Vincent Lau

(236) 867-7290 | Vancouver, Canada | [Website] | [Email]

SUMMARY

Experienced Data Engineer with a strong background in designing and optimizing data pipelines, maintaining cloud-based AWS database, and ensuring data integrity. Skilled in ETL data processes, as well as in collaborating with co-workers and customers to understand data needs. Proven ability to transform raw data into valuable insights to support data-driven decision-making.

TECHNICAL STACK

Programming Languages: Python	Shell	SQL	**Concept**: ETL	BI	Scrum Development
Databases: MySQL	DB2	Postgre	**Tools**: Linux	Airflow	Tableau	PowerBI	Kettle	Azure

PROFESSIONAL EXPERIENCE

Data Engineer FreeLancer, Glocal, Vancouver, CA   **May 2024 – Present**

Led a volunteering **dashboard** project aimed at highlighting the scale and impact of the volunteer network, showcasing key demographics and metrics to attract more volunteers and secure additional funding and support.

- Streamlined data preparation by cleaning and normalizing datasets using **Excel, Power Query, and Python Pandas**, resulting in accurate key metrics on dashboard and normalizing **85%** of previously inconsistent data.
- Developed an intuitive **Power BI** dashboard with real-time data updates and cloud integration, enhancing user accessibility and resulting in a **40%** increase in dashboard clicks.

Database Administrator, Defend Technology, Vancouver, CA   **Jan 2023 – May 2024**

Designed and managed the database architecture for a dating app, creating **30+** well-structured table schemas, achieving high query efficiency and system performance.

- Built 10+ data pipelines by collaborating with business departments to understand their needs, using Google Cloud Dataflow and Airflow for efficient ETL operations, resulting in quicker data transformation and a 16% improvement in data processing efficiency.
- Developed and optimized complex **SQL** queries and created reasonable **index**, improving query speed by 35%.

Data Engineer Team Lead, Rainer Technology, Beijing, CN   **Oct 2019 - Jun 2021**

Led a team of 3 data engineers in designing and implementing scalable data solutions to support various business functions. Discussed customer needs and documented their requirements for projects, ensuring all data solutions met customer expectations.

- Created data mapping documentation based on customer requirements.
- Designed and implemented scalable data models, increasing data query performance by 35%.
- Optimized existing databases, resulting in a 20% reduction in storage costs and improved system performance utilizing **DB2**.
- Used **DB2** to write **stored procedures** to extract and transform data, enhancing data processing accuracy and efficiency.
- Designed and optimized **ETL** pipelines using **Python** and **SQL**, reducing data processing time by 30%.
- Automated data loading and reporting processes, decreasing manual effort by 80% and improving accuracy using **Shell** scripting.
- Led efforts to migrate local data infrastructure to **Azure** cloud services, enhancing scalability and reducing maintenance costs.

BI Engineer Internship, Huaxia Diantong IT, Beijing, CN   **June 2019 - Oct 2019**

- Managed **BI** dashboards showcase government key metrics, supporting management adjustment.
- Established data warehouses to ensure data consistency from business systems to dashboards, meeting government data standards.

Data Engineer Technical Lead Internship, Guangxi Normal University, Guilin, CN   **Jan 2018 - May 2019**

Led a team of 2 student interns in developing and optimizing data engineering solutions, providing mentorship and guidance to new team members, fostering their professional growth and development in data engineering practices.

- Collaborated with cross-functional teams to identify and address data needs, resulting in improved data quality and accessibility.
- Utilized **MySQL** database for data storage and management, ensuring data integrity and reliability.
- Developed and maintained **ETL** processes, enhancing data extraction and transformation accuracy by 20% using **Kettle**.
- Implemented **Airflow** for scheduling and orchestrating complex data workflows, improving data processing automation.
- Reduced database downtime by 30% by implementing a robust backup and recovery system utilizing **MySQL** and **Azure** services.

EDUCATION

University of Nottingham   **Nottingham, UK**
M.S. in Artificial Intelligence   **Jul 2021 - Dec 2022**

- Graduated with Merit

Guangxi Normal University   **Guilin, China**
B.S. in Computer Science   **Sep 2016 - Jul 2020**

- GPA: 3.53/4.0