**JSIS 595 / LAW B 554**

**Assignment 1: Graphic and Numeric Presentation of Data**

**Name: Lauren Hwayoung Lee**

1. **Type of Variables:**

For each of the following items, indicate what type of variable it is measuring?

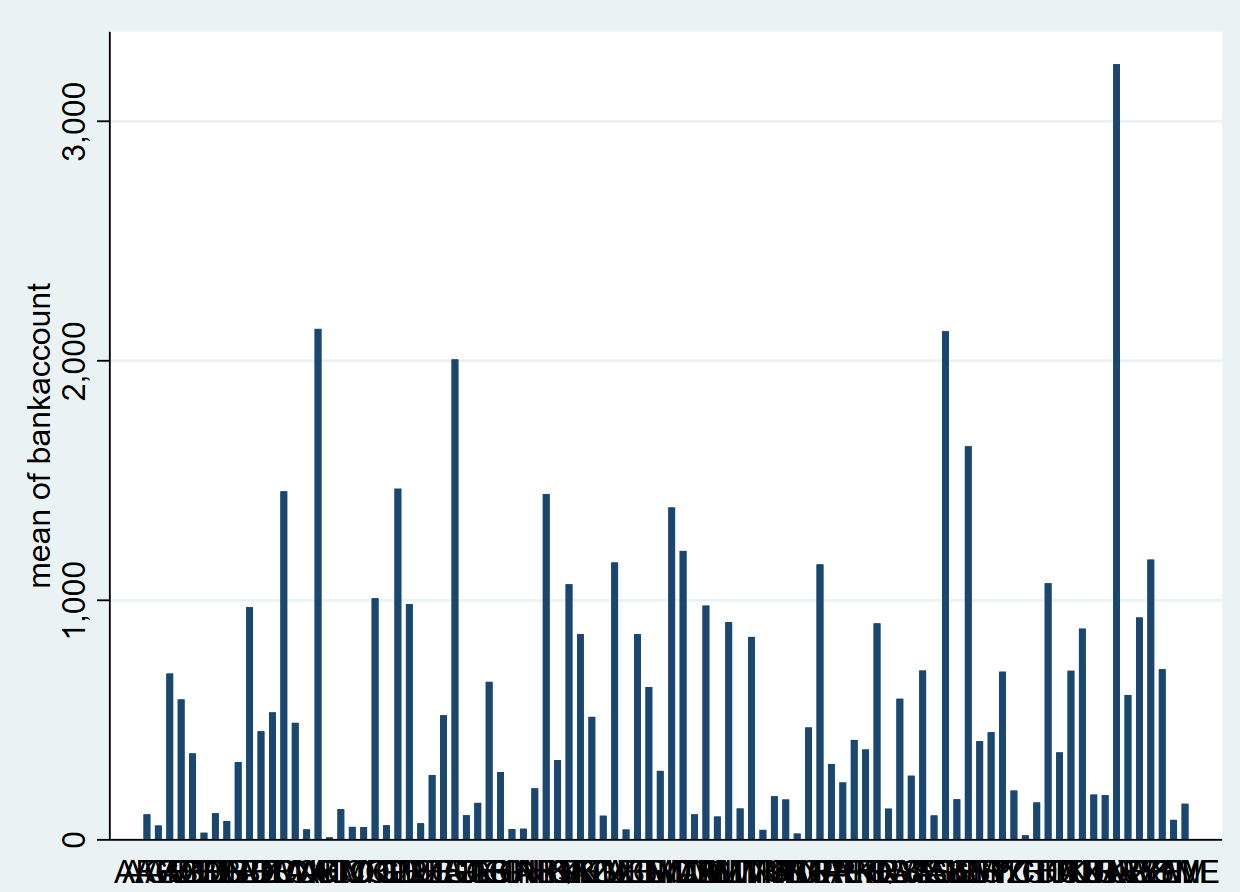
* 1. The speed of runners in a race timed with a stopwatch; **Quantitative variable (interval-speed)**
  2. A zoologist counts the number of tigers, lions, and elephants she saw in a designated conservation area; **Quantitative variable (interval – number of animals)**
  3. A convenient store takes an inventory of all items at the end of the month; **Quantitative variable (interval-number of inventory)**
  4. During an election season, a political poll asks voters to rank the running candidates from their favorite to their least favorite; **Quantitative variable (ordinal-favorability)**
  5. During an election season, a political poll asks voters about their party affiliation; **Categorical variable (party affiliation)**

1. **Choices of Graphs**

Download the following data set from the Canvas website: **HW1\_Bank Acc\_GDPpc.dta**. This is a cross-country dataset with variables such as (1) number of bank accounts per 1000 adults, from the Global Financial Development database, and (2) GDP per capita (constant 2011 US$) from the World Development Indicators database, covering all available countries and 20 years from 1998 to 2018. Variable (1) can be used to measure financial development, and (2) used to measure economic development.

Depending on different goals of your presentation, what kind of graph would you choose to present the data? Please write the name of the graph on the line, and make a plot using the data you just downloaded.

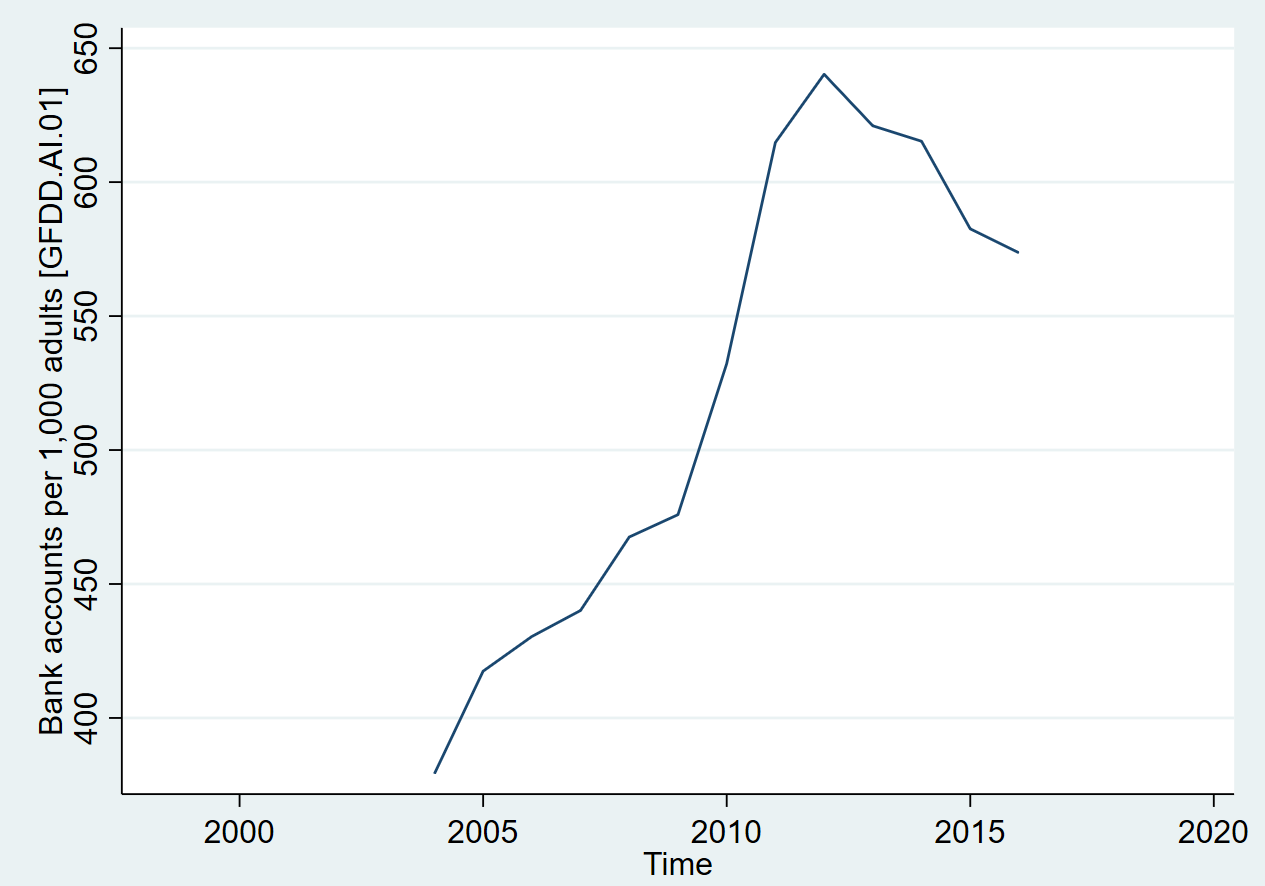
* 1. To show the distribution of countries in terms of their levels of financial development in year 2010 **Bar Chart**



**drop if missing(bankaccount)**

**graph bar bankaccount if year(2010), over(countrycode)**

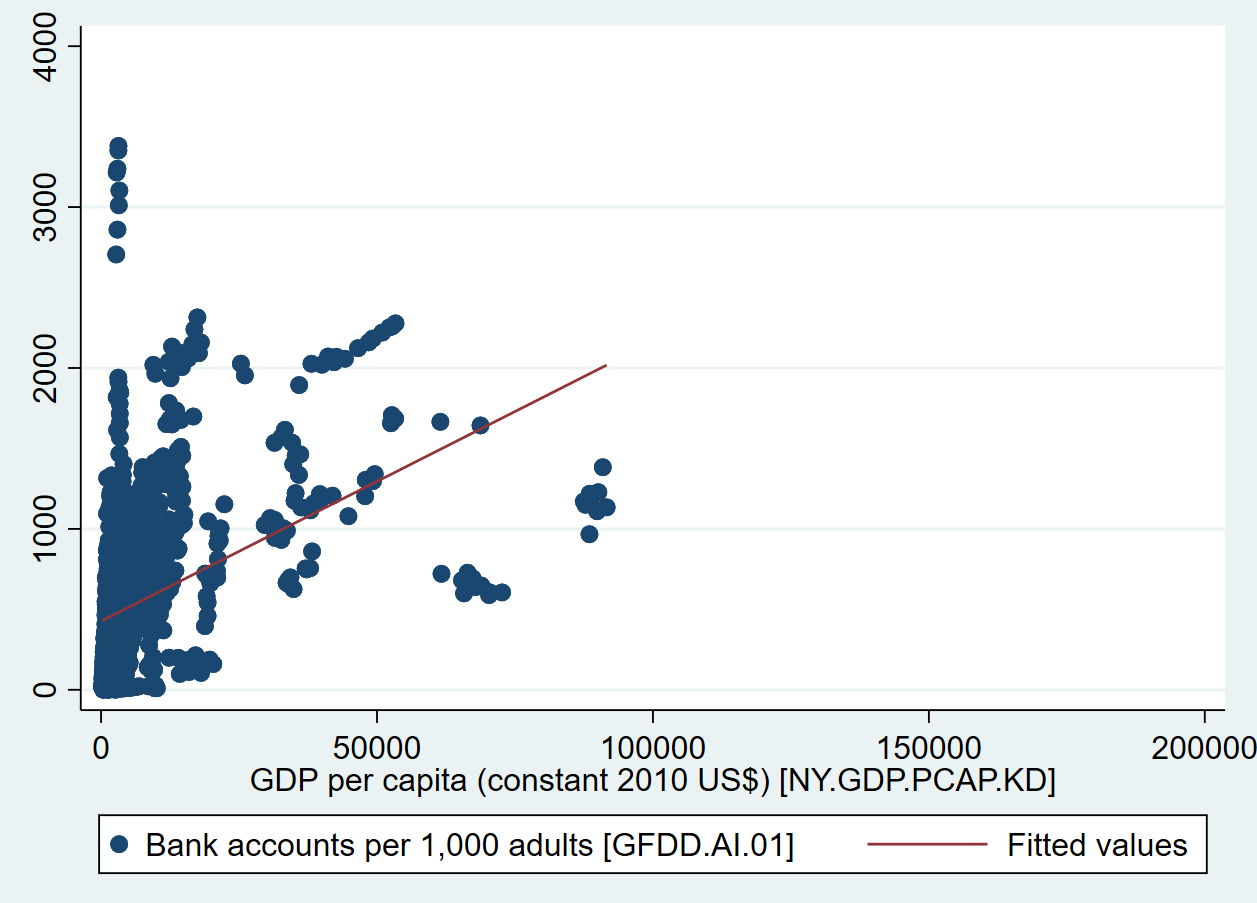
* 1. To show how financial development has evolved in Brazil from 1998 to 2018 **Histogram**

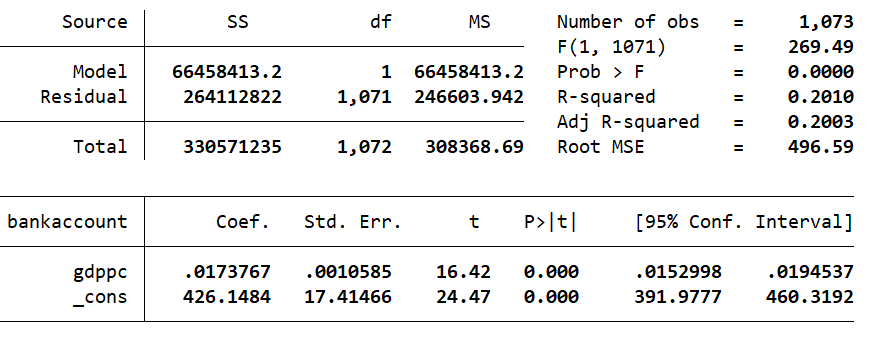


**keep if countryname=="Brazil"**

**twoway line bankaccount year**

* 1. To show the correlation between financial and economic development **Scatterplot**





**twoway scatter bankaccount gdppc || lfit bankaccount gdppc**

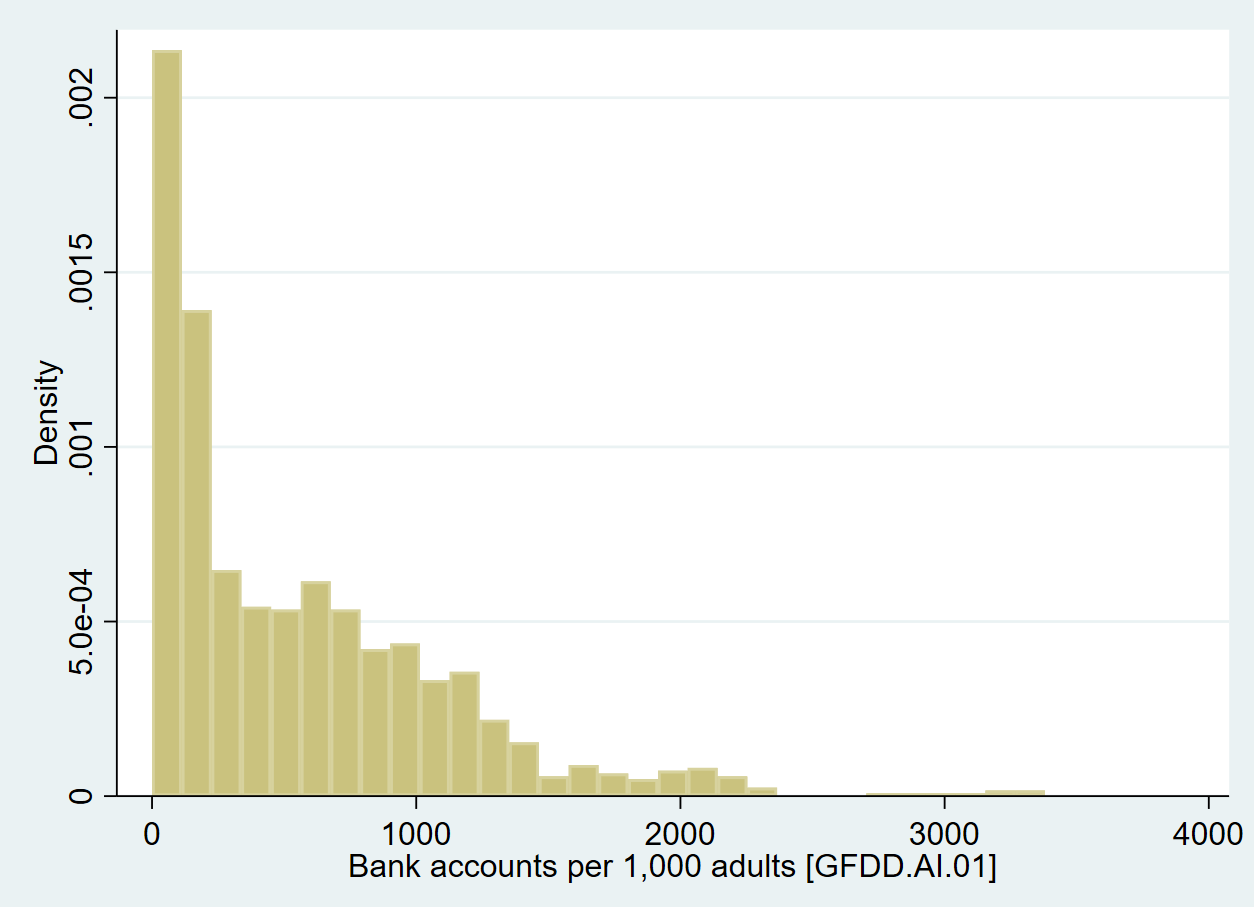
**reg bankaccount gdppc**

* 1. To show the distribution of financial development in year 2010 for middle-level income countries (you can define the middle-level income countries as the middle 50%)
  2. To show the time trend of country-level inequality in financial development.

1. **Choices of Statistics**

Use the same dataset as the previous question. Depending on different goals of your presentation, what kind of statistics would you choose to present the data? Please write the name of the statistics on the first line, the Stata code or Excel code you will use to calculate the statistics on the second line, and calculate the statistics on the third line.

* 1. To show how spread out financial developments are among the countries in year 2010.
* **Name of the statistics: Histogram**
* **Stata code: histogram bankaccount if year(2010)**
* **Statistics and Calculation**



* 1. To show what is a middle-level income country’s GDP per capita like in the year of 2010.
* **Name of the statistics**
* **Stata code: graph matrix forest year**
* **Statistics and Calculation**

1. **Basic Statistics**

The score on a religiosity scale were obtained for 46 adults. For the following simple frequency distribution, calculate (a) the mode, (b) the median, (c) the mean:

|  |  |
| --- | --- |
| Religiosity score | frequency |
| 10 | 6 |
| 9 | 2 |
| 8 | 4 |
| 7 | 7 |
| 6 | 8 |
| 5 | 3 |
| 4 | 5 |
| 3 | 2 |
| 2 | 7 |
| 1 | 2 |
|  | N= 46 |

**If religiosity score is x, frequency is f,**

1. **Mode: Religiosity 6 has the maximum frequency (8).**
2. **Median: 46(n)2 = 23. 23rd observation is 6.**
3. **Mean: 26446**
4. **Basic statistics**

On a measure of authoritarianism, seven countries scored as follows:

6, 3, 2, 1, 6, 7, 10

Calculate the (a) range, (b) inter-quartile range, (c) variance and standard deviation

1. **Range: Biggest value (10) – Smallest value (1) =**
2. **Inter-quartile Range: Upper quartile (7) - Lower quartile (2).**
3. **Variance and standard deviation:**
4. **Global Deforestation**

Please download the following data from the Canvas website: **HW1\_Deforestation.dta**. Use graphs, statistics, and logical argument to answer the following questions.

1. What is the world trend of deforestation during this period? Make a plot to show your point.
2. Is there a correlation between deforestation and economic development across countries?
3. Is there a correlation between deforestation and economic development within a country over time?
4. Is the pattern shown in (2) the same as that in (3)?
5. In (1), you probably used the average value of forest area of all countries to show the trend. This is a good indicator as a first step. But to further study the issue of global deforestation, do you think there could be problems with this indicator? How about big countries such as US and China, and small countries such as Aruba and Singapore?
6. **Contract Enforcement around the World**

Please download the following data from the Canvas website: **HW1\_DoingBusiness.dta**

Use the above datasets to answer the following questions:

1. Is the world converging or diverging in terms of the easiness of contract enforcement? Make a plot(s) to support your point.
2. Is there a gender difference in the easiness of starting business around the world in year 2010? Use some statistics to support your argument.
3. Are poor countries associated with weaker contract enforcement? Use both plots and statistics to support your point.
4. Are there any outliers in (3)? Who are the outliers? And why do you think these outliers deviate from the global pattern? Pick one outlier country to analyze.
5. Merge the Doing Business data with the Financial Development data in question 2. Is there any correlation between financial development and the easiness of starting a business?