

Juana Wong

✉ wongjuanaa@gmail.com | 🌐 wong-ja.github.io | 🐙 github.com/wong-ja | 💼 linkedin.com/in/wongjuanaa

EDUCATION

The City College of New York, CUNY

B.S. Computer Science, Mathematics Minor

New York, NY

Expected December 2025

- GPA: 3.3 / 4.0
- Selected Coursework: Introduction to Data Science, Machine Learning, Python for Business Analytics, Data Structures, Algorithms, Database Systems, Scientific Programming, Elements of Probability Theory, Financial Mathematics

PROFESSIONAL EXPERIENCE

The Lohman Laboratory, The City College of New York

Web Designer / Full Stack Web Developer

New York, NY

Nov 2024 – present

- Technologies: TypeScript, JavaScript, React, Next.js, Tailwind CSS, PostgreSQL, Prisma ORM, NextAuth, Vercel.
- Designed and developed an interactive web application for academic professionals to host, filter, create, insert, and manage (CRUD operations) comprehensive taxonomic and geographic data on butterfly mimicry for over 19,000 species, supporting professional classification and research workflows.

CUNY Tech Prep

Data Science Fellow

New York, NY

Jul 2025 – present

- Collaborated and implemented a capstone project using **Python**, **pandas**, and **NumPy** for exploratory data analysis (EDA) and machine learning insights, deploying interactive dashboards with **Streamlit** to communicate results on live sites.

PROJECTS

🔗 Specialty Candy Shop: Data-Driven Plan to Prevent Inventory Shortages | Final Report 🔗 | Presentation 🔗

- Integrated **SQL Server** data tables spanning logistics, sales, targets, and product information, using **Python** libraries such as **Pyodbc**, **Pandas**, **NumPy**, **Matplotlib** and **Seaborn**, finding optimization strategies and critical operational issues.
- Designed a range-based inventory strategy using **Holt-Winters exponential smoothing (statsmodels)** to forecast demand with fixed seasonality, guiding long-term planning on inventory and capacity decisions to a 5-year operational lead time.

🔗 Random Forest Analysis of Health & Poverty Dynamics, U.S. (2015-2025) | Final Report 🔗 | Presentation 🔗

- Synthesized and cleaned 35,000+ multi-source US health, poverty and population records (Google PDFM, County Health Rankings) using **Python**, **Pandas**, **NumPy** & **DuckDB**, with complex joins, data mappings and missing data remediation.
- Built and evaluated multi-target **Random Forest** regression models (**scikit-learn**), using imputation and feature importance to identify smoking rates as the top predictor of poor mental & physical health days; visualized results with **Matplotlib** & **Seaborn**.

🔗 Butterfly Species Classification Using Transfer Learning with ResNet50 | Final Report 🔗 | Presentation 🔗

- Achieved a 69.3% training and 72.4% validation accuracy by training a custom **convolutional neural network** (SimpleButterflyCNN) model using **PyTorch** and **Torchvision** to classify 9,000+ butterfly images across 75 species.
- Applied transfer learning with the **ResNet50**, significantly improving classification results to a 96.7% training and 91.8% validation accuracy (Precision: 0.92, Recall: 0.91, F1-score: 0.91, Support: 1300), amplifying generalization accuracy.

TECHNICAL SKILLS

Languages: C++, Python, SQL, HTML/CSS, JavaScript, TypeScript/Tailwind CSS, Kotlin

Data Analysis & Visualization: Pandas, NumPy, Matplotlib, Seaborn, Plotly, Streamlit

Machine Learning & Statistical Tools: Scikit-learn, TensorFlow, PyTorch, Statsmodels

Databases & Data Management: Firebase, PostgreSQL, Supabase

Collaboration & Development Tools: Git/Github, Visual Studio Code, Jupyter Notebooks, Android Studio

Certifications: CodePath | Certificate in Android Development (Summer 2025) 🔗