```
C:\Users\acrpr\OneDrive\Documents\cs23localab\CS176\CaseStudy-DataPreprocessing\.venv\Scripts\python.exe C:\Users\acrpr\OneDrive\Documents\cs23localab\CS176\CaseStudy-DataPreprocessing\preprocessing.py
========== GROUP 3: CASE STUDY - DATA PREPROCESSING ==========

~ SECTION ON DESCRIPTIVE ANALYTICS AND VISUALIZATION ~

--- NUMERICAL STATISTICS ---
Age Column:
count    600.0000
mean      42.3950
std       14.4249
min       18.0000
25%       30.0000
50%       42.0000
75%       55.2500
max       67.0000
Name: age, dtype: float64

Income Column:
count      600.0000
mean     27524.0312
std      12899.4682
min       5014.2100
25%      17264.5000
50%      24925.3000
75%      36172.6750
max      63130.1000
Name: income, dtype: float64

Children Column:
count    600.0000
mean       1.0117
std        1.0568
min        0.0000
25%        0.0000
50%        1.0000
75%        2.0000
```

```
max         3.0000
Name: children, dtype: float64

----------------------------------------

RELEVANT AGE VALUES:

Age Range: 49
Age Variance: 208.0791
Age Sum: 25437
Age Count: 600

----------------------------------------

--- SEGREGATION OF CUSTOMER ACCOUNT TYPE ---
Savings Account Count: 414
Current Account Count: 455

----------------------------------------

--- PIVOT REPORT: MARRIED VS. NUMBER OF CHILDREN ---
          children
married
NO          1.0833
YES         0.9747

----------------------------------------

--- GROUPED MEANS ---
                        age      income  children
pep married car
NO  NO      NO    36.5476  21758.4948    1.5000
            YES   39.2619  23635.4655    1.5000
    YES     NO    40.1270  25031.1746    0.8333
            YES   41.6552  26355.4868    1.0086
YES NO      NO    44.3833  30565.4300    0.8000
            YES   45.9833  31752.5300    0.7833
    YES     NO    43.3947  28293.1428    1.0526
            YES   46.7308  32145.5318    1.0769

----------------------------------------
```

```
 78
 79 --- PATTERN ANALYSIS: AGE AND PEP ---
 80 pep
 81 NO     40.0982
 82 YES    45.1277
 83 Name: age, dtype: float64
 84
 85 --- PATTERN ANALYSIS: NUMBER OF CHILDREN, PEP, AND
    MARRIAGE ---
 86 pep  married
 87 NO   NO         1.5000
 88      YES        0.9174
 89 YES  NO         0.7917
 90      YES        1.0649
 91 Name: children, dtype: float64
 92
 93 ----------------------------------------
 94
 95 ~ END OF DESCRIPTIVE ANALYTICS AND VISUALIZATION ~
 96
 97 ~ SECTION ON DATA TRANSFORMATION ~
 98
 99 --- NORMALIZED INCOME COLUMN ---
100     income  income_Normalized
101 0  17546.0             0.2156
102 1  30085.1             0.4314
103 2  16575.4             0.1989
104 3  20375.4             0.2643
105 4  50576.3             0.7840
106
107 ----------------------------------------
108
109 --- BINNED INCOME COLUMN ---
110     income income_Binned
111 0  17546.0           Low
112 1  30085.1        Medium
113 2  16575.4           Low
114 3  20375.4        Medium
115 4  50576.3          High
116
117 ----------------------------------------
```

```
118
119 --- DUMMY VARIABLES FOR REGION ---
120         id  age     sex  ...  region_RURAL
    region_SUBURBAN  region_TOWN
121 0  ID12101   48  FEMALE  ...         False
                False        False
122 1  ID12102   40    MALE  ...         False
                False         True
123 2  ID12103   51  FEMALE  ...         False
                False        False
124 3  ID12104   23  FEMALE  ...         False
                False         True
125 4  ID12105   57  FEMALE  ...          True
                False        False
126
127 [5 rows x 17 columns]
128
129 ================================================
130
131
132 Process finished with exit code 0
133
```