Using the crosstab Function to Create Contingency Tables

1. Task Description

The goal of this task was to demonstrate how to use the crosstab function in the pandas library to create contingency tables with multiple categorical variables. A contingency table is a useful tool in data analysis for summarizing the frequency distribution of variables and exploring the relationships between them.

2. Screenshot of Output

```
Items              Shirt  Shoes
Gender Age Group
Female 20-30           0      2
       30-40           1      0
       40-50           0      1
Male   20-30           1      0
       30-40           0      1
       40-50           1      1
```

3. Algorithm Used in Task

➢ **Function and Algorithms:**

- **Crosstab Function:**

  o The pandas.crosstab function is used to compute a simple cross-tabulation of two or more factors.

  o It provides the frequency of occurrences for combinations of values from two or more categorical variables.

  o In this task, the function was applied to the Gender and Age Group columns to compare against the Items column.

4. Add Report in Your Task Zip File

The task report has been added to the zip file. This includes:

1. The Python script (Data Science & Machine Learning_Task_4.ipynb).

2. A text version of this report (Task_4_Report.pdf).

Prepared by
Karan Gohil