

# Zhuoxinran(Zoe) Li

514-814-3577 | [zoezrli.github.io](https://zoezrli.github.io) | [zhuoxinran.li@mail.utoronto.ca](mailto:zhuoxinran.li@mail.utoronto.ca) | [linkedin](#) | Toronto, ON

## EDUCATION

### University of Toronto

BA, Major in Human Geography, Minor in Geographic Information Systems

Toronto, ON

Sep 2023 – May 2025 (exp)

### Concordia University

Major in Urban Studies

Montréal, QC

Jan 2021 – May 2023

## SKILLS

**Techs:** ArcGIS Pro, Python (*Pandas, Geopandas, Numpy, Matplotlib, Folium*), QGIS, SQL, AutoCAD, SketchUp, Adobe Illustrator, Adobe Photoshop, Microsoft Office

**Languages:** Mandarin, English, French

**Interests:** Gerontology, Disability, Public Health, Transportation, Visualization Design, Art History

## PROJECTS

### Data Visualization for Potential Prediction

May - Jun 2024

- Trained a **XGBoost** to predict energy, with test performance of 0.112 mean squared error and 0.992  $R^2$ .
- Illustrated the importance of 14 features on the well-trained **XGBoost**, using scatter plot in Matplotlib.
- Calculated **Pearson correlation** among 14 features and visualized it using Heatmap in Matplotlib.
- Manuscript "Identification of Cu-N<sub>2</sub> Sites for Zn-Air Batteries in Harsh Electrolytes: Computer Virtual Screening, Machine Learning, and Practical Application" in submission.

### Clustering Geolocation Data Intelligently in Python

Apr - May 2024

- Initiated the clustering process by utilizing the **KMeans** to cluster 837 geolocation data entries of taxi ranks into 70 clusters, achieving a silhouette score of 0.64.
- Improved clustering quality by applying **DBSCAN** to filter out outliers, resulting in 51 clusters with an enhanced silhouette score of 0.92; however, the model is better suited for uniform density within clusters.
- Refined clustering performance with **HDBSCAN** to address varying cluster densities, yielding 61 clusters with a silhouette score of 0.77.
- Finalized the clustering process by leveraging **KNeighborsClassifier** to evaluate and analyze remaining outliers, ensuring comprehensive data integrity.

### Geospatial Data Visualization using Python and Folium

Mar - Apr 2024

- Categorized 10,001 taxi trip start points by their three 'CALL TYPE's using **Folium** in Porto, Portugal.
- Extracted day and hour data of each trip from the 'TIME STAMP' column using a lambda function.
- Visualized traffic volumes in bar charts using **Matplotlib**, identifying 9 AM as the peak hour.
- Highlighted four primary routes frequently used during the 9 AM rush hour through detailed mapping.

### Analyzing Neighbourhood Quality of Life with GIS

Oct - Dec 2023

- Utilized **ArcGIS Pro** to map and analyze 77 Chicago neighborhoods with specific socioeconomic factors (health, culture, education, green space, and crime). Each analysis map assigns a unique socioeconomic impact value to each neighborhood.
- Normalized each socioeconomic factor using the **Max-Min** in the **Field Calculator** to ensure comparability across different metrics and facilitate accurate livability scoring.
- Finalized livability score for each neighborhood using the **Weighted Sum Model** in ArcGIS Pro, with scores ranging from 15.6 to 79.7.
- Created a map of overall livability in Chicago neighborhoods, identifying five communities with the highest livability scores (70-80), primarily located in the northern part.

### Geovisualization with Python

Nov - Dec 2023

- Mapped Toronto's population with **Geopandas** and **Matplotlib**, highlighting demographic trends.
- Visualized **LiDAR data** (3D point clouds) of the University of Toronto, St. George campus using **Laspy**, **Numpy**, and **Matplotlib**, providing detailed insights into the campus's structural layout.
- Animated Ontario tornado tracks using **Geopandas**, **Pandas**, **Numpy**, **Plotly**, and **Shapely**.

## PUBLICATIONS

---

- **Identification of Cu-N<sub>2</sub> Sites for Zn-Air Batteries in Harsh Electrolytes: Computer Virtual Screening, Machine Learning, and Practical Application.** C. Xu, K. Li, S. Liu, J. Xu, S. Sharma, J. Zhang, B. Mao, H. Chen, H. Zhang, H. Xu, B. Luo, H. Zhao, **Z. Li**, Z. Huang, J. Wang, K. Xi, C. Fu, Y. Zhao, G. Chai, G. He, I. Parkin, H.L [\[In Submission to Energy and Environmental Science\]](#)
- **ProteinDT: A Text-guided Protein Design Framework.** S. liu, Y. Li, **Z. Li**, A. Gitter, Y. Zhu, J. Lu, Z. Xu, W. Nie, A. Ramanathan, C. Xiao, J. Tang, H. Guo, A. Anandkumar. [\[In Submission to Nature Machine Intelligence\]](#)
- **Unsupervised Discovery of Steerable Factors When Graph Deep Generative Models Are Entangled.** S. liu, C. Wang, J. Lu, W. Nie, H. Wang, **Z. Li**, B. Zhou, J. Tang. [\[Transactions on Machine Learning Research 2024\]](#)
- **Symmetry-Informed Geometric Representation for Molecules, Proteins, and Crystalline Materials.** S. liu, W. Du, Y. Li, **Z. Li**, Z. Zheng, C. Duan, Z. Ma, O. Yaghi, A. Anandkumar, C. Borgs, J. Chayes, H. Guo, J. Tang. [\[NeurIPS Dataset and Benchmark 2023\]](#)

## EXPERIENCE

---

**Graphic Designer / Certified Travel Consultant** Apr, 2019 - Apr, 2021

- Achieved a 90% approval rating from the company manager and a 95% client satisfaction rate for the quality and impact of trips plans, reflecting consistent high performance and customer contentment.
- Designed trip posters for both domestic (Canada) and international trips using **Adobe Illustrator** and **Adobe Photoshop** to attract target audience and sell products.

**Art Teaching Assistant** Jan, 2019 - Dec, 2019

- Instructed five students aged 6-12 in self-expression, creativity, and critical thinking about their artwork.
- Assisted in organizing and setting up art exhibitions to showcase students' work, thereby enhancing their confidence, public presentation skills, and appreciation for art.

## CERTIFICATIONS

---

### Coursera:

- IBM Data Engineering Professional Certificate (In Progress; [Course Description](#))
- Sustainable Cities and Communities Specialization ([Course Description](#))

### Udemy:

- Psychology of Older Age/Gerontology ([Course Description](#))