

Jake Haines

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

Education.

Statistics, B.S. at North Carolina State University

August 2021 – May 2025. Raleigh, North Carolina.

- + Member of Data Analytics Club, Automotive Club, Collegiate Competitive E-sports Team
- + Tutor for Chinese natives learning English across the world

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

August 2019 – May 2021. Asheville, North Carolina.

- + Undergraduate research on traffic flow in Spring 2021
- + Minor in mathematics
- + Dean's List in Fall 2019 and Spring 2020

Research in Traffic Flow

January 2021 – May 2021

- + Collected resident parking traffic data via Docker containerized TensorFlow object detection model on Raspberry Pi
- + ETL in Visual Studio SSDT (SSIS) and Microsoft SSMS (SQL), cleaned and augmented 7mil rows of raw data to 5k meaningful rows
- + Data mining, analysis, and visualization in Tableau and presented at UNC Asheville and Wake Forest University

Coursework

Computer Science: Computer Vision, Database Management Systems, Data Structures, Intro to R, Intro to SAS, Natural Language Processing, Object-Oriented Programming, Python/R for Data Sci, Software Eng, Systems I

Statistics: Applied Data Analysis, Linear Algebra, Mathematical Stats I, Research Methods I, Statistical Comp & Data Management I

Experience.

Myndmap Co-founder & Data Engineer

April 2023 – Present. London, UK.

- + Built and managed scalable AWS infrastructure (PostgreSQL RDS, EC2, Lambda)
- + Authored 10 internal research papers for feature development
- + Received over 500 waitlist signups in first month of publicity and featured on *There's An AI For That*

Tesla Data Engineering Intern – Chassis

August 2022 – December 2022. Palo Alto, CA.

- + Authored over 100 SQL queries to transform/augment over 10mil rows from over 100k vehicles and analyzed (Spark, pandas, plotly)
- + Managed 3 sensor data pipelines querying eight months of granular data from over 100k vehicles using Airflow
- + Created interactive real-time dashboard w/ 10 visualizations on sensor data from Airflow pipeline (plotly, Tableau/bokeh-based API)

Tesla Research & Development Intern – Vehicle Technology

May 2022 – August 2022. Fremont, CA.

- + Built 6 proof-of-concepts for future vehicle technologies that will innovate passenger experience and increase accessibility
- + Collected, validated, transformed and analyzed vitals sensing, CAN, geospatial, network, and IoT sensor data in Python (pandas, numpy, plotly, geojson, folium, datashader, GPIO, Mapbox)
- + Designed and conducted 2 empirical UX studies (pandas, scipy, scikit-learn)
- + Wrote detailed and high-level documentation and produced clear visuals for projects, code, databases, and datasets

Eagle Parts & Products Data Engineering Intern

May 2021 – September 2021. Augusta, GA.

- + Restructured over 10k manually entered rows of inventory data, cleaned and parsed data in Python (pandas, regex, usaddress)
- + Transformed data for new inventory management software format, automating data entry and reducing time required by 50%

Skills & Interests.

Interests: audio, automotive, neuroscience, aviation, music production, travel, software, data operations, psychology, research

Database tools: Apache Airflow, Apache Spark, AWS (EC2, RDS, IAM), Hadoop, Microsoft SQL Server, MySQL, PostgreSQL, Splunk

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

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

Software: ClickUp, DataGrip, DataSpell, Docker, Excel, Git, GitHub, Heroku, JIRA, Jupyter, PyCharm, Sharepoint, SQL Server Integration Services, Tableau, Visio

Languages: English (native), German (limited working proficiency)