# Jack Li

Data Analyst
Jack.li.2000@outlook.com | www.linkedin.com/in/jack-li-2000 | Jack-Li-2000.github.io/JackLiPortfolio/

## Skills

**Python:** Highly proficient with pandas, numpy, matplotlib, seaborn, scikit, scipy, etc.
**Excel:** Highly proficient with pivot tables, formulas, xlookup, conditional formatting, charts, data cleaning.
**SQL:** Experienced with query calls, views, joins, conditional statements, database management, etc.
**Data Visualization:** Familiar with telling stories using interactive dashboards in Tableau and PowerBI.
**Statistical Methods:** Highly proficient with supervised learning methods, time series, hypothesis testing, etc.
**Cloud computing:** Experienced with Google Cloud Platform and Microsoft Azure.

## Work History

### Reconciliation Data Analyst, OneDataTree

(01/2024 – Present, Toronto ON)

- Developed dashboard automation pipeline for data ingestion, processing, and reporting, achieving **75% reduction in daily time commitment** in **SQL and Python**.
- Implemented cross-database validation report to maintain data accuracy and reliability between internal and client records significantly **reducing discrepancies by 43% in SQL, Python, and Excel.**
- Conducted financial predictive forecasting, classification, and ad hoc analytics, leading to a **15% increase in revenue reporting accuracy with Python**.

### Data Science Intern, Teck Resources

(08/2021 – 06/2022, Vancouver BC)

Machine Failure Classification

- Performed data cleaning, conveyed statistical inferences, and conducted exploratory analysis with **Python and SQL**, resulting in a **22% reduction in data processing time**.
- Retrieved relevant features for classification using forward stepwise and ElasticNet feature selection, **improving model precision by 15% and reducing overfitting**.

Engine Life Cycle Prediction

- Predicted equipment lifecycle trends to inform replace vs. maintain decisions using regression models, **resulting in a 15% reduction** in maintenance costs and a **10% increase in equipment uptime**.
- Displayed insights with analytics dashboards and visualizations in **PowerBI**.

## Projects

### Wordnet Similarity Recognition Analysis

Conducted WordNet similarity recognition analysis to explore NLP data structure similarities with human semantic relationships. Leveraged data from the Semantic Priming Project to assess human reaction time and accuracy in response to stimuli. Analyzed correlation between WordNet path similarity and human word-pair recognition performance, performing hypothesis and significance testing in Python.

### Toronto AirBNB Popularity Insights Dashboard

Developed an interactive Tableau dashboard analyzing data from November 2022 to November 2023 for the Greater Toronto Area. Provided insights into popular neighborhoods, pricing trends, and accommodation types, enhancing decision-making and strategic planning.

## Education

### Honours B.Sc. in Statistics & Computational Cognitive Science

(University of Toronto, Toronto Ontario)

Practical knowledge in data regression models, predictive data techniques, classifiers, time series analysis, dimensionality reduction, data visualization, and extracting key insights from complex data sets. In-depth understanding of data structures and programming.