FRANCIS NEAL ALTARES

Toronto, ON | +1 (437) 559-4509 | altares.fneal@gmail.com
https://www.linkedin.com/in/francisnealaltares | https://github.com/FNeal10

PROFILE SUMMARY

Data Engineer with extensive experience building high-performance data solutions across telecom and IT environments. Specialist in architecting scalable pipelines that transform complex, multi-source data into actionable business insights. Expert at bridging legacy database systems with modern cloud stacks, with a proven track record of automating workflows and optimizing large-scale data ecosystems for reliability and speed.

SKILLS

Core Data Engineering
Azure Data Factory, Snowflake, dbt, Airflow, PySpark, Python (Pandas/NumPy), SQL

Cloud & Storage:
Azure Data Lake, AWS S3, PostgreSQL, MongoDB

Pipeline Automation
CI/CD (GitHub Actions), Docker, Bash/Shell Scripting, Flask REST API

Data Modelling & SQL
Star Schema, Snowflake Schema, Stored Procedures, Functions, Triggers, Views, CTEs, and Window Functions

Business Intelligence & Reporting
SSRS, IBM Cognos, and Power BI

STRATEGIC PROJECTS

End-to-End Cloud Data Pipeline (Azure, Snowflake, dbt)
- **Engineered a custom ingestion framework** using **Python (Selenium)** and **Flask REST APIs**, designed for seamless adaptation to **AWS (S3)** or **Azure (ADLS)** landing zones.
- **Architected a multi-environment workflow**, utilizing **PostgreSQL** for local development and **Snowflake** as the production data warehouse to optimize compute costs.
- **Developed modular ELT models in dbt**, implementing automated data testing and documentation to ensure high-quality datasets for downstream analytics.
- **Integrated a Lakehouse pattern** by landing raw data in ADLS Gen2 and automating the schema migration and loading process.
- **Deployed containerized pipeline components** using Docker and GitHub Actions (CI/CD) for serverless execution on Azure Container Instances.

Sentiment Analysis of YouTube Comments (Academic Project)
- **Scalable Data Ingestion:** Developed a system to pull metadata and **1 million comments** from 1,000 YouTube videos, managing API limits and large-scale data flow.
- **Distributed Processing:** Used **PySpark** to build a batch workflow that cleaned and organized the data, making it 10x faster than traditional processing methods.
- **Text Refinement:** Applied **NLP techniques** (like cleaning and tokenization) to turn messy social media comments into high-quality data ready for analysis.
- **Structured Architecture:** Organized the data into layers (**Raw, Processed, and Analytical**) to ensure the system was easy to debug and maintain.

WORK HISTORY

IT Support and Automation
Variety Village Children's Charity of Ontario | Scarborough, ON | JULY 2024 – January 2026
- **Engineered** complex SQL queries, stored procedures, and views to retrieve, analyze and transform data for reporting and analytics
- **Optimized** CRM-to-SharePoint integrations, ensuring data consistency for stakeholder reporting.
- **Automated** SharePoint and Microsoft 365 processes, including account provisioning, list updates, and workflow management, reducing manual effort and errors by **40%.**

Operating System and Database Administration Lead (L3)
Globe Telecom Inc. | Manila, Philippines | MARCH 2022 – April 2023
- **Reduced database incidents by 60%** through proactive monitoring, log analysis, and performance tuning of production systems.
- **Standardized database architecture** and design patterns, streamlining operations and improving system scalability
- **Resolved critical L3 outages** for high-availability systems (99.9% uptime) via deep-dive root cause analysis.
- **Optimized change management** in ServiceNow, accelerating issue resolution and improving cross-team data reliability.

Application, Backend and Data Developer
Globe Telecom Inc. | Manila, Philippines | MARCH 2012 – March 2022
- **Engineered end-to-end ETL pipelines** using **Python**, **Bash**, and **SQL** to ingest data from APIs, FTPS, and JSON into central storage, streamlining multi-source data integration.
- **Architected relational schemas and data models**, optimizing indexing and partitioning strategies to support scalable, high-performance analytics.
- **Developed advanced SQL objects** (Stored Procedures, CTEs, Triggers) to automate complex data cleaning and transformation logic.
- **Improved data reliability by 25%** by implementing automated validation frameworks and quality checks in Python.
- **Boosted query performance** on large-scale telecom datasets through rigorous execution plan analysis and query tuning.
- **Built self-service data tools** using **C# ASP.NET**, enabling business units to access critical reports without manual IT intervention.
- **Ensured system stability** for mission-critical data systems through 24/7 production support and rapid incident resolution.

CERTIFICATIONS
- Microsoft Data Fundamentals DP-900, Microsoft, July 2024
- Google Data Analytics Professional Certificate, Coursera, February 2024
- IBM DevOps and Software Engineering Certificate, Coursera, December 2023

EDUCATION
LAMBTON COLLEGE, Toronto, Canada
Postgraduate Certificate in Artificial Intelligence and Machine Learning

MAPUA INSTITUTE OF TECHNOLOGY, Makati, Philippines
Bachelor of Science in Information Management

MAPUA INSTITUTE OF TECHNOLOGY CENTER, Makati, Philippines
Associate in Management Systems Technology