```
"C:\Users\acrpr\OneDrive\Documents\Local Code Files\CS176\CaseStudy-DataPreprocessing\.venv\Scripts\python.exe" "C:\Users\acrpr\OneDrive\Documents\Local Code Files\CS176\CaseStudy-DataPreprocessing\DFTest.py"
========== GROUP 3: CASE STUDY - DATA PREPROCESSING ==========

~ SECTION ON DESCRIPTIVE ANALYTICS AND VISUALIZATION ~

--- NUMERICAL STATISTICS ---
--- Age Column Statistics ---
       Age Stats
count   600.0000
mean     42.3950
std      14.4249
min      18.0000
25%      30.0000
50%      42.0000
75%      55.2500
max      67.0000

--- Income Column Statistics ---
       Income Stats
count      600.0000
mean     27524.0312
std      12899.4682
min       5014.2100
25%      17264.5000
50%      24925.3000
75%      36172.6750
max      63130.1000

--- Children Column Statistics ---
       Children Stats
count        600.0000
mean           1.0117
std            1.0568
min            0.0000
25%            0.0000
```

```
50%             1.0000
75%             2.0000
max             3.0000

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

RELEVANT AGE VALUES:

   Age Range  Min Age  Max Age  Median Age
0         49       18       67        42.0

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

--- SEGREGATION OF CUSTOMER ACCOUNT TYPE ---
   Savings Account Count  Current Account Count
0                    414                    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

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

```
--- PATTERN ANALYSIS: AGE AND PEP ---
     Avg Age by PEP
pep
NO          40.0982
YES         45.1277

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

--- PATTERN ANALYSIS: NUMBER OF CHILDREN, PEP, AND MARRIAGE ---
              Avg Children
pep married
NO  NO             1.5000
    YES            0.9174
YES NO             0.7917
    YES            1.0649

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

~ END OF DESCRIPTIVE ANALYTICS AND VISUALIZATION ~

~ SECTION ON DATA TRANSFORMATION ~

--- NORMALIZED INCOME COLUMN ---
    income  income_Normalized
0  17546.0             0.2156
1  30085.1             0.4314
2  16575.4             0.1989
3  20375.4             0.2643
4  50576.3             0.7840

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

--- BINNED INCOME COLUMN ---
    income income_Binned
0  17546.0          Low
1  30085.1       Medium
2  16575.4          Low
3  20375.4       Medium
4  50576.3         High

```

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

--- DUMMY VARIABLES FOR REGION ---
         id  age     sex  ...  region_RURAL
   region_SUBURBAN  region_TOWN
0  ID12101   48  FEMALE  ...         False
              False        False
1  ID12102   40    MALE  ...         False
              False         True
2  ID12103   51  FEMALE  ...         False
              False        False
3  ID12104   23  FEMALE  ...         False
              False         True
4  ID12105   57  FEMALE  ...          True
              False        False

[5 rows x 17 columns]

==================================================


Process finished with exit code 0

```