KAI-YIN CHAN

Los Angeles, CA | chankaiyin24@gmail.com | github.com/carolyntw | linkedin.com/in/kaiyinchan | kaiyinchan.com | (213) 618-7101

SKILLS

Programming Languages: Python, R, SQL
Data Visualization: Tableau, Power BI, Excel, Plotly, Streamlit
Tools And Frameworks: AWS (EC2, S3, Athena, Glue), GCP (BigQuery), Databricks, Git, Jupyter Notebook, Databases (MySQL, MongoDB)
Soft Skills: Business Communication, Teamwork, Presentation, Time Management

WORK EXPERIENCE

Ark Biotech | Boston, MA   May 2023 – August 2023
Software Engineer Intern
- Halved dashboard loading times leveraging Python and developed tools to enhance cross-functional data analysis capabilities
- Extracted over 100K+ operational time series data points across 9 metrics from TimescaleDB using SQL and Python (Pandas), enabling ad-hoc graph development for contamination identification
- Engineered a real-time anomaly detection system using Databricks and TimescaleDB after identifying data inconsistencies with Python (NumPy), boosting data reliability and operational agility
- Designed and deployed a data automation tool with Databricks and Google Sheets API to streamline data extraction and analysis, achieving a 15% improvement in data quality

NEC Taiwan Ltd. | Taipei, Taiwan   October 2020 – April 2021
System Engineer
- Led customer education initiatives on system usage, achieving a 97% user satisfaction rate and fostering strong stakeholder relationships
- Cut project delivery time by 14% with Agile methods and clear documentation, boosting productivity and project management efficiency
- Reduced system downtime by 10% through detailed data analysis and resolution of discrepancies
- Coordinated with 4 cross-functional teams on system design and deployment, aligning with business goals and fostering teamwork

Taipei Medical University Hospital | Taipei, Taiwan   July 2019 – August 2019
Data Science Intern
- Refined 2M+ patient records using SQL String/Window functions and Subqueries for data cleansing, ensuring reliable inputs for analysis
- Constructed a Logistic Regression model in Python (Numpy, Scikit-learn) to evaluate metabolic syndrome risk, guiding treatment decisions
- Co-developed an automated notification system for high-risk osteoporosis patients by applying Python and SQL, optimizing care efficiency

First Securities Inc. | Taipei, Taiwan   February 2019 – June 2019
Data Analyst Intern
- Spearheaded development of an R Shiny data visualization tool, enhancing SQL data retrieval and interdepartmental reporting, accelerating business intelligence workflows by 15%
- Enhanced decision-making and report clarity by crafting Power BI reports and visualizations in collaboration with securities analysts

EDUCATION

University of Southern California | Los Angeles, CA   January 2022 – December 2023
Master of Science in Applied Data Science
- Core Courses: Machine Learning, Data Mining, Database Systems, Research Methods and Analysis for User Studies

Fu Jen Catholic University | Taipei, Taiwan   September 2015 – June 2019
Bachelor of Arts in Economics (Minor: Statistics and Information Science)

PROJECTS

YouTube Video Trends Analytics   December 2023
- Created an automated ETL pipeline for YouTube analytics using Apache Airflow on AWS EC2, improving data process efficiency
- Managed large datasets with Amazon S3, providing scalable storage solutions for in-depth analytics
- Designed a Tableau dashboard enabling data-driven decision-making with insights into viewer engagement and content popularity

Decoding Lifestyle's Influence on Success Metrics   November 2023
- Investigated lifestyle and success correlations using Python, focusing on statistical methods like t-tests and regression
- Identified sleep hours as a key predictor of success, demonstrating skills in data preprocessing, analysis, and insight generation
- Ensured study reliability through power analysis, achieving a precision level of over 95%, strengthening the validity of research findings

Sales & Customer Insights Dashboard   October 2023
- Orchestrated a comprehensive Business Intelligence solution using AWS (S3, Glue, Athena) and Tableau for e-commerce data analysis
- Analyzed trends in e-commerce datasets, uncovering correlations and visualizing insights with charts