Rosa Pham

Forest Lake, Queensland

📞 0468480813   ✉ rosapham.de@gmail.com   🏠 rosapham.vercel.app   github.com/rosapham   🔗 linkedin.com/in/rosa-pham

Work Experience

PTE Master   Sydney, Australia
Data Engineer   *October 2023 - Current*

- Designed and implemented automated ETL pipelines using Apache Airflow, managing data wrangling, orchestration, migration, and loading into Azure Data Lake from trusted educational sources, reducing manual processing time and ensuring high data quality.
- Developed ELT pipelines using Azure Data Factory and Azure Databricks within a Lakehouse architecture, enhancing data consistency and integrity, achieving significant reductions in data processing times and improved accuracy of analytical outcome.
- Optimised PostgreSQL and Cassandra databases, utilising advanced indexing and query fine-tuning techniques, improving query performance by 20%, while ensuring high availability and fault tolerance in Cassandra clusters.
- Transformed data using dbt CLI to meet business requirements, employing SQL for data modeling and schema design, leading to significant time savings in documentation and data model creation.
- Built a Data Analytics platform on Azure Synapse Analytics, leveraging SQL to analyse and meet business requirements, providing actionable insights and strategic decision-making.
- Conducted detailed analysis, generated insights, and visualisations using Power BI.
- Implemented Cloud Infrastructure as Code using Docker, ensuring optimal uptime and compliance with best practices for multiple commercial applications.

Ehrenberg-Bass Institute – The University of South Australia   Adelaide, Australia
Data Science Assistant   *October 2022 – September 2023*

- Extracted data based on specific patterns using Python and efficiently stored the processed data in Snowflake, ensuring seamless integration and accessibility for downstream processes.
- Performed extensive data cleaning using Python, including resolving missing values, rectifying misspellings, and ensuring data consistency, which optimised data quality and increased accuracy, directly supporting reliable analytics and reporting.
- Developed an interactive application using Streamlit focused on the 'NBD-Dirichlet Model of Consumer Buying Behaviour', improving understanding and accessibility for business and research teams.
- Built a user-friendly chatbot for the Institute's website using Python, leveraging Azure OpenAI and Azure Cognitive Search to enhance user experience and information retrieval.

Skills

Cloud	Microsoft Azure (Azure Data Lake, Azure SQL DB, Azure Data Factory, Azure Databricks, Azure Synapse Analytics, Azure Key Vaults, Azure OpenAI, Azure Cognitive Search), AWS (Amazon S3, Amazon Redshift, Amazon Ethena, AWS Glue, AWS IAM), Snowflake
Programming	Python (Pandas, Numpy, scikit learn, seaborn, Fast API), R, SQL
Database	PostgreSQL, Cassandra, MySQL
Visualisation	Power BI, Tableau, Plotly, Matplotlib, ggplot2
Miscellaneous	Apache Airflow, Spark, dbt, Docker, Machine Learning, Github
Soft skills	Problem solving, Decision making, Teamwork, Time management, Documentation

Education

The University of South Australia   Adelaide, Australia
Postgraduate Diploma in Data Science

University of Science   Ho Chi Minh, Vietnam
Bachelor of Information Technology

Professional Certificates

- Udacity Nanodegree for Data Engineering in Azure.
- Udacity Nanodegree for Data Engineering in AWS.
- IBM Certificate for Professional Data Science.
- Microsoft Certificate for Azure Data Fundamentals.
- Microsoft Certificate for Power BI Data Analyst Associate.

Awards and Achievements

- Receiving the "Chancellor's Letter of Commendation" award for top 5% of postgraduate students in the University of South Australia.