

# HONGCHAO HU

Mobile : (224) 600-6852 | Email : Jason.Hu168@outlook.com | Location : Worcester, MA 01604

LinkedIn: [www.linkedin.com/in/hohu](https://www.linkedin.com/in/hohu) | Website: <https://hongchaohu.github.io/>

## SKILLS

- **Programming Tools:** Python (Pandas, SageMath, Seaborn, Scikit-learn), Java, JavaScript
- **Data Management Tools:** MySQL, SQL Server, MapReduce infrastructure, Spark, HBase, NoSQL Databases
- **Data Science & Modeling:** Exploratory Data Analysis (EDA), Feature Engineering, Clustering, Predictive Modeling, A/B Testing, Statistical Analysis
- **Data Visualization:** Tableau, Power BI, Interactive Dashboards, Data Storytelling
- **Systems & Product Analytics:** User Segmentation, Retention Analysis, Metrics Definition, Data Validation
- **Project & Process:** Agile / Waterfall, Jira, BPMN, ARIS, Sequence Diagrams, Requirements & Data Modeling

## EDUCATION

### Worcester Polytechnic Institute

Aug 2025 - May 2027

- *Master of Science in Data Science*

### The University of California San Diego

Sep 2020 - Jun 2024

- *Bachelor of Science in Mathematics and Computer Science*
- *Minor in Business Economics*

## WORK EXPERIENCE

### Thaddeus Resource Center

Jan 2026 - Present

#### *Data Development Intern | San Dimas, CA | Hybrid Remote*

- Collected, cleaned, and integrated program, donor, and grant datasets using **Python and NoSQL storage**; performed exploratory data analysis to assess program outcomes and funding trends.
- Optimized scalable data workflows using **Spark** with **HBase-backed** datasets for multi-source operational data.
- Built interactive data visualizations and dashboards using Tableau to communicate key metrics, trends, and insights to executive and operations stakeholders. Guided grant prioritization and resource allocation.

### Tanbii

Oct 2025 - Jan 2026

#### *CRM Data Analyst Intern | Nashville, TN | Remote*

- Processed Terminal and in-game behavioral data using **Statbot and MySQL** to uncover retention patterns; applied predictive analytics and clustering to segment user cohorts, supporting lifecycle strategies.
- Translated data analytical insights into weekly and sprint new activity release for optimizing onboarding flow, enhanced in-game sustainability engagement, user segmentation, and in-app conversion rates.
- Built interactive **Tableau and Jupyter dashboards** to visualize user adoption metrics with development decisions.

### Mega IT

Jun 2024 - Jan 2025

#### *Data Engineer | San Jose, CA*

- Performed exploration data analysis using **Python (Pandas, SageMath)** to identify product usage inefficiencies with 94% accuracy and optimize feature rollout timing, decreased developing-wide operational costs by 23%.
- Conducted **ETL Process**, extracted and modeled telemetry data from 5 disparate vendor sources.
- Through **Arise**, monitored feature-level tracking metrics for platform usage; authored actionable roadmap insights.

### Mobalytics

May 2023 - Sep 2023

#### *Data Scientist Intern | Santa Clara, CA*

- Built a client attrition random forest model using Jupyter Lab, improved quarterly retention by 12 bias points.
- Led a geospatial analysis and created forecasting models by **Scikit-learn and SciPy** to predict market trends and optimize monetization strategy; present the model by **MS Power BI**. Eliminated erroneous deliveries by 7%
- Used Python and Excel to segment users and evaluate A/B test results on feature launches and platform changes.

## PROJECT

### Microservice Data Base API Development | UC San Diego

Dec 2022 - May 2024

- Authored Database Microservice API BRD for streamlining data retrieval and reducing third-party dependency.
- Designed and deployed user interaction flows and retrieval pipelines using **ARIS Express and MongoDB**.

## CERTIFICATIONS

- Microsoft **Power BI Data Analyst** Professional Certificate
- Professional Scrum Master (**PSM II**) | Professional Scrum Product Owner (**PSPO II**) by Scrum.org
- AWS Certified Machine Learning – Specialty