

# JJ Gong

## Data Engineer

CPA turned data engineer interested in automating processes & using data to enhance business insights & strategies.

 [gong-jj.com](http://gong-jj.com)  [jjgong7@gmail.com](mailto:jjgong7@gmail.com)  (510) 517-0077  New York City  [linkedin.com/in/justin-gong](https://linkedin.com/in/justin-gong)  [github.com/jjgong7](https://github.com/jjgong7)

## EDUCATION

### Master of Science in Analytics

Georgia Institute of Technology

05/2019

### Bachelor of Science in Managerial Economics

University of California, Davis

06/2014

## WORK EXPERIENCE

### Data Engineer

Pulse Analytics App | 06/2021 – Present | New York, NY

- Automate data extraction & collection processes through web scraping or API calls and perform data transformation, cleaning, and metric calculations on HTML/XML/JSON/CSV/XLSX files using Python/SQL/NoSQL.
- Develop data pipeline and data warehouse architecture in AWS Cloud (using S3, Athena, ECS, Comprehend, Airflow) and deploy Python applications in Docker containers.
- Implement Natural Language Processing in Python and AWS to extract dates of product approvals & indications.
- Perform fuzzy string matching on providers and indication names to improve mapping accuracy from data extraction.
- Calculate performance metrics and reports in Python to automate processes previously done by consultants in Excel.
- Collaborate with engineering, product team, and consultants to expand and improve existing data quality/breadth.

### Data Engineer/BI Engineer

Ghostery, Inc. | 04/2020 – 06/2021 | New York, NY

- Implemented daily Python data cleaning and data pipelining scripts in AWS Lambda & EC2 for data warehousing.
- Generated SQL queries in AWS Athena and Superset for ad-hoc analyses and automated daily metrics reporting.
- Analyzed metrics/KPIs to optimize marketing campaigns, improve the conversion funnel, and guide business strategy.
- Presented weekly and monthly performance reports on business insights to leadership and key stakeholders.
- Developed Apache Superset BI dashboard from scratch and maintained 600 charts/visualizations and 50+ dashboards.
- Collaborated with product, engineering, and marketing teams to design a/b tests and manage existing/new telemetry.

### Data Scientist

Kurv (Contract) | 11/2019 – 04/2020 | Remote

- Developed Python scripts for e-commerce clients to analyze customer purchasing history to upsell/cross-sell products.
- Implemented market basket analysis and collaborative filtering (ALS) on implicit order/purchasing data to make personalized product recommendations for customers, improving product recommendations by 5%.
- Performed sentiment analysis on COVID related tweets in Jupyter Notebooks and observed a positive trend.

### Associate Accountant (CA Certified Public Accountant License: #130384 - Inactive)

Devereaux, Kuhner and Co. LLP | 07/2014 – 02/2017 | San Francisco, CA

- Reviewed income statement, balance sheet, cash flow statements, and financial data of various companies.
- Prepared federal and state income tax returns for individuals, corporations, partnerships, and trusts.

## VOLUNTEER & PROJECTS

INFORMS Pro Bono Analytics – Young Audiences | 04/2021 – Present

- Ensure pay equality for YA's artists through data cleaning, analyzing past trends, and identifying important features.

## SKILLS

### Software and Programming Languages:

Python (Pandas, NumPy, Matplotlib), R (dplyr, ggplot2), SQL (SQLite, MySQL), NoSQL (MongoDB), JavaScript (D3), Git, Jupyter Notebooks, Tableau, Apache Superset, AWS (EC2, Lambda, Athena, S3, Comprehend, Airflow), Excel