Yanni Zhang

📞 (607) 262-3235 ✉ christinazyn996@gmail.com 🐙 MyPortfolio ✈ Open to relocation

EDUCATION

Cornell University — December 2023
Master of Professional Studies in Information Science (STEM) — *GPA: 4.07/4.3*
- Relevant Course: Business Data Analysis, Business Intelligence System, Data-driven Marketing, Human-Computer Interaction, Data Mining and Machine Learning, Statistical Methods
- Honors: Grace Hopper Celebration Sponsoring Awards

London School of Economics and Political Science — July 2023
Master of Science in Geographic Data Science

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: Bloomberg Market Concepts, 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+ lab usage 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 enabled data automation. 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 Service* — Feb. – May 2023
Data Scientist — *Remote*
- Leveraged SQL to manipulate over **100 million** transaction and user records through joins, sub-queries, and window functions. Developed Python visualizations and 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 90% test accuracy using **XGBoost**.
- Crafted insights and advised the Marketing Team with a list of prospect users labeled with propensity scores, visualized important features, and 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 metrics and indicators (KPIs) for the Materials Waste Management Toolkit.
- Developed reproducible code to analyze each brand and built up waste inference models using supervised machine learning algorithms, achieving an **80%** accuracy with **Random Forest**.
- Enhanced data visibility through prototyping 5 Power BI dashboards, facilitating actionable insights and enabling the detection of waste anomalies. Contributed to real-time strategic decision-making process, resulting a 100% increase in internal users.

EXPERIENCE

Cornell College of Computing and Information Science — Aug – Dec. 2023
Graduate Teaching Research Specialist for INFO5001 Computing for Information Science (R) — *Ithaca, NY*
- Led a collaborative effort with instructors and supervised undergraduate TAs to facilitate recitations, assess coursework, and address inquiries from over 50 students.
- Spearheaded the management of the class electronic mailing list, oversaw updates to the class website, and offered administrative support to the instructor of record.