

Yanni Zhang

☎ (607) 262-3235 ✉ christinazyn996@gmail.com 📁 MyPortfolio 🌐 GitHub 🔗 LinkedIn

EDUCATION

Cornell University

December 2023

Master of Professional Studies in Information Science

GPA: 4.07/4.3

- Relevant Course: Business Data Analysis, Business Intelligence System, Human-Computer Interaction, Object-Oriented Programming and Data Structures, Statistical Methods
- Honors: Grace Hopper Celebration Sponsoring Awards, Graduate Teaching Research Specialist

London School of Economics and Political Science

July 2023

Master of Science in Geographic Data Science

- Relevant Course: Geographical Information Systems, Managing and Visualising Data, Machine Learning

University College London

October 2021

Bachelor of Arts in History, Politics, and Economics

First Class Honours

SKILLS

Technical: SQL, Python (Pandas, NumPy), R (Tidyverse, Shiny, ggplot2), JavaScript, HTML/CSS

Tools: Excel (Pivot Tables, XLookup, VBA), Tableau, Microsoft PowerBI, Figma, QGIS, Geodata

Certifications: Microsoft Certified Azure AI Fundamentals, Tableau Desktop Specialist

COLLABORATED PROJECTS

Cornell Biotechnology Resource Center | *Lab Business Intelligence System*

Sept. – Dec. 2023

Business Intelligence Engineer

Ithaca, NY

- Led a strategic analysis of 200+ attributes in **SQL Developer** and designed an Entity Relationship Diagram to identify key metrics, achieving a 15% improvement in data accuracy.
- Utilized **WhereScape RED** to create dimensional data models with 10+ dimensions and 100+ attributes. Optimized the **ETL** pipeline and achieved a 20% reduction in processing time.
- Presented insights to the Management Team through 2 **Tableau Dashboard** visualizations. Addressed blind spots in reservation rules, resulting in a 20% improvement in scheduling efficiency.

Chamberlain Group LLC | *Propensity to Connect with In-Garage Delivery*

Feb. – May 2023

Data Scientist

Remote

- Utilized SQL and Python to manipulate over **100 million** transaction and user data. Conducted statistical analysis to categorize users into high, medium, or low-frequency engagement levels.
- Engineered **over 100 features** and implemented supervised machine learning pipelines predicting user segmentation on **50 million** records (10GB) in **Databricks**. Achieved a test accuracy of up to 90% using **XGBoost**.
- Reported to the Marketing Team with a list of prospect users labeled with propensity scores, visualized important features, and crafted customized engagement recommendations for different levels.

PepsiCo Inc. North America | *End-to-End Materials Waste Diagnostic*

Sept. – Dec. 2022

Data Analyst

Remote

- Conducted **qualitative interviews** with 10+ supply chain managers across regions to design key performance indicators (KPIs) and metrics for the Materials Waste Management Toolkit.
- Developed waste inference models using supervised machine learning algorithms, achieving an **80%** accuracy rate with **Random Forest**. Enabled effective identification and prediction of waste, contributing to dynamic and real-time decision-making.
- Implemented actionable findings and **waste anomaly detection** by prototyping 5+ **Power BI dashboards**. Provided visibility on data and led to a 100% increase in internal users.