

# Jack Li

Data Scientist | Data analyst

Jack.li.2000@hotmail.com | linkedin.com/in/jack-li-2000 | <https://Jack-Li-2000.github.io/JackLiPortfolio/>

---

## Skills

**Python:** Highly proficient with pandas, numpy, matplotlib, seaborn, scikit, scipy, etc.

**R:** Highly proficient with ggplot, glm, general scripting, visualizations.

**Excel:** Proficient with pivot tables, formula, xlookup, conditional formatting, charts, data cleaning, etc.

**SQL:** Experienced with query calls, views, joins, conditional statements, etc.

**Data Visualization:** Familiar with telling stories using dashboards and graphics in Tableau and PowerBI.

**Statistical Methods:** Highly proficient with regression, classification, time series, inferential and descriptive statistics.

**Agile and Scrum:** Familiar with Agile and Scrum ceremonies.

**Cloud computing:** Familiar with Google cloud platform and Azure.

---

## Professional Experience

### Data Science Intern @ Teck Resources

(2021 – 2022, Vancouver Canada)

- Performed data cleaning, conveyed statistical inferences and exploratory analysis with python and SQL.
- Retrieved features from various models using forward stepwise, lasso, and elastinet feature selection.
- Tested/deployed various regression models to predict equipment lifecycle, find trend, best, worst case.
- Tested/deployed classification models to identify equipment failure from noisy data.

---

## Projects

### Toronto Real Estate Regression Analysis & Prediction

Using available independent data such as list price, # of rooms, etc, and dependent data like sold price to create regression model to model relationships and predict price, evaluated with RMSE in Python.

### Wordnet Similarity Recognition Analysis

Using available data for human reaction time and accuracy based on response to stimuli, analyse correlation between Wordnet path similarity and human task performance with statistical testing in Python.

### AirBNB Toronto Location Expense Dashboard

Interactive Tableau dashboard featuring location and price data for the greater Toronto area as well as concentration of listings and historical pricing between mark and July of 2023.

### Global COVID-19 EDA & Visualization

MS SQL database manipulation and analysis of global and localized infection, mortality, vaccination, and vaccination survival rates by country, continent, and worldwide. Data imported, visualized as Tableau dashboard.

### Interactive Bike Sales Analysis & Dashboard in NA and Europe

Excel analytics of features contributing to bike sales for NA and Europe displayed as an interactive dashboard.

---

## Education

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

(University of Toronto, Canada)

Relevant skills include practical knowledge in data regression models, predictive data techniques, classifiers, time series, dimensionality reduction, data visualization, and drawing key information to deliver insights on complex data sets. Also, in depth knowledge with data structures and programming.