverse fields with machine intelligence. I believe I am uniquely equipped to bridge the gaps between scientific disciplines and deliver on the promise of data science in health research. My preferred tools are R and Python, open source programming languages kept on the cutting edge by their active and supportive communities. Through research and teaching, I am constantly improving my ability to obtain, tidy, explore, transform, visualize, model, and communicate data. I aim to utilize my technical skills and science background to become a leader among the next generation of multidisciplinary data scientists.

Recent Experience

Cancer Prevention Fellow

Rockville, MD

Biostatistics Branch, Division of Cancer Epidemiology and Genetics (DCEG), National Cancer Institute (NCI)

June 2017 - Present

- Integrates clinical, laboratory, epidemiologic, genomic, and medical imaging data
- Researches the effects of genes and exposures on cancer outcomes

Co-Chair

Bethesda, MD

Bioinformatics and Data Science Department, Foundation for Advanced Education in the Sciences (FAES)

January 2015 - Present

- Co-administers an academic program with over 20 faculty members
- Currently teaches two graduate data science courses:
  - *Introduction to Python* (BIOF309)
  - *Applied Machine Learning* (BIOF509)
- Designed, planned, organized and taught graduate practical courses including:
  - *Introduction to Text Mining* (BIOF395)
  - *Pharmacometric Analyses in Clinical Trials using R* (BioTech84)
  - *Junior Scientist Training Program* (BioTech42)
  - *Flow Cytometry: Principles and Methods* (BioTech23)
  - *Immunochemistry and Monoclonal Antibody Production* (BioTech8)

Instructor

Bethesda, MD

The Carpentries: Software, Data, and Library Carpentry

April 2017 - Present

- Teaches hands-on workshops, such as the 2019 NIH Genomics Data Carpentry Workshop

Open Source Software Developer and Package Reviewer

Remote

pyOpenSci

May 2019 - Present

- Reviews scientific Python packages submitted to the pyOpenSci community
- Developed and maintains three Python packages:
  - *Nbless*: (de)construct, convert, and execute Jupyter Notebooks
  - *Rmdawn*: (de)construct, convert, and render R markdown
  - *Gitone*: combine multiple git version controls steps into one

Data Science Content Developer

Remote

DataCamp

October 2018 - Present

- Developed and maintains an online course called *Creating Robust Python Workflows*

Adjunct Professor

Washington, DC

George Washington University (GWU)

January 2015 - May 2015

- Taught an undergraduate lecture and lab course: *Introductory Biology: The Biology of Organisms* (BISC 1112)
- Organized, led and invited speakers for an undergraduate seminar, *Women and Leadership* (WLP1111)

Graduate Education

MPH, Epidemiologic and Biostatistical Methods for Public Health and Clinical Research

Baltimore, MD

Johns Hopkins University School of Public Health (JHSPH)

May 2018

- Data Science Certificate

PhD, Tumor Biology

Washington, DC

Georgetown University-NIH Graduate Partnerships Program (GU-NIH GPP)

May 2015

- Teaching Certificate

MS, Biotechnology

Krakow, Poland

Jagiellonian University

June 2009