许华龙

高级数据工程师

菲律宾 | +63 9155511196 | brian.xu.work@gmail.com | linkedin.com/in/brianxu1997

个人简介

拥有 4 年以上经验的高级数据工程师，擅长使用 Python、SQL 和 Snowflake 构建并自动化数据管道。专注于数据质量、性能优化，以及交付可靠、可直接用于分析的数据。

工作经历

高级数据工程师，Escalent 2022 – 至今

- 构建并维护可扩展的数据管道，用于摄取、转换、验证和发布关键业务数据集。
- 自动化数据交付流程，将交付时间最多缩短 99%，并减少人为错误。
- 使用 Python、Pandas 和 Polars 处理大规模结构化与半结构化数据集，性能提升约 50%。
- 在 Snowflake 中编写并优化 SQL，用于数据转换、对账以及生成可直接用于分析的数据集。
- 设计自动化数据质量检查、异常检测和监控机制，识别出约 90% 的数据问题。
- 通过标准化和自动化部署流程，将部署时间缩短 67%。
- 使用开源 Laya 模型开发情感分析项目，对文本数据进行情感分类与评分。
- 使用 Python 创建数据可视化，供分析师查看数据概况、质量趋势和关键指标。

初级 Python 开发工程师，Torchbox 2020 – 2022

- 使用 Python、Django 和 Wagtail CMS，为 NHS、Oxfam 和 M+ 等客户构建全栈 Web 解决方案。
- 为高流量平台开发后端功能、API 集成和 CMS 定制。
- 参与 Wagtail 开源生态系统的贡献。

教育背景

Information and Communications Technology Academy 2016 – 2020

计算机科学学士（软件工程方向），以优异成绩毕业（Cum Laude）

证书

Data Engineering Professional Certificate, Snowflake
Associate Data Engineer in SQL, DataCamp
Associate Data Engineer in Snowflake, DataCamp

技能

数据工程：ETL/ELT 数据管道、工作流自动化、数据质量与验证、性能优化
语言与查询：SQL, Python, Pandas, Polars
可视化：Python (Matplotlib, Seaborn, Plotly)
数据库：Snowflake, PostgreSQL, MySQL
框架：Django, Flask, Streamlit
网络爬虫：Scrapy, Selenium, Playwright
云与系统：AWS (S3, EC2), Linux
工具：Git, CI/CD, Autosys

语言

英语 • 中文 • 菲律宾语