**Azure DevOps and Synapse Coding Challenge (19-12-2024)**

Done by: S.YAZHINI

**Given Challenge:**

**Buiild an etl pipline with azure synapse with dataflow running on it**

**Points on ETL Pipeline with Synapse and Notebooks:**

1. **Environment Setup:**
   * Created an Azure Synapse workspace with attached Azure Data Lake Storage for data staging and storage.
   * Configured Synapse's managed identity for secure access to linked resources.
2. **Data Ingestion:**
   * Utilized Synapse pipelines to orchestrate data loading directly from the source to staging storage.
3. **Transformation Using Notebooks:**
   * Designed and executed a Synapse notebook for data transformation.
   * Implemented Spark-based operations for data cleaning, aggregation, and formatting within the notebook.
   * Integrated the notebook into a Synapse pipeline for streamlined execution.
4. **Pipeline Orchestration:**
   * Built a Synapse pipeline to automate the ETL workflow.
   * Configured the pipeline to trigger and execute the Synapse notebook as part of the ETL process.
5. **Scheduling and Automation:**
   * Scheduled pipeline execution based on time intervals or custom triggers.
6. **Monitoring and Optimization:**
   * Leveraged Synapse Studio’s monitoring features to track pipeline runs and performance.
   * Optimized Spark jobs and pipeline configurations for cost and speed.

**Creating Azure Synapse:**

**Steps:**











