

David Nady

Data Engineer | BI Engineer

+20 10 124 79330 | davidnady4yad@gmail.com | linkedin.com/in/david-nady | github.com/david11133 | david-portfolio.io

EDUCATION

Minya University

Bachelor of Science in Computer Science

Egypt, MN

Aug. 2021 – Jun. 2025

Relevant Coursework: Database Design, Information Retrieval, File Processing, Algorithms, Data Structures, Operating Systems, System Analysis, Software Engineering

PROJECTS

Softcart E-commerce Data Pipeline | *dbt, Snowflake, MongoDB, MySQL, Spark, Airflow*

Jul 2024

- Engineered and implemented robust ETL pipelines to efficiently synchronize and process data across transactional, NoSQL, and data warehousing systems, achieving a 50% reduction in data processing time.
- Designed and developed a PostgreSQL staging warehouse, seamlessly transitioning data to Snowflake for comprehensive reporting and analytics.
- Leveraged Hadoop and Spark for advanced big data analytics, significantly enhancing sales forecasting accuracy and enabling data-driven decision-making.

Uber ETL Data Pipeline | *Snowflake, PostgreSQL, Python, Mage, and Looker Studio*

Apr 2024

- Architected and deployed a sophisticated data model that ensured high levels of data consistency and significantly optimized query performance, resulting in faster data retrieval times.
- Spearheaded the development of an advanced ETL pipeline using MAGE, boosting data integration efficiency by 35% and ensuring seamless data flow across systems.
- Orchestrated the consolidation and migration of disparate data sources from multiple CSV files, effectively reducing data redundancy by 20% and streamlining the data architecture.
- Engineered a comprehensive and dynamic dashboard with Looker Studio, enabling in-depth data analysis and providing actionable insights through interactive visualizations and real-time data updates.

E-commerce DWH | *Python, SQL, Docker, Mage, Snowflake, Postgres, and Power BI*

Mar 2024

- Aggregated and integrated multiple CSV datasets from Kaggle, enhancing data coverage by 30% and providing a richer dataset for analysis.
- Executed comprehensive ETL (Extract, Transform, Load) processes using MAGE, effectively preparing and cleaning data for advanced analytics.
- Architected and deployed a scalable data warehouse leveraging Snowflake and PostgreSQL, ensuring efficient data management and retrieval.
- Developed an insightful and interactive dashboard with Power BI to visualize key metrics and performance indicators, driving data-driven decision-making.

Netflix ETL Data Pipeline | *PostgreSQL, Python, pandas, psycopg2, sqlalchemy*

Dec 2023

- Developed and implemented a robust ETL pipeline, integrating diverse Netflix data sources from Kaggle into a centralized PostgreSQL table, enhancing data accessibility and analysis efficiency.
- Utilized pandas for data transformation, including the removal of redundant columns and merging of titles, ensuring data coherence and reducing inconsistencies.
- Leveraged psycopg2 for seamless data insertion and updates, enabling continuous data flow and maintaining database integrity.
- Conducted thorough auditing and tracking of the ETL process, generating detailed reports to ensure data accuracy and transparency throughout the pipeline.

TECHNICAL SKILLS

Languages: Python, Java, C#/C++, SQL (Postgres), R

Databases & Data Warehousing: PostgreSQL, MySQL, MS SQL Server, MongoDB, Snowflake

Cloud: AWS (S3, RDS, Redshift, Glue), GCP (Google Cloud Storage, Cloud SQL, BigQuery Dataflow)

Developer Tools: Git, Github Actions, Docker, Google Cloud Platform, VS Code, Visual Studio, Jupyter Notebook

Libraries: Apache Airflow, PySpark, pandas, NumPy, SQLAlchemy, BeautifulSoup, Selenium