

assignment-7

July 14, 2025

```
[0]: # 7.1 Read the CSV file
df = spark.read.csv("/Volumes/workspace/default/skit_usecases/
↳Am5Uzv06R2CXcb4NRunM_Employee_Attrition.csv", header=True, inferSchema=True)
df.display()
```

```
[0]: # 7.2 Filter the data (employees with high attrition risk and less job
↳satisfaction)
df_Transformed = df.filter((df.Attrition == 'Yes') & (df.JobSatisfaction < 3))
df_Transformed.display()
```

```
[0]: # 7.3 Overwrite mode for saving the data at desired path
df_Transformed.write.format("delta").mode("overwrite").
↳partitionBy("Department").save("/Volumes/workspace/default/skit_usecases/
↳high_attrition_risk_employees")
```