

## Task 1

**Objective:** Deploying spark code in production environment

**Code:**

```
if __name__ == "__main__":
    #Initialize the spark session
    spark = SparkSession \
        .builder \
        .appName("pyspark script job") \
        .getOrCreate()

    #Do all ETL here
    df1 = spark.read.csv("/home/s_kante/spark/data/agent_system.csv", header=True)
    df1.show()

    #Once done, stop the spark session
    spark.stop()
```

**How to execute:**

spark-submit <file name>.py