

ESTHER-ALEXANDRA NGALEBA-ZE

730 Dovercourt Road, Toronto, ON
647-916-6123 | ngaleba@gmail.com

TRANSFORMING BUSINESS WITH DATA STRATEGY, CLEANSING, ETL, AND AUTOMATION

I am a passionate data practitioner with over 5 years of experience in the data management field, specializing in optimizing and building scalable data systems. With a strong background in database management, data orchestration, and workflow automation, I have worked across various roles, including data developer, data analyst, and CRM database specialist. I am skilled in working with a wide range of database management systems, designing data pipelines and implementing efficient processes to collect, clean, and transform data for actionable business insights. My extensive experience across numerous facets of data management allows me to collaborate effectively with cross-functional teams and stakeholders to deliver high-quality data solutions that drive digital transformation and support business goals.

CORE SKILLS & COMPETENCIES

- | | | |
|--|--|---|
| • Data Management, Modeling, Analysis & Visualization | • Salesforce.com | • REST API |
| • Data Quality Assurance | • SAP HANA | • MySQL |
| • Data Pipelines, Conversion, Migration & Integration | • SAP S/4HANA | • SQL |
| • Data engineering | • SAP Data Services | • NoSQL |
| • Database Configuration & Cleaning | • Docker | • Python |
| • Big data | • Tableau | • PySpark |
| • Cloud Computing | • Microsoft Power BI | • HTML |
| • Report & Dashboard Development | • Microsoft SQL Server | • Linux Command |
| • Financial & Accounting Concepts | • Microsoft Visual Studio Code | • Shell script |
| • Agile Project Management | • Snowflake | • Elasticsearch |
| • Budgetary Control | • DBeaver | • Kibana |
| • Resource Management | • Jupyter Notebook | • Jira |
| • ETL | • Databricks | • Confluence |
| • Microsoft Dynamics 365 | • Apache (Spark, Spark Streaming, Parquet, Hadoop, Hive, Airflow, Kafka, Superset, Lucene, Nifi) | • Microsoft (Excel, Access, Word, PowerPoint, Outlook) |
| • AWS (EC2, S3, Lambda, EMR, Glue, Athena, CloudWatch, DynamoDB, MSK, CLI) | • Debezium | • Windows 11, Linux, Ubuntu, iOS |
| | • Zookeeper | • Process Improvement |
| | • Web Scraping | • Problem Solving & Organizational Strong Interpersonal, Verbal & Written Communication |

PROFESSIONAL EXPERIENCE

- Data Cleanser Specialist | [IPEX \(Via PrecisionERP\)](#)** **Mar 2023 – Nov 2024**
- Overview: Contributed to data preparation and profiling efforts to ensure data quality for a large-scale data migration project, where cleansed and validated data were transferred to a new SAP platform.
- **Data Cleansing & Transformation:** Leveraged advanced Excel functions (VLOOKUP, INDEX, MATCH, IF combined with AND/OR) to process customer and material master data, including customer contact information and product (raw, semi-processed, and finished goods) dimensions ensuring alignment with SAP standards. The cleansed data was subsequently uploaded into a staging area platform, which was later integrated into SAP.
 - **Data Quality & Governance:** Collaborated with the data governance team to maintain and validate a city master list (containing cities, regions, provinces, states, and countries). Ensured the data matched predefined SAP codes and was accurate, ensuring consistent data across multiple systems and preventing discrepancies in SAP migration.
 - **Data Cleansing Standards & Automation:** Assisted in the development of data cleansing rules to enforce consistency across datasets and improve overall SAP warehouse efficiency. This included defining standard processes for data validation, ensuring scalable, automated data processing workflows.
 - **Data Analysis & Problem Resolution:** Conducted data analysis to identify data quality issues, assess their scope, and develop effective solutions. Collaborated with relevant teams to implement fixes and optimize data pipelines for the SAP migration.
 - **Issue Tracking & Collaboration:** Used Jira to monitor data quality issues and assign tasks to relevant team members, ensuring quick and efficient issue resolution. This improved team collaboration and ensured that the data cleansing process was synchronized with SAP project milestones and data migration schedules.

ESTHER-ALEXANDRA NGALEBA-ZE

730 Dovercourt Road, Toronto, ON
647-916-6123 | ngaleba@gmail.com

Data Engineering Intern | [WatchCharts \(via Weclouddata\)](#)

Jan 2024 – March 2024

Overview: Designed and deployed end-to-end data pipelines to scrape and transform luxury watch listings from Chrono24, Crown & Caliber, and Analog Shift using Python-based tools (Selenium, Scrapy, BeautifulSoup). Orchestrated workflows with Apache Airflow (Docker) on AWS EC2 with Microsoft VS Code, storing cleaned data in S3 and PostgreSQL (via DBeaver), and integrated it into WatchCharts' pricing intelligence system alongside marketplace data from platforms like eBay.

SAP Data Services Developer | [SAAQ \(Via VisionSoft Canada Inc.\)](#)

Nov 2021 – Aug 2022

Overview: Designed and implemented scalable data pipelines and ETL workflows that support the integration and transformation of large datasets within an enterprise data ecosystem.

- Data Extraction & Integration: Extracted data from legacy systems and integrated it into a centralized data warehouse using SAP Data Services. Configured ETL jobs to efficiently transfer data and ensure smooth migration from legacy systems to the new platform.
- Data Pipeline Development: Developed and optimized complex data workflows, to meet business requirements, leveraging SQL and SAP Data Services Designer (GUI) along with its built-in functions. This ensured the correct sequencing of ETL (extraction, transformation, and loading) operations across multiple data sources, enabling seamless and accurate data flow.
- Data Transformation & Validation: Designed data flows that replicated ETL processes through SQL-based transformations, validating and cleansing incoming data to meet business rules. Developed data validation checks to ensure only accurate, consistent, and relevant data was loaded into target systems.
- Automation & Workflow Orchestration: Automated data processing tasks and orchestrated workflows to streamline recurring data operations, reducing manual interventions and improving pipeline efficiency.
- Testing & Debugging: Used SQL and Microsoft SQL Server to rigorously test and validate the robustness of data pipeline components, ensuring the stability and reliability of the data pipeline infrastructure. Performed unit tests on individual jobs and integration tests to verify seamless interaction between modules within the data pipeline.
- System Integration: Executed IDOCs and configured job orchestration to ensure seamless data transfer between systems, specifically integrating SAP S/4 HANA into the larger data environment, enhancing system interoperability.

Database Administrative Assistant | [OMERS \(Via Hays\)](#)

Oct 2021 - Jul 2022

Overview: Led data cleansing initiatives and designed processes to improve data quality and operational efficiency within the Microsoft Dynamics CRM system, contributing to more accurate data management and reporting.

- Data Cleansing: Cleansed and updated over 60K CRM records, collaborating with business stakeholders to ensure the accuracy and relevance of the data.
- Data Deduplication & Consolidation: Developed and executed processes to identify and merge duplicate and obsolete CRM records, improving the integrity and usability of the database.
- Collaboration for CRM Reporting: Worked closely with developers to design and configure custom CRM reports that provided real-time insights into data health.
- Data Quality Monitoring & Dashboard Development: Supported the creation and maintenance of an Excel data quality dashboard, enabling stakeholders to track the progress of the data cleaning project and measure improvements over time.
- Documentation & Process Improvement: Documented the data cleaning procedures and best practices, creating reusable guides to streamline future data management efforts. Developed templates and workflows for ongoing data maintenance, ensuring long-term consistency and quality.

CRM Database Specialist | [Porsche Cars Canada Ltd \(Via Adecco\)](#)

Jan 2021 – Jul 2021

Overview: Led database cleansing efforts to optimize customer data management within an SAP-powered CRM platform, ensuring data quality, system efficiency, and actionable insights for marketing and sales teams.

- Data Quality Management & Data Cleaning Project Planning: Designed and implemented a comprehensive data cleaning strategy for the CRM system, identifying data issues, establishing resolution methods, and defining quality metrics. Automated the monitoring of data quality, ensuring visibility and timely reporting for stakeholders and enabling ongoing improvements.
- Project Leadership & Resource Management: Managed a cross-functional team through all stages of the data cleaning project, utilizing project management methodologies to allocate tasks, estimate costs, and manage resources. Ensured project milestones and deadlines were met while staying within budget, optimizing both human and technical resources.
- Data Collection & Preparation: Created documentations, to gather sales related information from dealers, and used them, in collaboration with affiliated agencies, to update the CRM virtually eliminating data issues associated with approx. 20K fields within the database which positively affected dealers' KPIs.
- Data Integration: Updated 100K+ clients' information in SAP by integrating external data sources to update empty fields, enhance marketing communication and sales compensation process.

ESTHER-ALEXANDRA NGALEBA-ZE

730 Dovercourt Road, Toronto, ON
647-916-6123 | ngaleba@gmail.com

- **Data Pipeline Optimization & Workflow Automation:** Re-engineered notification workflows to automate and streamline the flow of customer inquiries from online platforms to Porsche Centers. Optimized processes to reduce lag time and enhance real-time data updates, improving customer engagement.
- **Advanced Analytics & Segmentation:** Leveraged SAP HANA to perform client segmentation and generate actionable insights for decision-making.
- **Reporting & KPI Dashboards:** Developed dashboards with Excel and PowerPoint that showcased KPIs and work progress related to the data cleaning project. Presented reports and insights to management, sales teams, and external stakeholders, facilitating informed decision-making.
- **System Configuration & Data Enrichment:** Configured CRM fields to enrich data and improve lead status tracking, enhancing sales forecasting and performance analysis.
- **Documentation & Knowledge Sharing:** Created comprehensive documentation and guides in English and French, detailing CRM best practices and usage tips. Published guides through company newsletters and shared insights through presentations to sales teams, promoting better data management practices and improving overall data quality.

Data Analyst I | [GE Grid Solutions \(Via Abacus Service Corporation\)](#)

Aug 2020 – Nov 2020

Overview: Meticulously analysed, stored, and structured data related to installed products within Salesforce, ensuring it was organized for easy access and retrieval. This effort was aimed at supporting the broader organization's objectives, enabling more informed decision-making and helping drive the achievement of key business goals.

- **Data processing:** Performed data analysis, data cleansing, and generated reports on large datasets.
- **Report generation:** Developed reports using Salesforce and MS Excel Pivot Tables to track data updates and cleaning progress.
- **Data extraction:** Developed data extraction strategies from multiple sources and tools to capture relevant data.
- **Stakeholder collaboration:** Collaborated with regional IB leads to evaluate data requirements, identify gaps, and assess progress towards data-driven objectives.
- **Data update documentation:** Created document showcasing procedures to follow to accurately update data.

Data Input Associate | [LinkedIn \(Via Swoon\)](#)

Feb 2020 – May 2020

Overview: Researched and gathered data from multiple sources to update and restructure account hierarchies in Salesforce, ensuring accurate, up-to-date, and actionable customer data for improved sales efficiency.

Bilingual Data Associate | [Amazon Canada \(Via Collabera\)](#)

Feb 2019 – Jan 2020

Overview: Contributed to language solutions that power voice-controlled products.

CERTIFICATION & EDUCATION

Applied Data Engineering Diploma | WeCloudData

Apr 2023 – Oct 2023

Relevant Coursework: Programming Foundations, Cloud Computing, Docker and Kubernetes, RDBMS and Data Warehouse, Data Lake, NoSQL Database, and Data Integration and Data Pipelines.

Midterm Project: Created a data pipeline architecture designed to ingest data from multiple files using Snowflake to extract data of interests from different sources, S3 buckets to store data extracted by Snowflake for further manipulation, Visual Studio Code and ec2 for transforming extracted data using Pyspark and cloud computing, EMR for facilitating processing of large data sets on the cloud, and parquet files for storing large volume of complex data. The data was then visualized with Power BI using Glue and Athena to make this visualization possible. In the architecture created, Cloudwatch was used to program Lambda to scan S3 buckets and send a signal to Airflow when the lambda judged that a specific day's data is ready.

Final Project: Built a real-time application that collects GPS data, via an open API, streamed by IoT devices located on buses, by using Apache NiFi, CDC (using Debezium), Kafka (using MSK), Spark Streaming, Docker, and MySQL. Showcased the application using Superset.

IBM AI Engineering Professional Certificate | IBM (Via Coursera)

Sept 2022 - Present

Excel to MySQL: Analytic Techniques for Business Specialization | Duke University (Via Coursera)

May 2019 – Apr 2020

Relevant Coursework: Business Metrics for Data-Driven Companies, Mastering Data Analysis in Excel, Data Visualization and Communication with Tableau, and Managing Big Data with MySQL.

Capstone Project: [Development of Model Testing Tableau Dashboard](#)

Bachelor of Commerce (Honors), Minor in Economics | McMaster University

Sept 2008 – Nov 2012

ESTHER-ALEXANDRA NGALEBA-ZE

730 Dovercourt Road, Toronto, ON
647-916-6123 | ngaleba@gmail.com

VOLUNTEERING

Treasurer (Volunteering) | [Visionary Toastmasters](#)

Apr 2020 – Present

Overview: Responsible for Club financial policies, procedures, and controls.

LANGUAGE PROFILE

English | French | Spanish | Japanese

PROFILES | PORTFOLIOS | PERSONAL PROJECTS

- LinkedIn Url: <https://www.linkedin.com/in/alexandrangalebaze>
- Portfolio: <https://ngaleba.wixsite.com/my-portfolio>
- GitHub: <https://github.com/ProcessOptimizingDataAnalyst>