

Julie Zhuo Yang

j.zhuo.yang@gmail.com | 647-949-2566 | North York, ON

Data & Fact Driven Analyst

- Distinguished professional with experience in **data visualization** & Data Science.
- Solid Business analytical skills in Health Care, Inventory, Retail & Banking
- Experienced in all project phases with Waterfall and Agile Methodologies
- Consulting skills on BRD, tackling complex challenges & finding computational solutions

Power BI
Tableau
Python
SQL, Hive
SAS, Deep Learning
Databricks, PySpark
Shell Script, APIs
Snowflake DW

CRM & Campaign
Sales & Marketing
Data Cleansing
ETL, Hadoop
Data Science
AWS, Azure
Google Analytics
DB Modelling

Beam Data, Toronto – Data Science Consultant

May 2022 - Present

- | | | |
|------------|---|---|
| Consulting | — | Multiple project experiences across Automation, Business Intelligence, Cloud Computing and Data Science |
| | ▪ | Facilitated ongoing communication and relationships with stakeholders to translate project requirements into end- to-end data automation solutions and recommendations. |
| Automation | ▪ | Build automated end-to-end pipeline collecting google ranking data through web-scraping. Adopted Ngram analysis to discover new keywords. Ingested data into Metabase and built interactive dashboards. |
| | ▪ | Generated business insight dashboard for SEO performance analyses through calls to the Google Analytics API |
| Dashboard | ▪ | Designed database entities and ensured real-time data accessibility for Metabase Dashboard using SQL queries |
| | ▪ | Created a python based recommendation system, exceeding the client's expectations. |

Environment: Amazon AWS, EMR Cloud, S3, EC2, Hadoop, Metabase, QuickSight, Snowflake, Google Analytics

BioBasic, Markham – Business Intelligence Analyst

Jun 2020 - Mar 2022

- | | | |
|------------|---|--|
| PBI | — | Mainly focused on Power BI Dashboard Design and ML Predictive Analyses |
| Dashboard | ▪ | Developed Customer Services Operation Performance Dashboard using Power BI |
| | ▪ | Automated Weekly Sales Summary Report through Scheduling, Refreshing and Delivering |
| Data | ▪ | Wrote efficient code and queries in SAS programming to assess, clean, validate and analyze large datasets for reporting and quoting purposes |
| Cleansing | ▪ | Co-developed sales forecasting ML model for estimating future sales to optimize inventory level Monitoring |
| | ▪ | Trained and optimized hyper-parameters in XGBoost, extra tree regressor and linear model |
| Predictive | ▪ | Designed primers and gene amplification reaction process using SnapGene. |
| Analysis | | |

Environment: Power BI, SAS, Azure, Snowflake, Python, SnapGene, Tableau

PROJECTS EXPERIENCE – Cloud BigData - Deep Learning - Web Scraping - Automation

- | | | |
|------------|---|---|
| Spark | — | Hands on Experience on Web data collection, Machine Learning, Cloud Computing and Automation Design |
| Processing | ▪ | Big Data: Collected and processed massive tweet data for marketing proposed to access datalake using Scala, PySpark and SparkSQL in DataBricks; Performed sort, join, aggregations, filter and other transformations against unstructured data and delta tables in RDD and Dataframe |
| | ▪ | Deep Learning: Developed a prototype machine learning algorithm to identify various lung cancers across hundreds of CT scan images using FastAI. The algorithm performed with 96% accuracy |
| Machine | ▪ | Web scraping: Automated end-to-end process - collecting google ranking data through web-scraping. Adopted Ngram analysis to discover new keywords. Ingested data into Metabase and built interactive dashboards. |
| Learning | | |
| Web | | |
| Scraping | | |

Environment: Databricks, Scala, PySpark, SparkSQL, Python, Metabase, FastAI, HTML, CSS

EDUCATION

- **Master of Medical Science** **Queen's University** Jan 2018 – Dec 2019
Medical information technology, health research methodology, statistics, project management principles and data management
- **Bachelor of Science & Business** **University of Waterloo** Sep 2010 – May 2015
One-of-a-kind degree that equipped learners with both scientific know-how and valuable business skills

CERTIFICATE – IBM Data Science Professional

Oct 2019 – Dec 2020

Data Analysis and Visualization with Python, Data Science Methodologies, Database and SQL for Data Science

INTEREST: **iHealth app development** – **AI Art** **Science Fiction Books** **PADI Certified Water Scuba Diver**