e-ISSN: 2809-7610

Village Potential Statistics (PODES): Visualization of Schools in Jambi Province with Statistical Programming (R)

Fadhlul Mubarak[1], Atilla Aslanargun[2], Vinny Yuliani Sundara[3],
[1]Department of Sharia Economic, UIN Sulthan Thaha Saifuddin Jambi, Indonesia
[2]Department of Statistics, Eskisehir Technical University, Turkey
[3]Department of Mathematics Education, UIN Sulthan Thaha Saifuddin Jambi, Indonesia
e-mail: mubarakfadhlul@gmail.com, aaslanar@eskisehir.edu.tr,
vinnyyulianisundara@uinjambi.ac.id

Abstract

One of the primary data that can be used in research is village potential statistics (PODES). The data was obtained based on research in a certain period by the Statistics Indonesia (BPS). This study aims to visualize the percentage of schools in each city/district in Jambi Province using R programming based on PODES data in 2014 and 2019. In this study, we not only visualize but also how to build attractive graphics and arrange them starting from windows, graphic size, dimensions, color, horizontal axis, vertical axis, and others. Of course, the graph produced in this study is different from the basic plot found in the R program, although the process carried out is also more complicated. From 2014 to 2019, in general, within a period of 5 years there has been an increase in the number of schools in each city/district in Jambi Province. However, from the university level, the number decreased. In 2014 the number of universities in Jambi City was 32 but in 2019 the number decreased to 24. There are even interesting things in Kerinci and Tebo district. In 2014 there were no universities listed, while in 2019 there were 3. This also affects the percentage of education level in each city or district.

Key words: attractive graphic, percentage, R Programming, Statistics Indonesia (BPS)

PRELIMINARY

The village potential statistics (PODES) is one of the primary data that can be used for research. The data was obtained based on research in a certain period by the Central Statistics Agency, Indonesia (BPS). Research related to PODES data includes Fitriansyah & Nuryakin (2021), Hardiyanto (2020), Paais (2021), Sari & Handayani (2020), Kaban et al. (2022), etc. This research is limited to PODES data in Jambi Province. Jambi Province consists of 2 cities including Jambi City and Sungai Penuh. In addition, there are 9 districts including Batang Hari, Bungo, Kerinci, Merangin, Muaro Jambi, Sarolangun, Tanjung Jabung Barat, Tanjung Jabung Timur, and Tebo district. Even the PODES data in Jambi Province has been used by Junaidi et al. (2020), Hasan et al. (2022), Permatasari & Larasati (2022), Trilaksono & Sukartini (2020), Hakim et al. (2021), etc. However, these studies do not discuss related to visualization or graphics.

The visualization process in this research we use statistical software such as R. This year, research related to graph analysis has been published by Wang & Wang (2022), Borisov (2022), Diaz-Diaz & Estrada (2022), Tiddi & Schlobach (2022), Shin et al. (2022), etc. Apart from being free, the software is always developing to date, especially in terms of packages. Research using R programming includes Saavedra-Nieves (2021), Abdollahi et al. (2022), Snitker et al. (2022), Atkins et al. (2022), Li et al. (2022), etc. However, these studies did not use PODES data.

METHODOLOGY

In this research, we use R programming to visualize the data. We not only visualize graphs but also how to build attractive graphs and arrange them starting from windows, graph size, dimensions, colors, horizontal axis, vertical axis, and others based on the PODES data. The PODES data that we used in this study were 2014 and 2019. So, the difference in the percentage of schools in each City or District for 2014 and 2019 can be seen. The graph is clearly different if we use the basic "plot" function available in the R program.

RESULTS AND DISCUSSION

The form of the questionnaire in the 2014 and 2019 PODES research is quite different. However, the number of cities or districts recorded is still the same. Some of the questions that appeared in the 2014 survey did not appear in the 2019, and vice versa. Even so, data on the number of educational facilities that are the focus of this research are still well available. The number of schools in each city or district in Jambi Province in 2014 and 2019 is provided in Tables 1 and 2. These tables are the result of the pivot from the original data. Several types of levels of public or private educational facilities are combined. Also, the same school level as general, vocational, and religious are combined.

Table 1. Number of schools in Jambi Province (2014)

City/Districts	Kindergarten	Elementary S.	Junior High S.	Senior High S.	University
Batang Hari	149	259	77	36	3
Bungo	144	316	82	52	6
Jambi	159	265	90	81	32
Kerinci	106	265	73	28	0
Merangin	216	386	111	53	4
Muaro Jambi	182	298	106	48	2
Sarolangun	198	255	102	61	3
Sungai Penuh	48	78	14	13	8
Tanjung Jabung Barat	108	261	106	53	4
Tanjung Jabung Timur	75	253	74	44	0
Tebo	181	289	94	49	4

In general, within a period of 5 years there has been an increase in the number of schools in each city/district in Jambi Province. If viewed in more detail, the level of education has also increased, for example in Batang Hari District. In 2014 the number of kindergartens was 149 and the number increased to 181 in 2019. Similar to the number of universities in the district, in 2014 there were 3 and in 2019 there were 2 additional universities.

Next, we can see the development of the number of schools in Jambi City. The number of kindergartens in Jambi City was 159 in 2014 and became 233 in 2019. However, from the university level there was a decrease in the number. In 2014 the number of universities in Jambi City was 32 but in 2019 the number decreased to 24.

There are even interesting things in Kerinci and Tebo district. In 2014 there were no universities listed while in 2019 there were 3.

Table 2. Number of schools in Jambi Province (2019)

City/Districts	Kindergarten	Elementary S.	Junior High S.	Senior High S.	University
Batang Hari	181	347	90	47	5
Bungo	197	262	101	64	8
Jambi	233	284	103	101	24
Kerinci	126	280	74	33	3
Merangin	276	368	134	72	5
Muaro Jambi	250	298	117	60	4
Sarolangun	244	260	121	71	3
Sungai Penuh	47	83	20	15	8
Tanjung Jabung Barat	92	269	125	62	4
Tanjung Jabung Timur	73	277	78	45	2
Tebo	174	299	117	71	3

We built a visualization of the percentage of the number of schools in each of these cities/districts by using R programming. We did the data input and the percentage calculation process manually so that it was easier for readers to understand and execute the R programming syntax in this study. The education levels in Indonesia are Kindergarten, Elementary School, Junior High School, Senior High School, and University. We store these educational levels in the R program with the name "levels". Furthermore, the percentage of education level in each city/district has been saved with the name "percent". If we use a simple dot graph with the R syntax below, the results will look like Figure 1. The resulting image appears to have several drawbacks including being less attractive because there is no label on the horizontal axis, providing less information because the education level is not listed, even the size of the graph is too wide so that if it is reduced then the description of the graph becomes unclear, etc.

```
R>levels<-c("Kindergarten","Elementary School", "Junior High School", "Senior High School",
          "University")
R>percent<-c(28, 49, 15, 7, 1)
R>plot(percent)
```

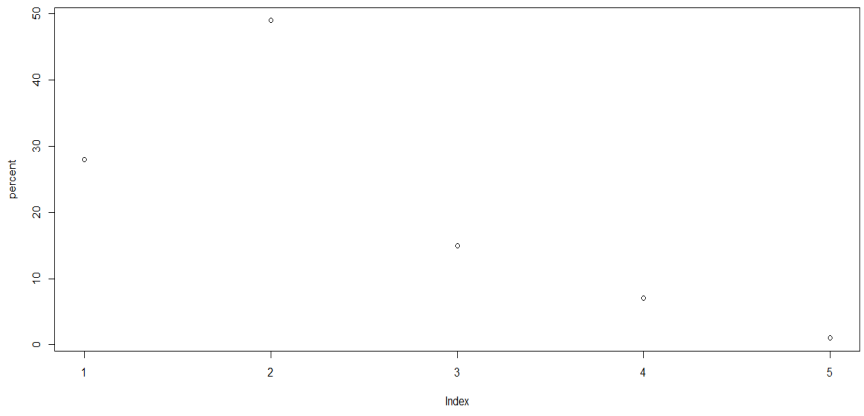
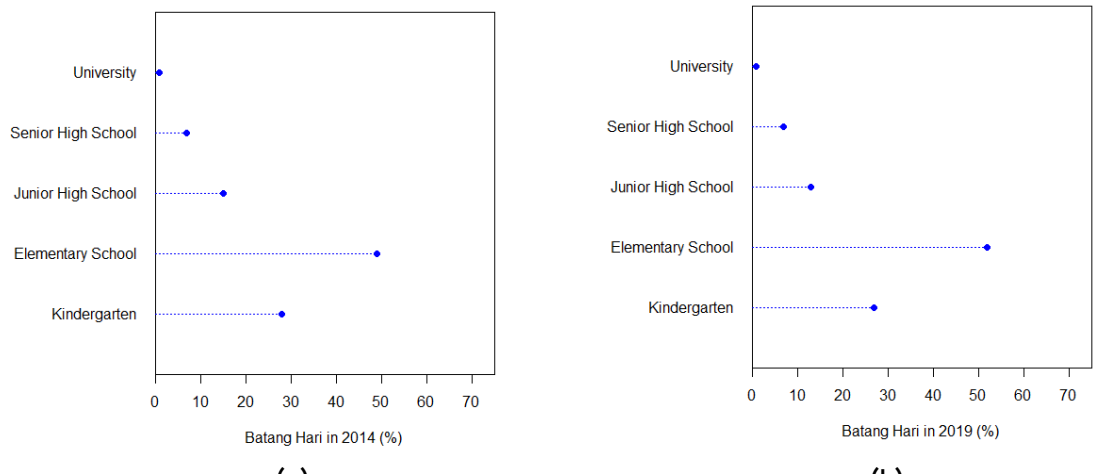


Figure 1. Percentage of schools in Batang Hari District using the basic plot function in 2014

Based on these shortcomings, we modify the plot function in R programming. We use graphics.off, windows, and par functions to set the background display of the graph. The window size for the chart has width of 15 and height of 15 where the point size is 12 because we are using the dot chart. We set the parameters where the "fin" function is for the figure region dimensions where the width is 5 inches and the height is 5 inches. The “pin” function is used for the current plot dimensions where it is 5 inches and the height is 5 inches. Next is the “mai” function for the numerical vector of the form where the bottom is 0.875 inches, the left is 1 inch, the top is 0.125 inches, and the right is 0.25 inches. Next, we create a dot graph for the percentage of education level in each city/district with the “plot” function. The dot graph is not equipped with a blue dotted line. The x-axis in the graph is the percentage of schools from each city/district while the y-axis is the level of education.

```
R>graphics.off()
R>windows(width=15, height=15, pointsize=12)
R>par(fin=c(5,5), pin=c(5,5), mai=c(0.875, 1, 0.125, 0.25),xaxs="i")
R>plot(percent, 1:5, type="n", xaxt="n", yaxt="n", xlim=c(0,75), ylim=c(0,6), xlab='Batang Hari in
      2014 (%)', ylab='', xaxs="i", yaxs="i")
R>for (i in 1:5) lines(x=c(0,percent[i]), y=c(i,i), lty=3, col="blue")
R>points(x=percent, y=1:5, pch=19, cex=1.0, col="blue")
R>axis(1, at=10*(0:7.5), labels=TRUE, tick=TRUE, outer=FALSE)
R>axis(2, at=1:5+0.12, labels=levels, tick=FALSE, outer=FALSE, las=2, hadj=1, padj=1)
```



(a) (b)

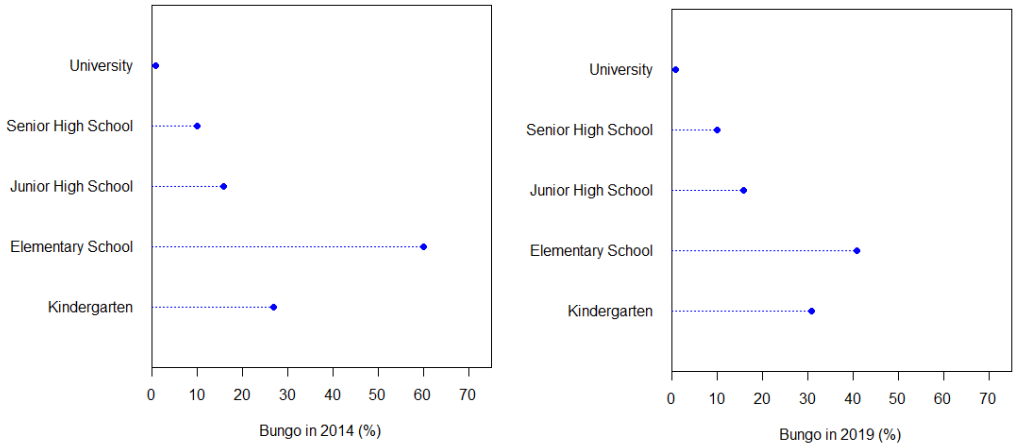
Figure 2. Percentage of schools in Batang Hari District using the modified plot. (a) in 2014 and (b) in 2019

Based on the syntax of the R program for visualizing schools in Batang Hari District appears as above, so, the graph like Figure 2. Then the visualization of the percentage of schools for other cities/districts only replaces the value of "percent" and the label on the xlab description as in Table 3. So that the results of each visualization of the percentage of school level in each city/district is shown in Figure 3 – Figure 12. Based on these graphs the dominating school level in each city/District is elementary school either in 2014 or 2019. Meanwhile, the school level with the lowest percentage is university. The availability of junior high school schools also needs to be increased because the percentage is very far from

elementary school. The consequence if not added is the large number of dropouts from elementary school graduates. If readers want to try this program, please visit the link in Figure 13.

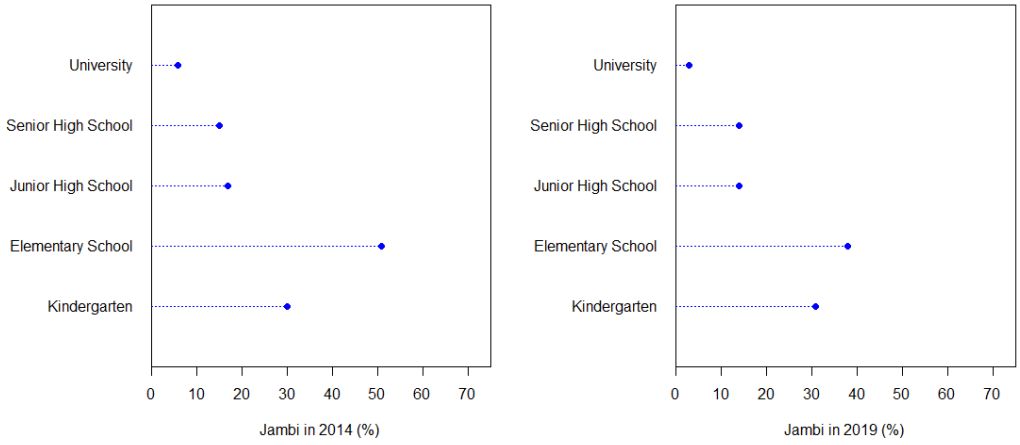
Table 3. Change the syntax of the R program for each city/district

City/Districts	Percent	Xlab
Bungo	c(27, 60, 16, 10, 1)	Bungo in 2014 (%)
Jambi	c(30, 51, 17, 15, 6)	Jambi in 2014 (%)
Kerinci	c(20, 51, 14, 5, 0)	Kerinci in 2014 (%)
Merangin	c(41, 74, 21, 10, 1)	Merangin in 2014 (%)
Muaro Jambi	c(35, 57, 20, 9, 0)	Muaro Jambi in 2014 (%)
Sarolangun	c(38, 49, 19, 12, 1)	Sarolangun in 2014 (%)
Sungai Penuh	c(9, 15, 3, 2, 2)	Sungai Penuh in 2014 (%)
Tanjung Jabung Barat	c(21, 50, 10, 10, 1)	Tanjung Jabung Barat in 2014 (%)
Tanjung Jabung Timur	c(14, 48, 14, 8, 0)	Tanjung Jabung Timur in 2014 (%)
Tebo	c(35, 55, 18, 9, 1)	Tebo in 2014 (%)
Batang Hari	c(27, 52, 13, 7, 1)	Batang Hari in 2019 (%)
Bungo	c(31, 41, 16, 10, 1)	Bungo in 2019 (%)
Jambi	c(31, 38, 14, 14, 3)	Jambi in 2019 (%)
Kerinci	c(24, 54, 14, 6, 1)	Kerinci in 2019 (%)
Merangin	c(32, 43, 16, 8, 1)	Merangin in 2019 (%)
Muaro Jambi	c(34, 41, 16, 8, 1)	Muaro Jambi in 2019 (%)
Sarolangun	c(35, 37, 17, 10, 0)	Sarolangun in 2019 (%)
Sungai Penuh	c(27, 48, 12, 9, 5)	Sungai Penuh in 2019 (%)
Tanjung Jabung Barat	c(17, 49, 23, 11, 1)	Tanjung Jabung Barat in 2019 (%)
Tanjung Jabung Timur	c(15, 58, 16, 9, 0)	Tanjung Jabung Timur in 2019 (%)
Tebo	c(26, 45, 18, 11, 0)	Tebo in 2019 (%)



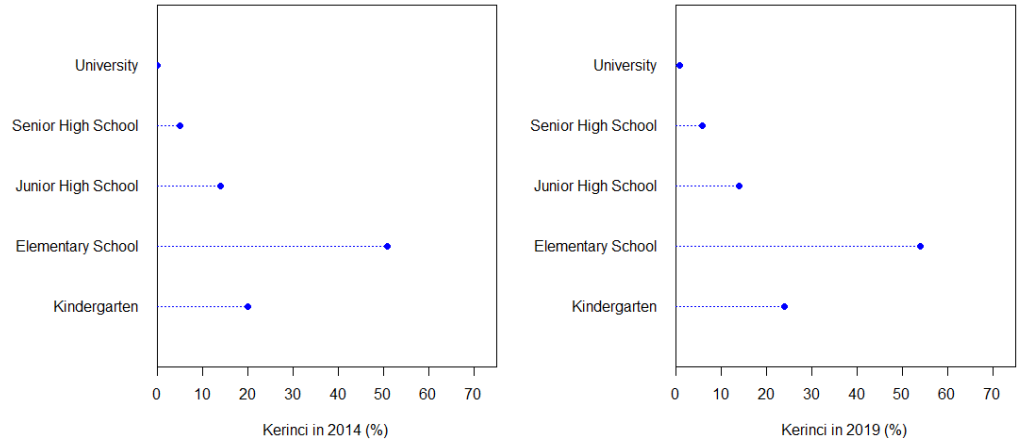
(a) (b)

Figure 3. Percentage of schools in Bungo District using the modified plot. (a) in 2014 and (b) in 2019



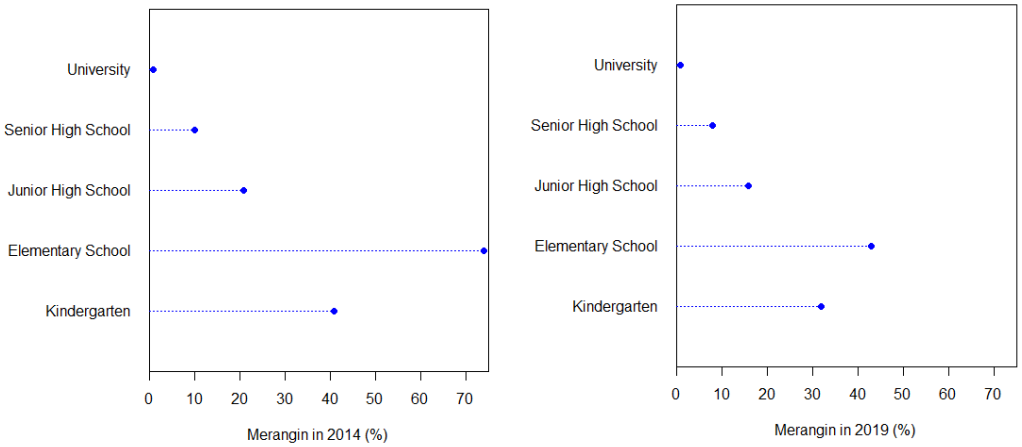
(a) (b)

Figure 4. Percentage of schools in Jambi City using the modified plot. (a) in 2014 and (b) in 2019



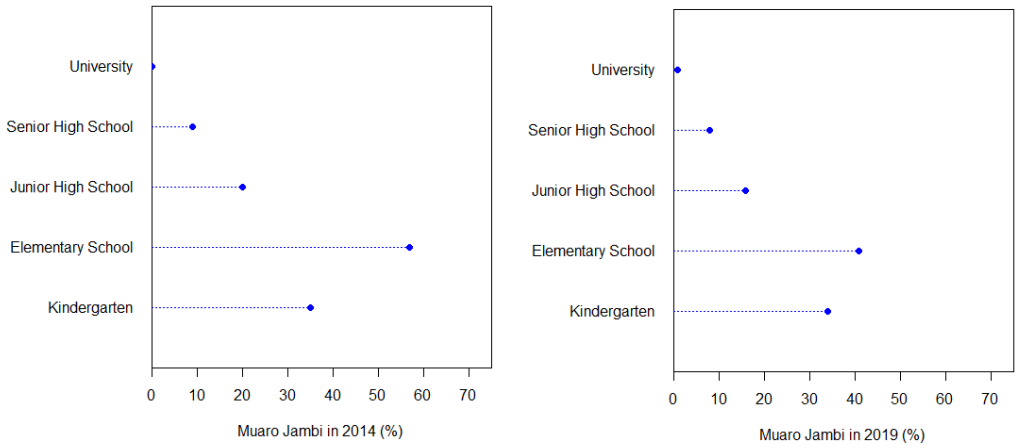
(a) (b)

Figure 5. Percentage of schools in Kerinci District using the modified plot. (a) in 2014 and (b) in 2019



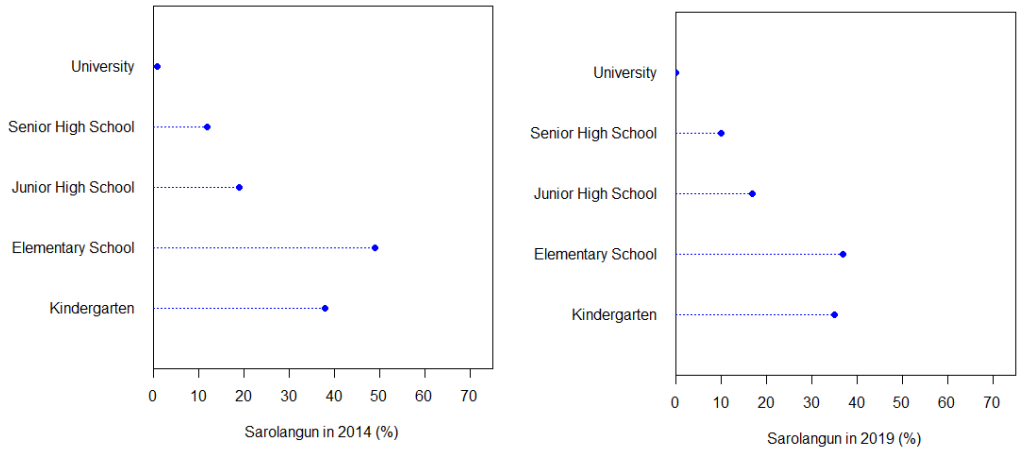
(a) (b)

Figure 6. Percentage of schools in Merangin District using the modified plot. (a) in 2014 and (b) in 2019



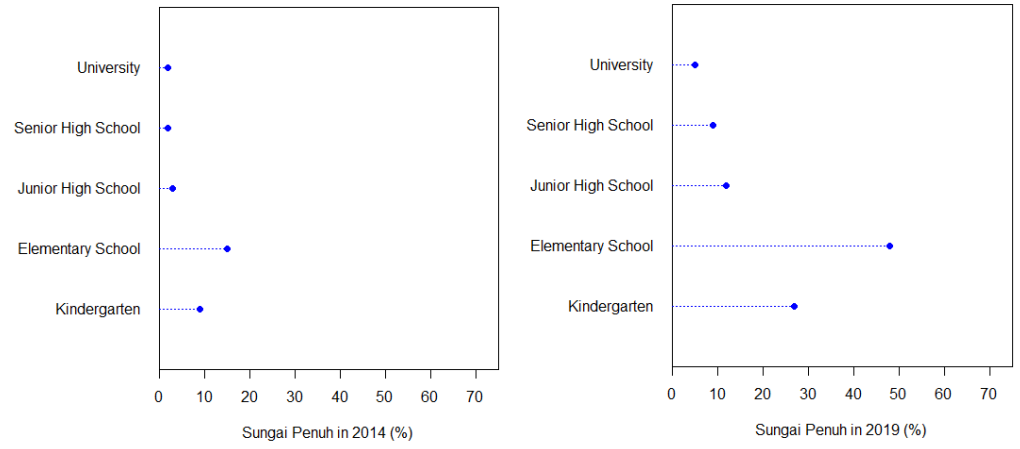
(a) (b)

Figure 7. Percentage of schools in Muaro Jambi District using the modified plot. (a) in 2014 and (b) in 2019



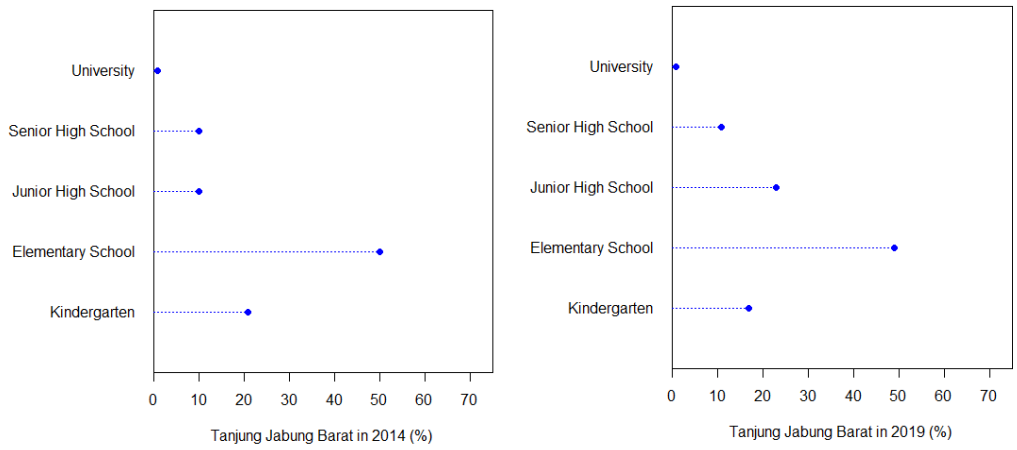
(a) (b)

Figure 8. Percentage of schools in Sarolangun District using the modified plot. (a) in 2014 and (b) in 2019



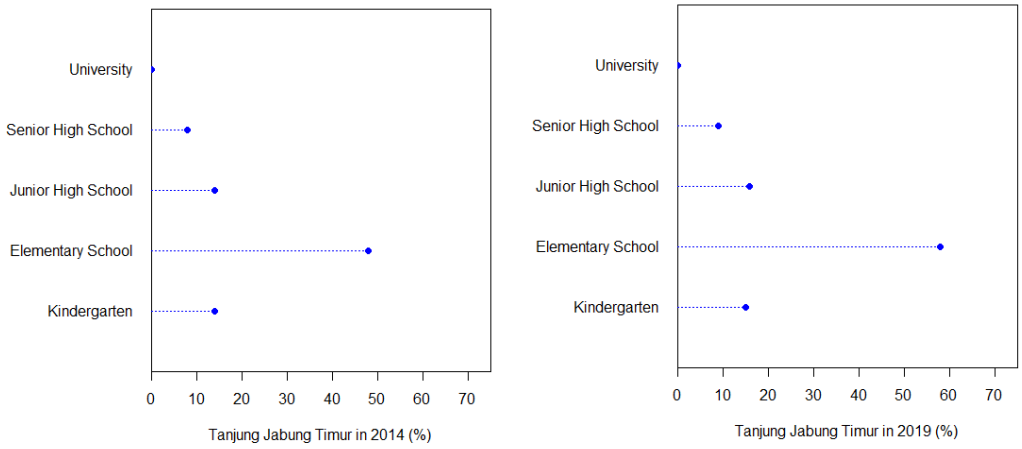
(a) (b)

Figure 9. Percentage of schools in Sungai Penuh City using the modified plot. (a) in 2014 and (b) in 2019



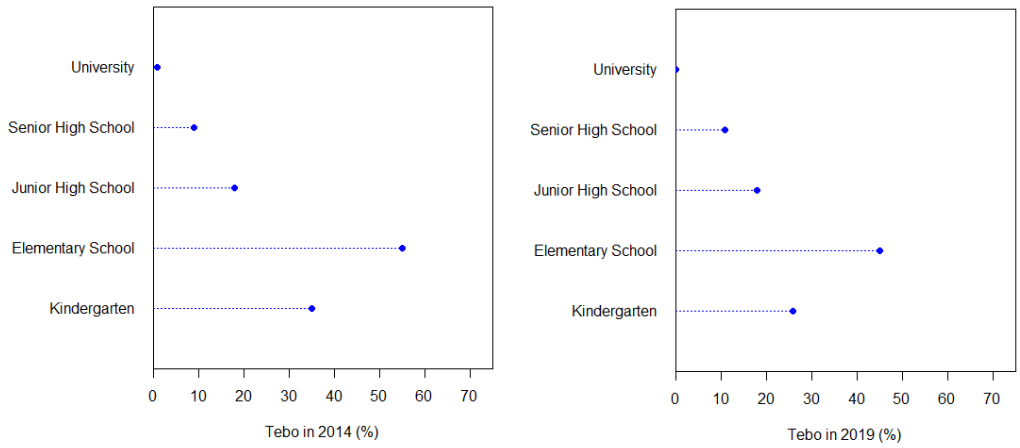
(a) (b)

Figure 10. Percentage of schools in Tanjung Jabung Barat District using the modified plot. (a) in 2014 and (b) in 2019



(a) (b)

Figure 11. Percentage of schools in Tanjung Jabung Timur District using the modified plot. (a) in 2014 and (b) in 2019



(a) (b)

Figure 12. Percentage of schools in Tebo District using the modified plot. (a) in 2014 and (b) in 2019

Based on Figure 2 and Figure 8, the percentage of Senior High School and University in Batang Hari and Sarolangun Regencies in 2014 and 2019 is constant. Meanwhile, other levels of education experienced a slight change. Only the percentage of universities in Bungo, Merangin, and Tanjung Jabung Timur regencies remains (Figure 3, Figure 6 and Figure 11). In addition, the percentage of each level of education in Bungo, Merangin, and Tanjung Jabung Timur regencies has changed. In Jambi City only the percentage at the Junior High School level remained constant in 2014 and 2019 (Figure 4). Meanwhile, at other levels in Jambi City, there are changes. Also, based on Figure 5 and Figure 7, the percentage of Senior High Schools in Kerinci and Muaro Jambi districts in 2014 and 2019 is constant. Meanwhile, other levels of education have changed. Only the percentage of Elementary Schools and Universities in Sungai Penuh City and Tanjung Jabung Barat District is constant (Figure 9 and Figure 10). In addition, the percentage of each level of education in Sungai Penuh City and Tanjung Jabung Barat District has changed. Based on Figure 12, even in Tebo District, all levels of education experienced a percentage change.

Figure 13. R programming syntax

CONCLUSION

The village potential statistics (PODES) data which is one of the primary data from the survey period of the Central Statistics Agency, Indonesia (BPS) can be used for research purposes. The research conducted does not have to be in the form of a statistical model, for example graphic analysis. We provide solutions for readers to modify graphics that are in R programming. Also, we modify label parts, horizontal axis, vertical axis, colors, etc. So that these modifications can be used to visualize the data, one of which is the PODES data. Podes data used in this study is Jambi Province in 2014 and 2019. In addition, the graphic visualization displayed consists of each city/district in Jambi Province. There are 2 cities in Jambi Province, namely Jambi City and Sungai Penuh. In addition, there are 9 regencies including Batang Hari, Bungo, Kerinci, Merangin, Muaro Jambi, Sarolangun, Tanjung Jabung Barat, Tanjung Jabung Timur, and Tebo regencies. In general, the number of schools in 2019 in each city/district has increased compared to 2014. This also affects the change in the percentage of schools in each city/district.

Differences between simple graphs and modified graphs are available based on the PODES data. The choice of using graphs is of course an option for researchers. The choice of using a simple programming syntax so as to produce simple graphics as well or using a more complex syntax but giving more attractive results. This research certainly has the

potential to be further developed. Use of other graphs such as bars, histograms, scatters, and others that have the opportunity to be used. There are so many PODES data variables that can be used in addition to the number of schools so that further research can use these variables. Furthermore, the percentage of the number of schools in this study is only at the City/District level. So that lower levels such as sub-districts and even sub-districts/villages can be used so that the results of further research are more detailed.

REFERENCES

Abdollahi, M., Farjad, B., Gupta, A., & Hassan, Q. K. (2022). CMIP6-D&A: An R-based software with GUI for processing climate data available in network common data format. *SoftwareX*, *18*, 101044. https://doi.org/10.1016/j.softx.2022.101044

Atkins, J. W., Stovall, A. E. L., & Silva, C. A. (2022). Open-Source tools in R for forestry and forest ecology. *Forest Ecology and Management*, *503*, 119813. https://doi.org/10.1016/j.foreco.2021.119813

Borisov, D. I. (2022). Analyticity of resolvents of elliptic operators on quantum graphs with small edges. *Advances in Mathematics*, *397*, 108125. https://doi.org/10.1016/j.aim.2021.108125

Diaz-Diaz, F., & Estrada, E. (2022). Time and space generalized diffusion equation on graph/networks. *Chaos, Solitons & Fractals*, *156*, 111791. https://doi.org/10.1016/j.chaos.2022.111791

Fitriansyah, F., & Nuryakin, C. (2021). Desa Digital dan Tingkat Literasi Keuangan Aparatur Desa: Studi Kasus Kabupaten Aceh Tamiang. *Jurnal Ekonomi Dan Pembangunan Indonesia*, *21*(2), 220–234. https://doi.org/10.21002/jepi.v21i2.1076

Hakim, L., Juwita, A. H., Bintariningtyas, S., & Guritno, D. C. (2021). Poverty and SMEs: A New Pradigm of SDGs Development. *Optimum: Jurnal Ekonomi Dan Pembangunan*, *11*(2), 224–231. https://doi.org/10.12928/optimum.v11i2.3943

Hardiyanto, E. (2020). KOMBINASI METODE ENSEMBLE, CFS DAN POHON KEPUTUSAN UNTUK PREDIKSI KINERJA PETUGAS STUDI KASUS: SURVEY PODES BADAN PUSAT STATISTIK. *Joutica*, *5*(1), 337–345. https://doi.org/10.30736/jti.v5i1.390

Hasan, M. F., Fadhil, I., Fahmid, M. M., & Ahmad, T. (2022). Impact of the European Union Regulations on Indonesian Oil Palm Smallholder Farmers. *International Journal of Oil Palm*, *5*(1), 1–15. https://doi.org/10.35876/ijop.v5i1.69

Junaidi, J., Yulmardi, Y., & Hardiani, H. (2020). Food crops-based and horticulture-based villages potential as growth center villages in Jambi Province, Indonesia. *Journal of Critical Reviews*, *7*(9), 514–519. https://doi.org/10.31838/jcr.07.09.102

Kaban, P. A., Nasution, B. I., Caraka, R. E., & Kurniawan, R. (2022). Implementing night light data as auxiliary variable of small area estimation. *Communications in Statistics - Theory and Methods*, 1–18. https://doi.org/10.1080/03610926.2022.2077963

Li, P., Wei, L.-Q., Pan, Y.-F., & Zhang, Y.-M. (2022). dQTG.seq: A comprehensive R tool for detecting all types of QTLs using extreme phenotype individuals in bi-parental segregation populations. *Computational and Structural Biotechnology Journal*, *20*, 2332–2337. https://doi.org/10.1016/j.csbj.2022.05.009

Paais, L. S. (2021). Keragaman Agama, Etnis, Bahasa, dan Pembangunan Desa. *Journal of Regional and Rural Development Planning*, *5*(2), 77–90. https://doi.org/10.29244/jp2wd.2021.5.2.77-90

Permatasari, N., & Larasati, W. (2022). Perbandingan Metode SAE EBLUP dan SAE HB Pada Pendugaan Area Kecil (Studi Kasus Pendugaan Kemiskinan di Provinsi Jawa

Timur). *Jurnal Statistika Dan Aplikasinya*, *6*(1), 96–108. https://doi.org/10.21009/JSA.06109

Saavedra-Nieves, A. (2021). Statistics and game theory: Estimating coalitional values in R software. *Operations Research Letters*, *49*(1), 129–135. https://doi.org/10.1016/j.orl.2020.12.005

Sari, R. K., & Handayani, D. (2020). Pemanfaatan Pelayanan Kesehatan pada Anak Indonesia: Pengaruh Kemiskinan dan Karakteristik Ibu. *Media Kesehatan Masyarakat Indonesia*, *16*(3), 305–316. https://doi.org/10.30597/mkmi.v16i3.9709

Shin, E., Yoo, S., Ju, Y., & Shin, D. (2022). Knowledge graph embedding and reasoning for real-time analytics support of chemical diagnosis from exposure symptoms. *Process Safety and Environmental Protection*, *157*, 92–105. https://doi.org/10.1016/j.psep.2021.11.002

Snitker, G., Moser, J. D., Southerlin, B., & Stewart, C. (2022). Detecting historic tar kilns and tar production sites using high-resolution, aerial LiDAR-derived digital elevation models: Introducing the Tar Kiln Feature Detection workflow (TKFD) using open-access R and FIJI software. *Journal of Archaeological Science: Reports*, *41*, 103340. https://doi.org/10.1016/j.jasrep.2022.103340

Tiddi, I., & Schlobach, S. (2022). Knowledge graphs as tools for explainable machine learning: A survey. *Artificial Intelligence*, *302*, 103627. https://doi.org/10.1016/j.artint.2021.103627

Trilaksono, T., & Sukartini, N. M. (2020). Kaitan Karakteristik Perangkat Desa Dengan Indeks Pembangunan Desa Di Indonesia. *Jurnal Sains Sosio Humaniora*, *4*(1), 194–204. https://doi.org/10.22437/jssh.v4i1.9916

Wang, Z., & Wang, J. (2022). Analytical graphs to describe operating status of industrial alarm variables. *Control Engineering Practice*, *118*, 104961. https://doi.org/10.1016/j.conengprac.2021.104961