

Noa Bendit-Shtull

noa-codes.github.io / noa.f.bs@gmail.com / 718.873.3572

I am an experienced data scientist with a passion for diversity in tech, efficient code, and effective communication. I care about *how* work gets done, and lead by example.

ETSY

SENIOR APPLIED SCIENTIST II ■ 2022 - PRESENT

Apply machine learning methods to build structured data for Etsy's Knowledge Base. Drive process improvements in team code reviews, documentation, and development environment.

OSCAR HEALTH

STAFF DATA SCIENTIST ■ 2022

Designed and implemented best practices and tooling to support machine learning work at Oscar.

SENIOR DATA SCIENTIST ■ 2020 – 2022

Built machine learning models to automate insurance processes and manage medical costs.

- ▶ Applied NLP methods to electronic health record data to develop vector representations of medical concepts.
- ▶ Automated the identification and prioritization of unit cost, utilization, and routing issues across 25 insurance markets.
- ▶ Developed and led training on technical skills (Git, Arcanist, Looker) and non-technical skills (giving and receiving feedback).
- ▶ Provided mentorship and guidance to more junior data scientists on programming best practices, statistics, causal inference, and command-line tools.

LINKEDIN

DATA SCIENCE INTERN ■ SUMMER 2019

Developed a clustering workflow using Apache Hive, Spark MLlib, and Scikit-learn to identify cohorts of high-value LinkedIn Premium Subscribers.

CORNERSTONE RESEARCH

ANALYTICS CONSULTANT ■ 2016 – 2018

Provided consulting assistance and embedded support to 30 internal teams (3 to 50+ team members, across 8 offices) in all aspects of data analytics and programming.

- ▶ Built a web data pipeline using Python, Docker, AWS, and PostgreSQL to process 100k documents.
- ▶ Designed ETL pipeline for terabytes of financial market data using SAS, SQL, and IBM Netezza, improving turn-around time by 10x.
- ▶ Developed a data analytics training program for 200 analysts, including training in version control, SAS, R, Netezza, and SQL.

ANALYST ■ 2014 – 2016

Conducted financial and economic analyses in SAS and Excel, to support leading industry experts in securities, antitrust, and patent infringement cases.

EDUCATION

STANFORD UNIVERSITY

2018 – 2020

M.S. in Statistics: Data Science

UNIVERSITY OF PENNSYLVANIA

2010 – 2013

B.A. in Economics, Summa Cum Laude

SKILLS

- ▶ Languages: Python, R, Scala, C++
- ▶ Scripting & Query Languages: Bash, SQL
- ▶ Environments: Linux, Docker, GCP
- ▶ ML Frameworks: scikit-learn, XGBoost, PyTorch
- ▶ Big Data: BigQuery, Spark, Hive, Hadoop
- ▶ Communication: Writing, Training, Mentoring

INTERESTS

Swing dance (lindy hop)
Biking
Cooking and baking
Education
Microeconomics
Glassblowing