

# JOLIE TRAN

[jolietran525@gmail.com](mailto:jolietran525@gmail.com) | [linkedin.com/in/jolietran525](https://www.linkedin.com/in/jolietran525) | [jolietran525.github.io](https://jolietran525.github.io)

## LANGUAGES AND TECHNOLOGY

---

- Data: PL/SQL, Python (NumPy, Pandas, GeoPandas, scikit-learn, matplotlib, seaborn), MATLAB, R, Excel
- Web Development: HTML, CSS, JavaScript, LeafletJS, MapboxGL, ChartJS, JSON
- Tools: Visual Studio Code, QGIS, Tableau

## EXPERIENCE

---

**GIS Analyst** Washington State Transportation Center (TRAC) Jun 2023 – Present

### ORCA Analysis

- Implemented robust methods in PostgreSQL to effectively rectify wrongly encoded stops within the ORCA data based on GTFS, ensuring data integrity and accuracy.
- Created an interactive [web interface](#) with tree map/billboard dashboards, seamlessly integrated with ORCA data, providing a comprehensive visualization of ridership at each stop in adherence to the GTFS standard.

### OpenSidewalks (Transportation Data Equity Initiative)

- Developed a method for automating sidewalk attributes collection from State and City DOT and feed into the sidewalk data in OpenStreetMap (OSM) using PostgreSQL and PostGIS.
- Designed sets of rules in SQL for handling data integration challenges from disparate sources, identifying poorly drawn segments, and generating missing OSM sidewalk networks.
- Tested and fine-tuned queries across multiple neighborhoods in Seattle (WA), achieving an accuracy rate exceeding 90%. Established a baseline for scaling in Portland (OR) and Baltimore (MD).
- Created a [web interface](#) with HTML, CSS, JavaScript, and Leaflet for transparent public review, fostering continuous enhancements in sidewalk data precision.

**Social Media Research, Intern** Humanities Data Science Summer Institute @ UW Jun 2023 – Aug 2023

- Collaborated with a multidisciplinary team to conduct a comprehensive analysis of social media conversations on Twitter, focusing on the influence of historical figures in political movements and historical events.
- Analyzed an extensive dataset of 3 million tweets spanning from 2007 to 2023, applying advanced data processing techniques using Pandas and Dask libraries in Python. This efficient analysis revealed key trends and patterns.
- Implemented a mixed-method research approach, combining quantitative data analysis with qualitative assessment to uncover nuanced patterns and trends in social media conversations.
- Contributed valuable insights and created compelling data visualizations for an ongoing book project led by a professor at UW, enhancing the research and its potential impact.

## EDUCATION

---

**B.A. in Geography: Data Science** University of Washington Sept 2021 – Mar 2024

- In-major GPA: 4.0. Cumulative GPA: 3.94
- Coursework: Data Structures and Algorithm, Database Management, Machine Learning, Web Development

## PROJECTS

---

**Interactive Influenza Disease Map** Web GIS Development Coursework Sept 2023 – Dec 2023

- Initiated the data collection and cleaning from reputable sources (CDC, NCHS), created a robust influenza dataset.
- Led a team of five members, optimizing task delegation for efficient project execution.
- Developed an intuitive [web map interface](#) with interactive dashboard from JavaScript and CSS that allows users to investigate the influenza weekly and annual statistics on both national and state level with ease.

**Database Management Projects** Database Management Coursework Sept 2022 – Dec 2022

- Performed basic to advanced SQL queries to analyze a dataset of more than 1,000,000 records.
- Developed a database using pymysql that resulted in a 50% improvement in query performance.
- Manipulated a large semi-structured dataset (80,000+ lines of text) in JSON using NoSQL (AsterixDB) to find insights into the business questions.