Visually Do Economics Data Analysis by Mathematica #1 Introduction

2nd December, 2018

Gakushuin University, Faculty of Economics
Prof. Yukari SHIROTA

The data
cited from https://www.bps.go.id/subject/52/produk-domestik-regional-bruto--lapangan-usaha-.html#subjekViewTab3

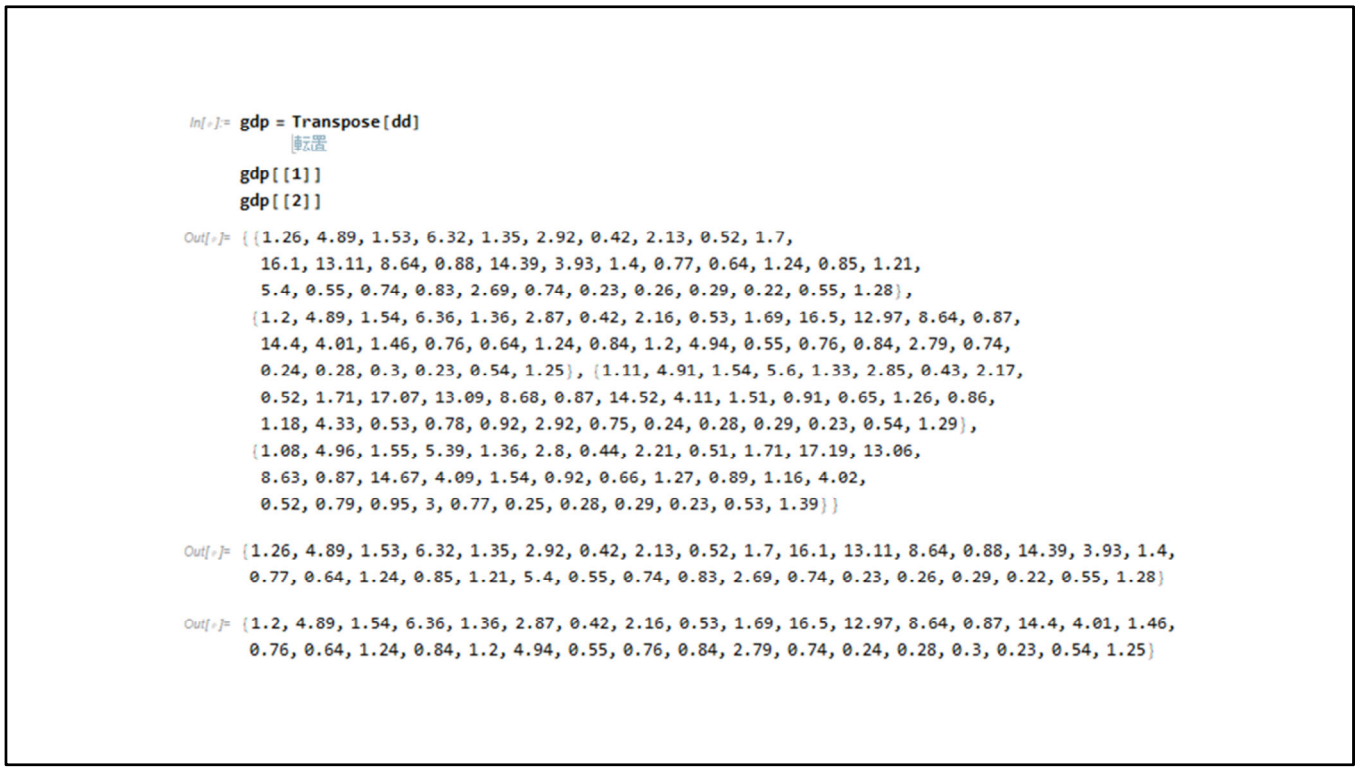
8	[2010 Version] Distribution of GRDP to Total GRDP of 34 Provinces at Current Market Prices by Province, 2010-2017 (Percent)	03 Aug 2018

Updated on 03 August, 2018

Let’s study Mathematica programming using the Indonesia economics data above mentioned.

01 ListPlot

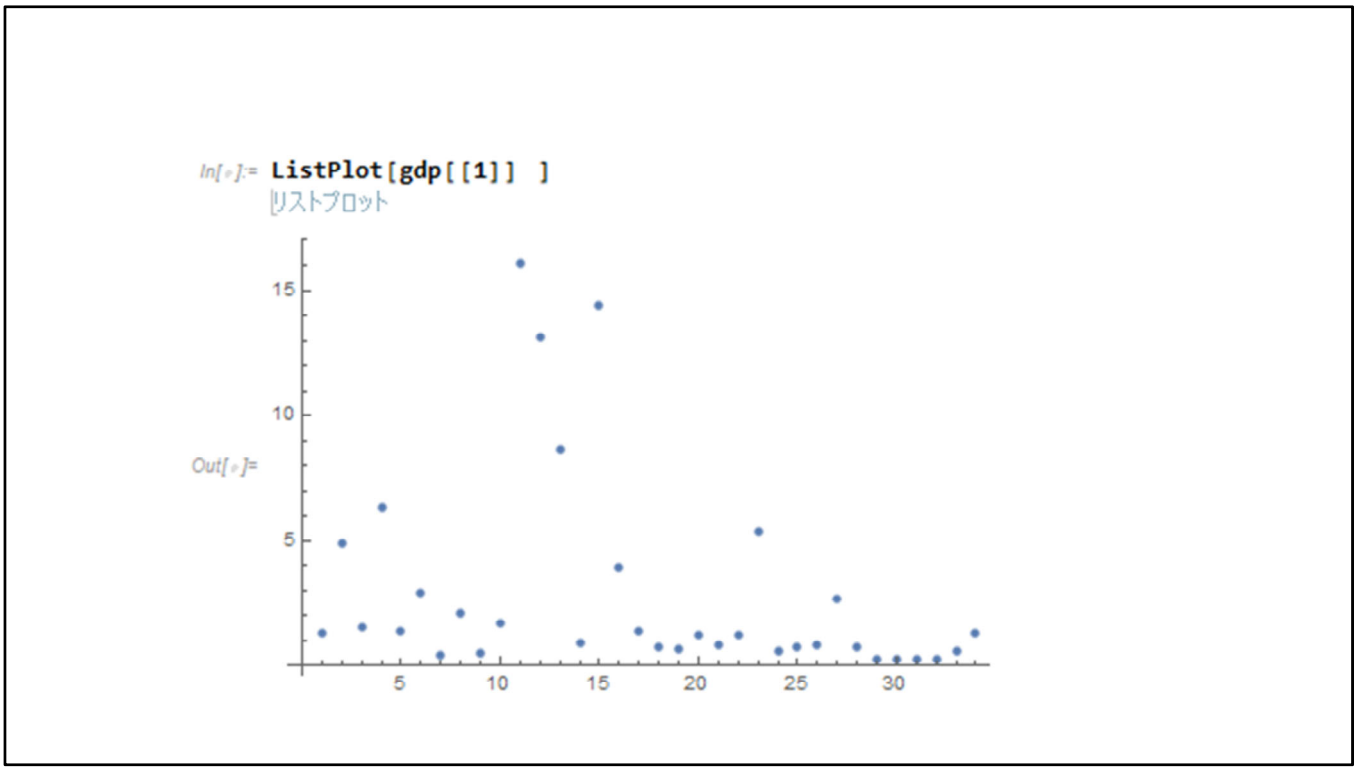
First let's try the command ListPlot.
In Mathematica, do not forget to execute the command line(s) by SHIFT and ENTER at the same time.



See the cdf program. “dd” is defined as 2 dimensional table as follows;

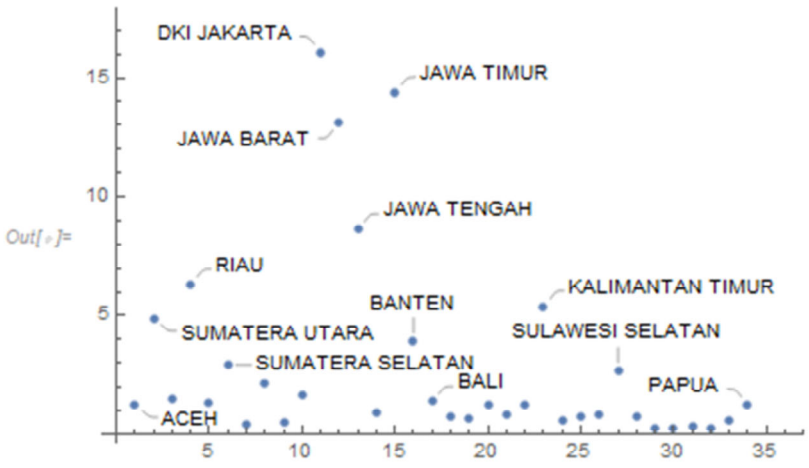
dd={{1.26,1.2,1.11,1.08},{4.89,4.89,4.91,4.96},{1.53,1.54,1.54,1.55},{6.32,6.36,5.6,5.39},{1.35,1.36,1.33,1.36},{2.92,2.87,2.85,2.8},{0.42,0.42,0.43,0.44},{2.13,2.16,2.17,2.21},{0.52,0.53,0.52,0.51},{1.7,1.69,1.71,1.71},{16.1,16.5,17.07,17.19},{13.11,12.97,13.09,13.06},{8.64,8.64,8.68,8.63},{0.88,0.87,0.87,0.87},{14.39,14.4,14.52,14.67},{3.93,4.01,4.11,4.09},{1.4,1.46,1.51,1.54},{0.77,0.76,0.91,0.92},{0.64,0.64,0.65,0.66},{1.24,1.24,1.26,1.27},{0.85,0.84,0.86,0.89},{1.21,1.2,1.18,1.16},{5.4,4.94,4.33,4.02},{0.55,0.55,0.53,0.52},{0.74,0.76,0.78,0.79},{0.83,0.84,0.92,0.95},{2.69,2.79,2.92,3},{0.74,0.74,0.75,0.77},{0.23,0.24,0.24,0.25},{0.26,0.28,0.28,0.28},{0.29,0.3,0.29,0.29},{0.22,0.23,0.23,0.23},{0.55,0.54,0.54,0.53},{1.28,1.25,1.29,1.39}};

In Mathematica, {…} represents a list or a table. A table of a table is a 2 dimensional table.
To indicate the first element of the table, please use the index element like gdp[[1]] or gdp[[2]]. You can get the index element. The index mark is like[[jj]].



To list the 2013 GDP data, please input ListPlot[gdp[[1]]]. The x axis shows the provinceID from 1 to 34.

```
In[ ]:= ListPlot[gdp[[1]] → province ]
        リストプロット
```



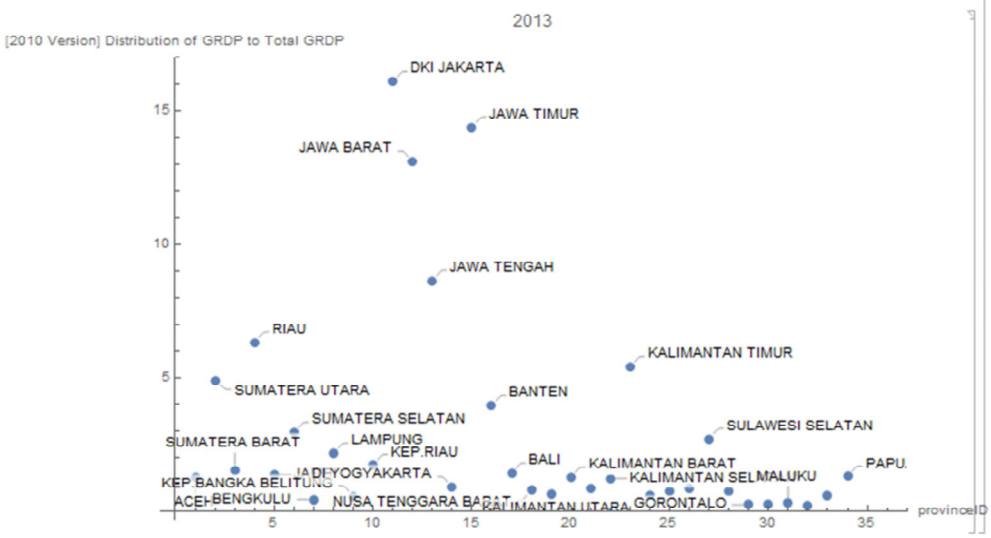
Let's add the province name by gdp[[1]]->province.
In advance I set the province as follows:

province={"ACEH","SUMATERA UTARA","SUMATERA BARAT","RIAU","JAMBI","SUMATERA SELATAN","BENGKULU","LAMPUNG","KEP.BANGKA BELITUNG","KEP.RIAU","DKI JAKARTA","JAWA BARAT","JAWA TENGAH","DI YOGYAKARTA","JAWA TIMUR","BANTEN","BALI","NUSA TENGGARA BARAT","NUSA TENGGARA TIMUR","KALIMANTAN BARAT","KALIMANTAN TENGAH","KALIMANTAN SELATAN","KALIMANTAN TIMUR","KALIMANTAN UTARA","SULAWESI UTARA","SULAWESI TENGAH","SULAWESI SELATAN","SULAWESI TENGGARA","GORONTALO","SULAWESI BARAT","MALUKU","MALUKU UTARA","PAPUA BARAT","PAPUA"};

Please watch out the last mark ";" of the above command. The mark ";" means not display the results.

```
In[ ]:= ListPlot[gdp[[1]] → province ,
        リストプロット
        AxesLabel → {"provinceID", "[2010 Version] Distribution of GRDP to Total GRDP"} ,
        軸のラベル
        PlotLabel → "2013", ImageSize → 700 ]
        プロットラベル          画像サイズ
```

Out[ ]=



To get better visualization, some options are offered in ListPlot.
Axeslabel, PlotLabel, and ImageSize.
The y-axis shows the percentage of each province GDP. DKI Jakarta has the biggest figure which is over 15 %.

Table for making the list

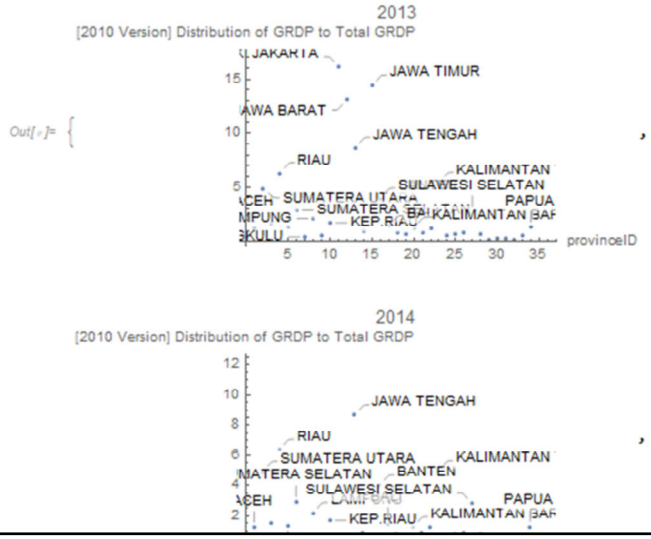
```
In[ ]:= years = Table[kk, {kk, 2013, 2016}]
                リストを作成

Out[ ]= {2013, 2014, 2015, 2016}
```

Then let’s make the year number list by Table.
Table[kk, {kk, 2013, 2016}] means to make the table of variable kk in which kk starts from 2013 to 2016 by step 1.
The result table is input to the variable “years” there.

Table of ListPlot

Table[
ListPlot[…],
{jj, 1, 3}]

```
In[ ]:= Table[ ListPlot[gdp[[jj]] → province ,
          リス…  リストプロット
        AxesLabel → {"provinceID", "[2010 Version] Distribution of GRDP to Total GRDP"} ,
                    軸のラベル
        PlotLabel → years[[jj]], ImageSize → 400 ], {jj, 1, 3}]
                    プロットラベル                画像サイズ
```



In Mathematica, you can make a table of PlotLists.
Wrap up the PlotList command by Table[…]. The index of Table is 1, 2, 3.
{jj, 1, 3} moves the variable jj from 1 to 3 by step 1.
In the above program, years[[jj]] is changed.

Definition of your function

```
pair[data1_, data2_] := Table[{data1[[k]], data2[[k]]}, {k, 1, Length[data1]}]
                        リストを作成                                  長さ
```

Let’s study the definition of your function. In this case, I defined the function “pair”. The “pair” function has two parameters which are data1 and data2. To that is a parameter, “_” is added.
The mark “:=“ means the definition of a function.

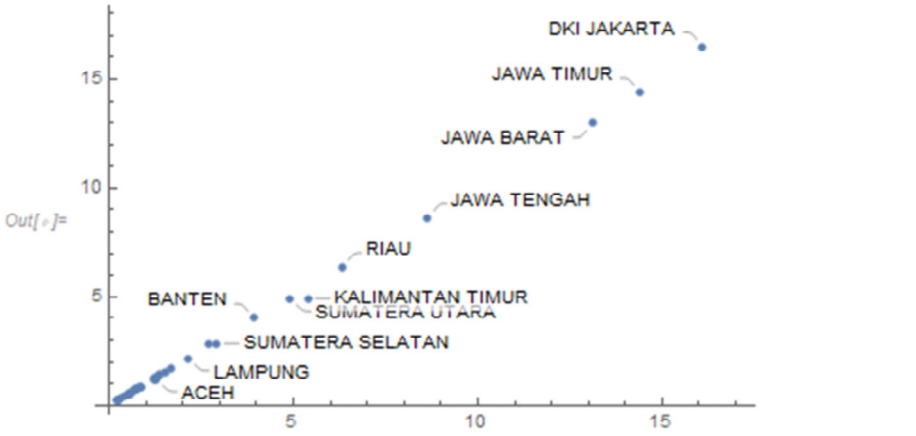
The function “Length” returns the length of the table. So variable k moves from 1 to the length of the given data1.

```
In[ ]:= gdp[[1]]
Out[ ]= {1.26, 4.89, 1.53, 6.32, 1.35, 2.92, 0.42, 2.13, 0.52, 1.7, 16.1, 13.11, 8.64, 0.88, 14.39, 3.93, 1.4,
  0.77, 0.64, 1.24, 0.85, 1.21, 5.4, 0.55, 0.74, 0.83, 2.69, 0.74, 0.23, 0.26, 0.29, 0.22, 0.55, 1.28}

In[ ]:= gdp[[2]]
Out[ ]= {1.2, 4.89, 1.54, 6.36, 1.36, 2.87, 0.42, 2.16, 0.53, 1.69, 16.5, 12.97, 8.64, 0.87, 14.4, 4.01, 1.46,
  0.76, 0.64, 1.24, 0.84, 1.2, 4.94, 0.55, 0.76, 0.84, 2.79, 0.74, 0.24, 0.28, 0.3, 0.23, 0.54, 1.25}

In[ ]:= pair[gdp[[1]], gdp[[2]]]
Out[ ]= {{1.26, 1.2}, {4.89, 4.89}, {1.53, 1.54}, {6.32, 6.36}, {1.35, 1.36}, {2.92, 2.87}, {0.42, 0.42},
  {2.13, 2.16}, {0.52, 0.53}, {1.7, 1.69}, {16.1, 16.5}, {13.11, 12.97}, {8.64, 8.64},
  {0.88, 0.87}, {14.39, 14.4}, {3.93, 4.01}, {1.4, 1.46}, {0.77, 0.76}, {0.64, 0.64}, {1.24, 1.24},
  {0.85, 0.84}, {1.21, 1.2}, {5.4, 4.94}, {0.55, 0.55}, {0.74, 0.76}, {0.83, 0.84}, {2.69, 2.79},
  {0.74, 0.74}, {0.23, 0.24}, {0.26, 0.28}, {0.29, 0.3}, {0.22, 0.23}, {0.55, 0.54}, {1.28, 1.25}}

In[ ]:= dataP = pair[gdp[[1]], gdp[[2]]]
Out[ ]= {{1.26, 1.2}, {4.89, 4.89}, {1.53, 1.54}, {6.32, 6.36}, {1.35, 1.36}, {2.92, 2.87}, {0.42, 0.42},
  {2.13, 2.16}, {0.52, 0.53}, {1.7, 1.69}, {16.1, 16.5}, {13.11, 12.97}, {8.64, 8.64},
  {0.88, 0.87}, {14.39, 14.4}, {3.93, 4.01}, {1.4, 1.46}, {0.77, 0.76}, {0.64, 0.64}, {1.24, 1.24},
  {0.85, 0.84}, {1.21, 1.2}, {5.4, 4.94}, {0.55, 0.55}, {0.74, 0.76}, {0.83, 0.84}, {2.69, 2.79},
  {0.74, 0.74}, {0.23, 0.24}, {0.26, 0.28}, {0.29, 0.3}, {0.22, 0.23}, {0.55, 0.54}, {1.28, 1.25}}
```

So, please try the pair function.
The input data are gdp[[1]] and gdp[[2]].
The pair function returns a set of pair data of gdp[[1]] and gdp[[2]]. Namely, each province's 2013 GDP% and 2014 GDP%.

ListPlot of pair data

```
In[ ]:= ListPlot[pair[gdp[[1]], gdp[[2]]] → province, PlotRange → All]
        リストプロット                                プロット範囲   すべて
```



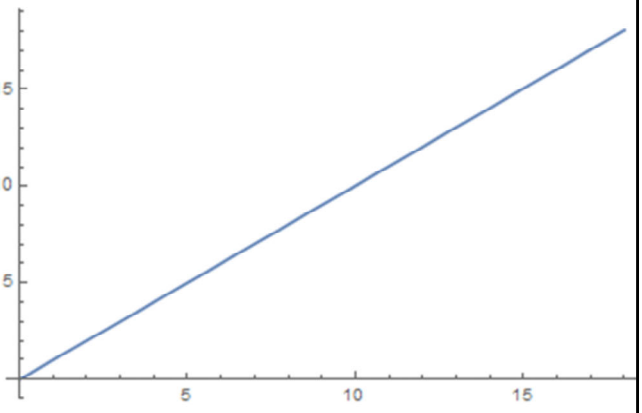
Let's see the pair data visually.
ListPlot can plot the 2 dimensional data, too.
The option of ListPlot PlotRange->All makes it so that all plot points are visible.

Linear Regression

```
In[ ]:= regm = LinearModelFit[dataP, x, x]
        線形モデルフィット

        Normal[regm]
        通常の式に変換

Out[ ]= FittedModel[ -0.0141751 + 1.00512 x ]

Out[ ]= -0.0141751 + 1.00512 x
```

```
In[ ]:= Plot[regm[x], {x, 0, 18}]
        プロット
```

Out[ ]=



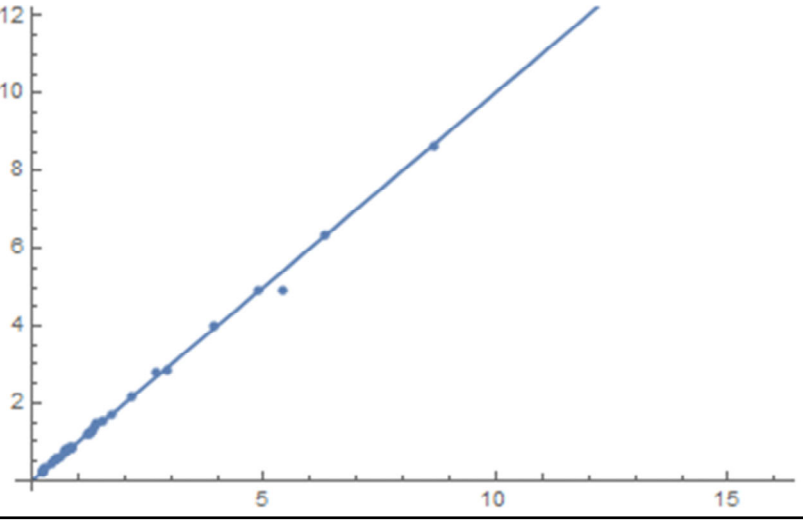
Let’s conduct a linear regression on the data. The command is LinearModelFit. The resultant expression can be displayed by Normal[…]. The decline is 1.00512.

Let’s plot the line using the command Plot. The regm[x] is the line expression and the range is indicated like {x, 0, 18}.

Showing many plots such as ListPlot and Plot

```
In[ ]:= Show[ListPlot[dataP], Plot[regm[x], {x, 0, 18}]]
        示す  リストプロット        プロット
```

Out[ ]=

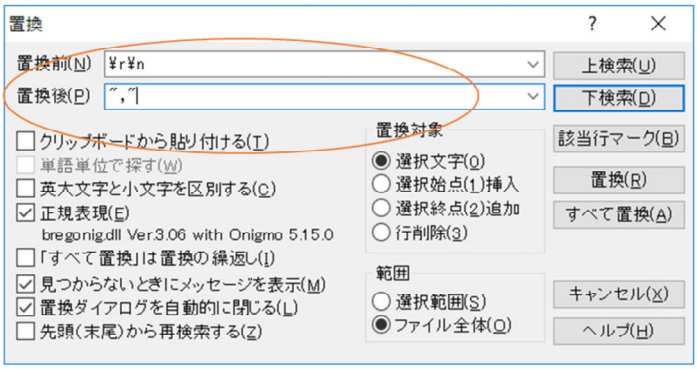


I would like to see the regression line and the points at the same time. Then please use Show. The command Show make many plot commans results at the same graph. In the above case, (1) is ListPlot and (2) Plot.

Preprocessing of the input data using an editor

When you use real economics data, it takes some preprocessing of the data so that you can input the data to Mathematica program.
Suppose that the input data is given in Excel.

Preprocessing of the input data using Sakura editor (Any editor would be OK)

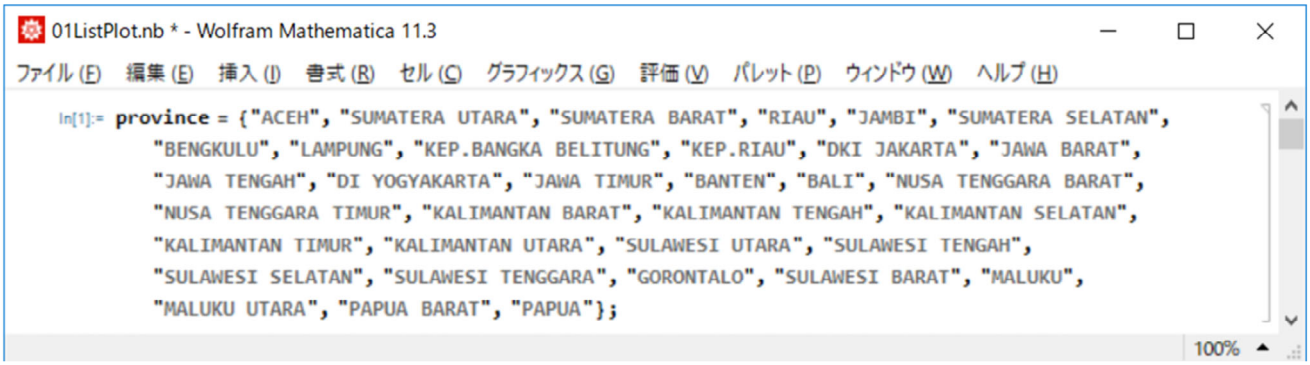


- ACEH","SUMATERA UTARA","SUMATERA BARAT","RIAU","JAMBI","SUMATERA SELATAN","BENGKULU","LAMPUNG","KEP. BANGKA BELITUNG","KEP. RIAU","DKI JAKARTA","JAWA BARAT","JAWA TENGAH","DI YOGYAKARTA","JAWA TIMUR","BANTEN","BALI","NUSA TENGGARA BARAT","NUSA TENGGARA TIMUR","KALIMANTAN BARAT","KALIMANTAN TENGAH","KALIMANTAN SELATAN","KALIMANTAN TIMUR","KALIMANTAN UTARA","SULAWESI UTARA","SULAWESI TENGAH","SULAWESI SELATAN","SULAWESI TENGGARA","GORONTALO","SULAWESI BARAT","MALUKU","MALUKU UTARA","PAPUA BARAT","PAPUA","","

Some editor is needed for the preprocessing. The editor has to have regular expression functions. Any editor is OK so far as it has regular expression functions.
In the above case, I changed the province name list.
ENTER is removed
Terminate each province name by “” and “,”.

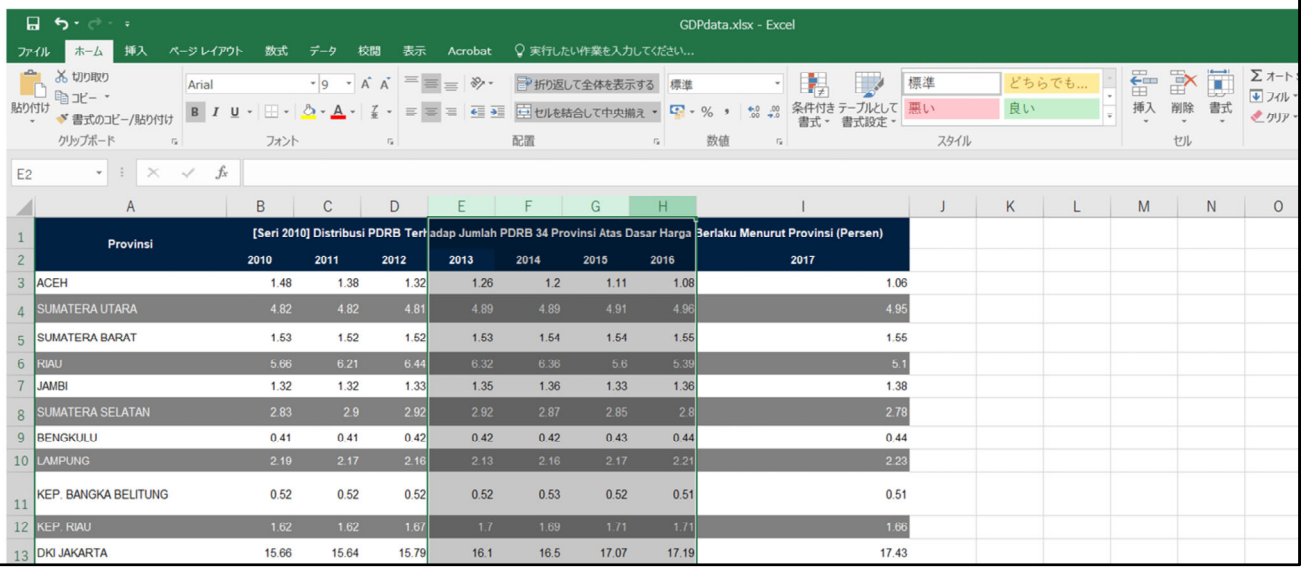
The regular expression ENTER ¥r¥n to “,” makes the column list to one line.
You will copy and paste the line to Mathematica.

Hand modification



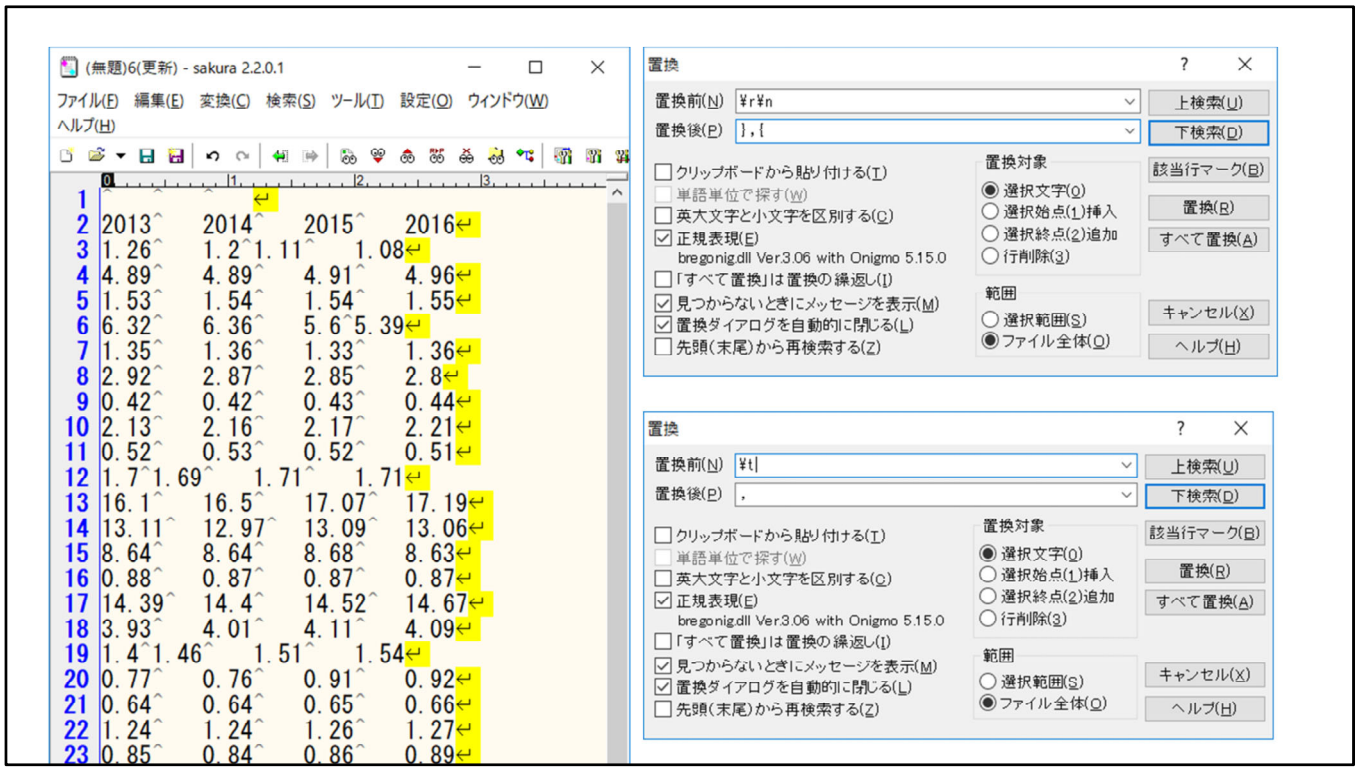
Some modification of the input line is needed so that finally you can get the above expression.

Figures from 2013 to 2016 to Mathematica

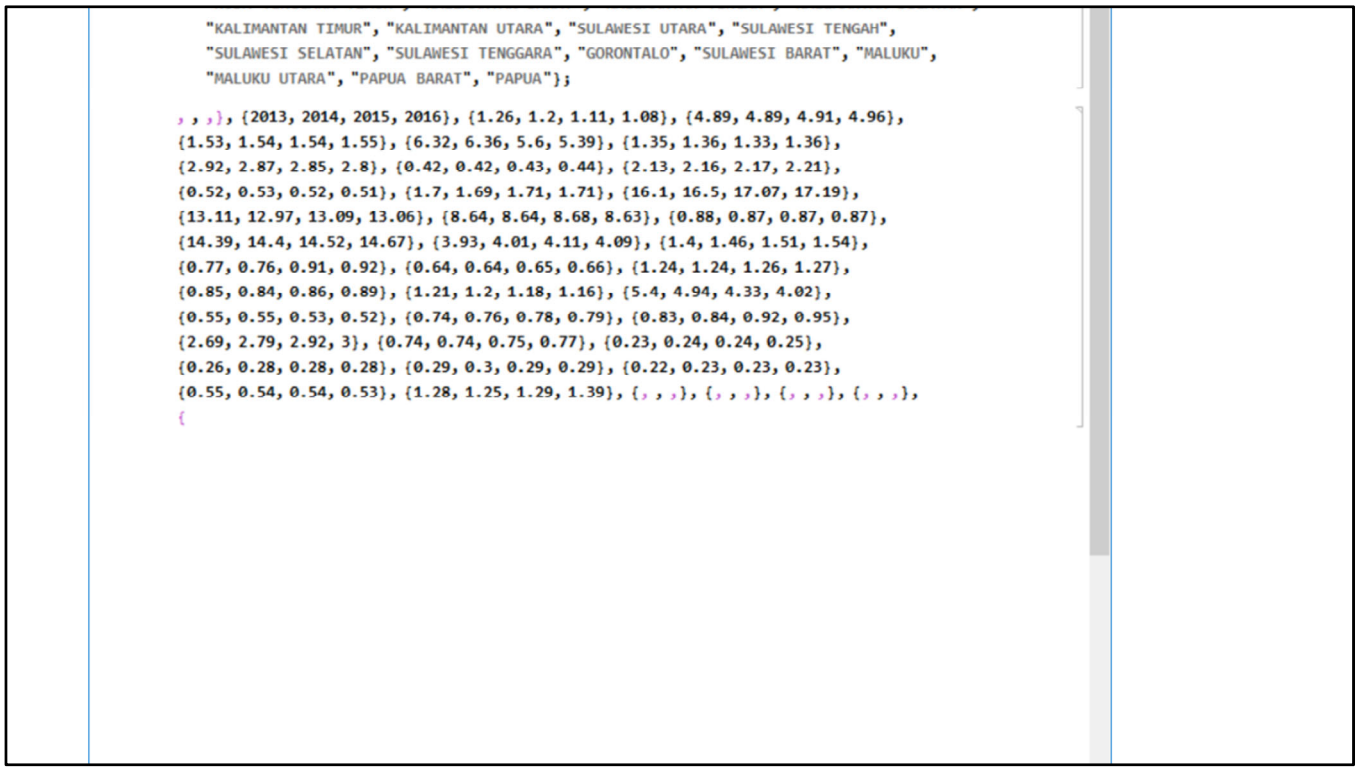


	Provinsi	[Seri 2010] Distribusi PDRB Terhadap Jumlah PDRB 34 Provinsi Atas Dasar Harga Berlaku Menurut Provinsi (Persen)							
		2010	2011	2012	2013	2014	2015	2016	2017
3	ACEH	1.48	1.38	1.32	1.26	1.2	1.11	1.08	1.06
4	SUMATERA UTARA	4.82	4.82	4.81	4.89	4.89	4.91	4.96	4.95
5	SUMATERA BARAT	1.53	1.52	1.52	1.53	1.54	1.54	1.55	1.55
6	RIAU	5.66	6.21	6.44	6.32	6.36	5.6	5.39	5.1
7	JAMBI	1.32	1.32	1.33	1.35	1.36	1.33	1.36	1.38
8	SUMATERA SELATAN	2.83	2.9	2.92	2.92	2.87	2.85	2.8	2.78
9	BENGKULU	0.41	0.41	0.42	0.42	0.42	0.43	0.44	0.44
10	LAMPUNG	2.19	2.17	2.16	2.13	2.16	2.17	2.21	2.23
11	KEP. BANGKA BELITUNG	0.52	0.52	0.52	0.52	0.53	0.52	0.51	0.51
12	KEP. RIAU	1.62	1.62	1.67	1.7	1.69	1.71	1.71	1.66
13	DKI JAKARTA	15.66	15.64	15.79	16.1	16.5	17.07	17.19	17.43

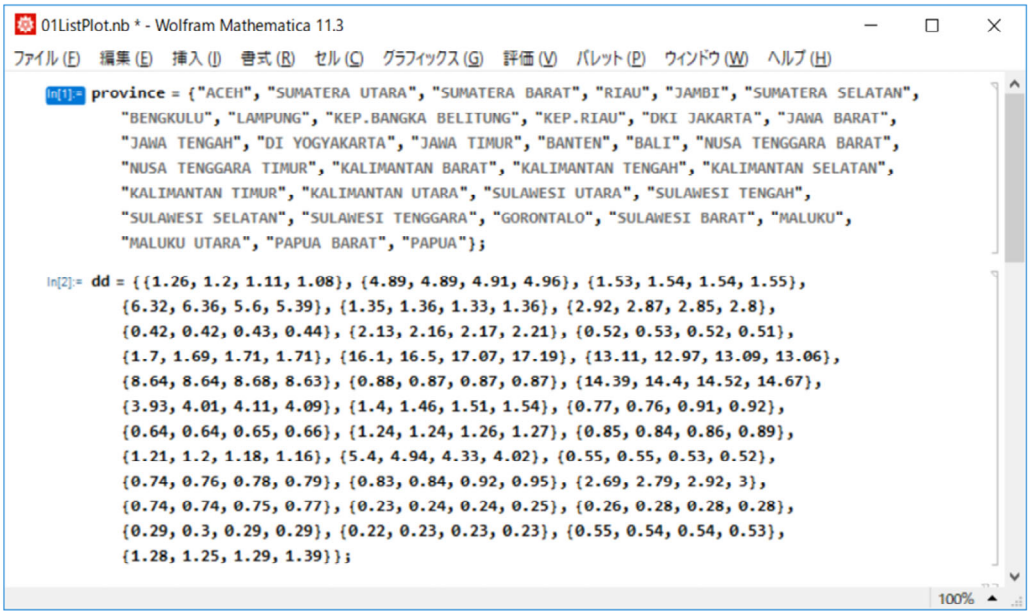
Then the GDP data of 2013 to 2016 will be moved to Mathematica. The selected area in EXCEL is 2013 to 2016 columns.



You will copy the data to the editor.
On the editor,
ENTER to “},{“ and
TAB to “,”
Then, you can get the 2 dimensional table data as a line.



This is the moved line data. Then you will make some modification.



The “dd” is the final data line.