Name :- Radrapati venkatesh

SID :- 2461075

Subjet :- **Machine Learning in Finance**

Week :- 2 logbook

```python
import pandas as pd

df = pd.read_csv("/content/adult_data_mini.csv")

print("=== ORIGINAL DATA (first 5 rows) ===")
print(df.head(), "\n")

sid = 2461075
n = sid % 10
if n == 0:
    n = 10

def reduce_hours(x):
    return x - n

df['reduced_hours'] = df['hours-per-week'].apply(reduce_hours)

grouped = df.groupby(['relationship',
'reduced_hours']).size().reset_index(name='count')

# Display results
print("=== DATA AFTER APPLYING reduce_hours FUNCTION (first 5 rows) ===")
```

```python
print(df.head(), "\n")


print("=== GROUPED SUMMARY BY relationship AND reduced_hours ===")
print(grouped.head(), "\n")


print(f"Reduction applied per person: {n} hours (SID = {sid})")
```

```
3   Married-civ-spouse    Handlers-cleaners          Husband   Black     Male
4   Married-civ-spouse       Prof-specialty             Wife   Black   Female

   capital-gain  capital-loss  hours-per-week  native-country  Answer
0        2174.0           NaN              40   United-States   <=50K
1           NaN           0.0              13   United-States   <=50K
2           NaN           NaN              40   United-States   <=50K
3           NaN           NaN              40   United-States   <=50K
4           0.0           NaN              40            Cuba   <=50K

=== DATA AFTER APPLYING reduce_hours FUNCTION (first 5 rows) ===
   age         workclass   fnlwgt  education  education-num  \
0   39         State-gov    77516  Bachelors             13
1   50  Self-emp-not-inc    83311  Bachelors             13
2   38           Private   215646    HS-grad              9
3   53           Private   234721       11th              7
4   28           Private   338409  Bachelors             13

        marital-status         occupation     relationship    race     sex  \
0        Never-married       Adm-clerical    Not-in-family   White    Male
1   Married-civ-spouse    Exec-managerial          Husband   White    Male
2             Divorced  Handlers-cleaners    Not-in-family   White    Male
3   Married-civ-spouse  Handlers-cleaners          Husband   Black    Male
4   Married-civ-spouse     Prof-specialty             Wife   Black   Female

   capital-gain  capital-loss  hours-per-week  native-country  Answer  \
0        2174.0           NaN              40   United-States   <=50K
1           NaN           0.0              13   United-States   <=50K
2           NaN           NaN              40   United-States   <=50K
3           NaN           NaN              40   United-States   <=50K
4           0.0           NaN              40            Cuba   <=50K

   reduced_hours
0             35
1              8
2             35
3             35
4             35

=== GROUPED SUMMARY BY relationship AND reduced_hours ===
    relationship  reduced_hours  count
0        Husband              8      1
1        Husband             35      4
2        Husband             40      1
3        Husband             75      1
4  Not-in-family             11      1

Reduction applied per person: 5 hours (SID = 2461075)
```

```python
print(grouped.head(), "\n")

print(f"Reduction applied per person: {n} hours (SID = {sid})")
```

=== ORIGINAL DATA (first 5 rows) ===
```
   age          workclass  fnlwgt  education  education-num  \
0   39          State-gov   77516  Bachelors             13
1   50   Self-emp-not-inc   83311  Bachelors             13
2   38            Private  215646    HS-grad              9
3   53            Private  234721       11th              7
4   28            Private  338409  Bachelors             13

        marital-status          occupation    relationship   race     sex  \
0        Never-married        Adm-clerical  Not-in-family  White    Male
1   Married-civ-spouse     Exec-managerial        Husband  White    Male
2             Divorced   Handlers-cleaners  Not-in-family  White    Male
3   Married-civ-spouse   Handlers-cleaners        Husband  Black    Male
4   Married-civ-spouse      Prof-specialty           Wife  Black  Female

   capital-gain  capital-loss  hours-per-week  native-country  Answer
0        2174.0           NaN              40   United-States   <=50K
1           NaN           0.0              13   United-States   <=50K
2           NaN           NaN              40   United-States   <=50K
3           NaN           NaN              40   United-States   <=50K
4           0.0           NaN              40            Cuba   <=50K
```

=== DATA AFTER APPLYING reduce_hours FUNCTION (first 5 rows) ===
```
   age          workclass  fnlwgt  education  education-num  \
0   39          State-gov   77516  Bachelors             13
1   50   Self-emp-not-inc   83311  Bachelors             13
2   38            Private  215646    HS-grad              9
3   53            Private  234721       11th              7
4   28            Private  338409  Bachelors             13

        marital-status          occupation    relationship   race     sex  \
0        Never-married        Adm-clerical  Not-in-family  White    Male
1   Married-civ-spouse     Exec-managerial        Husband  White    Male
2             Divorced   Handlers-cleaners  Not-in-family  White    Male
3   Married-civ-spouse   Handlers-cleaners        Husband  Black    Male
4   Married-civ-spouse      Prof-specialty           Wife  Black  Female

   capital-gain  capital-loss  hours-per-week  native-country  Answer  \
0        2174.0           NaN              40   United-States   <=50K
1           NaN           0.0              13   United-States   <=50K
2           NaN           NaN              40   United-States   <=50K
3           NaN           NaN              40   United-States   <=50K
4           0.0           NaN              40            Cuba   <=50K
```