**Project: Developing a Prototype of Data Pipeline**

Applied Data Science and Analytics

Data Engineering 2:Big Data Architecture

Prof. Dr. Frank Schulz and Prof. Ennurri Cho

19 December 2024

**Abstract**

In the fast-paced domain of financial markets, data plays a critical role in decision-making. This project focuses on developing a robust cloud-based data pipeline to integrate stock market data from the Alpha Vantage API with Google Trends data, providing deeper insights into investor behaviour and market movements. By leveraging GCP resources, the pipeline automates data ingestion, transformation, and storage while ensuring scalability and reliability. The methodology involves designing ETL workflows, implementing analytics using Python and SQL, and integrating automated data validation and trend correlation analyses. The pipeline effectively processes large volumes of data, reducing manual effort by 70%, and initial results demonstrate promising insights, enhancing the predictive power for stock market trends.