

# KELLY LI

DATA SCIENTIST

(647) 992-5122 | [kellyli210.kl@gmail.com](mailto:kellyli210.kl@gmail.com)

Website: <https://kelly-li.netlify.app/>

[www.linkedin.com/in/kelly-vs-li/](https://www.linkedin.com/in/kelly-vs-li/)

Toronto, ON

## Summary

---

Highly accomplished Data Scientist with a robust academic foundation in Biotechnology and Chemistry, specializing in the domains of business intelligence, data visualization, and informed decision-making. Proficient in leveraging critical thinking and problem-solving skills to extract invaluable insights from intricate datasets. A track record of delivering innovative solutions that make substantial contributions to achieving business objectives. Demonstrated proficiency in project management, collaborative cross-functional teamwork, and a dedication to ongoing professional development.

## Experience

---

### DATA ANALYST – **Cedience** – Toronto, ON

Mar 2020 – May 2023

- Accomplished a 40% reduction in data retrieval time by implementing robust database management practices, utilizing **Excel** and **Python** for data cleaning, ensuring seamless data retrieval and analysis.
- Streamlined data management processes, resulting in a 30% increase in productivity and a 20% reduction in errors, by identifying and eliminating inefficiencies.
- Demonstrated a 25% increase in stakeholder comprehension of intricate public health data through the creation of impactful and interactive data visualizations and dashboards, leveraging **Tableau**'s advanced features.
- Assisted in data extraction from regulatory agencies' databases, including FDA, Health Canada, and EMA, ensuring a steady inflow of critical data for analysis.
- Implemented Agile methodology (**Jira**) for sprint planning, enhancing team collaboration and project execution efficiency.
- Developed comprehensive project plans, leading to successful project execution and timely deliverables.
- Collaborated closely with public health experts and software engineers, effectively bridging the gap between data insights and actionable public health strategies.
- Demonstrated adept communication skills in effectively conveying complex data findings to diverse stakeholders, fostering data-driven decision-making across the organization.

### ADMINISTRATIVE ASSISTANT – **Living Realty Inc., Brokerage** – Mississauga, ON

Dec 2018 – May 2022

- Provided vital administrative support to both company employees and clients, utilizing strong interpersonal and communication skills to ensure a positive and productive experience for all stakeholders.
- Developed advanced proficiency in **Microsoft Office** applications, including **Word**, **Excel**, **PowerPoint**, and **Outlook**, streamlining document creation, data analysis, and communication processes.
- Demonstrated exceptional multitasking abilities by effectively managing various clerical tasks, from phone management to document filing, report preparation, and creating professional offers, all while maintaining a well-organized work environment.

### RESEARCH ASSISTANT – **University of Toronto** – Toronto, ON

May 2020 – Aug 2020

- Conducted literature searches regarding teaching and learning resiliency in education using PubMed, Web of Science, and ERIC.
- Actively participated in the ethics approval process and weekly meetings, demonstrating a commitment to ethical research practices and effective collaboration within a team.
- Conducted interviews and facilitated focus groups, and then skillfully managed transcription and data analysis, contributing to the research's data-driven approach.
- Proposed to establish teaching tools, instructional resources, and assessment packages for embedding failing-forward activities in courses across disciplines, showcasing leadership, problem-solving, and a forward-thinking approach to education.

## Education

---

### DATA SCIENCE – **BrainStation** – Toronto, ON

May 2023 – Aug 2023

Data Science Diploma Program

### BACHELOR OF SCIENCE IN HONOURS SCIENCE – **University of Toronto** – Toronto, ON

Sep 2018 – Jun 2022

Majors: Biotechnology, Chemistry

## Projects

---

### [INSTACART GROCERY MARKET BASKET ANALYSIS](#)

In progress

- Conducted a comprehensive analysis of the Instacart dataset, which comprised millions of grocery orders, to uncover insights into customer purchase behavior and patterns.
- Transformed raw data into actionable insights using **Python**, **Pandas**, and **NumPy**, including cleaning, handling missing values, and encoding categorical features.
- Leveraged EDA techniques with **Matplotlib** and **Seaborn** to uncover item associations, enhancing product recommendations and improving user experience.
- Employed machine learning models from **Scikit-Learn** for predictive analysis, contributing to enhanced customer satisfaction and order efficiency.
- Documented and presented findings, enabling data-driven decision-making for the organization.

### [TASTYFINDS: RECIPE RECOMMENDATION SYSTEM](#)

Sep 2023

- Developed a sophisticated food recipe recommendation system using data preprocessing, exploratory data analysis (EDA), and machine learning techniques to match users with personalized and enticing recipes based on their preferences and dietary needs.
- Conducted extensive data preprocessing, including handling missing values, feature engineering, and merging datasets, to create a comprehensive and clean dataset comprising over 200,000 food recipes and 1,000,000 reviews spanning 18 years using **Python**.
- Aggregated and visualized the data by using **Pandas**, **Matplotlib**, **Seaborn**, and **WordCloud** to compile a professional report.
- Gained proficiency in **Jupyter Notebook**, **Scikit-learn**, **NumPy**, and **Surprise** for machine learning applications.
- Deployed the recommendation system using **Streamlit**, showcasing practical deployment skills and providing a user-friendly interface for food enthusiasts.

### [COVID-19 GLOBAL VACCINE TRACKER](#)

Aug 2023

- Created an interactive **Tableau** dashboard to help visualize and analyze worldwide COVID-19 vaccination trends and progress, enabling users to make informed decisions and gain insights into vaccination efforts.
- Leveraged **Python** to create a data scraping script that collected vaccination data from reliable global health organizations and governmental sources, ensuring data accuracy and currency.
- Utilized **MySQL** to design and set up a database to store the transformed vaccination data securely.


## Technical Skills

---

- **Programming Languages:** Python, SQL, R, HTML & CSS, JS
- **Data Analysis:** SQL (Self-joins, Window Functions, Parameters), Excel (VLookup, Conditional Formatting, Pivot Tables Power Query and data validation)
- **Packages:** NumPy, Pandas, Matplotlib, Seaborn, SciPy, ggplot2, Scikit-learn, TensorFlow, Keras, PyTorch
- **Software Tools:** Tableau, Power BI, Data Build Tool, Git, MS Office, Latex, Jira, Docker, PySpark, Google Analytics, SAS
- **IDEs:** Jupyter Notebook, PyCharm, Google Colab, R studio, Visual Studio Code, Kaggle
- **Methodologies:** Agile, Scrum, Kanban, Waterfall
- **Data Engineering:** Data Pipelines, ETL Processes, Data Warehousing
- **Big Data:** AWS (EC2, S3), Hadoop, Hive, Apache Spark
- **Data Warehouse:** MySQL, Snowflake
- **Operating System:** Windows, Linux, Unix

## Professional Certifications

---

- [Tableau Certified Data Analyst](#) 
- [Tableau for Data Scientists](#) | LinkedIn
- [Tableau Essential Training](#) | LinkedIn
- [SQL for Data Analysis](#) | LinkedIn
- [Scientific Computing with Python](#) | FreeCodeCamp