**Data Transformation Activity**

# Part 1: Queries Used

## 1. Create an empty table, “employee\_sales”

CREATE EXTERNAL TABLE employee\_sales

(

Attrition string,

Department string,

JobSatisfaction int,

MonthlyIncome int

);

## 2. Check that steps have worked (used after each transformation)

SELECT \* FROM employee\_sales LIMIT 10;

\*\* To check the ordering of the data (#6), a limit of 200 was used, because there were many job satisfaction ratings of 4.

## 3. Load the data from “employee” into “employee\_sales”

INSERT OVERWRITE TABLE employee\_sales

SELECT Attrition, Department, JobSatisfaction, MonthlyIncome

FROM employee;

## 4. Round the monthly income to the nearest $1000

INSERT OVERWRITE TABLE employee\_sales

SELECT Attrition, Department, JobSatisfaction, ROUND(MonthlyIncome, -3) AS MonthlyIncome

FROM employee\_sales;

## 5. Filter the data to see only the Sales department

INSERT OVERWRITE TABLE employee\_sales

SELECT \*

FROM employee\_sales

WHERE Department LIKE "%Sales%";

## 6. Order the data from highest to lowest job satisfaction

INSERT OVERWRITE TABLE employee\_sales

SELECT \*

FROM employee\_sales

ORDER BY JobSatisfaction DESC;

# Part 2: Screenshot of the Resulting Table

