

Joseph Lin San Francisco, CA • thejosephlin93@gmail.com • (408) 646-3982 • josephdata.com • in/josephlin11

Multi-faceted Data Scientist with over four years of hands-on experience that includes engineering, data analysis, visualization, automation, and tool development to scale data infrastructure and empower well-informed decision making within data-driven organizations.

EXPERIENCE

Data Scientist, Doximity - San Francisco, CA

Nov 2018 - present

Doximity is the largest network of medical professionals in the US.

- ❖ **Architecture:** Utilizing test-driven development, re-designed and productionized Agent Based Model operating as an artificial ad network for clients to publish on the Doximity platform. Developed simulation to make assessments on potential features i.e. decision science studies. Design user personalization model to tailor content for targeted audiences. Developed [Outlier Detection](#) framework and repository to centralize company data health-checks to alert stakeholders of abnormal system behavior. Utilizing Kimball's data warehouse design, engineered ETL's (PySpark, Pandas, SQLAlchemy, SQL) deployed on Airflow to create dimensional, fact, report tables in Snowflake.
- ❖ **Experimentation:** Scoped, planned, and analyzed all A/B tests touching the Ads products. Ensure tests follow principles of experimental design (randomization, sample size, predetermined success metrics, duration to statistical significance). Built dashboards, performed ad hoc ad data analysis (i.e. targeting, engagement, performance metrics), created supply and demand forecasts, and established Looker infrastructure for the Ads team.
- ❖ **Optimization:** Translate business requirements into ad network strategies (bidding, budgeting, KPI maximization).
- ❖ **Leadership:** DS lead responsible for all analytics, prioritization, and predictive modeling for Doximity Ads, their leading revenue generator. Review technical on-boarding assignments, conduct analytical evaluation of DS hires, as well as train new DS members of the team.

Data Analyst, MZ - Palo Alto, CA

Mar 2018 - Sept 2018

MZ is a global leader in mobile gaming with titles that include Game of War, Mobile Strike, and Final Fantasy XV: A New Empire.

- ❖ **Analysis:** Supported research science, marketing analytics, data platform, and executive team initiatives ranging from fraud detection improvements, optimizing client budgets, modelling partner KPI predictors, and executing ad hoc deep-dive requests for the CRO.
- ❖ **Automation:** Established data pipelines for new clients to meet analytic requirements for Marketing Operations team. Coordinated cross-functionally to ensure reporting SLA's are met for stakeholders.
- ❖ **Forecasting:** Productionize Holt-Winters/ARIMA modules which forecast KPI's for use on workbooks and dashboards across analytics teams. Deployed alerting tool which integrated Google, Slack, MySQL, Vertica API's to share projections with leadership and identify inconsistencies/gaps and operations teams.

Business Intelligence Analyst, Genentech - South San Francisco, CA

July 2015 - Feb 2018

Genentech is a leading biotechnology company that discovers, develops, manufactures, and commercializes medicines.

- ❖ **Modernization:** Developed proprietary scheduling software that coordinated 1900+ production runs among 50+ researchers and reduced scheduling logistics time by 70%. Managed data visualization workbooks for 27 cost centers to monitor \$14M materials budget. Implemented an outlier classification system to notify management of spending anomalies, which led to \$1.2M in first year savings.

SKILLS

Databases: Snowflake, AWS S3, Vertica, Oracle 11g, MySQL

Data Analysis: Apache Spark, Python (NumPy, Pandas, Scikit-learn, SQL Alchemy), PostgreSQL, Excel

Visualization: Looker, Tableau, Superset, D3.js, Seaborn, matplotlib, plot.ly, bokeh

Other Languages: JavaScript, Google Script, Shell, HTML

Concepts/Tools: Agent Based Models, Feature Engineering, Outlier Detection, Time Series Forecasting, Dimensional Modeling, Test Driven Dev, Performance Marketing, AdNetworks (SSP, DSP, SANs), Auction/Bidding Systems, Airflow

EDUCATION & CERTIFICATIONS

B.S. Biochemistry/Cell Biology, University of California, San Diego - La Jolla, CA (2015)