### ****Candidate Name****: Nachiketbhai Prajapati

**GMS Data Engineering Challenge**

**Introduction**

This solution combines comprehensive data cleaning, transformation, and preparation processes in Azure Synapse Analytics and Azure Databricks. The aim was to ensure high-quality data for analysis and to meet the business objectives effectively. Each step is aligned with the goals of fraud detection, customer retention, operational efficiency, regional insights, and policy optimization.

**Objective 1: Fraud Detection**

**Solution Summary**:

* The solution identifies high-risk claims by evaluating patterns such as high claim amounts, rejected payments, frequent claims by the same customer, and claims from high-risk regions.
* Conditions such as claims filed within 30 days or claims exceeding 5 times the premium amount are flagged as potentially fraudulent.

**Alignment with Objective**:

* Helps the company mitigate financial losses by identifying suspicious claims early.
* Provides insights into fraudulent behavior based on claim history, payment status, and regional risks.

**Objective 2: Customer Retention Analysis**

**Solution Summary**:

* Analyzed the claim history and payment behavior of customers to classify them as "Engaged" or "At Risk."
* Key metrics such as LastClaimDate and PaymentStatus were used to identify customers at risk of churning.

**Alignment with Objective**:

* Empowers the business to take targeted actions, such as offering loyalty programs or addressing customer grievances, to improve retention rates.
* Focuses on reducing churn by identifying and engaging at-risk customers.

**Objective 3: Operational Efficiency**

**Solution Summary**:

* Measured processing times by calculating the difference between PaymentDate and ServiceDate.
* Classified claims into categories such as "Highly Efficient," "Moderately Efficient," "Average Efficiency," or "Inefficient" based on defined thresholds.

**Alignment with Objective**:

* Highlights bottlenecks in claim processing workflows.
* Helps streamline operations by identifying areas where delays occur, improving overall efficiency and customer satisfaction.

**Objective 4: Region-Wise Insights**

**Solution Summary**:

* Grouped claims data by Region and ClaimType to provide insights into regional claim trends.
* Metrics such as total claims, total claim amounts, and average claim amounts were calculated for each region and claim type.

**Alignment with Objective**:

* Enables better resource allocation by understanding claim distribution across regions.
* Helps in adjusting premiums or deploying staff to high-claim areas based on regional trends.

**Objective 5: Policy Optimization**

**Solution Summary**:

* Evaluated the performance of insurance policies by calculating the ClaimsRatio (total claims amount divided by the premium amount).
* Classified policies into categories such as "Highly Profitable," "Moderately Profitable," "Risky," or "Underperforming" based on their claims ratio.

**Alignment with Objective**:

* Improves profitability and customer satisfaction by identifying underperforming policies.
* Provides insights to redesign or adjust policies based on their performance metrics.

**Conclusion**

The solution effectively addresses all objectives by leveraging clean, reliable data and robust analytical methods. The results enable actionable insights that support fraud detection, customer retention, operational efficiency, regional analysis, and policy optimization. These insights form a strong foundation for driving business growth and improving decision-making.