

HAROLD ESCARRILLA

A collaborative IT professional with extensive and wide experience in database management, application development, data engineering, data warehouse, data analysis, and business intelligence. Successfully handled various roles and tasks in a resource-constrained projects with aggressive timelines. Proven ability to quickly learn new technologies. Passionate learner and always keeping up to date with new IT technologies and trends through external trainings and certifications.

PROFESSIONAL EXPERIENCE

Tabularasa HealthCare

April 2019 - Present

Senior Data Engineer – Google Certified Professional

- Build the infrastructure required for extraction, transformation and loading of data from different sources using SQL and AWS technologies
- Create and maintain optimal data pipeline architecture
- Automate manual processes and optimizing report delivery
- Work with data scientists to strive for greater functionality in our data systems

Citigroup, New York, NY

Oct 2013 – April 2019

IT Business Lead Analyst – Senior Data Warehouse Developer

Development lead of the Compliance Data Warehouse and eventually became application manager of a reporting application using Cognos Analytics

- Designed and built ETL processes using shell scripts and Sybase IQ stored procedures
- Data sourcing from various groups within Citi as well as external data providers such as Reuters to load into Compliance Data Warehouse
- Managed and implemented the migration of a reporting application that uses Compliance Data Warehouse to IBM Cognos Analytics
- Gathered requirements from Compliance business users to create or modify compliance and regulatory reports
- Led code review and coordinated development activities between onshore and offshore teams
- Data analysis to resolve data issues reported by Compliance business users
- Assist business users in formulating and tuning performance of SQL queries

Morgan Stanley, New York, NY

Feb 2006 - Oct 2013

Data Engineer - Lead ETL Developer

As a member of the Data Team, design and develop ETL processes for data mart to support various reporting groups in Compliance IT

- Interacted with business analysts and application managers to gather requirements
- Authored source-to-target data mapping specifications
- Developed scripts and utilities using Unix shell scripts, Perl and Q for ETL processes
- Designed, coded and implemented file handler utility to perform validation checks to ensure data files provided by various data sources are valid before further processed
- Developed DB2 stored procedures to extract data. Develop and tune SQL queries for ETL processes
- Troubleshoot complex data issues and recommend, test and implement solutions
- Subject-matter expert in market data and products reference data

Merrill Lynch, New York, NY

Jan 2000 - Feb 2006

Development, maintenance and production support of a global financial data repository and workflow system

- Developed programs for extraction, transformation and loading of data into the global repository using T-SQL and Informatica
- Redesigned, built, and implemented the global batch cycle job flow using Autosys scheduling software to ensure data is available on time
- Ensure the successful and timely completion of the processing cycle by resolving data-related or infrastructure-related problems

TECHNICAL SKILLS

Languages: Python, SQL, Shell, Perl, Q, T-SQL, PL/SQL

Databases PostgreSQL, Sybase IQ, Netezza, DB2, Sybase, Oracle, SQL Server, CouchDB

Software: Spark, Cognos Analytics, RStudio, GitLab, GitHub, Jenkins, BitBucket, Autosys, Informatica

Platforms: Google Cloud Platform (Pub/Sub, Dataflow/Beam, Dataproc, BigQuery), Unix/Linux, AWS

EDUCATION

Bachelor of Science in Mathematics

Adamson University, Manila, Philippines

CERTIFICATIONS AND TRAININGS

Data Engineering on Google Cloud Platform Specialization (Coursera) - [Certificate Link](#)

- Google Cloud Platform Big Data and Machine Learning Fundamentals
- Leveraging Unstructured Data with Cloud Dataproc on Google Cloud Platform
- Serverless Data Analysis with Google BigQuery and Cloud Dataflow
- Serverless Machine Learning with Tensorflow on Google Cloud Platform
- Building Resilient Streaming Systems on Google Cloud Platform

IBM Applied Data Science Specialization (Coursera)

- Python for Data Science - [Certificate Link](#)
- Data Analysis with Python - [Certificate Link](#)
- Data Visualization with Python – On-going
- Applied Data Science Capstone – On-going

Big Data Specialization (Coursera)

- Introduction to Big Data - [Certificate Link](#)
- Hadoop Platform and Application Framework - [Certificate Link](#)
- Introduction to Big Data Analytics - [Certificate Link](#)
- Machine Learning With Big Data - [Certificate Link](#)

R Programming - [Certificate Link](#)