

Manny Argueta

[LinkedIn](#) | 818-624-3060 | mannyxargueta@gmail.com | [GitHub](#)

WORK EXPERIENCE

Poseida Therapeutics

San Diego, CA

Data Scientist – Data Engineering

July 2022-July 2024

- Designed and implemented automated data pipelines using Python/NumPy/pandas/SQL to centralize disparate data from CSV, JSON source formats, and web APIs, saving \$90k in consulting fees and an estimated \$120k per year with automation (I won a company SPOT Award for their impact).
- Built and maintained scalable cloud-based data pipelines for business leaders using AWS (Lambda, EC2, S3, Glue, Athena, Step Functions) for efficient data storage, processing, and delivery which resulted in 92% increase in data delivery frequency.
- Developed interactive dashboards (Plotly, Power BI) for process development, manufacturing, donor screening departments and their management teams, providing rapid updates on production progress and allowing them to make fast, data-driven decisions for products.
- Collaborated effectively with cross-functional teams to analyze manufacturing data for root cause identification and yield prediction using machine learning algorithms (PCA, PLS regression).

Boehringer Ingelheim

Fremont, CA

Scientist I - Process Analytics and e-Documentation

March 2020-March 2022

- Leveraged experimental design and statistical analysis, including A/B testing methodologies comparing scaling effects on product quality, providing actionable insights for clients and a minimum of 10% increase in product yield.
- Created customized data visualizations using Spotfire that provided real-time updates for our production cycles and completed an integrated data pipeline from lab instruments to analysis tools.
- Automated data collection from lab instruments to spreadsheets using Python and VBA, as a part of our e-documentation team, saving 20 billable hours monthly for my team and minimizing errors.

EDUCATION

University of California, Riverside

Riverside, CA

MS in Engineering, Data Science (GPA: 3.9)

December 2021

- Capstone Project: Developed a machine learning model (Python, Pandas, Scikit-learn, H2O.ai) using 12,000 patient EHR records to predict in-hospital mortality risk with high accuracy.

University of California, Davis

Davis, CA

BS in Pharmaceutical Chemistry, minor in Human Physiology

June 2017

PROJECTS

Court Reservation Bot

October 2024

- Used Python, Selenium, Git and crontab to build an automation bot to reserve pickleball courts in a competitive area

TECHNOLOGY & TOOLS

Programming Languages: Python, SQL (PostgreSQL)

Cloud Technologies: AWS Certified Cloud Practitioner (Lambda, EC2, S3, Glue, Athena, Step Functions)

Data Analysis, Data Visualization: NumPy, Pandas, Scikit-learn, Jupyter, Plotly, Matplotlib, Power BI

Other: Data Engineering (AWS), Selenium, Git, Kanban, Agile Software Development