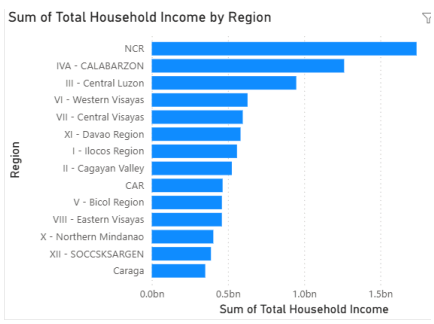
Family Income and Expenditure

By Christian Joseph Ostaga & Elaiza Mae Flores Lim



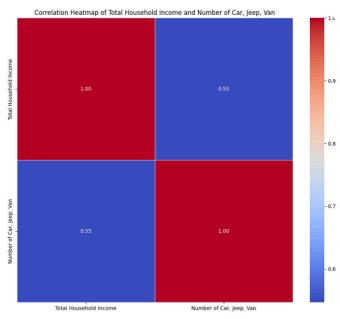
INSIGHT

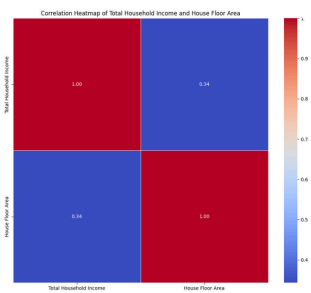
Base on the visualization result we can see that there is a big difference of household income per region showing that on every region base on my research The income disparities across regions in the Philippines can be attributed to a combination of factors, including economic policies, wealth distribution, social systems, and cultural norms. and also The differences in income per region are primarily driven by a combination of factors, including geographical location, infrastructure, education, and government policies. These disparities can significantly impact local communities and economies, affecting not only the economic wellbeing of residents but also their health, education, and overall quality of life. and the one that has the highest total of household income is from the NCR Region with the total of 1,738,159,488.

INSIGHT

Total Household Income And Number of Car , Jeep ,Van

Base on the result we there is a correlation between Total Household Income And Number of Car , Jeep ,Van because if a household owning at least one vehicle, and of owning multiple vehicles, increases significantly as income levels rise.





INSIGHT

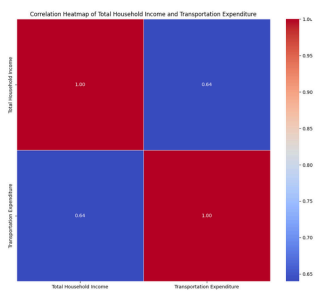
Total Household Income And House Floor Area

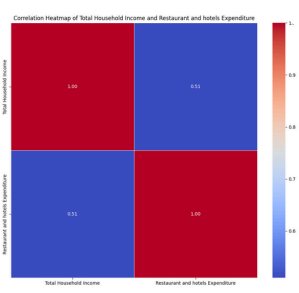
Base on the result we there is a correlation between Total Household Income And House Floor Area a higher income directly translates to greater purchasing power and the ability to finance more expensive properties, which generally have larger floor areas.

INSIGHT

Total Household Income And Transportation Expenditure

Base on the result we there is a correlation between Total Household Income And Transportation Expenditure a larger households with more members tend to have higher transportation needs and therefore higher expenditures.





INSIGHT

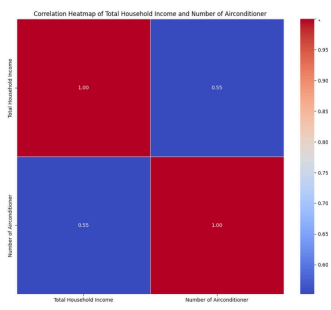
Total Household Income And Restaurant and Hotel Expenditure

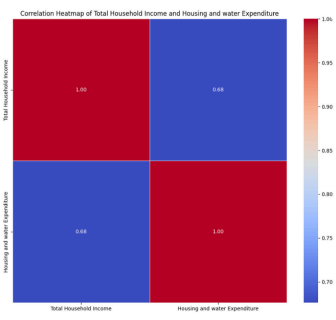
Base on the result we there is a correlation between Total Household Income And Restaurant and Hotel Expenditure if an Individuals with higher incomes tend to spend more on eating out, both in terms of the number of times they dine out and the amount they spend per meal and spending on hotels and holidays, though this spending is often still in the early stages of growth for low-income groups and has considerable potential for increase.

INSIGHT

Total Household Income And Number of Air Conditioner

Base on the result we there is a correlation between Total Household Income And Number of Air Conditioner a higher income households are significantly more likely to own air conditioning, especially more effective systems like central air this shows that even if they were in a lower income family they still bought and use air conditioner and higher income family definitely has a lot of air conditioner on their houses.





INSIGHT

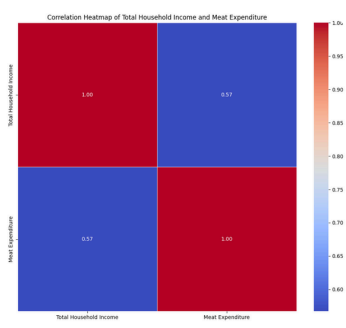
Total Household Income And Water Expenditure

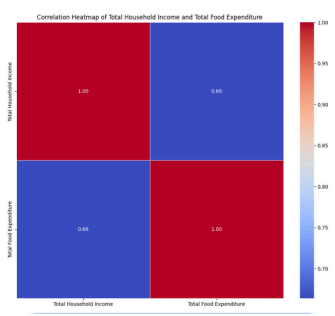
Base on the result we there is a correlation between Total Household Income And Water Expenditure because households are more likely to have better, more reliable water infrastructure, such as private piped water connections, while lower-income households may depend on less reliable sources.

INSIGHT

Total Household Income And Meat Expenditure

Base on the result we there is a correlation between Total Household Income And Meat Expenditure because Households in low and middle income brackets often increase their consumption of more accessible and cheaper meats, such as poultry (chicken) and pork.





INSIGHT

Total Household Income And Total Food Expenditure

Base on the result we there is a correlation between Total Household Income And Total Food Expenditure means Higher income leads to greater overall spending on food, both at home and away from home.

INSIGHT

Based on the result we can all see that the data that the csv contains is legit and accurate based on the Decision Tree with the total of 100% accuracy .

Prediction Using Linear Regression

```
[12]: from sklearn.linear_model import LinearRegression

[13]: linear_model = LinearRegression()
      linear_model.fit(X,y)

[13]: LinearRegression
      ▸ Parameters

[14]: linear_model_predisct = linear_model.predict(X)

[15]: r2 = r2_score(y, linear_model_predisct)
      print("R² Score:", r2)
      print(f"The model variation is at {r2:.1%}")

      R² Score: 0.6789288962667498
      The model variation is at 67.9%
```

INSIGHT

Based on the result in Linear Regression there is 67.9% that the CSV data is still containing an accurate data because still the result is still high but not high like what the result we got on the Decision Tree.

Decisition Tree

```
[39]: from sklearn.tree import DecisionTreeRegressor

      my_model = DecisionTreeRegressor()
      my_model.fit(X, y)

[39]: DecisionTreeRegressor
      ▸ Parameters

[40]: predictions = my_model.predict(X)

[41]: predictions

[41]: array([117048.,  67766.,  61609., ...,  45882.,  81416.,  78195.],
            shape=(41544,))

      TESTING THE ACCURACY OF PREDICTIONS

[11]: from sklearn.metrics import r2_score

      r2 = r2_score(y, predictions)
      print("R² Score:", r2)
      print(f"The model variation is at {r2:.1%}")

      R² Score: 1.0
      The model variation is at 100.0%
```