

(610) 757-7971
Philadelphia, PA
wangcharles.633@gmail.com

Charles Wang

West Chester University

charleswangprofessionalsite.com
github.com/charles633-wcu
linkedin.com/in/charles-zhou-wang

EDUCATION

B.S Computer Science, *West Chester University* | GPA: 3.72

Aug. 2024 — Present

Relevant Coursework, *Big Data Engineering, Data Structures & Algorithms, Data Science, Computer Systems*

SKILLS

Programming	Python, Java, JavaScript (ES6+), SQL
Data Analytics	Apache Spark, PySpark, Tableau, Power BI, Matplotlib, Data Visualization, ETL Pipelines, Pandas, SciPy
Backend & Web	FastAPI, Bootstrap, Spring Boot, RESTful APIs, SQLite, SQLAlchemy, Microservices, Flask, Django
Cloud & DevOps	AWS, Apache APISIX, Docker, CI/CD, Postman, Containerization
Methodologies	Agile/Scrum, SDLC, Git Version Control
Artificial Intelligence	Claude CLI, Claude Code, LangGraph, RAG Concepts, Recommender Systems (ALS Collaborative Filtering), Context Engineering

PROJECTS

NFL Elo Analytics Platform

nflelodashboard.com

- Built and shipped a live, full-stack sports analytics platform that turns a full NFL season of raw game data into automated team power rankings and an interactive Monte Carlo playoff simulator, serving fans with data-driven forecasts.
- Engineered a batch Elo rating engine in Python computing weekly ratings with home-field advantage, a FiveThirtyEight-style margin-of-victory multiplier, and an OLS margin-prediction model, then validated playoff odds across 100,000 Monte Carlo bracket simulations.
- Designed an in-browser matchup and playoff-bracket simulator that samples realistic final scores via the Box-Muller Normal transform, computing win probability, Elo edge, and projected scores client-side from JSON embedded at build time.
- Architected a containerized multi-service backend (two FastAPI services behind an Apache APISIX gateway) with Docker Compose, deployed on AWS EC2, and published a Jinja2-generated static site to S3 + CloudFront, backed by a pytest suite covering build, simulation, and API wiring.

Crime & Weather Visualization

crimeweatherdashboard.com

- Developed a full-stack interactive platform that uncovers environmental correlations in urban safety data, applying regression analysis to 20 years of NYC records to visualize the relationship between temperature and incident patterns.
- Implemented an ETL pipeline ingesting time-series data from the National Weather Service and NYC crime datasets, performing schema normalization and cleaning prior to storage.
- Architected a containerized, layered system separating frontend, backend services, and data storage, deployed on AWS EC2 with content distributed via CloudFront.
- Built an interactive JavaScript frontend allowing users to dynamically access backend services to explore trends and adjust regression analysis across selectable date ranges.

Scalable Amazon Recommendation Engine

https://github.com/charles633-wcu/spark-amazon-recommender

- Built a scalable recommendation pipeline that transformed 4.62 million Amazon reviews into personalized top-10 product recommendations, reducing the dataset by 79% and achieving a test RMSE of 1.28.
- Engineered a PySpark ETL pipeline to clean, deduplicate, filter, and validate sparse review data using explicit quality checks before storing 953,017 processed ratings in compressed Apache Parquet format.
- Developed and tuned an ALS collaborative-filtering model with Spark MLlib, evaluating latent-factor rank and regularization configurations against an 80/20 train-test split.
- Containerized the Python 3.11, Apache Spark 3.5.1, and Java 17 environment using Docker Compose, enabling reproducible execution across operating systems.

Powerlifting Performance Analysis

kaggle.com/code/charles633/statistical-patterns-in-competitive-powerlifting

- Analyzed large-scale competitive powerlifting data to evaluate performance differences by gender, bodyweight class, and lift type.
- Conducted hypothesis-driven statistical analysis using Welch's t-test and Welch's one-way ANOVA to assess differences in Wilks scores and total weight lifted across groups.
- Modeled relationships between bodyweight and squat, bench press, and deadlift performance using Pearson correlation and linear regression, identifying lift-specific sensitivity to bodyweight changes.

EXPERIENCE

Bartender & Server

Teca

May. 2019 — Mar. 2020

Newtown Square, PA

- Served 50+ guests per shift in a fast-paced restaurant environment while maintaining strong customer service, order accuracy, and professionalism. Developed strong communication, teamwork, and time management skills.

Patient Transport Volunteer

Paoli Hospital

May. 2013 — Aug. 2019

Paoli, PA

- Transported 10+ patients per shift safely and efficiently across hospital departments, helping support timely procedures, discharges, and overall hospital operations.