

Juhyeok (Jay) Sin

Milpitas, CA, 95035 | 408 839 3121 | juhyeok.sin91@gmail.com
<https://www.linkedin.com/in/jay-sin> | <https://github.com/JuhyeokSin>

PROFESSIONAL EXPERIENCE

Cuberg Inc. (Acquired by Northvolt)

San Leandro, CA

Data Scientist

Sep 2021 – Present

- Implemented advanced statistical models, including linear regression, the Gaussian distribution, and the Empirical Cumulative Density Function (ECDF), within the battery cycling channel management system. Accurately forecasted channel availability and achieved a 20% increase in channel utilization, enhancing resource allocation efficiency.
- Developed an Anomaly Detection model to detect and classify the process deviations and instantly alert stakeholders. Implemented in a human-in-the-loop architecture, enabling continuous improvement of detection and classification accuracy, while empowering engineers to draw causal conclusions from the data.
- Designed and implemented a data collection with ETL and analytical dashboard to analyze and infer the frequency of part replacement in a manufacturing machine, reducing part replacement frequency by 5-10 times.
- Built a Python script that automatically populates the Statistical Process Control (SPC) dashboards in Looker, ensuring a consistent and predictable level of process quality.
- Support the manufacturing team by organizing data marts and creating dashboards in Looker, such as first pass yield and failure Pareto chart, to track key performance indicators (KPIs).
- Established a PostgreSQL data model for integrating procurement data from an API source and implemented Looker Explore for streamlined KPI dashboarding.

Apple Inc.

Cupertino, CA

Data Analyst

Mar 2020 – Aug 2021

- Wrote a Python script with REST API to scrap image, video data, and corresponding metadata and delivered a few thousands of fully QA'd images and videos weekly.
- Created a visually informative dashboards, enabling quantification of the amount of image and video data across different categories in Tableau Desktop and Tableau Server.
- Managed image and video data QA processes by coordinating with Data Science, Data Annotation, Data Collection teams.
- Contributed to the Exposure Notification project, Apple/Google's efforts to contain COVID-19 pandemic by planning and conducting survey research.

SKILLS

Data Analysis: Python (Pandas, Numpy, Plotly, Scipy, Scikit-Learn, Stats), SQL (Trino, PostgreSQL, MySQL), Statistical Process Control, ANOVA, Machine Learning

Dashboard/BI Tools: Looker, LookML, Streamlit, Grafana, Tableau

Data Model/Pipeline: Python (Prefect and Pydantic), ETL, PySpark

Data Storage/Management: Google Cloud, Dropbox, Bigquery

Developer Tools: Conda, Docker, Git, Unix

Documentation: Atlassian (Jira and Confluence), Lucid Chart, MS Office, G-suite

EDUCATION

M.S. in Analytics

Jan 2024 – Present

Georgia Institute of Technology

B.S. in Applied Mathematics, Minor in Computer Science

Dec 2019

San José State University

Cum Laude