1. Submit to JESTEC (1 Aug 2021)
2. First Revision (7 Sept 2021)
3. Submit Revised Paper (13 Sept 2021)
4. Accepted Submission (28 Sept 2021)
5. Paper Publish (05 Oct 2023)

1. Submit to JESTEC (1 Aug 2021)

Yusrila Y Kerlooza

<<kerlooza@email.unikom.ac.id>>

[JESTEC] Manuscript ID: JESTEC_16_5_2021_Yusrila Y Kerlooza–Submission Received

1 message

Editorial Office <<jjestec@gmail.com>>   1 August 2021 at 09:39
To: Yusrila Y Kerlooza <<kerlooza@email.unikom.ac.id>>

Dear Author,

Thank you very much for uploading the following manuscript to the submission system. One of our editors will be in touch with you soon.

Journal Name	Journal of Engineering, Science and Technology
Manuscript ID	JESTEC_ 16_5_2021_Yusrila Y Kerlooza
Authors	Yusrila Y. Kerlooza, Dian Dharmayanti, Adam Mukharil Bachtiar, Taufiq Nugraha
Title	Data Visualization for Education Domain at Dinas Pendidikan Jawa Barat

If you have any questions, please do not hesitate to contact editorial office.

Kind regards,

Editorial Office
Journal of Engineering, Science and Technology
E-Mail: jjestec@gmail.com

*** This is an automatically generated email ***

DATA VISUALIZATION FOR EDUCATION DOMAIN AT DINAS PENDIDIKAN JAWA BARAT

D Dharmayanti, A M Bachtiar, T Nugraha

Informatics Engineering Department, Universitas Komputer Indonesia, Jalan Dipatiukur Nomor 112-116, Bandung, Jawa Barat, Indonesia

*dian.dharmayanti@email.unikom.ac.id

Abstract. Based on the results of interviews to one of the employee in *Sub Bagian Perencanaan dan Pelaporan*, there were problems such as receiving information from the visualization used today not as expected. Its happen because the information from the visualization used today doesnt support one thing related to education equity. This research is about to helping *Sub Bagian Perencanaan dan Pelaporan* to understanding the case of education equity. Statistical methods, data mining, and concept design of data visualization to provide appropiate and usable information. After design and refine of existing visualization, the visualization outcome is implemented to website-based prototypes. For the results of prototypes already efficient and effective, prototypes tasted by Usability Testing. The result of usability testing reach 95.45%, which means that design used has achived a minimum goal. From these result, prototypes of data visualization can helping *Sub Bagian Perencanaan dan Pelaporan* to understand information about education equity.

1. Introduction

Increasing human resources by providing all educational facilities is one way for the Indonesian nation to become a superior country [1]. Education is one of the benchmarks of a nation to create quality human resources. Local governments require an education distribution program based on applicable laws. Based on interviews that have been conducted with the Planning and Reporting Subdivision in the West Java Education Office stated that education justice is available in four national education missions, educational services, affordability of education services, quality or quality of education services, and equality of education services [2].

Based on the results of interviews and observations found problems with the presentation of information related to the current distribution of education. Presentations in the form of statistical graphs have not been able to explain in detail information on equal distribution per district and sub-district area. Besides the information obtained is also not in accordance with the wishes of policy makers in the Office of Education. The large number of indicators in the distribution of education causes errors in concluding information and requires a long time to get one information. Misinformation can result in the focus of program plans issued the following year often not on target.

Based on current problems, data visualization can be used to make large and complex information understandable quickly. Visualization can also be useful to improve insight effectively and optimally [3]. The results of visualization will be efficient and effective in accordance with the information requirements requested by the *Sub Bagian Perencanaan dan Pelaporan*. Data visualization can run well if the method of delivering information from the form of visualization to the recipient of information in accordance with the answers to be submitted [4]. The purpose of visualization is that users can easily understand and interpret a large and complex collection of information [5]. Visualization is expected to help the *Sub Bagian Perencanaan dan Pelaporan* monitor various indicators for education equity.

2. Research methodology

Methodology used was inspired by 8 stages of visualization in the book "Visualizing Data" by Ben Fry [6]. These stages are adjusted to the research so as to produce stages that can be seen in **Error! Reference source not found.**.
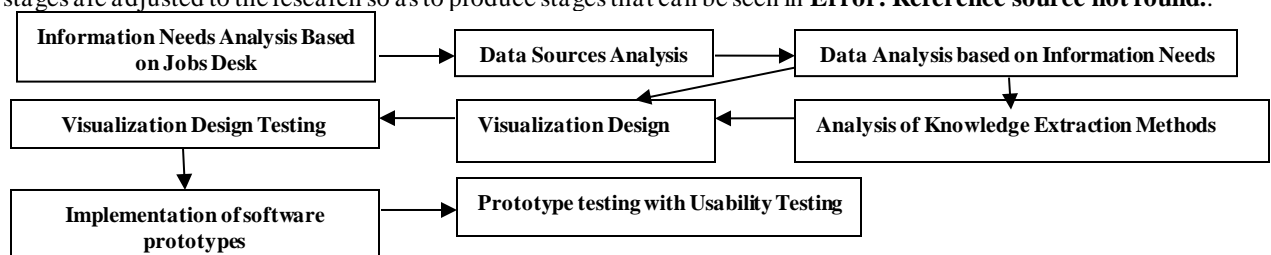

Figure 1. Research methodology

All stages in Figure 1 are done in stages. At the stage of designing visualization consists of 6 stages in more detail, namely understanding the context, determining the appropriate form of presentation, eliminating clutter, giving attention to focus on the user, then improving visualization by thinking as a designer and finally telling the visualization results [7]. And at the last stage of testing using the usability testing method. This method is evaluative in that the user will be given a task to find out the level of success in completing a given task.

3. Results and discussion
In this section, we will explain in detail each stage and research of the research.

3.1. Information Needs Analysis Base on Jobs Desk
At this stage, an analysis of the jobs desk at Sub Bagian Perencanaan dan Pelaporan of the West Java Provincial Education Office. Sub bagian perencanaan dan pelaporan has 3 major jobs such as collect data, planning and evaluation. Not all jobs require the results of data visualization, only two of the three jobs that require visualization.

3.2. Data Sources Analysis
The data used in this research came from the *Dinas Pendidikan Jawa Barat and Badan Pusat Statistik*. Data obtained from the *Dinas Pendidikan* is the output of the system in Microsoft Excel documents which consist of various attributes. Data obtained from the Badan Pusat Statistik is data taken from several Microsoft Excel documents. The following data used in this study can be seen in **Table 1**.

Table 1 Used Tables

Data	Table
Population data	The population of the school people table
	Development table
	Population table based on age
	Table of government performance indicators
Education data	Education table

3.3. Data Analysis based on Information Needs
The data mapping phase is the use of attributes related to the problem to be resolved. The attribute is formed by a collection because of the many of value used [7]. The following are the results of the data mapping as follows:
1. To obtain information on the number and percentage of increase or decrease in students repeating in the sub-district area, it is ideal to use pd_tkt_mengulang.
2. To find out the students the number and percentage of the number of increase or decrease in dropout students in the sub-district area using pd_tkt_putus_sekolah.
3. To get information on the number and percentage of the number and percentage of increase or decrease in students who did not graduate from the school in the sub-district area using pd_tkt_tdk_lulus.
4. To obtain information on the ratio of teachers and students in the sub-district area using pd_rasio_guru_siswa.
5. To obtain information on the percentage of increase or decrease in teacher worth teaching in the sub-district area using guru_lyk.
6. To obtain information on the percentage of increase or decrease in certification teachers in the sub-district area using a certification teacher.
7. To find out the development of the quality of teachers from time to time to see the lowest developments based on sub-district and district areas using the development of teachers.
8. To find out the total quality of damaged infrastructure in that year per sub-district area using means_rusak.
9. To find out the total quality of infrastructure that was damaged in that year per sub-district area using means_rusak.
10. To find out about schools without spaces such as libraries, schools, laboratories, labkom in one school in the district using school facilities.
11. To find out the percentage of gross participation of the community towards the government targets in the regency area using apk_target.
12. To find an affordable area, ideal education services are in the sub-districts using affordable areas.
13. To see the number of sub-districts that do not have a school in a district, they use no-school.
14. To see the percentage of pure community participation in the government targets in the district area using apm_target.
15. To see the distribution of the level of teacher needs that require teachers in an area using tkt_butuh_guru. To see the distribution of the area that has been evenly distributed using the equation

3.4. Analysis of Knowledge Extraction Methods

This stage is a process where the selection of knowledge extraction methods is based on information needs. Not all information needs can be delivered directly, at this stage the selection of methods for information needs is done through exploratory analysis or explanatory analysis. Information needs number 3, 11, 14 using explanatory analysis method. The rest of the information needs using exploratory knowledge extraction methods [7][8]. Following is the calculation of information needs:

1. Information on the number and percentage of increase or decrease in students repeating

After the calculation of this year and the previous year, the results of these calculations are reduced to see the increase or decrease that occurred. Following are the results of the increase or decrease that occurred in Baleendah Subdistrict:

$$Repeating\ figures\ this\ year - Figures\ Repeating\ Previous\ Years = 0,21 - 0,22\%$$
$$Repeating\ figures\ this\ year - Figures\ Repeating\ Previous\ Years = -0,01\%$$

It can be concluded that the Repeating Year is 0.01 better than the previous year but it is still not ideal because there are still students who repeat.

2. Student information the number and percentage of the number of students who have dropped out of school

$$Dropout\ rates\ this\ year = \frac{74}{10055} * 100\%$$
$$Dropout\ rates\ this\ year = 0,73\%$$
$$Previous\ Year\ dropout\ rate = \frac{92}{10932} * 100\%$$
$$Previous\ Year\ dropout\ rate = 0,84\%$$
$$Dropout\ rates\ this\ year - Previous\ Year\ dropout\ rate = -0,11\%$$

It can be concluded that the dropout rate for this year is 0.11 better than the previous year but still not ideal.

3. Information on the number and percentage of the number and percentage of increase or decrease in students who did not graduate from school

$$Graduate\ Numbers\ This\ year = \frac{3015}{3015} * 100\%$$
$$Graduate\ Numbers\ This\ year = 100\%$$
$$Number\ Not\ Passed = 0\%$$
$$Previous\ Year\ Graduates\ Figures = \frac{3532}{3532} * 100\%$$
$$Previous\ Year\ Graduates\ Figures = 100\%$$
$$Previous\ Year\ Graduates\ Figures = 0\%$$
$$Numbers\ Not\ Passing\ This\ Year - Numbers\ Not\ Passing\ Previous\ Years = 0\%$$

It can be concluded that the number of not graduating from school this year is the same as the previous year and ideal

4. Information on the ratio of teachers and students in the subdistrict area.

$$\frac{S}{G} = \frac{The\ number\ of\ students}{Total\ number\ of\ teachers} * 100\%$$
$$\frac{S}{G} = \frac{10055}{481} * 100\%$$
$$\frac{S}{G} = 20,90$$

Rasio siswa:guru = 20 : 1 ideal

5. Information on the percentage of increase or decrease in decent teachers in the sub-district area

$$\%GL = \frac{Number\ of\ authorized\ teachers}{Total\ number\ of\ teachers} * 100$$
$$\%GL = \frac{160+300}{161+320} * 100\%$$
$$\%GL = 96\%$$

$$\%GL\ Previous\ year = \frac{Number\ of\ authorized\ teachers}{Total\ number\ of\ teachers} * 100$$
$$\%GL\ Previous\ year = \frac{132+331}{143+340} * 100\%$$
$$\%GLTahunSebelumnya = 95\%$$

Only 96% of teachers are eligible, up 1% from the previous yea

6. Information on the percentage of increase or decrease in certification teachers in the sub-district area using a certified teacher.

$$GS = \frac{Number\ of\ certified\ teachers}{Total\ number\ of\ teachers} * 100\%$$
$$\%GS = \frac{132+287}{161+320} * 100\%$$
$$\%GS = 87,11\%$$
$$\%GS\ Previous\ year = \frac{Number\ of\ certified\ teachers}{Total\ number\ of\ teachers} * 100$$
$$\%GS\ Previous\ year = \frac{82+92}{143+340} * 100\%$$
$$\%GSTahunSebelumnya = 36\%$$

Only 87.11% of teachers who received certification rose 51.11%

7. Regional information affordable ideal education services are in the sub-district area.

$$KPUS = \frac{Specific\ age\ population}{An\ area}$$

$$KSek = \frac{Certain\ level\ schools}{An\ area}$$
$$KPUS = \frac{4962}{41,933}$$
$$KPUS = 118,33$$
$$KSek = \frac{20}{41,933}$$
$$KSek = 0,476$$
$$DTSek = \frac{22}{7} * 10^2 * 0,476$$
$$DTSek = 149,89$$
$$DT = \frac{37189,34}{149,89}$$
$$DT = 248,1$$

The results of the summation were informed that there were only 248 children aged 16-18 years who could attend one secondary school. If 20 schools are available, 99.95% of children can enter the school.

8. To obtain information on the level of teacher needs per region

The method used here is grouping using the K-Means algorithm [9]. Variables used include number of teachers, number of schools, number of students. The desired group determination is 3. After several iterations have been obtained, the first cluster of centroid 1 has 15 members, centroid 2 has 11 members, and centroid 3 has 1 member. The implementation is done with the programming language R. From these results it can be seen the characteristics of the groups formed which are group 1 for groups that have less distribution of teacher distribution, group 2 has evenly distributed teacher distribution, and group 3 has the level of distribution of teachers who have fulfilled and excess

9. To see the spread of the area that is evenly distributed.

Variables used are percentage of apk, APM, student / teacher ratio, s / k ratio, student / school ratio. The desired group determination is 2. After several iterations are obtained the results of the first cluster of centroid 1 has 20 members, centroid 2 has 7 members. The results can be seen that group 1 is an area that is not evenly distributed and group 2 is an evenly distributed area

3.5. Design Visualization

At this stage the visualization of the information needs to be conveyed is carried out. In the design stage of transpiration visualization in storytelling books with data is the result of beginning with understanding the context of information needs, selecting visualization forms, eliminating existing clutter, focusing attention, and testing using acceptance testing techniques [10][11]. The following are some of the following designs:

a. Understanding context

The following is the context understood in this information need can be seen in **Table 2.**

Table 2. Understanding Context

WHAT	Knowing the number of students repeating has increased or decreased in order to determine one of the ideal indicators of student quality.
HOW	Showing the number of students who repeat compared to the previous year to see a comparison to the ideal or away from the ideal.

b. Selection of visualization

Following is the selection of forms that have been adapted to the required context can be seen in **Table 3**.

Table 3. Selection of Visualization

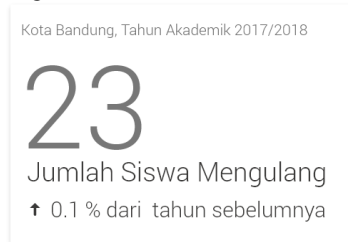| Visualization	Simple text and symbols
Reason	Can display one to two information needs simultaneously

c. Remove clutter

The following are the results of clutter removal which can be seen in **Table 4**

Table 4. Remove Clutter

Clutter	Reason
Alignment	Left-aligned usage to facilitate reading from left to right
White Space	Use of white space to provide space for labels and content so that they do not overlap.
Closure	There is no use of borders in this form of visualization
Similarity	Use of colors to indicate relationships.

The following are the results of the existing clutter removal can be seen in **Figure 1.**

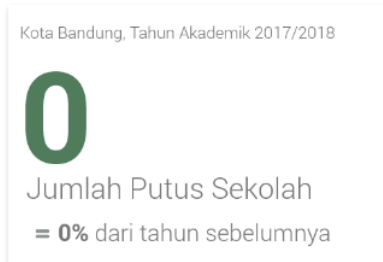
Figure 1. Result Remove Clutter

d. Focus attention

Attention focusing is done by giving different font colors and sizes. Following are the results of attention focusing can be seen in **Figure 2.**

Figure 2. Result Focusing Attention

Here are some visualization results that can be seen on **Figure 3**, **Figure 4**, **Figure 5**, **Figure 6**, and Figure **7.**

Figure 3. Result Visualization Dropout Students

Figure 4. Result Visualization Students Dont Graduate

Figure 5. Result Visualization Percentage APK

Figure 6. Result Visualization the area does not have a school
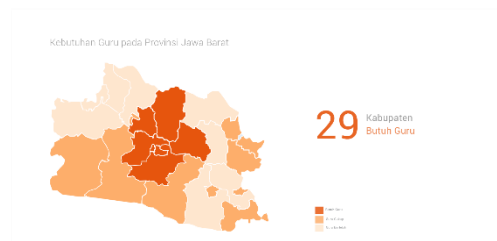
Figure 7. Result Visualization Education Equity

3.6. Visualization design testing

Tests conducted for participants who are experts in their fields [11]. The following are the results of the acceptance test which can be seen in Table 5.

Table 5 Visaulization Design Tetsing

Information Needs	Visualization	Acceptance results
How to know the number of students repeating an increase or decrease from the previous year in the sub-district area per academic year is ideal	Simple text and symbols	accepted
How to find out the number of students dropping out of school experienced an increase or decrease from the previous year in the sub-district area per academic year is ideal	Simple text and symbols	accepted
How to know the percentage of students graduating from school experiencing an increase or decrease from the previous year in the sub-district area per academic year is ideal	Simple text and symbols	accepted

3.7. Implementation of software prototypes

At this stage the software prototype is designed as well as supporting technology for prototype development. The following are the results of the design of a software prototype that displays the quality of students based on student quality indicators can be seen on **Figure 8**, **Figure 9, Figure 10, Figure 11**, **Figure 12**, and **Figure 13 [11].**
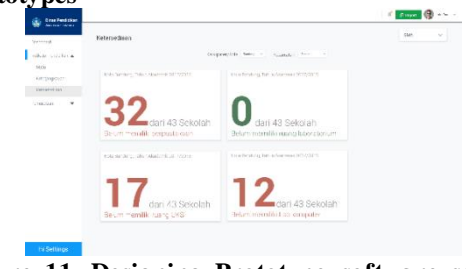

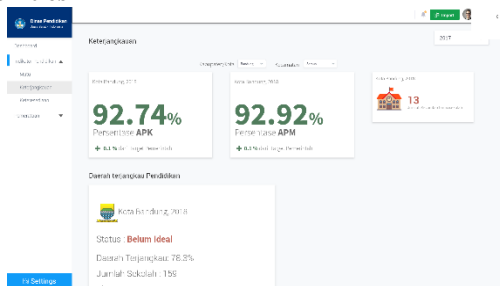**Figure 8. Designing student quality software prototypes**
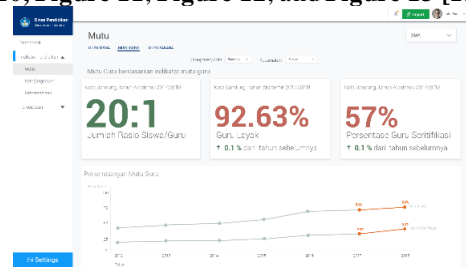
Figure 9. Designing teacher quality software prototypes
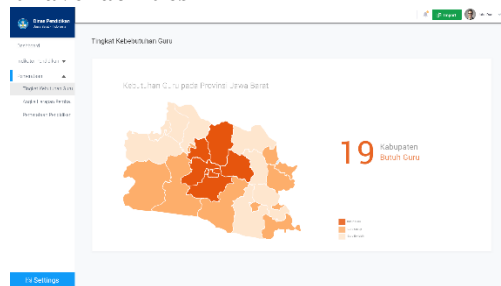

Figure 10. Designing prototype software quality facilities

Figure 11. Designing Prototype software schools dont have facilities

Figure 12. Affordable regional prototype design

Figure 13. Education Equity Design Prototypes

Conclusions

Based on the results of the implementation and testing that has been carried out on the educational data visualization system and the population in the education office of the province of West Java, it can be concluded that visualization makes it easy to see and understand the data of each information displayed. Data visualization can solve problems that occur and can meet the objectives of this study. However, there are still obstacles faced in visualizing population data that can be accessed in a limited way because of using open data. Therefore, suggestions for the development of data visualization systems in the future are to collaborate with the statistical center to access population data in more detail.

References
[1] P. Lestari, "Pentingnya Pendidikan Bagi Bangsa Indonesia," 2016. [Daring]. Tersedia pada: https://student.cnnindonesia.com/inspirasi/20160114104154-322-104215/pentingnya-pendidikan-bagi-bangsa-indonesia/. [Diakses: 19-Feb-2018].
[2] Dinas Pendidikan Provinsi Provinsi Jawa Barat, *Statistik Pendidikan Jawa Barat*. Jawa Barat, 2017.
[3] A. M. Bachtiar dan D. Dharmayanti, "Data Visualization at Trading Sector Using Government ' s Open Data," vol. 118, no. 20, hal. 523–530, 2018.
[4] N. A. Syaripul dan A. M. Bachtiar, "VISUALISASI DATA INTERAKTIF DATA TERBUKA PEMERINTAH PROVINSI DKI JAKARTA: TOPIK EKONOMI DAN KEUANGAN DAERAH," vol. 12, hal. 82–89, 2016.
[5] M. Khan dan S. Khan, "Data and information visualization methods, and interactive mechanisms: A survey," *Int. J. Comput. Appl.*, vol. 34, no. 1, hal. 1–14, 2011.
[6] C. N. Knaflic, *Storytelling With Data*, vol. 8, no. June. John Wiley & Sons, Inc, 2015.
[7] M. Diamond dan A. M. Pedagogies, "Data Visualization: An Exploratory Study into the Software Tools Used by Businesses.," *Eric*, vol. 18, hal. 1–7, 2017.
[8] N. Saquib, *Mathematica Data Visualization*. 2014.
[9] M. Shamim Reza, "Study of Multivariate Data Clustering Based on K-Means and Independent Component Analysis," *Am. J. Theor. Appl. Stat.*, vol. 4, no. 5, p. 317, 2015.
[10] R. H. G Ginde, S Saha, C Balasubramaniam, "Visualisation of massive data from scholarly Article and Journal Database : A Novel Scheme," *IJRCCE Conf. Proc.*, 2016.
[11] A. M. Tervakari, K. Silius, J. Koro, J. Paukkeri, dan O. Pirttilä, "Usefulness of information visualizations based on educational data," *IEEE Glob. Eng. Educ. Conf. EDUCON*, no. April, hal. 142–151, 2014.

2. First Revision (7 Sept 2021)

Yusrila Y Kerlooza

<kerlooza@email.unikom.ac.id**>**

[JESTEC] Manuscript ID: JESTEC_ 2021_Yusrila Y Kerlooza – Minor Revision

1 message

Editorial Office <jjestec@gmail.com>   7 September 2021 at 16:20
To: Yusrila Y Kerlooza <kerlooza@email.unikom.ac.id>

Dear Author,

Thank you for submitting the following manuscript to Journal of Engineering, Science and Technology:

Journal Name	Journal of Engineering, Science and Technology
Manuscript ID	JESTEC_ 16_5_2021_Yusrila Y Kerlooza
Authors	Yusrila Y. Kerlooza, Dian Dharmayanti, Adam Mukharil Bachtiar, Taufiq Nugraha
Title	Data Visualization for Education Domain at Dinas Pendidikan Jawa Barat

It has been reviewed by experts in the field and we request that you make major revisions before it is processed further. Please find the comments from reviewer in the attached file.

Please revise the manuscript according to the reviewers' comments and upload the revised file within 10 days. Use the version of your manuscript found at the above link for your revisions, as the editorial office may have made formatting changes to your original submission.

If the reviewers have suggested that your manuscript should undergo extensive English editing, please address this during revision. We suggest that you have your manuscript checked by a native English-speaking colleague or use a professional English editing service.

Do not hesitate to contact us if you have any questions regarding the revision of your manuscript or if you need more time. We look forward to hearing from you soon.

Kind regards,

Editorial Office
Journal of Engineering, Science and Technology
E-Mail: jjestec@gmail.com

DATA VISUALIZATION FOR EDUCATION DOMAIN AT DINAS PENDIDIKAN JAWA BARAT

YUSRILA Y. KERLOOZA*, DIAN DHARMAYANTI, ADAM MUKHARIL BACHTIAR, TAUFIQ NUGRAHA

Informatics Engineering, Faculty of Engineering and Computer Science,
Universitas Komputer Indonesia,
Jalan Dipatiukur Nomor 112-116, Bandung, Jawa Barat, Indonesia
*Corresponding Author : kerlooza@email.unikom.ac.id

Abstract

The purpose of this research is to make the proper visualization and can be use in Dinas Pendidikan Jawa Barat to looking education equity level in the West Java Province. This research is about to helping *Sub Bagian Perencanaan dan Pelaporan* to understanding the case of education equity. Based on the results of interviews to one of the employee in *Sub Bagian Perencanaan dan Pelaporan*, there were problems such as receiving information from the visualization used today not as expected. Its happen because the information from the visualization used today doesnt support one thing related to education equity. There are several methods used to produce the right form of information including statistics, data mining, and the application of data visualization rules. After design and refine of existing visualization, the visualization outcome is implemented to website-based prototypes. For the results of prototypes already efficient and effective, prototypes tasted by Usability Testing. The result of usability testing reach 95.45%, which means that design used has achived a minimum goal. From these result, prototypes of data visualization can helping Sub Bagian Perencanaan dan Pelaporan to understand information about education equity.

Keywords : Data Visualization, Education Equity, Statistics, Data Mining, Usability Testing, Acceptance Testing.

1. Introduction

Monitoring of equal distribution of education must be carried out to create equitable education throughout West Java Province. Monitoring is carried out to produce a focus on the education plan for the following year. The number of indicators of educational equity makes the information obtained is wrong and it takes a long time to get the meaning of the information [1]. Misinformation can result in the focus of the program plans issued in the following year being not on target. Visualization can also be useful to improve insight effectively and optimally [2]. The results of visualization will be efficient and effective in accordance with the information requirements requested by the Sub Bagian Perencanaan dan Pelaporan.

The results of interviews and observations found problems with the presentation of information related to the current distribution of education. Presentations in the form of statistical graphs have not been able to explain in detail information on equal distribution per district and sub-district area. Besides the information obtained is also not in accordance with the wishes of policy makers in the Office of Education. The large number of indicators in the distribution of education causes errors in concluding information and requires a long time to get one information.

Data visualization can run well if the method of delivering information from the form of visualization to the recipient of information in accordance with the answers to be submitted [3]. The purpose of visualization is that users can easily understand and interpret a large and complex collection of information [4]. Visualization is expected to help the *Sub Bagian Perencanaan dan Pelaporan* monitor various indicators for education equity.

2. Research methodology

Methodology used was inspired by 8 stages of visualization in the book "Visualizing Data" by Ben Fry. These stages are adjusted to the research so as to produce stages that can be seen in Fig. 1 [3].
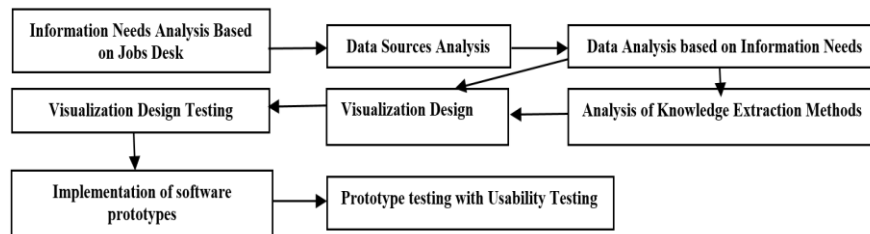
Fig. 1. Research methodology

All stages in Figure 1 are done in stages. At the stage of designing visualization consists of 6 stages in more detail, namely understanding the context, determining the appropriate form of presentation, eliminating clutter, giving attention to focus on the user, then improving visualization by thinking as a designer and finally telling the visualization results [3]. And at the last stage of testing using the usability testing method. This method is evaluative in that the user will be given a task to find out the level of success in completing a given task.

3. Results and discussion

The stages in the research are described in this section.

3.1. Information Needs Analysis Base on Jobs Desk

At this stage, an analysis of the jobs desk at Sub Bagian Perencanaan dan Pelaporan of the West Java Provincial Education Office. Sub bagian perencanaan dan pelaporan has 3 major jobs such as collect data, planning and evaluation. Not all jobs require the results of data visualization, only two of the three jobs that require visualization.

3.2. Data Sources Analysis

The data used in this research came from the *Dinas Pendidikan Jawa Barat and Badan Pusat Statistik*. Data obtained from the *Dinas Pendidikan* is the output of the system in Microsoft Excel documents which consist of various attributes. Data obtained from the Badan Pusat Statistik is data taken from several Microsoft Excel documents. The following data used in this study can be seen in Table 1.

Table 1. Used Tables

Data	Table
Population data	The population of the school people table
	Development table
	Population table based on age
	Table of government performance indicators
Education data	Education table

3.3. Data Analysis based on Information Needs

The data mapping stage is the use of attributes related to the problem to be resolved and the existing strategic information needs. These attributes are formed by a collection because of the number of values used [4]. The following is an example of the results of data mapping as follows :

1. To obtain information on the number and percentage of increase or decrease in students repeating in the sub-district area, it is ideal to use pd_tkt_mengulang.
2. To find out the students the number and percentage of the number of increase or decrease in dropout students in the sub-district area using pd_tkt_putus_sekolah.
3. To get information on the number and percentage of the number and percentage of increase or decrease in students who did not graduate from the school in the sub-district area using pd_tkt_tdk_lulus.
4. To obtain information on the ratio of teachers and students in the sub-district area using pd_rasio_guru_siswa.
5. To obtain information on the percentage of increase or decrease in teacher worth teaching in the sub-district area using guru_lyk.
6. To obtain information on the percentage of increase or decrease in certification teachers in the sub-district area using a certification teacher.
7. To find out the development of the quality of teachers from time to time to see the lowest developments based on sub-district and district areas using the development of teachers.
8. To find out the total quality of damaged infrastructure in that year per sub-district area using means_rusak.
9. To find out the total quality of infrastructure that was damaged in that year per sub-district area using means_rusak.
10. To find out about schools without spaces such as libraries, schools, laboratories, labkom in one school in the district using school facilities.

The mapping above is part of all the strategic information needs in this study to be mapped against the data used. The data is mapped based on data sources relating to the monitoring and dissemination of education in West Java.

3.4. Analysis of Knowledge Extraction Methods

This stage is a process where the selection of knowledge extraction methods is based on information needs. Not all information needs can be delivered directly, at this stage the selection of methods for information needs is done through exploratory analysis or explanatory analysis. Information needs number 3, 11, 14 using explanatory analysis method. The rest of the information needs using exploratory knowledge extraction methods [5,6]. Following is the calculation of information needs:

1. Information on the number and percentage of increase or decrease in students repeating.

After the calculation of this year and the previous year, the results of these calculations are reduced to see the increase or decrease that occurred. Following are the results of the increase or decrease that occurred in Baleendah Subdistrict:

$$Repeating\ figures\ this\ year - Figures\ Repeating\ Previous\ Years = 0{,}21 - 0{,}22\%$$
$$Repeating\ figures\ this\ year - Figures\ Repeating\ Previous\ Years = -0{,}01\%$$

It can be concluded that the Repeating Year is 0.01 better than the previous year but it is still not ideal because there are still students who repeat.

2. Student information the number and percentage of the number of students who have dropped out of school.

$$Dropout\ rates\ this\ year = \frac{74}{10055} * 100\%$$
$$Dropout\ rates\ this\ year = 0{,}73\%$$
$$Previous\ Year\ dropout\ rate = \frac{92}{10932} * 100\%$$
$$Previous\ Year\ dropout\ rate = 0{,}84\%$$
$$Dropout\ rates\ this\ year - Previous\ Year\ dropout\ rate = -0{,}11\%$$

It can be concluded that the dropout rate for this year is 0.11 better than the previous year but still not ideal.

3. Information on the number and percentage of the number and percentage of increase or decrease in students who did not graduate from school.

$$Graduate\ Numbers\ This\ year = \frac{3015}{3015} * 100\%$$
$$Graduate\ Numbers\ This\ year = 100\%$$
$$Number\ Not\ Passed = 0\%$$
$$Previous\ Year\ Graduates\ Figures = \frac{3532}{3532} * 100\%$$
$$Previous\ Year\ Graduates\ Figures = 100\%$$
$$Previous\ Year\ Graduates\ Figures = 0\%$$
$$Numbers\ Not\ Passing\ This\ Year - Numbers\ Not\ Passing\ Previous\ Years = 0\%$$

It can be concluded that the number of not graduating from school this year is the same as the previous year and ideal.

4. Information on the ratio of teachers and students in the subdistrict area.

$$\frac{S}{G} = \frac{The\ number\ of\ students}{Total\ number\ of\ teachers} * 100\%$$
$$\frac{S}{G} = \frac{10055}{481} * 100\%$$
$$\frac{S}{G} = 20{,}90$$

Rasio siswa:guru = 20 : 1 ideal

5. Information on the percentage of increase or decrease in decent teachers in the sub-district area.

$$\%GL = \frac{Number\ of\ authorized\ teachers}{Total\ number\ of\ teachers} * 100$$
$$\%GL = \frac{160+300}{161+320} * 100\%$$
$$\%GL = 96\%$$
$$\%GL \text{Previous year} = \frac{Number\ of\ authorized\ teachers}{Total\ number\ of\ teachers} * 100$$
$$\%GL \text{Previous year} = \frac{132+331}{143+340} * 100\%$$
$$\%GLTahunSebelumnya = 95\%$$

Only 96% of teachers are eligible, up 1% from the previous year.

6. Information on the percentage of increase or decrease in certification teachers in the sub-district area using a certified teacher.

$$GS = \frac{Number\ of\ certified\ teachers}{Total\ number\ of\ teachers} * 100\%$$
$$\%GS = \frac{132+287}{161+320} * 100\%$$
$$\%GS = 87,11\%$$
$$\%GS \text{Previous year} = \frac{Number\ of\ certified\ teachers}{Total\ number\ of\ teachers} * 100$$
$$\%GS \text{Previous year} = \frac{82+92}{143+340} * 100\%$$
$$\%GSTahunSebelumnya = 36\%$$

Only 87.11% of teachers who received certification rose 51.11%.

7. Regional information affordable ideal education services are in the sub-district area.

$$KPUS = \frac{Specific\ age\ population}{An\ area}$$
$$KSek = \frac{Certain\ level\ schools}{An\ area}$$
$$KPUS = \frac{4962}{41,933}$$
$$KPUS = 118,33$$
$$KSek = \frac{20}{41,933}$$
$$KSek = 0,476$$
$$DTSek = \frac{22}{7} * 10^2 * 0,476$$
$$DTSek = 149,89$$
$$DT = \frac{37189,34}{149,89}$$
$$DT = 248,1$$

The results of the summation were informed that there were only 248 children aged 16-18 years who could attend one secondary school. If 20 schools are available, 99.95% of children can enter the school.

8. To obtain information on the level of teacher needs per region

The method used here is grouping using the K-Means algorithm. Variables used include number of teachers, number of schools, number of students. The desired group determination is 3. After several iterations have been obtained, the first cluster of centroid 1 has 15 members, centroid 2 has 11 members, and centroid 3 has 1 member. The implementation is done with the programming language R.

From these results it can be seen the characteristics of the groups formed which are group 1 for groups that have less distribution of teacher distribution, group 2 has evenly distributed teacher distribution, and group 3 has the level of distribution of teachers who have fulfilled and excess.

9. To see the spread of the area that is evenly distributed.

Variables used are percentage of apk, APM, student / teacher ratio, s / k ratio, student / school ratio. The desired group determination is 2. After several iterations are obtained the results of the first cluster of centroid 1 has 20 members, centroid 2 has 7 members. The results can be seen that group 1 is an area that is not evenly distributed and group 2 is an evenly distributed area.

3.5. Design Visualization

At this stage the visualization of the information needs to be conveyed is carried out. In the design stage of transpiration visualization in storytelling books with data is the result of beginning with understanding the context of information needs, selecting visualization forms, eliminating existing clutter, focusing attention, and testing using acceptance testing techniques [7,8]. The following are some of the following designs:

a. Understanding context

The following is the context understood in this information need can be seen in Table 2.

Table 2. Understanding Context

WHAT	Knowing the number of students repeating has increased or decreased in order to determine one of the ideal indicators of student quality.
HOW	Showing the number of students who repeat compared to the previous year to see a comparison to the ideal or away from the ideal.

b. Selection of visualization

Following is the selection of forms that have been adapted to the required context can be seen in Table 3.

Table 3. Selection of Visualization

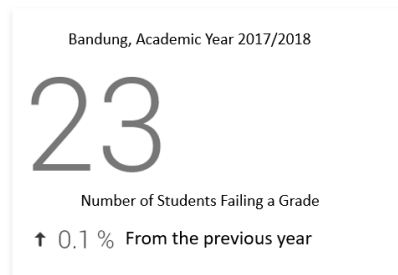|	
Visualization	Simple text and symbols
Reason	Can display one to two information needs simultaneously

c. Remove clutter

The following are the results of clutter removal which can be seen in Table 4.

Table 4. Remove Clutter

| Clutter	Reason
Alignment	Left-aligned usage to facilitate reading from left to right
White Space	Use of white space to provide space for labels and content so that they do not overlap.
Closure	There is no use of borders in this form of visualization
Similarity	Use of colors to indicate relationships.

The following are the results of the existing clutter removal can be seen in Fig. 2.
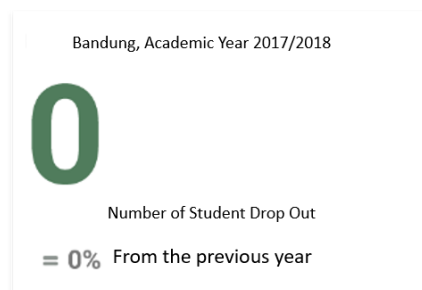

Fig. 2. Result Remove Clutter

d. Focus attention

Attention focusing is done by giving different font colors and sizes. Following are the results of attention focusing can be seen in Fig. 3.
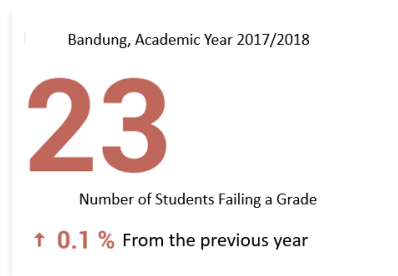

Fig. 3. Result Focusing Attention

Here are some visualization results that can be seen. Fig. 4 shows the results of the need for information on the number of school dropouts who have been given different colors and font sizes.


Fig. 4. Result Visualization Dropout Students

After removing the clutter, Fig. 5 is an emphasis on the number of students failing to pass within a certain period of time.
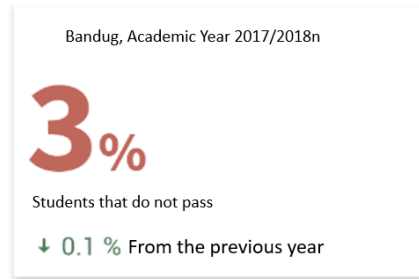
Fig. 5. Result Visualization Students Dont Graduate

The main focus in Fig. 6 is the percentage of the gross enrollment rate. This is indicated by the distinguished font size and color. The green color indicates the percentage of the target set by the government for that year.
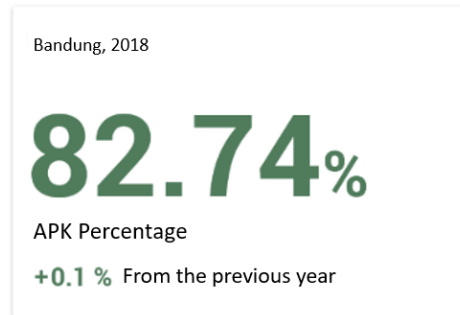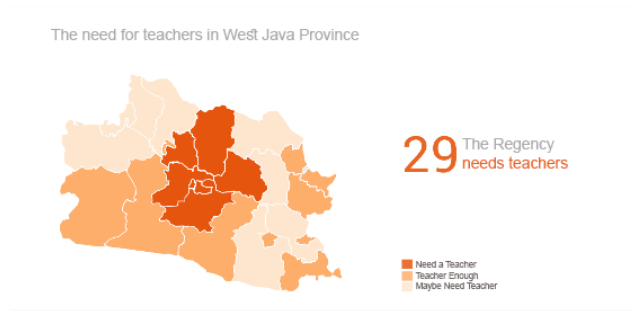

Fig. 6. Result Visualization Percentage APK

Fig. 7 is the result of visualization that has emphasized the information to be conveyed.

Fig. 7. Result Visualization the area does not have a school

In the information needs of teachers, a simple text visualization form and a choropleth map are used as shown in Fig. 8.

Fig. 8. Result Visualization Education Equity

The choice of simple text visualization is because it is suitable for displaying one or two pieces of information simultaneously. Choosing the form of choropleth map visualization for the division of regions based on geographic location. Combining these two types of visualization can add value to information in an information need.

3.6. Visualization design testing

Tests conducted for participants who are experts in their fields [9,10]. The following are the results of the acceptance test which can be seen in Table 5.

Table 5. Visaulization Design Tetsing

Information Needs	Visualization	Acceptance results
How to know the number of students repeating an increase or decrease from the previous year in the sub-district area per academic year is ideal	Simple text and symbols	accepted
How to find out the number of students dropping out of school experienced an increase or decrease from the previous year in the sub-district area per academic year is ideal	Simple text and symbols	accepted
How to know the percentage of students graduating from school experiencing an increase or decrease from the previous year in the sub-district area per academic year is ideal	Simple text and symbols	accepted

3.7. Implementation of software prototypes

At this stage the software prototype is designed as well as supporting technology for prototype development [9]. The following is a dashboard display that shows information related to the quality of education at the high school level. can be seen on Fig. 9.
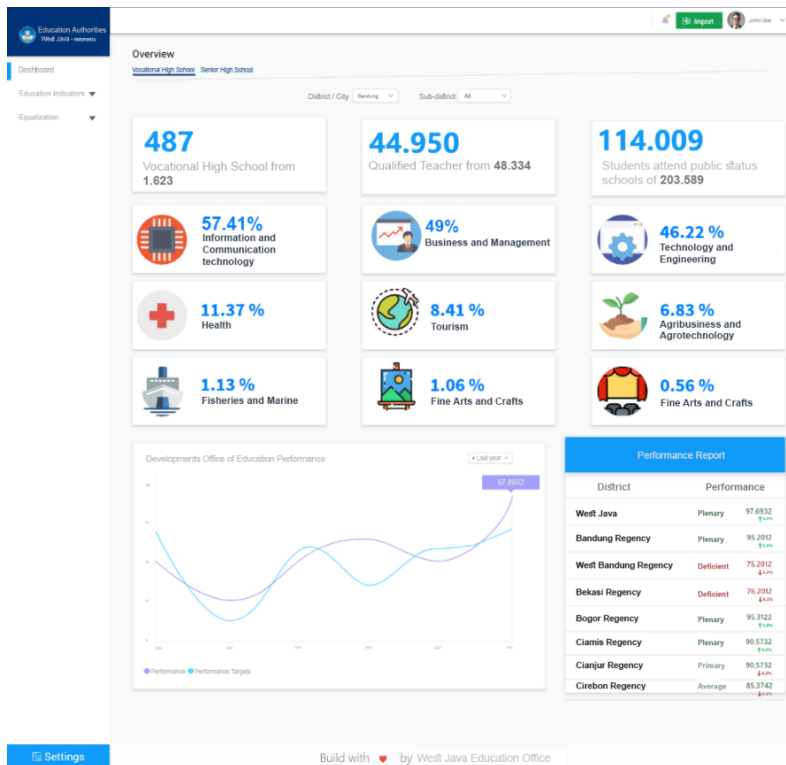
Fig. 9. Dashboard of the quality of education at the high school level.

Designing prototype software quality of facilities describes information related to the quality of educational facilities and infrastructure in the province of West Java (see Fig. 10).
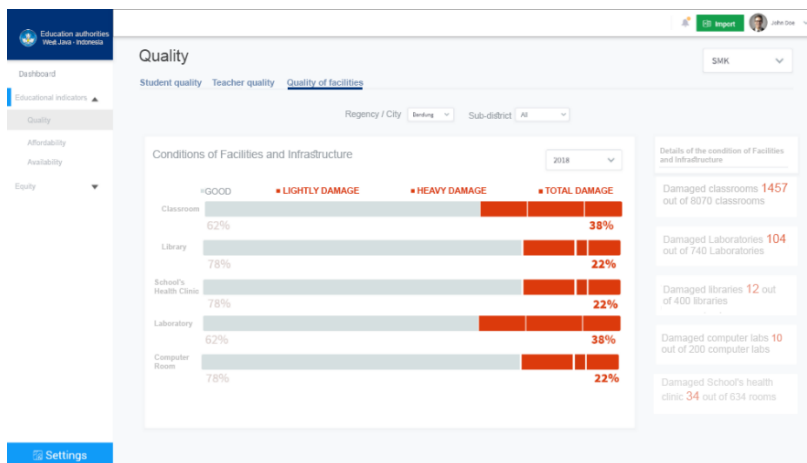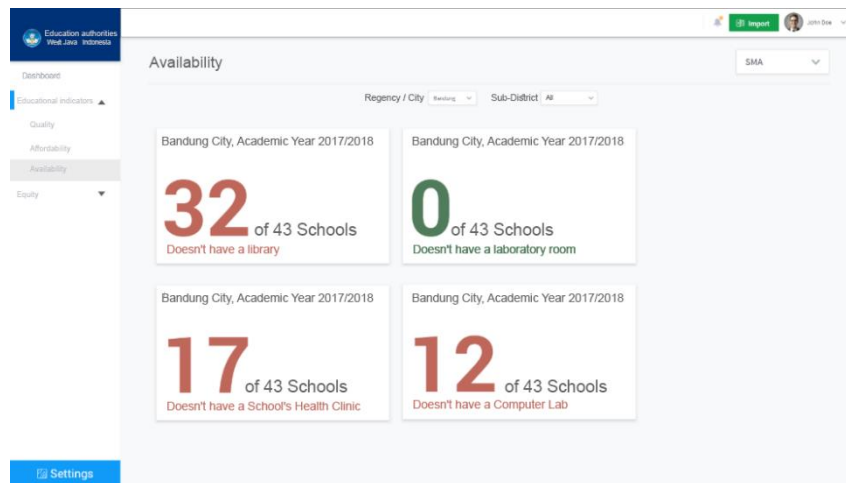

Fig. 10. **Designing prototype software quality of facilities**

Meanwhile, Designing prototype software availability of educational services describes information related to the availability of educational services from schools in West Java (see Fig. 11).

**Fig. 11. Designing prototype software of availability of educational services**

3.8 Testing software prototypes

Testing software prototypes aims to determine the design of the prototypes that have been made effective and efficient [11]. The following is a test scenario for each menu containing each visualization that can be seen in Table 5.

Table 5. Testing Prototypes

No	Task	Scenario	Expected results
1.	Enter the page that contains the visualization of the quality of vocational students in West Java	I am an employee of the planning and reporting subdivision who wants to know information about the quality of vocational students in West Java	Successfully enter the page that contains visualization of the quality of vocational students in West Java through the education indicator menu.
2	Enter the page that contains the visualization of the quality of vocational teachers in the province of West Java	I am an employee of the planning and reporting subdivision who wants to know information about teachers of vocational students in the province of West Java.	Successfully enter the page containing visualization of the quality of vocational teachers in the province of West Java through the education indicator menu.

Prototype testing can determine the success rate of prototype design. The following are the results of mapping efficiency testing based on tasks that can be seen on Fig. 12.
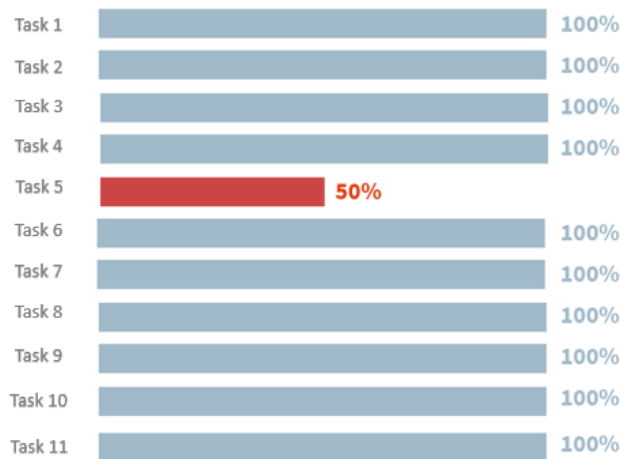
Fig. 12. Visualization Result Testing

From the test results it can be concluded that the effectiveness and efficiency in the completion of the task is 95.45% which can be said to be successful because achieving the minimum goal is 70%.

4. Conclusion

Based on the results of the implementation and testing that has been carried out on the educational data visualization system and the population in the education office of the province of West Java, it can be concluded that visualization makes it easy to see and understand the data of each information displayed. Data visualization can solve problems that occur and can meet the objectives of this study. However, there are still obstacles faced in visualizing population data that can be accessed in a limited way because of using open data. Therefore, suggestions for the development of data visualization systems in the future are to collaborate with the statistical center to access population data in more detail.

References

1. Dinas Pendidikan Provinsi Provinsi Jawa Barat, *Statistik Pendidikan Jawa Barat*. Jawa Barat, 2017.
2. Chen, M.; Ebert, D.; Hagen, H.; Laramee, R. S.; Van Liere, R.; Ma, K. L.; ...and Silver, D. (2008); Data, information, and knowledge in visualization. IEEE computer graphics and applications, 29(1), 12-19.
3. Fry, B. (2008). Visualizing data: Exploring and explaining data with the processing environment. " O'Reilly Media, Inc.".
4. Khan, M.; and Khan, S. S. (2011). Data and information visualization methods, and interactive mechanisms: A survey. International Journal of Computer Applications, 34(1), 1-14.
5. Verbert, K.; Duval, E.; Klerkx, J.; Govaerts, S.; and Santos, J. L. (2013). Learning analytics dashboard applications. *American Behavioral*

Scientist, 57(10), 1500-1509.

6. Cimpanu, C.; and Ferariu, L. (2012). Survey of data clustering algorithms, *Bul. Institutului Politeh. din Iasi, Tome LVIII(LXII)*, vol. 3, no. Lxii, pp. 23–42.

7. Morissette, L.; and Chartier, S. (2013). The k-means clustering technique: General considerations and implementation in Mathematica. *Tutorials in Quantitative Methods for Psychology*, 9(1), 15-24.

8. Ellis, G.; and Dix, A. (2007). A taxonomy of clutter reduction for information visualisation. IEEE transactions on visualization and computer graphics, 13(6), 1216-1223.

9. Amar, R.; and Stasko, J. (2004, October). BEST PAPER: A knowledge task-based framework for design and evaluation of information visualizations. In *IEEE Symposium on Information Visualization*. pp. 143-150.

10. Diamond, M.; and Mattia, A. (2017). Data visualization: An exploratory study into the software tools used by businesses. *Journal of Instructional Pedagogies*, 18.

11. Mifsud, J. (2015). Usability metrics–A guide to quantify the usability of any system. Usability Geek. [Online]. Available: https://usabilitygeek.com/usability-metrics-a-guide-to-quantify-system-usability/.

3. Submit Revised Paper (13 Sept 2021)

[JESTEC] Manuscript ID: JESTEC_ 16_5_2021_Yusrila Y Kerlooza –Revised VersionReceived

1 message

Editorial Office <jjestec@gmail.com>
To: Yusrila Y Kerlooza <kerlooza@email.unikom.ac.id>

13 September 2021 at 10:24

Dear Author,

Thank you very much for resubmitting the modified version of the following manuscript:

Journal Name	Journal of Engineering, Science and Technology
Manuscript ID	JESTEC_ 16_5_2021_Yusrila Y Kerlooza
Authors	Yusrila Y. Kerlooza, Dian Dharmayanti, Adam Mukharil Bachtiar, Taufiq Nugraha
Title	Data Visualization for Education Domain at Dinas Pendidikan Jawa Barat

A member of the editorial office will be in touch with you soon regarding progress of the manuscript.

Kind regards,

Editorial Office
Journal of Engineering, Science and Technology
E-Mail: jjestec@gmail.com

*** This is an automatically generated email ***

DATA VISUALIZATION FOR EDUCATION DOMAIN AT DINAS PENDIDIKAN JAWA BARAT

YUSRILA Y. KERLOOZA*, DIAN DHARMAYANTI,
ADAM MUKHARIL BACHTIAR, TAUFIQ NUGRAHA

Informatics Engineering, Faculty of Engineering and Computer Science,
Universitas Komputer Indonesia, Jalan Dipatiukur Nomor 112-116, Bandung,
Jawa Barat, Indonesia
*Corresponding Author: kerlooza@email.unikom.ac.id

Abstract

The purpose of this research is to make a proper visualization which can be used in West Java Provincial Education Board to look education equity level in the West Java Province. This research is to help the Planning and Reporting Sub Division understand the case of education equity. Based on the results of interviews to one of the employees in the sub division, there were problems such as receiving information from the visualization used today were not as expected. It happened because the information does not support education equity. There are several methods used to produce the right form of information including statistics, data mining, and the application of data visualization rules. After designing and refining the existing visualization, the outcome is implemented in website-based prototypes. The results of prototypes were efficient and effective, it was tested by Usability Testing. The result of usability testing reached 95.45%, which means that the design used has achived a minimum goal. From these results, prototypes of data visualization can help the sub division to understand information about education equity.

Keywords: Data Visualization, Education Equity, Statistics, Data Mining, Usability Testing, Acceptance Testing.

1. Introduction

Equal distribution of education monitoring must be carried out to create equitable education throughout West Java Province. Monitoring is carried out to produce a focus on the education plan for the following year. It is because there are a lot of educational equity indicators, the obtained information might be wrong and need a long time to obtain the true meaning of the information [1]. Misinformation can result in non-conformity of program plans in the following year. Visualization can also be useful to improve knowledge effectively and optimally [2]. The results of visualization will be efficient and effective in accordance with the information requirements requested by the sub division.

The results of interviews and observations discovered a problem with the presentation of information related to the current distribution of education. Presentations in the form of statistical graphs have not been able to explain information on equal distribution per district and sub-district area in detail. Besides, the information obtained was not in accordance with the wishes of policy makers in the Education Board. The large number of indicators in the distribution of education caused errors in concluding an information and it requires a long time to get one information[3].

Data visualization can run well if the method of delivering information from visualization to the recipient of information are in accordance with the answers to be submitted [4]. Visualization purpose is that users can easily understand and interpret a large and complex collection of information [5]. The visualization that is built is used to help the Planning and Reporting Subdivision in monitoring various indicators of educational equity.

2.   Research Methodology

Methodology used was inspired by eight stages of visualization in the book "Visualizing Data" by Ben Fry. These stages are adjusted to the research as produce stages that shown in Fig. 1 [4].
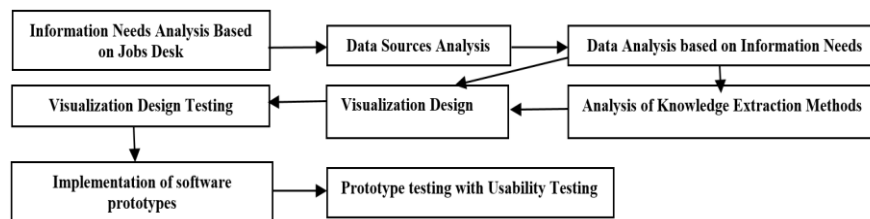
Fig. 1. Research methodology

All stages in Figure 1 were done in several stages. The stage of designing visualization consists of six   stages in more detail, namely understanding the context, determining the appropriate form of presentation, eliminating clutter, giving attention to focus on the user, then improving visualization by thinking as a designer and finally telling the visualization results [4]. The last stage of testing was using the usability testing method. This method is evaluative where the user will be given a task so that they can find out the success level in completing a given task.

3. Results and Discussion

3.1. Information Needs Analysis Base on Jobs Desk

At this stage, jobs desk analysis of Planning and Reporting Sub Division of the West Java Provincial Education Office was done. The sub division has three major jobs such as collecting data, planning, and evaluating the data. However, not all jobs require the results of data visualization but only two of the three jobs that require visualization.

3.2. Data Sources Analysis

The data used in this research came from the West Java Provincial Education Office and Statistics Indonesia. The data were obtained from the Education Office is the output of the system in Microsoft Excel documents which consist of various attributes. Data obtained from the Statistics Data is data taken from several Microsoft Excel documents. The following data used in this study is shown in Table 1.

Table 1. Used Tables

Data	Table
Population data	The population of the school people table
	Development table
	Population table based on age
	Table of government performance indicators
Education data	Education table

3.3. Data Analysis based on Information Needs

The data mapping stage used attributes related to the problem to be resolved and the existing strategic information needs. These attributes are formed by a collection because of the values number used [5]. The results of data mapping is shown as follows:

1. To obtain information on the number and percentage of increase or decrease in students repeating in the sub-district area, it is ideal to use pd_tkt_mengulang.
2. To find out the increase or decrease students number and percentage in dropout students in the sub-district area is using pd_tkt_putus_sekolah.
3. To get information on the increase or decrease students number and percentage of students who did not graduate from school in the sub-district area is using pd_tkt_tdk_lulus.
4. To obtain information on the ratio of teachers and students in the sub-district area is using pd_rasio_guru_siswa.
5. To obtain information on the increase or decrease percentage in teacher worth teaching in the sub-district area is using guru_lyk.
6. To obtain information on the increase or decrease percentage of certification teachers in the sub-district area is using a certification teacher.

7. To find out the teacher quality development from time to time and to see the lowest developments based on sub-district and district areas are using the development of teachers.
8. To find out the total quality of damaged infrastructure in that year per sub-district area is using means_rusak.
9. To find out the total quality of infrastructure that was damaged in a year per sub-district area is using means_rusak.
10. To find out about schools without spaces such as libraries, schools, laboratories, labkom in one school in the district is using school facilities.

The mapping above is part of all the strategic information needs in this study to be mapped against the data used. The data is mapped based on data sources related to the monitoring and dissemination of education in West Java.

3.4. Analysis of Knowledge Extraction Methods

This stage is a process where the selection of knowledge extraction methods based on information needs was done. However, not all information can be delivered directly. At this stage, the selection of methods for information needs was done through exploratory or explanatory analysis. Information needs number 3, 11, 14 use explanatory analysis method and the rest of the information needs was using exploratory knowledge extraction methods [6,7]. The calculation of information needs is shown as follows:

1. Information on the number and percentage of increase or decrease in students repeating.
   After the calculation of this year and the previous year, the results of these calculations were reduced to see the increase or decrease occurred. The results of the increase or decrease occurred in Baleendah Subdistrict are shown as follows:

$$Repeating\ figures\ this\ year - Figures\ Repeating\ Previous\ Years = 0{,}21 - 0{,}22\%$$
$$Repeating\ figures\ this\ year - Figures\ Repeating\ Previous\ Years = -0{,}01\%$$

It can be concluded that the Repeating Year is 0.01 better than the previous year. However, it is still not ideal because there are still students who have not graduated yet.

2. Information of total students and percentage of the number of dropout students are shown as follows.

$$Dropout\ rates\ this\ year = \frac{74}{10055} * 100\%$$
$$Dropout\ rates\ this\ year = 0{,}73\%$$
$$Previous\ Year\ dropout\ rate = \frac{92}{10932} * 100\%$$
$$Previous\ Year\ dropout\ rate = 0{,}84\%$$
$$Dropout\ rates\ this\ year - Previous\ Year\ dropout\ rate = -0{,}11\%$$

It can be concluded that the dropout rate for this year is 0.11 better than the previous year but still not ideal.

3. Information on the increase or decrease students number and percentage who did not graduate from school is shown as follows.

$$Graduate\ Numbers\ This\ year = \frac{3015}{3015} * 100\%$$
$$Graduate\ Numbers\ This\ year = 100\%$$
$$Number\ Not\ Passed\ = 0\%$$
$$Previous\ Year\ Graduates\ Figures = \frac{3532}{3532} * 100\%$$
$$Previous\ Year\ Graduates\ Figures = 100\%$$
$$Previous\ Year\ Graduates\ Figures = 0\%$$
$$Numbers\ Not\ Passing\ This\ Year\ - Numbers\ Not\ Passing\ Previous\ Years = 0\%$$

It can be concluded that the number of not graduating from school this year is the same as the previous year and is ideal.

4. Information on the ratio of teachers and students in the subdistrict area is shown as follows:

$$\frac{S}{G} = \frac{The\ number\ of\ students}{Total\ number\ of\ teachers} * 100\%$$
$$\frac{S}{G} = \frac{10055}{481} * 100\%$$
$$\frac{S}{G} = 20,90$$

Ratio of students:teachers = 20 : 1 ideal

5. Information on the percentage of increase or decrease of good teachers in the sub-district area is shown as follows.

$$\%GL = \frac{Number\ of\ authorized\ teachers}{Total\ number\ of\ teachers} * 100$$
$$\%GL = \frac{160 + 300}{161 + 320} * 100\%$$
$$\%GL = 96\%$$
$$\%GL\ Previous\ year = \frac{Number\ of\ authorized\ teachers}{Total\ number\ of\ teachers} * 100$$
$$\%GL\ Previous\ year = \frac{132 + 331}{143 + 340} * 100\%$$
$$\%GL\ Tahun\ Sebelumnya = 95\%$$

Therefore, we could see that only 96% of teachers are eligible. There is an increase of 1% from the previous year.

6. Information on the percentage of increase or decrease in certification teachers in the sub-district area.

$$GS = \frac{Number\ of\ certified\ teachers}{Total\ number\ of\ teachers} * 100\%$$
$$\%GS = \frac{132 + 287}{161 + 320} * 100\%$$
$$\%GS = 87,11\%$$
$$\%GS\ Previous\ year = \frac{Number\ of\ certified\ teachers}{Total\ number\ of\ teachers} * 100$$
$$\%GS\ Previous\ year = \frac{82 + 92}{143 + 340} * 100\%$$
$$\%GS\ Tahun\ Sebelumnya = 36\%$$

Therefore, it could be condluded that there is only 87.11% of teachers who received certification. This shows an increase of 51.11%.

7. Regional information affordable ideal education services in the sub-district area is shown as follows.

$$KPUS = \frac{Specific\ age\ population}{An\ area}$$
$$KSek = \frac{Certain\ level\ schools}{An\ area}$$
$$KPUS = \frac{4962}{41,933}$$
$$KPUS = 118,33$$
$$KSek = \frac{20}{41,933}$$
$$KSek = 0,476$$
$$DTSek = \frac{22}{7} * 10^2 * 0,476$$
$$DTSek = 149,89$$
$$DT = \frac{37189,34}{149,89}$$
$$DT = 248,1$$

The results of the summation were informed that there were only 248 children aged 16-18 years who could attend one secondary school. If 20 schools are available, 99.95% of children can enter the school.

8. To obtain information on the level of teacher needs per region

The method used here was grouping using the K-Means algorithm. Variables used include number of teachers, schools, and students. The desired group determination is 3. After several iterations have been obtained, the first cluster of centroid 1 has 15 members, centroid 2 has 11 members, and centroid 3 has 1 member. The implementation is done with the programming language R. From these results, it can be seen that the characteristics of the groups formed. Group 1 is for groups with less distribution of teacher distribution. Group 2 has evenly distributed teacher distribution. Group 3 has the level of distribution of teachers who have fulfilled and excess.

9. To see the spread of the area that is evenly distributed

Variables used are percentage of apk, APM, student / teacher ratio, and s / k ratio, student / school ratio. The desired group determination is 2. After several iterations are obtained, the results of the first cluster of centroid 1 has 20 members while centroid 2 has 7 members. The results can be seen that group 1 is an area that is not evenly distributed and group 2 is an evenly distributed area.

3.5. Design Visualization

At this stage, the visualization of the information that need to be conveyed was carried out. The design stage of transpiration visualization in storytelling books with data is the result from understanding the context of information needs, selecting visualization forms, eliminating existing clutter, focusing attention, and testing using acceptance testing techniques [8,9]. Some of the following designs are shown as follows:

a. Understanding context

The following table is the context understood in this information need is shown in Table 2.

Table 2. Understanding Context

WHAT	Knowing the number of repeating students has increased or decreased in order to determine one of the ideal indicators of student quality.

HOW	Showing the number of repeating students compared to the previous year to see a comparison the ideal or away from the ideal.

b. Selection of visualization

The following table is the selection of forms that have been adapted to the required context which shown in Table 3.

Table 3. Selection of Visualization

Visualization	Simple text and symbols
Reason	Can display one to two information needs simultaneously

c. Remove clutter

The following table is the results of clutter removal (See Table 4).

Table 4. Remove Clutter

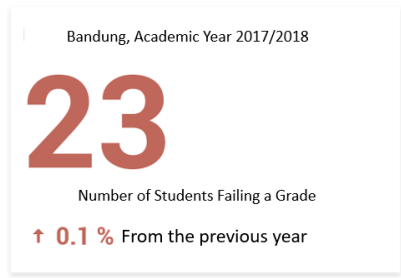| *Clutter*	**Reason**
Alignment	Left-aligned usage to facilitate reading from left to right
White Space	Use of white space to provide space for labels and content so that they do not overlap.
Closure	There is no use of borders in this form of visualization
Similarity	Colors used to indicate relationship

The following figure is the results of the existing clutter removal which is shown in Fig. 2.

Bandung, Academic Year 2017/2018

23

Number of Students Failing a Grade

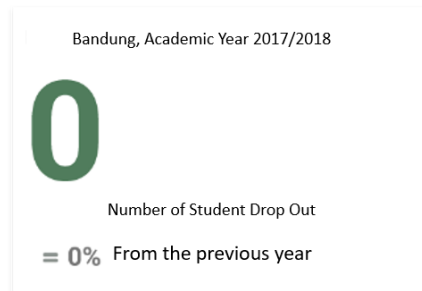↑ 0.1 % From the previous year
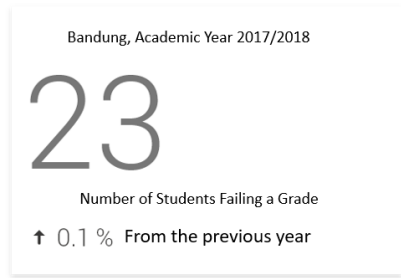
Fig. 2. Result Remove Clutter

d. Focus attention

Attention focusing was done by giving different font colors and sizes. The following figure is the results of attention focusing (See Figure 3).

Bandung, Academic Year 2017/2018

23

Number of Students Failing a Grade

↑ 0.1 % From the previous year

Fig. 3. Result Focusing Attention

Here are some visualization results. Fig. 4 shows the results of the need for information on the number of school dropouts who have been given different colors and font sizes.


Fig. 4. Result Visualization Dropout Students

After removing the clutter, Fig. 5 is an emphasis on the number of students failing to pass within a certain period of time.
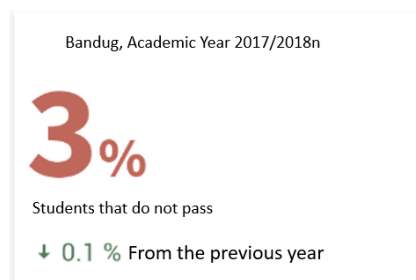
Fig. 5. Result Visualization Students Dont Graduate

The main focus in Fig. 6 is the percentage of the gross enrollment rate. This is indicated by the distinguished font size and color. The green color indicates the percentage of the target set by the government for that year.
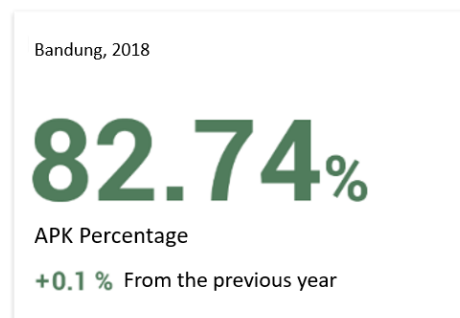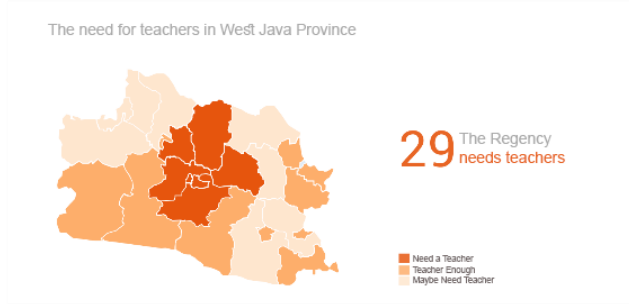

Fig. 6. Result Visualization Percentage APK

Figure 7 is the result of visualization that has emphasized the information to be conveyed.

Fig. 7. Result Visualization the area does not have a school

The information needs of teachers in a simple text visualization form and a choropleth map are shown in Fig. 8**.**

Fig. 8. Result Visualization Education Equity

The simple text visualization was chosen because it is suitable for displaying one or two pieces of information simultaneously. Choosing the form of choropleth map visualization for the division of regions was based on geographic location. Combining these two types of visualization can add value to information in an information need.

3.6. Visualization design testing

Tests conducted for participants who are experts in their fields [10,11]. The following table is the results of the acceptance test which is shown in Table 5.

Table 5. Visualization Design Testing

Information Needs	Visualization	Acceptance results
The number of students repeating experienced an increase or decrease from the previous year in the sub-district area per academic year is ideal	Simple text and symbols	accepted
The number of students dropping out of school experienced an increase or decrease from the previous year in the sub-district area per academic year is ideal	Simple text and symbols	accepted
The percentage of students graduating from school experienced an increase or decrease from the previous year in the sub-district area per academic year is ideal	Simple text and symbols	accepted

3.7. Implementation of software prototypes

At this stage, the software prototype was designed as well as supporting technology for prototype development [10]. The following figure is the dashboard display that shows information related to the quality of education at the high school level (See Fig. 9).
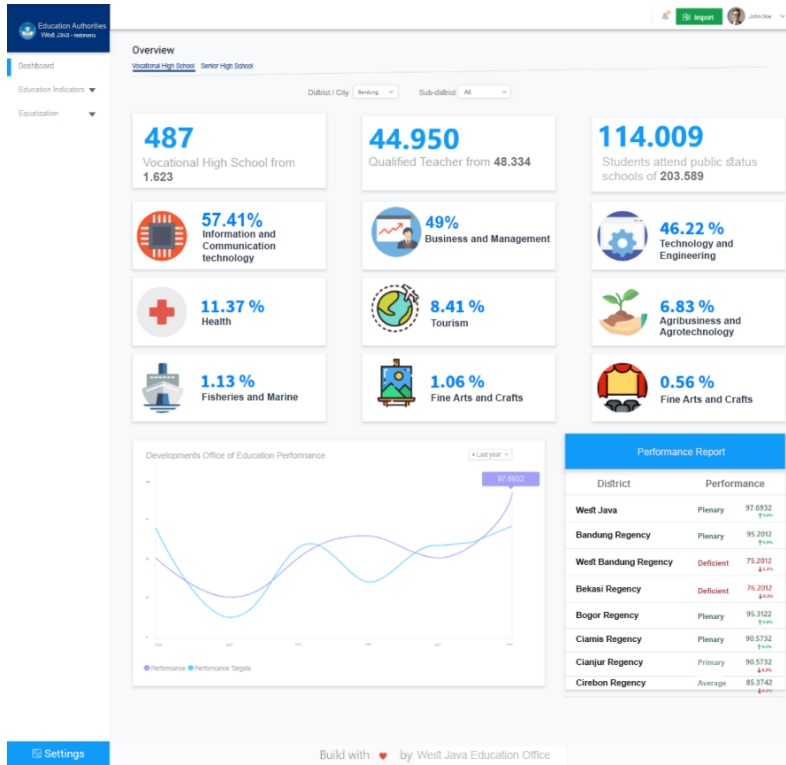
Fig. 9. Dashboard of the quality of education at the high school level.

Designing prototype software quality of facilities describes information related to the quality of educational facilities and infrastructure in West Java Province is shown in Fig. 10.
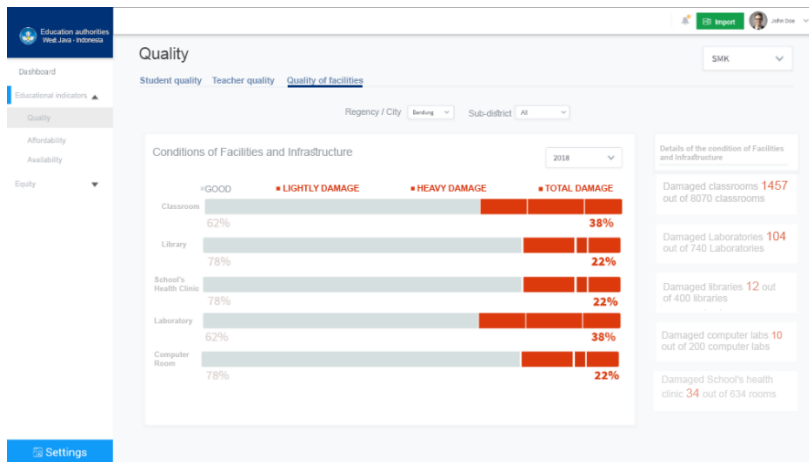
Fig. 10. **Designing prototype software quality of facilities**

Moreover, designing prototype software availability of educational services describes information related to the availability of educational services from schools in West Java (see Fig. 11).
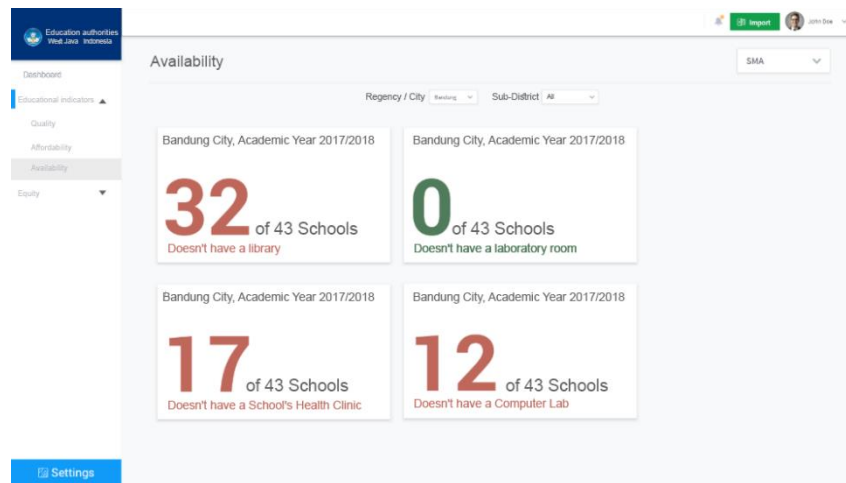

Fig. 11. Designing prototype software of availability of educational services

### 3.8. Testing software prototypes

Testing software prototypes aims at determining the design of the prototypes that have been made effectively and efficiently [12]. The following table is a test scenario for each menu that contain each visualization which shown in Table 5.

Table 6. Testing Prototypes

No	Task	Scenario	Expected results
1.	Enter the page that contains the visualization of the vocational students in West Java quality	I am an employee of the planning and reporting subdivision who wants to know information about	Successfully enter the page that contains visualization of the quality of vocational students in West Java

No	Task	Scenario	Expected results
		the quality of vocational students in West Java	through the education indicator menu.
2	Enter the page that contains the visualization of the quality of vocational teachers in the province of West Java	I am an employee of the planning and reporting subdivision who wants to know information about teachers of vocational students in the province of West Java.	Successfully enter the page containing visualization of the quality of vocational teachers in the province of West Java through the education indicator menu.

Prototype testing can determine the success rate of prototype design. The following figure is the results of mapping efficiency testing based on tasks that shown in Fig. 12.
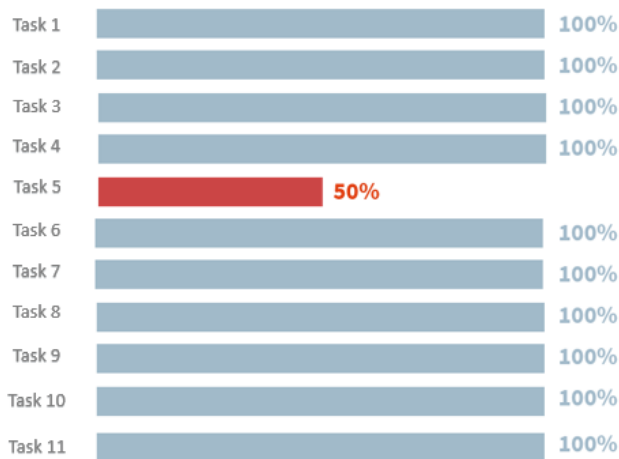
Fig. 12. Visualization Result Testing

From the test results, it can be concluded that the effectiveness and efficiency in the completion of the task is 95.45% which can be said as successful because achieving the minimum goal, which is 70%.

4. Conclusion

Based on the results of the implementation and testing, it can be concluded that visualization makes it easier to see and understand the data of each information displayed. Data visualization can solve problems that occur and can meet the objectives of this study. However, there are still obstacles faced in visualizing population data that can be accessed in a limited way because of using open data. Therefore, suggestions for the development of data visualization systems in the future are to collaborate with the statistical center to access population data in more detail.

References

1. Dinas Pendidikan Provinsi Provinsi Jawa Barat. (2017). Statistik pendidikan jawa barat. jawa barat, 2017. Retrieved February 15, 2021, from https://jabar.bps.go.id/subject/28/pendidikan.html.

2. Chen, M.; Ebert, D.; Hagen, H.; Laramee, R. S.; Van Liere, R.; Ma, K. L.; ...and Silver, D. (2008). Data, information, and knowledge in visualization. *IEEE Computer Graphics and Applications*, 29(1), 12-19.

3. Unterhalter, E. (2019). The many meanings of quality education: politics of targets and indicators in sdg 4. *Global Policy*, 10, 39-51.

4. Fry, B. (2008). *Visualizing data: exploring and explaining data with the processing environment*. Sebastopol, United States: O'Reilly Media, Inc.

5. Khan, M.; and Khan, S. S. (2011). Data and information visualization methods, and interactive mechanisms: A survey. *International Journal of Computer Applications*, 34(1), 1-14.

6. Verbert, K.; Duval, E.; Klerkx, J.; Govaerts, S.; and Santos, J. L. (2013). Learning analytics dashboard applications. *American Behavioral Scientist*, 57(10), 1500-1509.

7. Cimpanu, C.; and Ferariu, L. (2012). Survey of data clustering algorithms, *Buletinul Institutului Politeh. Din Lasi, Tome LVIII(LXII)*, 3(Lxii), 23–42.

8. Morissette, L.; and Chartier, S. (2013). The k-means clustering technique: general considerations and implementation in Mathematica. *Tutorials in Quantitative Methods for Psychology*, 9(1), 15-24.

9. Ellis, G.; and Dix, A. (2007). A taxonomy of clutter reduction for information visualisation. *IEEE Transactions on Visualization and Computer Graphics*, 13(6), 1216-1223.

10. Amar, R.; and Stasko, J. (2004). Best paper: a knowledge task-based framework for design and evaluation of information visualizations. *In IEEE Symposium on Information Visualization*. Austin, TX, USA, 143-150.

11. Diamond, M.; and Mattia, A. (2017). Data visualization: an exploratory study into the software tools used by businesses. *Journal of Instructional Pedagogies*, 18, 1-7

12. Mifsud, J. (2015). Usability metrics–a guide to quantify the usability of any system usability geek. Retrieved February 15, 2021, from https://usabilitygeek.com/usability-metrics-a-guide-to-quantify-system-usability/.

4. Accepted Submission (28 Sept 2021)

Gmail

Yusrila Y Kerlooza

<kerlooza@email.unikom.ac.id>

[JESTEC] Manuscript ID: JESTEC_ 16_5_2021_Yusrila Y Kerlooza –Accepted

1 message

Editorial Office <jjestec@gmail.com>   25 September 2021 at 11:05
To: Yusrila Y Kerlooza <kerlooza@email.unikom.ac.id>

Dear Author,

We are pleased to inform you that the following paper has been officially
accepted for publication:

Journal Name	Journal of Engineering, Science and Technology
Manuscript ID	JESTEC_ 16_5_2021_Yusrila Y Kerlooza
Authors	Yusrila Y. Kerlooza, Dian Dharmayanti, Adam Mukharil Bachtiar, Taufiq Nugraha
Title	Data Visualization for Education Domain at Dinas Pendidikan Jawa Barat

We will now make the final preparations for publication, then return the
manuscript to you for your approval.

If, however, extensive English edits are required to your manuscript, we will
need to return the paper requesting improvements throughout.

Kind regards,

Editorial Office
Journal of Engineering, Science and Technology
E-Mail: jjestec@gmail.com

5. Paper Publish (5 Oct 2023)

M Gmail

Yusrila Y Kerlooza

<kerlooza@email.unikom.ac.id**>**

[JESTEC] Manuscript ID: JESTEC_ 16_5_2021_Yusrila Y Kerlooza –Published Online

1 message

Editorial Office <jjestec@gmail.com>   5 October 2021 at 11:19
To: Yusrila Y Kerlooza <kerlooza@email.unikom.ac.id>

Dear Author,

We are pleased to inform you that your article:

Journal Name	Journal of Engineering, Science and Technology
Manuscript ID	JESTEC_ 16_5_2021_Yusrila Y Kerlooza
Authors	Yusrila Y. Kerlooza, Dian Dharmayanti, Adam Mukharil Bachtiar, Taufiq Nugraha
Title	Data Visualization for Education Domain at Dinas Pendidikan Jawa Barat

is available online:

https://jestec.taylors.edu.my/Vol%2016%20Issue%205%20October%202021/16_5_28.pdf

Please take a moment to check that everything is correct. You can reply to
this email if there is a problem. Note that at this stage we will not accept
further changes to the manuscript text.

Thank you for choosing our journal to publish your work, we look forward
to receiving further contributions from your research group in the future.

Kind regards,

Editorial Office
Journal of Engineering, Science and Technology
E-Mail: jjestec@gmail.com