

Mark Le

☎ 657-339-3225 | ✉ mark.mqle@gmail.com | 🔗 linkedin.com/in/minh-le-mark | 🌐 markqle.github.io | Open to relocate

TECHNICAL SKILLS

Programming Languages: Python, Spark, R, Django, JavaScript

Database Management Systems: MySQL, Elasticsearch, PostgreSQL, MongoDB, DataBricks, Snowflake, MS Access

Cloud Services: AWS (EC2, S3, Athena, QuickSight, Lambda), Azure, Databricks, BigQuery, Docker

Data & Workflow Tools: Power BI, Tableau, Alteryx, SPSS, scikit-learn, Pytorch, Tensorflow, Airflow, Kafka, n8n Automation, Jira, Statistical Analysis (t-test, chi squared), A/B Testing, Data Modelling (DBSchema, draw.io), Jenkin CI/CD, Postman

WORK EXPERIENCES

Data Management Student Assistant, California State University Long Beach

Feb 2025 – Present

- Managed Career Development Center database using MySQL to maintain 45K student records, ensuring 99% data accuracy and security.
- Engineered Python script for automated data entry for the WorkAbility IV program, and data validation resulting in 50% increase in productivity.
- Leverage data analysis tools with Python, Power BI and MS Access to extract insights to enhance efficiency and decision making for stakeholders.

Data Analyst, GeoComply Solution Inc.

Feb 2023 – Aug 2024

- Developed custom metrics for API utilization, including throttle control alert for usage exceeding 1 million requests per hour, and automated alerts in Python for continuous monitoring using Airflow. This reduced human error in system checks and accelerated daily operations by 15%.
- Built data pipelines with Python Kibana Elasticsearch integration tools using signals from GeoComply Core Solution internal databases to enrich 3 millions data points, improving by 40% more data for GeoGuard database.
- Conducted in-depth investigations and created large-scale ad-hoc reports using 100M+ data records on Databricks on system fraud detection per request for upper management and customer inquiries.
- Proactively monitored Jenkins automation cron jobs, performed testing and data quality assurance to ensure data integrity by identifying failed or unstable job runs during daily operations to achieve near 0% downtime.

Jr. Data Analyst, LF Global Tech Inc.

Sep 2021 – Dec 2022

- Developed interactive Business Intelligence dashboards with Power BI to enhance market research for company concept testing in the North American, Brazil and Mexico market, resulting in 25% increase in actionable insights and improved decision-making efficiency.
- Cleaned raw data coming from a third party partner and analyzed the 500,000 data points using Python.
- Conducted Statistical Analysis and ad hoc report using Python and Power BI with DAX and assisted the System Analyst team in data modeling and building Entity Relation Diagram and UML.

Business Intelligence Intern, EMC Business Solution

Oct 2020 – Apr 2021

- Leveraged internal customer review data using MySQL to identify product improvement opportunities with SWOT analysis and drive strategic expansion to uncover competitive insights and industry trends.
- Collaborating with business partners to enhance their integrated marketing, financial strategies for digital transformation.

PROJECTS

- **AdFraud Detection**, [GitHub](#): Developed a Decision Tree model to classify risky transactions and detect attribution fraud in mobile app marketing. Improved fraud detection accuracy to enhance marketing analytics and decision-making.
- **Redfin Web Scrape & Analysis**, [GitHub](#): Developed a Python web scraper to extract real estate data from Redfin and performed analysis to identify pricing trends and market insights in Toronto and Vancouver.
- **Bigram LLM**, [GitHub](#): Developed a Bigram Language Model to predict the next words based on the script of *Star Wars: Episode V*. Implemented text preprocessing, NLP transformer techniques to enable accurate next-word prediction.

EDUCATION & AWARD

Master in Information System (Data Science major) - California State University Long Beach

Aug 2024 - May 2026

Bachelor of Arts in Economics - University of British Columbia

Aug 2017 - May 2021

AWS Academy Cloud Foundations - AWS Academy Graduate

Aug 2024 - Dec 2024

3rd place - Alteryx South California Datathon 2024 (100 participants)

Oct 2024