Lesson 5 Introduction to Analyzing Data

Practice

1. Working with Data Items

a. Open the browser and sign in to SAS Viya.

b. Open the **GAINDS-Practice5.1** report from the **Courses/GAINDS/Practices** folder.

c. View the data items in the Data pane and answer the following questions:

What is the classification of **Employee ID**? **Manager at 1. level**?

Answer: _____

What does the **Frequency** data item represent?

Answer: _____

d. Change the classification for **Manager at 1. level** to **Category**.

e. Change the format for **Annual Salary** to **Dollar13.2**.

f. Rename the following data items:

Old name	New name
Employee ID	**ID**
Employee Name	**Name**
Manager at 1. level	**Manager ID**
Frequency	**Number of Employees**

Note: Click  (**Actions**) and select **Refresh EMPLOYEES_CLEAN** at the top of the Data pane to collapse the data item properties.

g. Save the report.

End of Practices

Practice

2. Exploring Data: Part 1

a. Open the browser and sign in to SAS Viya.

b. Open the **GAINDS-Practice5.2a** report from the **Courses/GAINDS/Practices** folder.

c. Create an automatic chart using the following data items:

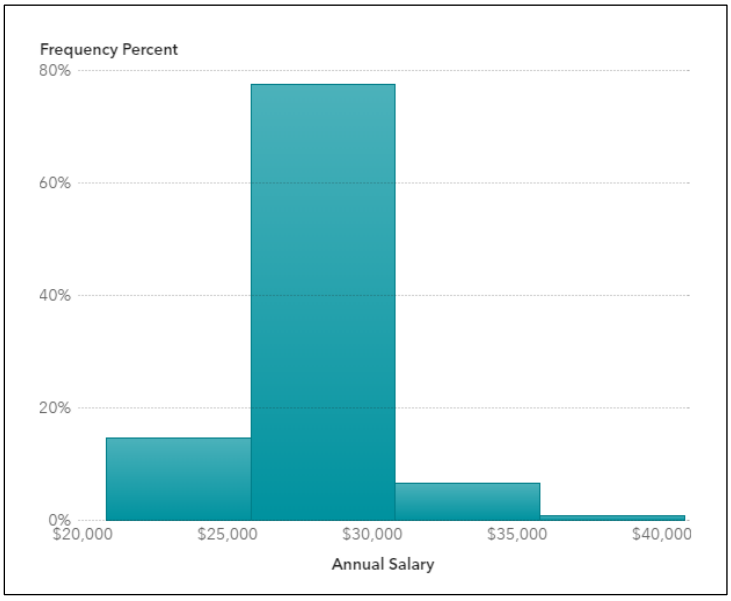
Annual Salary

Frequency Percent

d. Modify the following options for the automatic chart:

Name	Distribution of Salary
Bin range	Measure values
Set a fixed bin count	<selected>
Bin count	4

The automatic chart should resemble the following:



e. Maximize the histogram to answer the following question:

Into which range do the majority of salaries fall?

Answer: _____

Hint: After answering the question, click (**Restore**) in the upper right corner.

f. Add a bar chart on the right of the automatic chart by assigning the following data items to the specified roles:

Category	**Job Title**
Measure	**Annual Salary**
Group	**Department**

g. Specify **Total Salary by Job and Department** as the name of the bar chart.

The bar chart should resemble the following:

h. Answer the following questions:

In which department are a majority of our salary costs spent? For which job title?

Answer: _____

What could be some reasons why salary costs are so much higher for this group?

Answer: _____

i. Save the report.

End of Practices

Practice

3. Exploring Data: Part 2

a. Open the browser and sign in to SAS Viya.

b. Open the **GAINDS-Practice5.2b** report from the **Courses/GAINDS/Practices** folder.

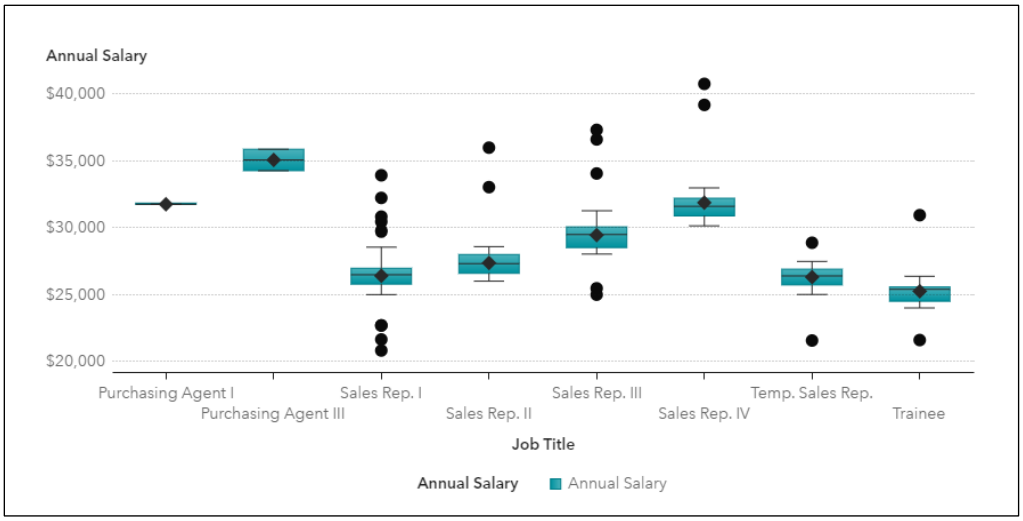
c. On Page 2, create a box plot by assigning the following data items to the specified roles:

Category	**Job Title**
Measures	**Annual Salary**

d. Modify the following options for the box plot:

Name	Salary Analysis by Job Title
Outliers	Show Outliers
Averages	<selected>

The box plot should resemble the following:



e. Maximize the box plot to answer the following questions:

Which job title has the highest average salary? The lowest?

Answer: _____

Which job title has the largest number of outliers?

Answer: _____

Hint: After answering the question, click  (**Restore**) in the upper right corner.

f. Save the report.

End of Practices

Practice

4. Creating Data Items

a. Open the browser and sign in to SAS Viya.

b. Open the **GAINDS-Practice5.3a** report from the **Courses/GAINDS/Practices** folder.

c. Create a new data item, **Employee Status**, by assigning the following labels to the values:

Employee Status (label)	Employee Termination Date (value)
Active	.
Retired	<all remaining values>

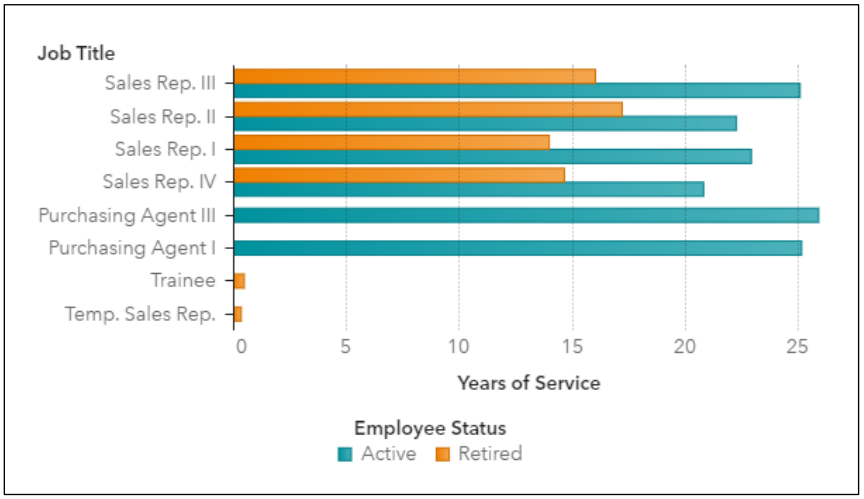
d. On Page 3, create a bar chart by assigning the following data items to the specified roles:

Category	**Job Title**
Measure	**Years of Service**
Group	**Employee Status**

e. Specify **Years of Service by Job Title and Status** as the name of the bar chart.

f. Change the aggregation for **Years of Service** to **Average**.

The bar chart should resemble the following:



g. Answer the following questions:

Which job title has the highest average years of service among active employees? Among retired employees?

Answer: _____

h. Save the report.

End of Practices

Practice

5. Applying Filters

a. Open the browser and sign in to SAS Viya.

b. Open the **GAINDS-Practice5.3b** report from the **Courses/GAINDS/Practices** folder.

c. Add a data source filter to filter for active employees in the Sales Department.

Note: Use the AND operator (in the Boolean group) to filter for multiple conditions. After the data source filter is applied, 429 observations should be returned.

d. Change the classification for **Employee Country** to **Geography ⇨ Country or Region ISO 2-Letter Codes**.

e. On Page 4, create a geo coordinate map by assigning the following data items to the specified roles:

Category	**Employee Country**
Size	**Total Profit**
Color	**Number of Employees**

The geo map should resemble the following:

f. Maximize the geo map to answer the following questions:

Management has decided that one possible criterion for promotion is profit generated. Which two countries generate the highest profit? Why do they have such high profits?

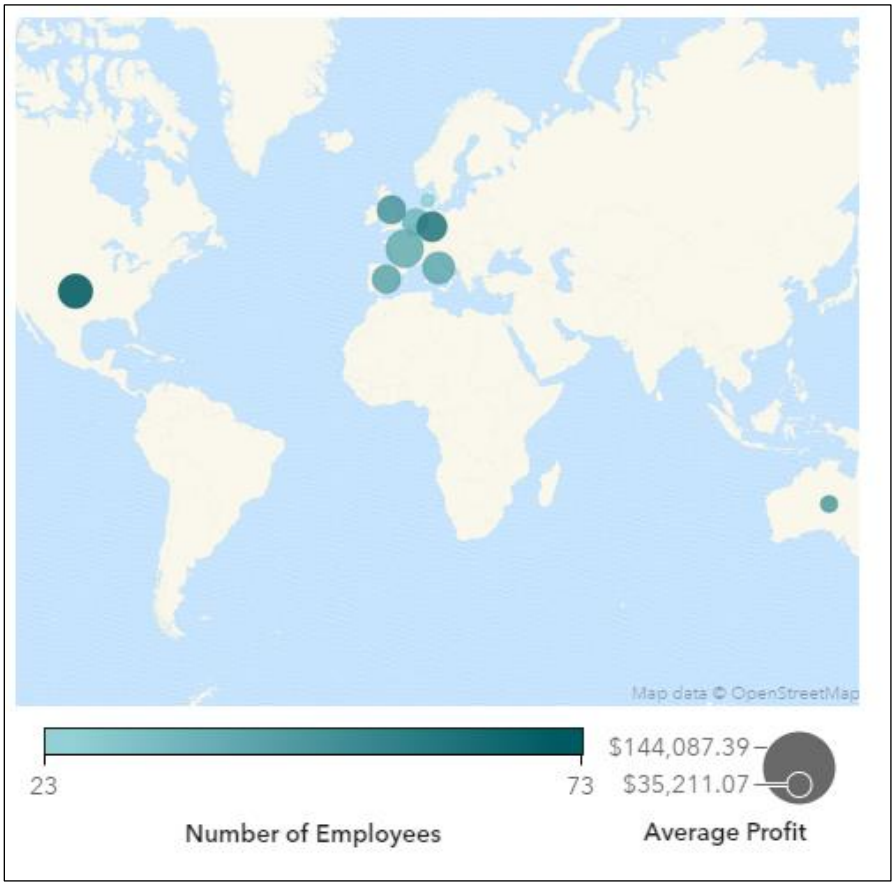
Answer: _____

Hint: After answering the questions, click (**Restore**) in the upper right corner.

g. In the geo map, specify **Average Profit** for the **Size** role.

h. Specify **Average Profit and Number of Employees by Country** as the name of the geo map.

The updated geo map should resemble the following:



i. Maximize the geo map to answer the following question:

Which country has the highest average profit? Highest number of employees?

Answer: _____

Hint: After answering the question, click (**Restore**) in the upper right corner.

j. Save the report.

End of Practices

Practice

6. Analyzing Data

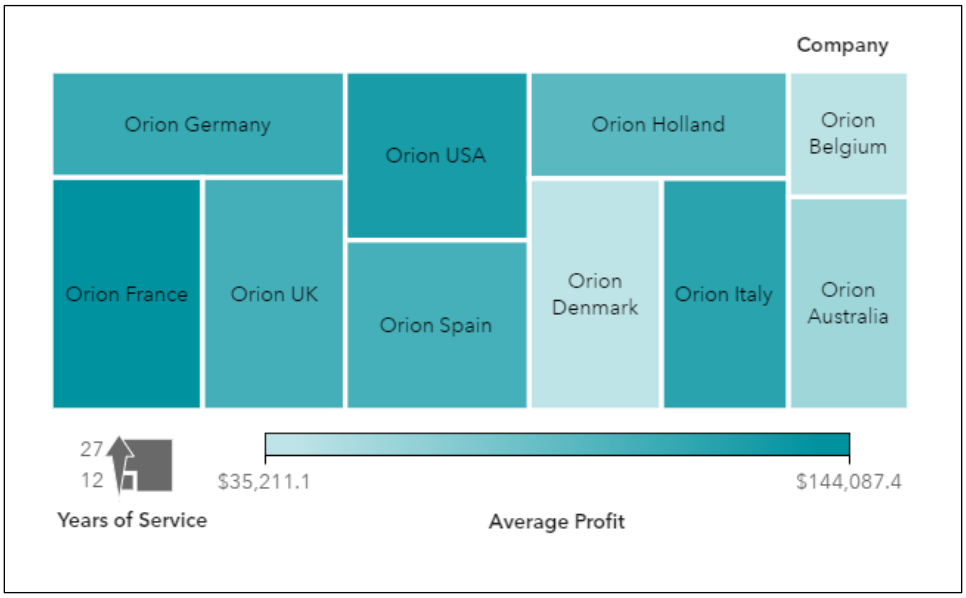
a. Open the browser and sign in to SAS Viya.

b. Open the **GAINDS-Practice5.4a** report from the **Courses/GAINDS/Practices** folder.

c. On Page 5, create a treemap by assigning the following data items to the specified roles:

Tile	**Company**
Size	**Years of Service**
Color	**Average Profit**
Data tip values	**Add Number of Employees**

The treemap should resemble the following:



d. Create a new hierarchy (**Employee Hierarchy**) that contains the following categories:

Company

Job Title

Group

e. In the treemap, specify **Employee Hierarchy** for the **Tile** role, and navigate through the hierarchy to answer the following questions:

Which two companies have the highest average profit generated (one possible criterion for promotion)?

Answer: _____

For these two companies, which job titles have the highest average years of service and average profit generated?

Answer: _____

f. Save the report.

End of Practices

Practice

7. Adding Data Analysis

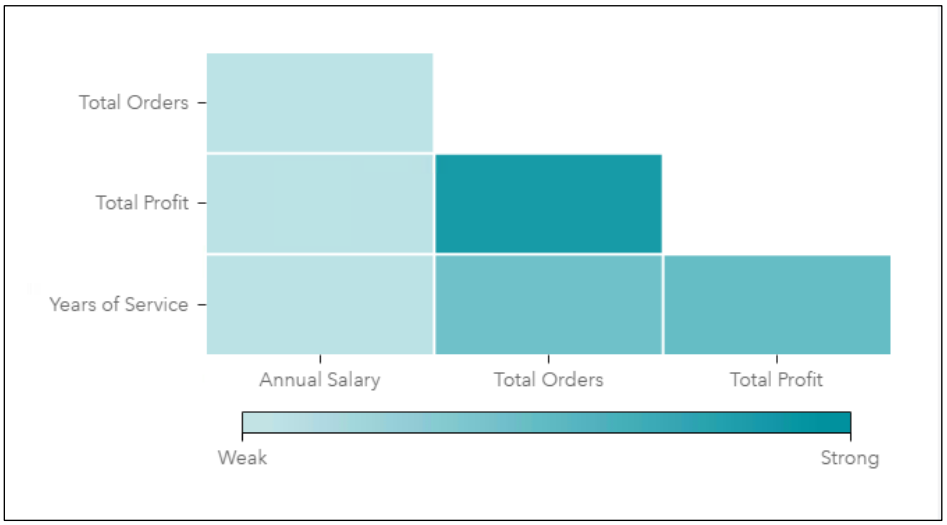
a. Open the browser and sign in to SAS Viya.

b. Open the **GAINDS-Practice5.4b** report from the **Courses/GAINDS/Practices** folder.

c. On Page 6, create a correlation matrix by assigning the following data items to the specified roles:

Measures	**Annual Salary**
	Total Orders
	Total Profit
	Years of Service

The correlation matrix should resemble the following:



d. Answer the following question:

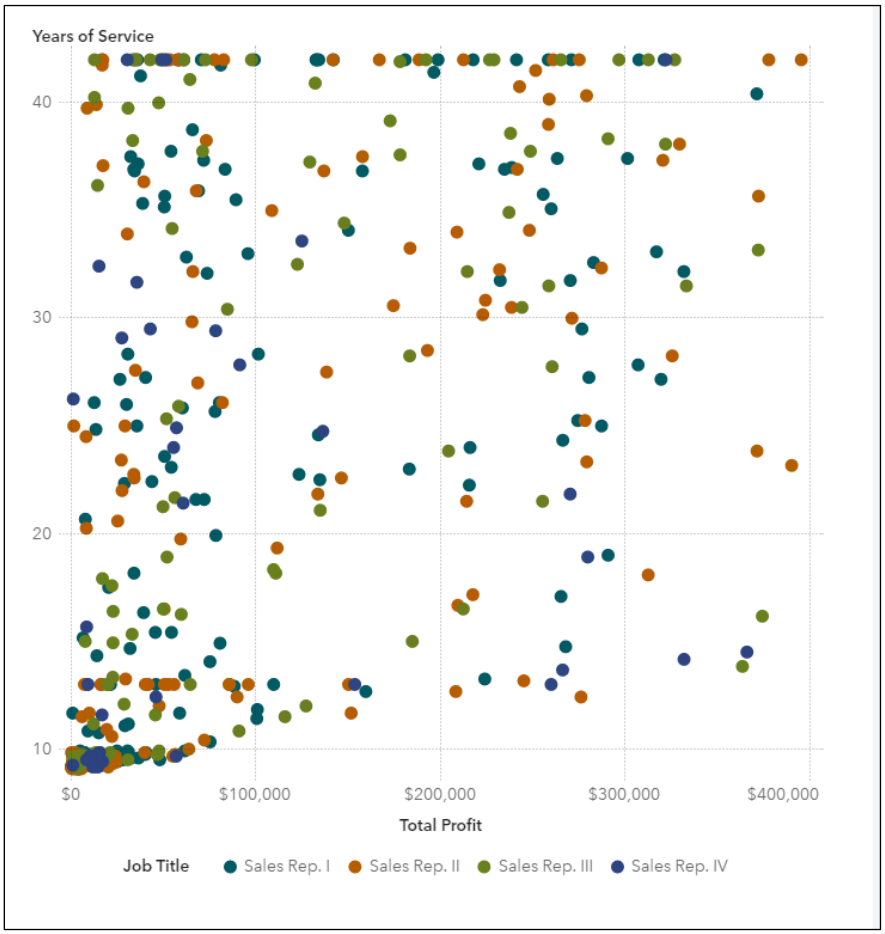
What is the degree of correlation between **Total Orders** and **Total Profit**?

Answer: _____

e. Create a scatter plot, on the right of the correlation matrix, by assigning the following data items to the specified roles:

Measures	**Total Profit**
	Years of Service
Color	**Job Title**

The scatter plot should resemble the following:



f. Answer the following question:

Using years of service and profit generated as promotion criteria, do you notice any differences between job titles?

Answer: _____

g. Save the report.

End of Practices