In Class Exercise 7

Define the ORION library.

1. **Creating a New Column by Recoding Ranges of Values**

   Create a query that groups employee salary values into categories.

   a. In the **InClassExercise7** project, add the **employee_master** table from the **ORION** library to the Problem 1 process flow.

   b. Use the Query Builder to create a new query named **Salary Range Query**.
   - Name the output table **Salary_Range**.
   - Include only the **Employee_ID** and **Employee_Hire_Date** columns.
   - Use the New Computed Column Wizard to create a new column named **Salary_Range** based on the values in the **Salary** column.

Replace	With
0 – 24999.99	Below $25K
25000 – 49999.99	$25K to $50K
50000 – 99999.99	$50K to $100K
100000 and above	Over $100K

   c. Run the query and view the results.

   Partial Results

	Employee_ID	Employee_Hire_Date	Salary_Range
1	120101	01-JUL-2008	Over $100K
2	120102	01-JUN-1994	Over $100K
3	120103	01-JAN-1979	$50K to $100K
4	120104	01-JAN-1986	$25K to $50K
5	120105	01-MAY-2004	$25K to $50K
6	120106	01-JAN-1979	$25K to $50K

   d. Use the One-Way Frequencies task to validate the results.
   - Specify **Salary_Range** as the analysis variable.
   - Create the report to show only the frequency and percentage values.
   - Rename the task to **Salary Range Frequencies**.

   e. Run the task and answer the following question:

   What salary range has the highest frequency?

2. Recoding a Column Based on a Condition

Create a query to group customers into High, Medium, and Low activity groups.

a. In the **InClassExercise7** project, add the **customers** table from the **ORION** library to the Problem 2 process flow.

b. Use the Query Builder to create a new query named **Activity Level Query**.

- Name the new table **ActivityLevel**.
- Include the columns **Customer_ID**, **Customer_Name**, and **Customer_Type** in the new table.
- Use the Replace Condition tab to create a new column named **ActivityLevel** by recoding **Customer_Type**. **ActivityLevel** should be *High* when **Customer_Type** contains the word *high*, *Medium* when **Customer_Type** contains the word *medium*, *Low* when **Customer_Type** contains the word *low*, and missing otherwise.

  Hint: When you specify the condition, be sure to uncheck the 'Match case' box before entering the text values **high**, **medium**, and **low** to ignore the case of the text in the **Customer_Type** column.

c. Run the query and view the results.

Partial Results

	Customer_ID	Customer_Name	Customer_Type	ActivityLevel
1	4	James Kvarniq	Orion Club members low activity	Low
2	5	Sandrina Stephano	Orion Club Gold members medium activity	Medium
3	9	Cornelia Krahl	Orion Club Gold members medium activity	Medium
4	10	Karen Ballinger	Orion Club members high activity	High
5	11	Elke Wallstab	Orion Club members high activity	High
6	12	David Black	Orion Club members medium activity	Medium
7	13	Markus Sepke	Orion Club Gold members low activity	Low
8	16	Ulrich Heyde	Internet/Catalog Customers	
9	17	Jimmie Evans	Orion Club members medium activity	Medium

d. Use the One-Way Frequencies task to validate the result.

- Specify **ActivityLevel** as the analysis variable.
- Create the report to show only the frequency and percentage values.
- Rename the task to **Activity Level Frequencies**.

e. Run the task and answer the following question:

What percentage of customers are in the Medium activity level?

3. Writing a CASE Expression with the YEAR Function

Create a query to group employees into seniority groups based on their hire year.

a. In the **InClassExercise7** project, add the **employee_master** table from the **ORION** library to the Problem 3 process flow.

b. Use the Query Builder to create a new query named **Seniority Query**.

- Name the output table **Seniority**.
- Include the columns **Employee_ID**, **Employee_Name**, and **Department** in the new table.
- Create a new column named **HireYear**, which is the year in which each employee was hired.
- Create a new column, named **SeniorityGroup**, that is based on the value of **HireYear**. Values should be assigned according to the following table:

Year of Hire	SeniorityGroup
<= 1990	1990 and earlier
1991-2000	1991 to 2000
>= 2001	2001 and later

c. Run the query and view the results.

Partial Results

	Employee_ID	Employee_Name	Department	HireYear	SeniorityGroup
1	120101	Patrick Lu	Sales Management	2014	2001 and later
2	120102	Tom Zhou	Sales Management	2000	1991 to 2000
3	120103	Wilson Dawes	Sales Management	1985	1990 and earlier
4	120104	Kareen Billington	Administration	1992	1991 to 2000
5	120105	Liz Povey	Administration	2010	2001 and later

d. Use the One-Way Frequencies task to validate the results.

- Specify **SeniorityGroup** as the analysis variable.
- Create the report to show the frequency and percentage values.
- Rename the task to **Seniority Group Frequencies**.

e. Run the task and answer the following question:

Which seniority group has the fewest employees?

4. Creating a User-Defined Format for Discrete Values

Use the Create Format task to create a user-defined format that displays full state names. Apply the format to the **State** column in the Summary Statistics task.

a. In the **InClassExercise7** project, create the Problem 4 process flow.

b. Open the Create Format task.

- Go to Tasks > Browse > expand **Data,** and double-click **Create Format**.
- Name the format **statefmt** and save it in the **Work** library, i.e., Currently Assigned Libraries is WORK.
- In the Define formats window, create the following labels for each data value:

Label	Value
California	CA, Ca
Florida	FL, Fl
Pennsylvania	PA, Pa
Other	All other values

Note: To assign the Other value, use the drop-down list in the **Values** field.

c. Run the task and view the log to see whether the format was successfully created.

d. Add the **employee_addresses** table from the **ORION** library to the Problem 4 process flow

e. Use the List Data task to create a report for the **employee_addresses** table.

- Filter the data to include only the employees in the United States.
- Display the **Employee_Name**, **Street_Number**, **Street_Name**, and **City** variables.
- Identify the rows of the report by **Employee_ID**.
- Group the rows of the report by the formatted value of **State**.

Hint: After assigning it under Group Analysis by, right-click on State > Properties. Change Format.

- *Print the number of rows in each state group.* (Explore the Options tab)
- Rename the task to **State Addresses Report**.

f. Run the task and answer the following question:

How many employees are in the state of California?

Partial Results

State Addresses Report

State=California

Employee_ID	Employee_Name	Street_Number	Street_Name	City
120656	Amos, Salley	3524	Calico Ct	San Diego
120759	Apr, Nishan	105	Brack Penny Rd	San Diego
121017	Arizmendi, Gilbert	379	Englehardt Dr	San Diego
121062	Armant, Debra	10398	Crown Forest Ct	San Diego
121049	Bataineh, Perrior	892	Birchland Dr	San Diego
120998	Benedicto, Tondelayo	2902	Gable Ridge Ln	San Diego