

Brittany Allen Data Enthusiast

Number & email removed for privacy – please connect via my contact page or LinkedIn
Oakland, CA • traintestbritt.github.io • [GitHub](#) • [LinkedIn](#) • [Medium](#)

TECHNICAL SKILLS

Languages | Frameworks: Python • SQL • PySpark • Psycpg2 • Pytest • Pandas • Scikit-learn

Database | Cloud: AWS S3 • RedShift • EC2 • PostgreSQL • Apache Cassandra • Heroku

Analytics | Visualization: Apache Spark • Google Analytics • Excel • Tableau • Matplotlib • Seaborn

Productivity: Git • Github • Docker • Zube

EXPERIENCE

2020 **Data Engineer** (covid-19 layoff) @ Compaas • San Francisco

- Wrote a script to generate data for a custom seed company which allowed our sales team to demo our product with consistent and realistic data
- Captured additional client data by adding new fields to existing data models via Django's ORM
- Wrote scripts to flag duplicates and block updates on production until reconciled while retaining existing data and future-proofing us from client's being able to duplicate pay raises with our software

2019 **Data & Analytics Strategist** @ Palantir Technologies • NYC

- Developed data pipelines leveraging Spark's distributed computing framework to clean, transform, and integrate geospatial aircraft (ads-b) data collected via the OpenSky REST API with open-source aircraft, runway, and airport data enabling our client in the U.S. DOT to perform safety analyses at scale
- Defined a dynamic ontology modeled off of two mission-critical problems – aircraft misalignments and go-arounds – which reflected how safety analysts and aerospace engineers thought about their world
- [Parsed XML files](#) to extract text from federal documents and stored it as a PySpark SQL DataFrame
- Led in-person client demos and created training materials on [Palantir Foundry](#)

2019 **Associate Data Science Instructor** (contract) @ General Assembly • NYC

- Co-taught a 12-week course for 30 Morgan Stanley employees, developed [supplemental course material](#) to reinforce learning, and hosted virtual and in-person 1:1 office hours to support students

2018 **Data Science Scholar** @ General Assembly • NYC

- Awarded a full scholarship into an intensive program with coursework on regression analysis, hypothesis testing, bias-variance tradeoff, classification modeling, clustering analysis, etc.

2017 **Digital Marketing Analyst** (contract) @ Prudential • NYC

- Recreated Tableau dashboards on Adobe Analytics Workspace and led trainings as part of a team-wide effort to make data and tools more accessible for stakeholders

2016 **Data Analyst** @ Vox Media • NYC

- Analyzed social and website data, [built and maintained dashboards](#), [wrote reports](#), and shared insights on a weekly, monthly, and ad-hoc basis

PROJECTS

Data Modeling with Postgres via Udacity  [link](#)

Focused on data modeling using the STAR schema and building an ETL pipeline to transfer data from JSON files

EDUCATION

Rutgers Business School
B.S. Marketing, 2016

General Assembly
Data Science, 2019

Udacity
Data Engineering, 2020