

Jake Haines

Raleigh, NC | +1 919 307 9605 | hello@jakehaines.net | github.com/hainesdata | linkedin.com/in/jhainesnc | jakehaines.net

Education.

Bachelor of Science, Statistics—North Carolina State University

August 2021 – May 2025. Raleigh, North Carolina.

- + **Coursework:** Intro to R, Intro to SAS, Linear Algebra, Mathematical Statistics I, Natural Language Processing, Python/R for Data Science, Statistical Computation & Data Management I
- + **Activities:** Data Analytics Club member, Automotive Club member, collegiate division e-sports team member

Transfer, Computer Science & Music Technology—University of North Carolina at Asheville

August 2019 – May 2021. Asheville, North Carolina.

- + **Coursework:** Applied Data Analysis, Computer Vision, Database Management Systems, Data Structures, Object-Oriented Programming, Research in Statistics (Spring 2021), Research Methods I

Experience.

Data Engineering Intern (Chassis)—Tesla

August 2022 – February 2023. Palo Alto, CA.

- + Authored over 100 SQL queries for ETL and analyzed big data on over 10mil rows from over 100k vehicles (Spark, pandas, plotly)
- + Managed 3 big data pipelines and queried 8 months of granular sensor data from over 100k vehicles in Airflow
- + Built dashboard w/ 10 visualizations on sensor data from Airflow pipeline (plotly, Tableau/bokeh-based API, PostgreSQL, Presto, Trino)

Research & Development Intern (Vehicle Technology)—Tesla

May 2022 – August 2022. Fremont, CA.

- + Built 6 NDA proof-of-concepts, 1 project publicly in development for all vehicle models
- + Data collection and ETL from vitals sensing, CAN, geospatial, and Raspberry Pi sensors in Python (pandas, numpy, plotly, Mapbox)
- + Research design and execution of 2 UX studies for early-stage NDA projects (pandas, scipy, scikit-learn)
- + Authored technical documentation, visuals, and Gantt chart for projects, code, databases, and datasets (Clickup, Sharepoint, Visio)

Co-founder & Data Engineer—Myndmap

April 2023 – Present. London, UK.

- + Built and managed AWS infrastructure from ground-up (PostgreSQL RDS, EC2, Lambda)
- + Authored 10 research papers for features allowing users access to research-backed executive dysfunction management strategies
- + Generated over 600 waitlist signups in first month of publicity and featured on *There's An AI For That*

Data Engineering Intern—Eagle Parts & Products

May 2021 – September 2021. Augusta, GA.

- + Data cleaning, parsing, and ETL on over 10k manually entered rows of inventory data in Python (pandas, regex, usaddress)
- + Restructured data for inventory data migration by automating data entry, reducing labor time required by 50%

Research in Traffic Flow (Statistics Department)—University of North Carolina at Asheville

January 2021 – May 2021

- + Collected traffic flow data via TensorFlow object detection model containerized in Docker on Raspberry Pi
- + ETL in Visual Studio SSDT (SSIS) and Microsoft SQL Server Management Studio (SSMS) on 7mil rows of data to 5k meaningful rows
- + Data mining, analysis, and visualization in Tableau and presented research at 2 universities

Skills.

Database tools: Airflow, AWS (EC2, RDS, IAM, Lambda), Hadoop, Microsoft SQL Server Management, Postgres, Presto, Spark, SSIS

Programming languages: Python, SQL, Bash, Linux, R

Python modules: BeautifulSoup, Keras, Matplotlib, Numpy, OpenCV, Pandas, Plotly, Pyspark, Scikit-learn, Tensorflow

Software: DataGrip, Docker, Git, GitHub, Jira, Jupyter, PyCharm, Tableau, Visio

Skills: big data, data science, data engineering, data visualization, machine learning, audio engineering, data processing

Interests: data engineering, software, data science, machine learning, music production, infrastructure, data operations, travel, research