**Lifecycle of spark**

->Submit Application

Initialization Phase

->When the application is submitted, Spark Program launched on a cluster node.

->The driver program initialize SparkContext which is responsible for coordinating the job executions.

Job and stage creation

-> Spark application is divided into one or more jobs. Each job consists of stages. A stage is created for each RDD transformation,That requires data shuffling

After this stage, the next phase is

DAG(Directed Acyclic Graph) Generation

-> Spark constructs a DAG that represents the logical execution plan

Task Scheduling

-> The scheduler takes the DAG and schedules tasks for executions

A task is a single operation applied to single partition