

Name of the Project – Credit Card Analytics on Cloud

Objectives/Vision

Consider that you have been hired as a Data Engineer in PK&G Finance. PK&G Finance is a Credit Card provider operating for the last 10 years catering to 100 Million Card Consumers. They have 1000 Merchant Partnerships and currently managing \$100 Billion Card Spend Management backed with multiple data sources. With rapid expansion planned for business, PK&G Finance have planned to transform their current Credit Card management infrastructure to cloud

Functional Requirements

1. Understand Data Engineering Discipline and can work on Azure Cloud.
2. Design and develop a complete ETL process based on the analysis of input data from various sources using ETL tools like Spark.
3. Get exposure to data migration into the cloud.
4. Understand monitoring and security aspects and implementation.
5. Understand Hadoop big data, Hadoop fundamental, and Data warehouse fundamental.
6. Overview of a real-time ETL pipeline with Kafka.
7. Overview of PySpark, if there is a scope in the respective use case.
8. Get exposure to data visualization and a corresponding tool like Tableau.

Non-Function Requirement:

1. Analyze the problem domain
2. Analyze the input data formats and design for data clean-up
3. Address data migration from the current system to the cloud (Can Include selecting and applying an ETL Tool)
4. Identifying the optimal cloud services aligned to the business commercial goal
5. Define and document your proposed architecture
6. Create and present a Minimal Viable Release (MVP) to the stake holders

Tools and Technologies to be used

- Data Extraction, Data transformation, data loading and migration tools, Hadoop, spark
- Visualization tool: Tableau
- Cloud components

