Data Analysis on Filipino Family Income and Expenditure

By: Johann Sebastian E. Catalla, BSCS-I

An output for the course CP102: Computer Programming 2

The dataset was sourced through kaggle from the Philippine Statistics Authority's (PSA) Family Income and Expenditure Survey (FIES) nationwide. The survey, which is undertaken every three (3) years, is aimed at providing data on family income and expenditure, including, among others, levels of consumption by item of expenditure, sources of income in cash, and related information affecting income and expenditure levels and patterns in the Philippines.

The Dataset contains more than 40k observations and 60 variables which is primarily comprised of the household income and expenditures of that specific household. For this analysis, I will be using Python's **pandas**, **numpy**, and **hvplot** libraries.

In [ ]:
```
import pandas as pd
import numpy as np
import hvplot.pandas
import holoviews as hv
from bokeh.models.formatters import NumeralTickFormatter
formatter = NumeralTickFormatter(format="0,0")
```

In [ ]:
```
dataset = pd.read_csv("Family Income and Expenditure.csv")
dataset
```

Out[ ]:

	Total Household Income	Region	Total Food Expenditure	Main Source of Income	Agricultural Household indicator	Bread and Cereals Expenditure	Total Rice Expenditure	Meat Expenditure	Total Fish and marine products Expenditure	Fruit Expenditure	...	Number of Refrigerator/Freezer	Number of Washing Machine	Nu Aircon
0	480332	CAR	117848	Wage/Salaries	0	42140	38300	24676	16806	3325	...	1	1	
1	198235	CAR	67766	Wage/Salaries	0	17329	13008	17434	11073	2035	...	0	1	
2	82785	CAR	61609	Wage/Salaries	1	34182	32001	7783	2590	1730	...	0	0	
3	107589	CAR	78189	Wage/Salaries	0	34030	28659	10914	10812	690	...	0	0	
4	189322	CAR	94625	Wage/Salaries	0	34820	30167	18391	11309	1395	...	1	0	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	
41539	119773	XII - SOCCSKSARGEN	44875	Enterpreneurial Activities	1	23675	21542	1476	6120	1632	...	0	0	
41540	137320	XII - SOCCSKSARGEN	31157	Enterpreneurial Activities	1	2691	1273	1886	4386	1840	...	0	0	
41541	133171	XII - SOCCSKSARGEN	45882	Enterpreneurial Activities	2	28646	27339	480	4796	1232	...	0	0	
41542	129500	XII - SOCCSKSARGEN	81416	Enterpreneurial Activities	1	29996	26655	2359	17730	2923	...	0	0	
41543	128598	XII - SOCCSKSARGEN	78195	Enterpreneurial Activities	1	43485	41205	1985	7735	2062	...	0	0	

41544 rows × 60 columns

In [ ]:

In [ ]:
```
# adding the food to income ratio heading
dataset["Food Percentage to Income"] = dataset["Total Food Expenditure"] / dataset["Total Household Income"]
# ratio of food to income
res = dataset.groupby(by="Region")[["Food Percentage to Income"]].mean().sort_values(by="Food Percentage to Income").round(2)

total_food = dataset.groupby("Region")[["Total Food Expenditure", "Total Household Income"]].mean().round(2).sort_values(by=["Total Food Expenditure", "Total
```

```
total_food.hvplot.barh(title="Regions in the Philippines sorted by \nAverage Yearly Food Expenditure and Average Income",color=["pink", "darkgreen"],\
                       fontscale=0.9, height=700,width=800, xformatter=formatter, line_color=None,\
                       ylabel="Yearly Food Expenditure per Family", xlabel="Region in the Philippines",\
                       fontsize={'xticks': 10, 'yticks': 10}, stacked=True, legend="top") + res.hvplot.barh(height=700)
```

Out[ ]:

In [ ]:
```
dataset.shape
```

Out[ ]:
```
(41544, 60)
```

Total income and food expense ratio

In [ ]:

In [ ]: