**PEIWEN ZHANG**

peiwen.zhang@mail.utoronto.ca|(+1) 226-978-1295|Toronto, ON, Canada

**PROFILE SUMMARY**

* Proficient in programming languages such as SQL, Python and VBA; database analytics tools such as Snowflake, DBT and data visualization tools such as Tableau, Power BI (DAX), and Sisense.
* Proficient in creating multiple kinds of data visualization dashboards.
* Experienced in designing and implementing data models, writing complex queries/stored procedures, indexing, and data pipelines.
* Solid background in Mathematics and Statistics. Good knowledge of various statistical analysis methods, Natural Language Processing and Machine Learning algorithms.
* Highly motivated and result-driven team player with excellent problem-solving and analytical skills.
* An effective communicator with excellent client-facing skills and written communication skills. Able to explain and document technical concepts for non-technical audiences.

**EDUCATION**

**Master of Information, Human-Centered Data Science & User Experience Design Concentrations** 2022/09 - Present

**University of Toronto** Toronto, ON

**Bachelor of Applied Mathematics & Statistics**  2016/09 - 2020/04

**University of Waterloo** Waterloo, ON

**WORK EXPERIENCE**

**Product & BI Analyst** 2021/09 - 2022/12

**Reebee** Kitchener, ON

* Pioneered the Product Health Dashboard and finalized the product metrics calculations logarithms including App Retention Rates, Bounce Rate, User Tiering, Active Users, etc. to support the product team monitoring the status of our product.
* Created dashboards to support client reporting, product AB testing, daily data observability checking, etc.
* Expertise in the dashboard building process in business requirements documents generation, data model development, dashboard creation and data accuracy validation.
* Translated business requirements into workable requirements at detailed production level using workflow diagrams, sequence diagrams, and use case modelling.
* Implemented metrics and filters and created advanced chart types using calculations to manipulate the data and enable user-driven parameters for visualization.
* Ensure that the data provided are validated and accurate throughout every working process.
* Expert in providing Ad Hoc analysis for the product, sales, and marketing teams.
* Tools and skills: Python, SQL, Sisense Dashboarding, DBT, Git, Google Suite, Trello, Jira, Confluence

**Data Analyst**  2020/07 - 2021/08

**KGS Research**  Toronto, ON

* Collected and analyzed survey responses from political campaigns across the United States, each with 50-100s of survey questions using SQL and Excel.
* Developed SQL procedures to automatically fix and report on data quality issues, improving efficiency by more than 80%.
* Assessed the correlation between survey questions, made recommendations to clients on survey design, reducing the number of questions required to answer by 20%.
* Collected and processed freeform survey verbatim data. Applied sentiment analysis on the survey verbatim and compared it with the Net Promoter Score from surveys.

**Data Analyst**  2018/05 - 2018/08

**China Mobile** *(Largest telco in the world by the total number of subscribers)* Henan, China

* Analyzed possible reasons for signal interference from cell towers and helped the network operations team determine the best location to place new cell towers.
* Created SQL procedures for migrating conversion data to mirror environments and validating data according to the new product and data model.
* Worked with business and technical contacts of 3rd parties Nokia and Huawei to finalize the billing and provisioning interface requirements.

**RELEVANT PROJECT**

**The Study of Factors Affecting Cancer Death Rate: Poverty, Public Insurance and Private Insurance | In Class | 2022/12**

* Analyzed the relationship between cancer deaths with poverty, private insurance, and public insurance using R.
* Responsible for programming the data cleaning process in dealing with null values and outliers, the data visualization using histogram and scatter plots, the OLS regression modelling section and the final report proofreading.
* Discovered that people with better access to healthcare insurance and healthcare insurance have a better chance of surviving cancer.

**Potato Classification |In Class | 2022/12**

* Applied ML classification methods using Python to optimize the potato classifications by maximizing the accuracy and minimizing the prediction error.
* Reduced the dimension of the images dataset and converted the dataset to numerical values to optimize the training time.
* Maximized the training set prediction accuracy to 96% using the KNN method and obtained an accuracy of 89% with a standard deviation of 0.11 after 5-fold cross-validations.