

# Anders Poirel

---

[anderspoirel.me](https://anderspoirel.me) | [andersjopo@gmail.com](mailto:andersjopo@gmail.com)

## WORK EXPERIENCE

**data scientist, data scientist intern** | [Joby Aviation](#) *Jul. 2021 - Present*

- developed distributed framework on top of PySpark for running analyses automatically upon availability of new flight data, resulting in 2x faster runs and 10x more concise code compared to previous solution
- accelerated CI pipelines between 5 and 20x by deploying local Spark through Docker, mocking data and parallelizing builds

**data scientist intern** | [Fullpower Technologies](#) *Jan. 2021 - Jun. 2021*

- ported Tensorflow models to AWS Inferentia, increasing inference throughput 2x at fixed cost
- ran comparative analysis of sleep studies with medical device specifications using Pandas, which was used in an FDA device approval submission packet

**research assistant** | [UC Santa Cruz Genomics Institute](#) *Sep. 2020 - Jun. 2021*

- derived novel EM solver for generating synthetic genotypes with [implementation in numba](#) 100x faster than equivalent Python/numpy code
- investigated generative model for genotypes, with results suggesting that state-of-the-art GAN architectures for genotype generation do not scale to large sequences

**research intern** | [Center for Research in Open Source Software](#) *Jul. - Sep. 2020*

- added ability to run Jupyter through Docker and Singularity backends in the [Popper](#) task automation engine
- developed automation workflows for computational research enabling end-to-end replication of a research project

**data scientist intern** | [Startup Genome](#) *Apr. - Jun. 2020*

- built analysis pipeline in Python and pandas for global survey data which reduced time to view new survey results from 1 hour to 5 minutes

## VOLUNTEER EXPERIENCE

**president, vice-president** | [Data Science @ UCSC](#) *Jan. 2019 - Jan. 2021*

- led student team to top 5% global scores in [several](#) machine learning [competitions](#) on Kaggle and Driven Data

## SKILLS

**programming:** Python (PySpark, PyYorch, scikit-learn, pandas, Plotly, Dash), SQL, R (tidyverse), Javascript (React)

**software:** Docker, Git, Databricks, Kubernetes, Grafana, Tableau

## EDUCATION

University of California, Santa Cruz *Jul. 2017 - Jun. 2021*  
B.A. Computer Science (Honors), B.A. Mathematics (Honors).  
Coursework towards an M.A in Mathematics.

## AWARDS

**The Transparency Project**

[1<sup>st</sup> place winner](#) of CruzHacks 2020. Data visualization website aggregating data from Facebook and Google to bring transparency to political ad spending in the 2020 campaign.

