

## Datakind April 2025 Challenge: Exploring International Economic Opportunity Data and Insights - the Philippines

### Philippines Overview<sup>1</sup>

#### Geography<sup>[25]</sup>

Land area (in km <sup>2</sup> )	300,000
Regions	17
Provinces	81
Cities	146
Municipalities	1,488

#### Population<sup>[26]</sup>

	2015	2020	2022	2023
Population, in millions <sup>[27]</sup>	101.0	109.0	110.5	112.9
Adult population (aged 15 +), in millions	69.3	75.8	78.5	79.7
Number of households, in millions	23.0	26.4	---	---

#### Economy<sup>[25]</sup>

GDP Y-o-Y growth (Fourth Quarter 2022 vs. Fourth Quarter 2023)	5.5%
GDP full year growth (2023)	5.5%
Inflation rate (September 2024)	1.9%
Unemployment rate (July 2024)	4.7%

#### Income and Poverty<sup>[25]</sup>

Average annual family income, in PhP (2021)	307,190
Minimum wage in NCR for non-agriculture, in PhP (Effective: July 2023)	610.00
Poverty incidence among population (1st sem, 2023)	22.4% <sup>B/</sup>
Poverty threshold, average monthly estimate (family of five), in PhP (2023)	13,797 <sup>B/</sup>

### Financial Inclusion in the Philippines

The World Bank identified financial inclusion as a “catalyst for achieving seven of the 17 Sustainable Development Goals.”<sup>2</sup> The Philippines has been working toward greater inclusion since 2015 with the launch of National Strategy for Financial Inclusion (NSFI). In 2016, Executive Order No. 208 created the Financial Inclusion Steering Committee (FISC) as the interagency governing body that leads the coordinated and collaborative approach in implementing the NSFI and provides guidance in the development of its various policies, regulations, supervisory

---

<sup>1</sup>[https://www.bsp.gov.ph/Media\\_And\\_Research/Financial%20Inclusion%20Dashboard/2023/FIDashboard\\_4Q2023.pdf](https://www.bsp.gov.ph/Media_And_Research/Financial%20Inclusion%20Dashboard/2023/FIDashboard_4Q2023.pdf)

<sup>2</sup><https://www.worldbank.org/en/topic/financialinclusion/overview>

frameworks, programs, and initiatives.<sup>3</sup> Since then, the Philippines have made substantial strides in financial inclusion, with adult bank account ownership climbing from 26.6% in 2011 to 51.4% in 2021, according to the World Bank Global Findex. Despite this progress, approximately 37.6 million Filipinos remain unbanked, placing the country among the top 10 globally with the highest number of unbanked adults.<sup>4</sup>

According to the United Nations Secretary-General's Special Advocate for Financial Health,

*This is due, in part, to geography. The country consists of around 7,000 islands, many of which are sparsely populated and difficult to reach. Furthermore, they often lack telecom connectivity. For financial institutions, it is therefore difficult and not lucrative to provide services. As a result, low-income communities, small-scale businesses, smallholder farmers, fishing communities, and women are left without secure, affordable financial services.*

### Datakind's Kick-off Challenge

By exploring publicly available data, Datakind is seeking to “paint a comprehensive picture of the financial and economic opportunity” that will enable communities support local decision-making based on insights that should lead to positive outcomes, and suggested the following steps:<sup>5</sup>

1. Download country-specific suggested datasets, as available.
2. Research additional datasets for that country that provide information on financial inclusion and economic opportunity, focusing focus on the most recently available datasets (e.g. anything from 2018 onwards).
3. Identify the unit of analysis for each dataset (e.g individual, household, county, region).
4. Identify the variables that are representative of or proxies for economic development and financial well-being.
5. Transform data to be at the same unit of analysis, and append datasets for the selected country.
6. Execute exploratory data analysis of the data.

For the Philippines in particular, the suggested unit of analysis was the province-level and data downloaded from the following sources:

- Family Income and Expenditure Survey (FIES) – Income, spending, race, gender (<https://psa.gov.ph/>)

Provides data on family income and expenditure which include among others levels of consumption by item of expenditure as well sources of income in cash and in kind for 2023,2021,2018. Volume II (Provincial and Key City/Municipality) - presents data at the **national, regional, provincial, and key city/ municipality** levels and contains 2.8 GB. Need to register with PSA before downloading.

---

<sup>3</sup> <https://www.bsp.gov.ph/Pages/InclusiveFinance/NSFI-2022-2028.pdf>

<sup>4</sup> <https://www.unsgsa.org/country-visits/scaling-financial-services-better-financial-health-philippines>

<sup>5</sup> <https://github.com/datakind/datakit-financial-inclusion-2025/discussions/1>

- Bangko Sentral ng Pilipinas (BSP) Financial Inclusion Reports – Banking access & financial literacy (<https://www.bsp.gov.ph/>)

Posts Financial Inclusion reports by the NSFI and FISC; Biennial Financial Inclusion Surveys of 1,200 respondents grouped into **multi-regional geographic** areas; National financial inclusion national data are available quarterly in PDFs (2016-2023) and interactive dashboards for national and individual regions quarterly for 2024 and later. Does not appear any data can be downloaded directly from this website.

- Philippine Population and Housing Census & Labor Force Survey (LFS) – (<https://psa.gov.ph/>)

The Philippine Statistics Authority (PSA) conducts the **Labor Force Survey** each month. The Labor Force Survey is a nationwide survey of approx. 47,000 households grouped by **region** that contains data on the demographic and socio-economic characteristics of the population. Work location is available by provinces and HUC's. Need to register with PSA before downloading.

- Department of Trade and Industry (DTI) Micro Small Medium Establishment Data – Small business growth & investment (<https://www.dti.gov.ph/resources/msme-statistics/>)

**Regional** data on Number of and Employment by Region is published in a PDF report for 2023. The data must be pasted into a text file, the variable names renamed for successful conversion into a csv file. Previous year reports can be found on amazon web services site: <https://dtiwebfiles.s3.ap-southeast-1.amazonaws.com/>

- National Housing Authority Data – Housing affordability (<https://www.nha.gov.ph/>)

Unable to find any relevant data.

## Step 1: Download Suggested Datasets

See above for a detailed description of the suggested datasets. Although there is an abundance of financial inclusion, demographic, banking, and economic data, they were not available for download nor granular enough for robust analysis; for example, annual Family Income and Expenditure Survey (FIES) datasets were available on the provincial and high urban city (HUC) level for hundreds of variables, but microlevel, demand-side financial inclusion data could only be found in the printed Financial Inclusion Surveys of 1,200 respondents that are conducted every three years and aggregated into broad geographic areas comprising multiple regions.

This proved to be a time-consuming task that was not easily automated nor replicated by employing APIs or executing data transformation code. There were a handful of instances where the data could be downloaded as a data frame, other times the data had to be manually pasted from an HTML or PDF table. But most of the time, the data was contained in Excel spreadsheets that were inconsistent in the unit of analysis names and order.

Additionally, downloading longitudinal data was a challenge: banking and FIES data was accessible, but 2015 census data was not.

## Steps 2 and 3: Research Additional Datasets and Identify the Unit of Analysis for Each Dataset

In its kickoff challenge, Datakind organizers suggested compiling provincial-level data<sup>6</sup>, but the available data proved to be less comprehensive than those segmented by regions. Furthermore, there are autonomous High Urban Cities (HUC) that do not necessarily report to a province, province, and they are often reported with provincial data.

Additional data - predominantly segmented by **region** and in the form of excel spreadsheets that did not conform to the dataframe format - were retrieved from the following sites:

- Annual Bank Physical Network and Deposit Accounts data compiled by Bangko Sentral ng Pilipinas: <https://www.bsp.gov.ph/SitePages/Statistics/Statistics.aspx>
- Selected data extracts from the Family and Expenditure Survey - 2018, 2021, 2023p – on family income, family expenditures, Gini coefficient, <https://psa.gov.ph/statistics/income-expenditure/fies>
- Selected data extracts from the Census of Population and Housing, which is conducted every five years on population, density, school attendance, land ownership, internet access, literacy rate, household characteristics: <https://psa.gov.ph/statistics/population-and-housing/stat-tables>

## Step 4: Identify the variables that are representative of or proxies for economic development and financial well-being (Refer to Appendix for list of variables)

### Financial Inclusion Metrics:

According to Mojica and Mapa (2017):

*While there has been a consensus on the importance of financial inclusion, efforts to measure it remain incomplete (Camara and Tuesta, 2014), the issue of its robust measurement is still outstanding (Amidžić et al., 2014), and there is no standard method by which it can be measured (Park and Mercado, 2015). The challenge comes from the fact that financial inclusion is a multidimensional phenomenon.<sup>7</sup>*

There is a clear consensus on “the need to treat financial inclusion as a multidimensional concept but little agreement on what its main dimensions are.”<sup>8</sup> Research has been conducted on constructing financial inclusion indices incorporating attempts to address this multi-dimensional challenge, but it appears there is no widespread agreement on the conceptual or methodological approach.

Recent definitions have coalesced around four dimensions for measuring financial inclusion:

---

<sup>6</sup> It should be noted that due the presence of autonomous High Urban Cities (HUC) that do not necessarily report to a province, data is often segmented by 118 province + HUC or HUC + local municipalities.

<sup>7</sup> Mojica, M.B.R. & Mapa, Dennis. (2017). An index of financial inclusion in the Philippines: Construction and analysis. *Philippine Statistician*. 66. 59-74. [https://www.psai.ph/docs/publications/tps/tps\\_2017\\_66\\_1\\_5.pdf](https://www.psai.ph/docs/publications/tps/tps_2017_66_1_5.pdf)

<sup>8</sup> Pesqué-Cela V, Tian L, Luo D, Tobin D, Kling G. Defining and measuring financial inclusion: A systematic review and confirmatory factor analysis. *J. Int. Dev.* 2021; 33: 316–341. <https://doi.org/10.1002/jid.3524>

- Access, e.g. number of bank ATMs / branches, post offices, mobile users, availability & variety of financial services
- use, e.g. deposit and loan data
- cost, e.g. fees, minimum account size, minimum loan size
- quality, e.g. consumer satisfaction, financial literacy and protection.

It should be noted that the sub-national data described in **Steps 1** and **Steps 2/3** above are aggregate not micro-level indicators that inform on individuals' access to or use of financial services. As discussed above, metrics that would be useful to explore and segmented by province or region, such as those collected for the Financial Inclusion Surveys – % uptake of financial services, types of account owned, frequency and type of financial transactions -- could not be accessed.

#### **Step 5: Transform data to be at the Same Unit of Analysis and Append Datasets.**

This proved to be a complex and arduous process that was not easily automated or replicated by employing APIs or executing code to convert the data tables into usable dataframes. There were a handful of instances where the data could be downloaded as a data frame; other times the data had to be manually pasted from an HTML or PDF table; but most of the time, the data was contained in Excel spreadsheets that contained column names that did not conform to dataframe variable naming conventions, rows with different units of analysis, blank lines, comment lines, totals lines. Because these were small files, Excel was the most efficient way to transform, clean, and standardize the data into a usable structure. Moreover, it appeared the data required minimal cleaning. Ideally, longitudinal data would be useful.

2 datasets were compiled:

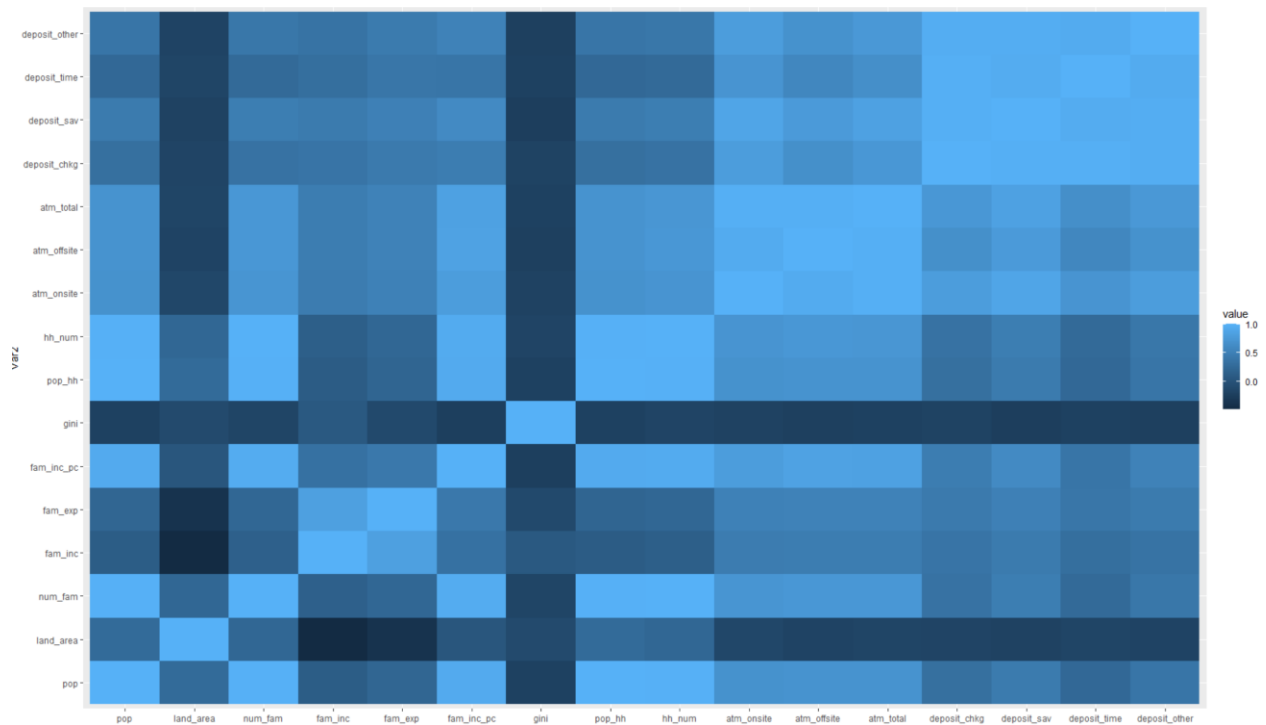
1. "philippine.2021fies.2020census.2021other.data.csv": 52 - 2020 census, 2021 income, 2020 internet usage, 2021 banking - variables for 17 regions
2. "philippines.census2020.feic2021.banking2021.data.province\_huc.csv": 19 - 2020 census, 2021 income, 2021 banking - variables for 106 province+HUCs.

It should be noted that the analyses of these datasets are unlikely to reveal any useful actionable insights, due to the exclusion of financial inclusion metrics.

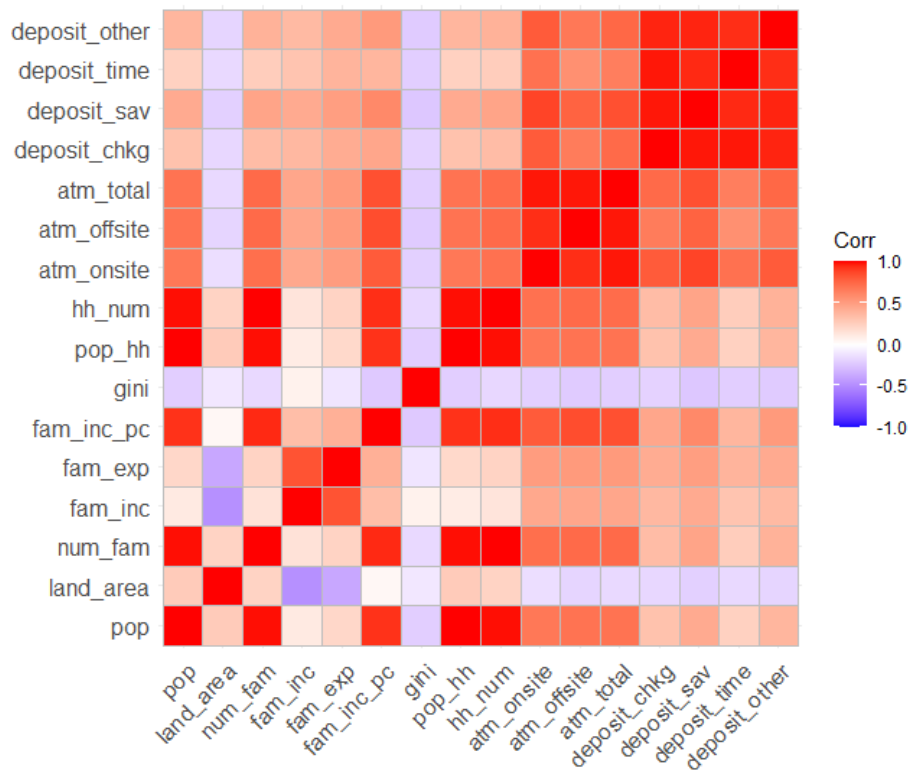
#### **Step 6: Execute exploratory data analysis of the data.**

Several data visualizations of cross-correlations were created in R (script saved in 'philippines.province\_huc.visualizations.R')

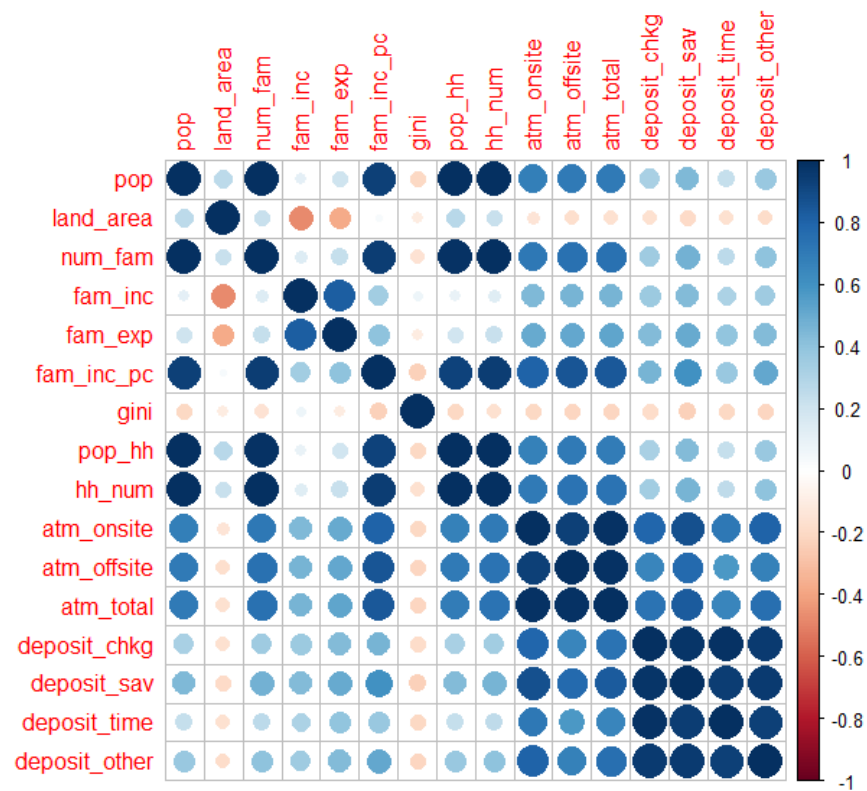
#1 – ggplot () function: hard to see



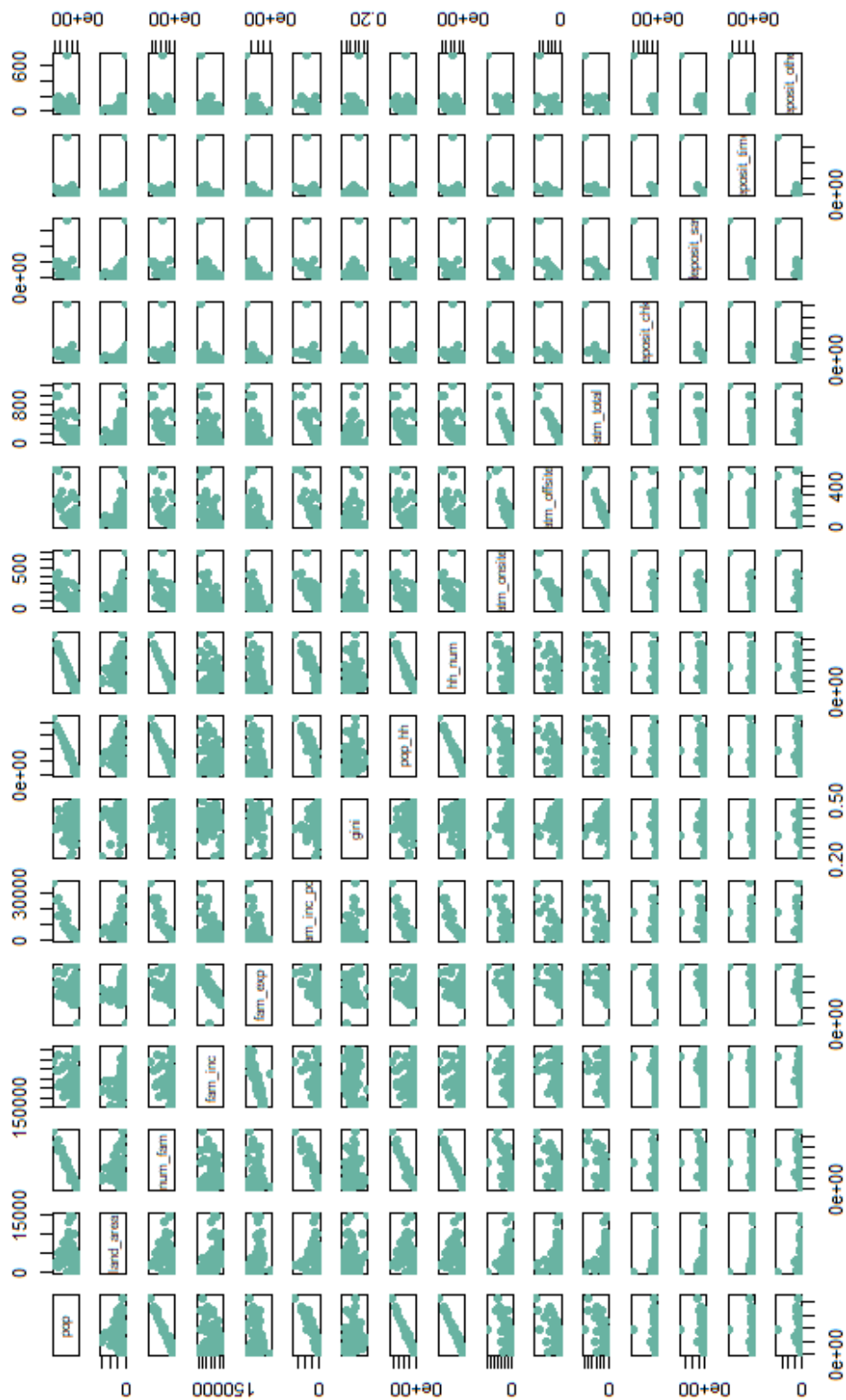
#2 – ggcorrplot() function: the color of the squares in the correlation matrix help visualize the correlations between each variable.



#3: corrplot() function: the color and size of the circles in the correlation matrix help visualize the correlations between each variable.



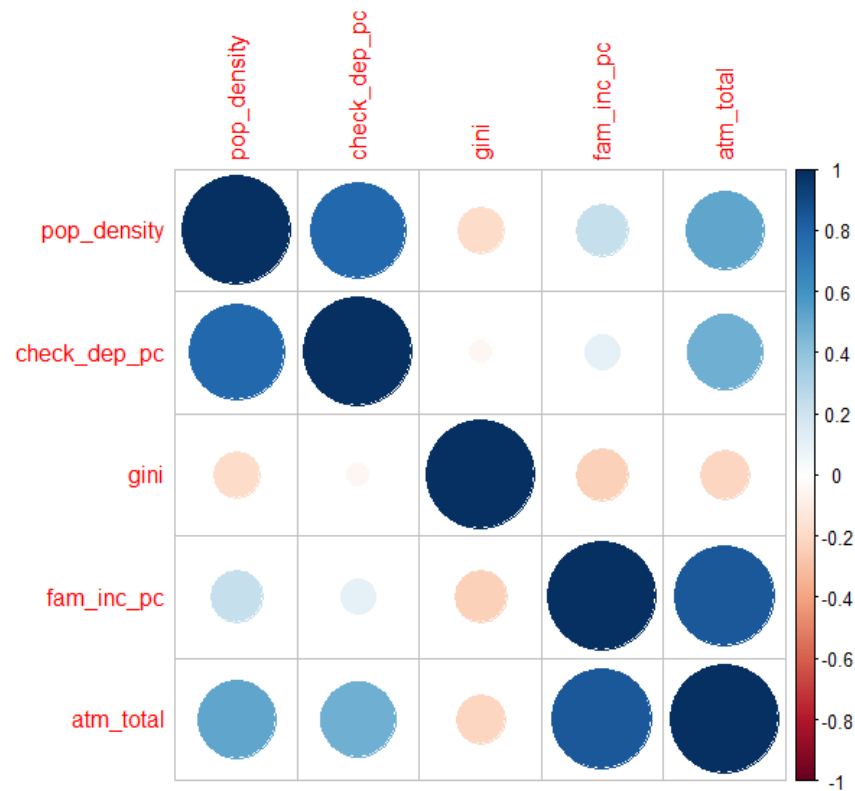
#4: plot() function: the scatter plot matrix shows all the pairwise scatterplots of the variables on in a matrix format.





#5: scatterplot matrix of:

- 2020 population per square kilometer by province
- Checking deposits per capita
- Gini coefficient
- Total # ATMs
- Median family income



## Appendix: Regional Dataset Variables: “philippine.2021fies.2020census.2021other.data.csv”

unit of analysis: region	variables	year	available by province+ HUC?
NATIONAL CAPITAL REGION	Gini_coeff	2021	yes
CORDILLERA ADMINISTRATIVE REGION (CAR)	GDP_percap_current_prices	2021	
REGION I (ILOCOS REGION)	ATM_universal	2021	yes
REGION II (CAGAYAN VALLEY)	ATM_thrift	2021	yes
REGION III (CENTRAL LUZON)	ATM_rural	2021	yes
REGION IV-A (CALABARZON)	Population	2020	yes
MIMAROPA REGION	Land_area	2013	yes
REGION V (BICOL REGION)	MICRO_est	2021	
REGION VI (WESTERN VISAYAS)	SMALL_est	2021	
REGION VII (CENTRAL VISAYAS)	MEDIUM_est	2021	
REGION VIII (EASTERN VISAYAS)	total_MSMEs	2021	
REGION IX (ZAMBOANGA PENINSULA)	LARGE_est	2021	
REGION X (NORTHERN MINDANAO)	MICRO_emp	2021	
REGION XI (DAVAO REGION)	SMALL_emp	2021	
REGION XII (SOCCSKSARGEN)	MEDIUM_emp	2021	
REGION XIII (CARAGA)	MSMEs_emp	2021	
BANGSAMORO AUTONOMOUS REGION IN MUSLIM MINDANAO (BARMM)	LARGE_emp	2021	
	HH_number	2020	yes
	HH_pop	2020	yes
	Median_family_income	2021	yes
	HH_refrigerator	2020	
	HH_stove	2020	
	HH_microwave	2020	
	HH_washer	2020	
	HH_air_conditioning	2020	
	HH_fan_cooling	2020	
	HH_radio	2020	
	HH_television	2020	
	HH_disc_player	2020	
	HH_audio_equipment	2020	
	HH_landline_phone	2020	
	HH_mobile_phone	2020	
	HH_tablet	2020	
	HH_personal_computer	2020	
	HH_motor_vehicle	2020	
	HH_motorcycle	2020	
	HH_bicycle	2020	
	HH_motor_boat	2020	
	HH_nonmotor_boat	2020	
	HH_with_internet	2020	
	HH_wired_BB	2020	
	HH_wireless_bb	2020	
	HH_satellite_bb	2020	
	hh_mobile_bb	2020	
	HH_use_internet_home	2020	
	HH_use_internet_work	2020	
	HH_use_internet_school	2020	
	HH_use_internet_otherhome	2020	
	HH_use_internet_public_place	2020	
	HH_use_internet_cafe	2020	
	HH_use_internet_private_est	2020	
	HH_use_internet_in_mobility	2020	

**“philippines.census2020.feic2021.banking2021.data.province\_huc.csv”: Province+HUC Dataset**  
**Variables**

region\_name  
unit of analysis: province\_huc  
pop\_2015  
pop\_2020  
land\_area\_sq\_km\_2013  
family\_num-2021  
avg\_fam\_inc\_2021  
avg\_fam\_exp\_2021  
fam\_income\_pc\_median-2021  
gini\_coeff\_2021  
HH\_pop\_2020  
HH\_num\_2020  
ATM\_onsite\_num\_2021  
ATM\_offsite\_num-2021  
ATM\_total\_num\_2021  
deposit\_demand\_2021\_mmpeso  
deposit\_savings\_2021\_mmpeso  
deposit\_time\_2021\_mmpeso  
deposit\_other\_2021\_mmpeso

## Additional References and Links

General data statistics

Philippine Statistics Authority: <https://psada.psa.gov.ph/query>

Philippine Statistics Authority OpenSTAT data platform: <https://openstat.psa.gov.ph/>

Philippine Social Science Council: <https://www.data.pssc.org.ph/docs/>

BSP Financial Inclusion Interactive Dashboard (quarterly, 2024 - ):

<https://financialinclusion.gov.ph/fi-dashboard/>

## Reports, Research Papers and Surveys:

Bangko Sentral ng Pilipinas (BSP) Financial Inclusion Dashboard Reports (quarterly, 2016-2023): <https://www.bsp.gov.ph/Pages/MediaAndResearch/FinancialInclusionDashboard.aspx>

BSP National Baseline Survey on Financial Inclusion (2015):

<https://www.bsp.gov.ph/Inclusive%20Finance/Financial%20Inclusion%20Reports%20and%20Publications/2015/NBSFIFullReport.pdf>

BSP 2021 Financial Inclusion Survey:

<https://www.bsp.gov.ph/Inclusive%20Finance/Financial%20Inclusion%20Reports%20and%20Publications/2021/2021FISToplineReport.pdf>

Pesqué-Cela V, Tian L, Luo D, Tobin D, Kling G. Defining and measuring financial inclusion: A systematic review and confirmatory factor analysis. J. Int. Dev. 2021; 33: 316–341.

<https://doi.org/10.1002/jid.3524>

Mojica, M.B.R. & Mapa, Dennis. (2017). An index of financial inclusion in the Philippines: Construction and analysis. Philippine Statistician. 66. 59-74.

[https://www.psa.gov.ph/docs/publications/tps/tps\\_2017\\_66\\_1\\_5.pdf](https://www.psa.gov.ph/docs/publications/tps/tps_2017_66_1_5.pdf)

BSP Consumer Finance Surveys (2021, 2018):

<https://www.bsp.gov.ph/SitePages/MediaAndResearch/ConsumerFinanceSurvey.aspx>