DA 6223 Exercise 2

Problem 1

Producing a One-Way Frequency Task Report

Use the One-Way Frequencies task to produce a report of employees by department.

a. Create a new Project and add a new Process Flow.

b. Then add the **employee_master** table.

c. Use the One-Way Frequencies task to analyze the Department column.

- Include a horizontal bar chart as part of the report.
- Add the title **Employee Count by Department** and delete the footnote.
- Modify the task properties to change the task label to **Dept Freq Report**.
- Generate the report and examine the results.

d. How many employees are in the Marketing Department?

e. Rename the Process Flow as **Problems 1, 2, 3**.

Problem 2

Creating a Grouped Frequency Report

Use the One-Way frequencies task to generate a count and percentage report of employees by job title. Group the results by department.

a. Use the One-Way Frequencies task to analyze **Job_Title**, grouped by **Department**.

- Include only frequencies and percentages.
- Order the output report by descending frequencies.
- Add the title **Employees by Job Title** and delete the footnote.
- Change the task label to **Jobs by Dept Freq Report**.
- Run the task and answer the following question:

b. In the Engineering Department, which job title has the most employees?

Problem 3

Using Multiple Group Values in a Listing Report

Use the List Data task to produce an employee report that lists employees by department and job title.

a. Open the **employee_master** table.
b. Use the Tasks pane to launch the List Data task.
c. Make the following role assignments:

List variables	**Employee_Name**, **Employee_Hire_Date**, **Salary**
Group analysis by	**Department, Job_Title**
Identifying label	**Employee_ID**

- Add the title Departmental Employee Report and delete the footnote.
- Modify the task properties to change the task label to Dept Employee List.
- Generate the report and answer the following question:

d. What is the highest salary for an Accountant II in the Accounts Department?

Problem 4

Filtering Data in the List Data Task

Use the List Data task on the **employee_addresses** table to produce a report. Display only employees located in San Diego.

a. Select the **employee_addresses** table and create a report with the List Data task and show the **Employee_Name, Street_Number, Street_Name,** and **Postal_Code** columns. Identify each row with the **Employee_ID** column.
   - Modify the properties for the columns so that the report displays the following labels: **ID, Name, Street Number, Street Name,** and **ZIP Code**.
   - Filter the data so that only the employees from San Diego are included in the report.
   - Change the title to **Employee List for San Diego** and delete the footnote.
   - Modify the task properties to change the task label to **San Diego List**.
   - Generate the report and answer the following question:

b. On what street does James Blackley live?

c. Rename the Process Flow as **Problem 5**.

d. Save the project as **LastName_FirstName_Exercise2.**

Upload your project under Exercise 2 and answer the questions on Canvas.

We will answer Problem 5 in class on Thursday, February 8, 2024. First, we will Finish the PROC SQL essentials lecture and then solve Problem 5 in class. Feel free to make an attempt before that if you feel comfortable with SQL.

Problem 5

Exploring the WORK.Employee_Payroll SAS Table

Note that you should first import **employee_payroll.csv** file as a new SAS dataset. Change the type of Employee_ID to character. Save the resulting dataset in the WORK library.

a. Add a new Process Flow, and create a new program named **Explore Employee Payroll** in it.

b. In the first SQL procedure step, add a **DESCRIBE TABLE** statement to see column attributes of the **WORK.Employee_Payroll** table. Highlight the step and run the selected code. Examine the log and results.

How many numeric variables does this data set have?

c. In the second SQL procedure step, add a **SELECT** statement and select all the columns from the **WORK.Employee_Payroll** table using an asterisk. Add the **OBS=10** data set option to the table to limit the report to 10 rows.

d. In the third SQL procedure step, delete the **OBS=10** and modify the **SELECT** statement to select the columns **Employee_Hire_Date, Employee_Gender** and **Salary** (exactly in this order). Highlight the step and run the selected code. Examine the log and results.

What is the order of the variables in the result table?

e. Start the fourth SQL procedure step by restricting the number of rows that SAS retrieves from **Employee_Payroll** table to 30 by using PROC SQL options. (Hint: Use `PROC SQL INOBS=30;`) Then, modify the **SELECT** clause in step d. by adding an **ORDER BY** clause. Sort the data in descending order based on Salary. (Hint: The ORDER BY command sorts the result set in ascending order by default. To sort the records in descending order, use the DESC keyword after Salary).

What is the highest salary printed in the result table?

f. Start the fifth SQL procedure step by restricting the number of observations in the output to 30. (Hint: Use `PROC SQL OUTOBS=30;`) Keep everything else the same as in step e. Examine the results.

What is the highest salary printed in the result table?

Are the results same as using INOBS=30?