# Michelle Wang

wang.michelle.yi@gmail.com | 917-847-8851 | www.linkedin.com/in/michelleyiwang | Mountain View, CA 94043

## SUMMARY

Data expert with 7 years of industry experience in building big data analytics pipelines and leading high-impact, consumer facing projects. Actively seeking a full-stack software engineering role to apply deep expertise in web development, machine learning, and data processing. Solid foundation in computer science and data structures, with strong proficiency in **Java, JavaScript, React, Python, SQL, and HTML/CSS.** Fast learner with hands-on experience in cloud platforms (GCP, AWS) and emerging AI tools like Cursor AI. Passionate about building scalable applications from the ground up.

## EDUCATION

**Columbia University**                                                                                                                   New York, NY
*M.S., Applied Analytics* (GPA: 4.1 / 4.0)                                                                              *Aug 2018-Dec 2019*
Relevant Courses: Machine Learning, SQL & Relational Database, Advanced Python for Data Analysis, Analytics with R

**University of Southern California**                                                                                         Los Angeles, CA
*B.S., Business Administration* (GPA: 3.5 / 4.0)                                                                      *Sep 2012-Dec 2014*

## SKILLS

**Languages**: Java, JavaScript, React, Node.js, jQuery, HTML, CSS, R, Python (Pandas, NumPy, Matplotlib, Seaborn)
**Frameworks**: AWS, GCP, Spring Boot, SQL/NoSQL(MySQL, PostgreSQL, MongoDB), LangChain, Docker, Cursor AI
**Expertise**: RESTful API, Web Application, Distributed Computing, CI/CD Integration, Big Data ETL, Agile/Scrum

## TECHICAL PROJECTS

**Staybooking: An online stay rental application based on React and Spring Boot**
- Designed and built a responsive web application using **React**; Bootstrapped the development with **Ant Design**.
- Developed and maintained **RESTful APIs** and **Spring Boot-based microservices**, enabling users to create, search, reserve, and manage stay listings.
- Leveraged **PostgreSQL**, **PostGIS**, and **Hibernate (JPA)** for storing user data and geospatial queries; used **Google Cloud Storage** for managing uploaded media files.
- Implemented stateless and scalable user access using **Spring Security** with **JWT** token-based authentication.
- Deployed the backend to **Google Cloud Run**, and the frontend to **AWS Amplify**, with **CI/CD integration** using **Gradle** and **Git Action** for automated builds and delivery.

**Twitch+: A personalized Twitch resources recommendation engine**
- Built a full-stack **Spring Boot** application for users to search Twitch streams, videos, and clips and receive personalized recommendations based on their favorites.
- Designed a web page with user-friendly UX using **React** and **Ant Design**; Stored game data in **MySQL** database hosted on **AWS RDS**; Conducted **CRUD operations** with **Spring Data JDBC**.
- Integrated the **Twitch API** using **OpenFeign** to fetch real-time Twitch content and built **RESTful APIs** for frontend-backend communication.
- Implemented **Spring Security** for user authentication and developed a **content-based recommendation system** using users' favorite collections.
- Optimized performance with Spring Caching + Caffeine and ensured reliability through unit testing.
- **Dockerized** the application and deployed it seamlessly on **AWS App Runner** for public access.

**Next AI: A full-stack web-based Q&A AI agent for PDF document queries**
- Created an interactive conversational UI leveraging **React** and **Ant Design**, enabling users to effortlessly upload and interact with PDF documents in real-time.
- Implemented **RESTful APIs** via **Express** and **Node.js** and optimized for high-performance request handling.
- Utilized an in-memory **vector store** to cache generated embeddings for efficient retrieval.
- Integrated **OpenAI's GPT-3.5 Turbo API** and **LangChain** technologies, resulting in a sophisticated AI Agent capable of document loading, splitting, storage, retrieval, and output functionalities.

**Global Terrorism Analysis: A PostgreSQL-powered analytics system for identifying terrorism patterns**

- Cleaned dataset with 180,000+ entries and 135 variables using **Python**; designed a relational schema in **3NF** with ER diagram including all integrity constraints and triggers.
- Extracted 23 key variables related to the project scope from original dataset, performed **data cleaning and aggregation** to transform the selected data, and staged the data on destination database by creating a connecting engine with **PostgreSQL server**; validated data consumed against initial reporting using **SQL** queries.
- Built an interactive dashboard visualizing targeted terrorism KPIs using **Metabase**, and summarized findings in a report presented to the United Nations Office of Counter-Terrorism.

**YouTubeTrends: A Python-based analytics project uncovering factors behind YouTube viewership**

- Implemented **data-cleaning** process on 40,000+ observations and 17 variables including both structured and unstructured data; transforming fragmented datasets into a unified dataset for streamlined analysis.
- Utilized **Pandas** and **Matplotlib** libraries to visualize data insights; plotted **correlation matrix** and visualized the result using **heatmaps** to identify statistical relationships between numerical and Boolean variables.
- Applied **sentiment analysis** on video titles, genres, and description keywords to identify and monitor viewer sentiment qualities and computed **word-clouds** to represent frequency counts and importance of words.

## WORK EXPERIENCE

**First Republic Bank**                                                                San Francisco, CA
*Data Analyst, Data Analytics and Automation*                                 *Nov 2022-Jun 2023*
**Scope: Engineered data tools and infrastructure to support Internal Audit's analytics strategy and automation.**

- **Agile Development**: Collaborated in Agile sprints with auditors and engineers to design and deliver internal audit analytics tools and automation pipelines.
- **ETL & Pipeline Engineering**: Built scalable ETL workflows and reusable data pipelines in Python and Snowflake, improving data accessibility and reliability.
- **Data Visualization**: Developed interactive Tableau dashboards to track collateral valuations and identify compliance gaps across thousands of real estate loans.
- **Code Optimization**: Refactored legacy scripts and standardized 1M+ records across 500+ fields to enhance maintainability and accelerate analysis.
- **End-to-End Delivery**: Contributed to full lifecycle of internal tools, from requirements gathering to testing and deployment, in a fast-paced environment.

**ATOMYZE LLC**                                                                              Greenwich, CT
*Data Scientist, Product*                                                              *May 2020-Nov 2022*
**Scope: Led full-stack development of blockchain products for digital asset trading, aligned with user requirements.**

- **Scrum Execution**: Collaborated with tech team in a Scrum environment to design and deliver smart contract-based solutions for tokenizing metal commodities.
- **Product Specs**: Authored detailed PRDs and aligned feature development with customer needs and business goals.
- **Web Development**: Built and launched the company's public website, integrating Google Analytics to track user behavior and optimize engagement; increased traffic by 15% and reduced bounce rate by 20%.
- **Project Management**: Managed product delivery using Jira, improving development velocity and quality through sprint planning and continuous integration.
- **Platform Testing**: Led benchmarking and testing with IBM Blockchain, identifying performance bottlenecks and driving UI/UX improvements based on user research.

**CITIC Asset Management Co., Ltd.**                                                    Beijing, China
*Senior Manager*                                                                        *Jul 2015-Jun 2018*
**Scope: Integrated statistical analysis of operational and financial data to provide investment consulting services.**

- **Financial Modeling**: Assisted in the development of annual business plans targeting healthcare industry by appraising 200+ investment projects using IRR and ERV models; signed off all validated projects and allocated on average $370 million in annual loans.
- **Business Intelligence**: Conducted metadata analysis of existing banking data and performed data categorization; incorporated automation into data capture and transformed data reporting system to incorporate metrics and KPIs.
- **Dashboard Design**: Built dashboards to analyze the competitive landscape and the impact of business plans; translated data into intelligence reports for executives.