Web-Based Application of Descriptive Statistics for Analysis of Concentration Measures in the Form Data of Frequency Distributions

I Gusti Agung Sadnyana Putra[1*], I Ketut Suja[2], and I Gusti Agung Ayu Uttami Vishnu Putri[3]

[1,2] Tourism Department, Politeknik Negeri Bali, Bali, Indonesia
[3] Information Systems Study Program, Stikom Institute of Technology and Business, Bali, Indonesia
*agungsadnyana@pnb.ac.id

Abstract. Statistical science consists of descriptive statistics and analytical statistics. The first part aims to carry out calculations on existing data, without making inferences or generalizations. Next, the second part is for interpretation and concluding the data. If we look at the processing process, data can be divided into ungrouped and grouped. Ungrouped is data that has not been arranged according to its groups, while grouped is data that has been classified based on a certain order, usually in the form of a frequency distribution table (FD). The processing of these two types of data statistically has very different characteristics. In this research, data processing was carried out for data of FD, especially regarding data centralization analysis. Statistical applications have been developed, but so far no statistical application has been found that processes FD. Apart from that, the application directly displays the final results without showing the calculations to obtain the final results. This research develops a descriptive statistical application for group data of FD and contains detailed steps for solving it. Waterfall is the method used in this research, consisting of requirements analysis, design, implementation, integration and testing, deployment, and maintenance. The results are in the form of concentration size analysis, distribution size analysis, and other analyses which will show in detail the steps for solving the problem by including the title, theory, formula, and detailed calculation steps to get the final result. The research results specifically presented in this article are analysis of centrality measures.

Keywords: Web Based Applications, Descriptive Statistics, Measures of Centralization, Frequency Distribution Data

1 Introduction

Statistical science consists of descriptive statistics and analytical statistics (Andjarwati et al., 2021; Ghozi & Sunindyo, 2015). The first part aims to carry out calculations on

A. A. N. G. Sapteka et al. (eds.), *Proceedings of the International Conference on Sustainable Green Tourism Applied Science - Social Applied Science 2024 (ICoSTAS-SAS 2024)*,
Advances in Economics, Business and Management Research 308,
https://doi.org/10.2991/978-94-6463-622-2_51

existing data, without making inferences or generalizations, and the second continues with analysis and conclusion (Syahfitri et al., 2023; Nengsih et al., 2022).

Almost all departments, both exact and non-exact, in various educational institutions, teach statistics courses (Nengsih et al., 2022; Rizki & Fauziddin, 2021). In non-exacts study programs, a common problem in studying statistics courses is that students are less interested in this course so learning outcomes in this course are generally relatively low (Rizki & Fauziddin, 2021; Dzakwan et al., 2021).

For ungrouped data, several statistical applications have been developed (Adinugroho & Wahyono, 2022; Nur et al., 2009; Shafa et al., 2024), but so far no statistical application has been found that processes grouped data (Saputra, 2013; Hakim & Kumadji, 1997). Apart from that, the application directly displays the final results without showing the calculations to obtain the final results. In terms of learning, this is not good because students do not receive a complete presentation on how to complete statistical analysis.

Based on this, we created a Descriptive Statistics Application which displays detailed problem-solving steps according to theory. The presentation of this course becomes more interesting and can increase learning outcomes.

2 Methodology

This research uses the waterfall method or linear sequential (Badrul, 2021; Arrifiyah, 2020; Windarti, 2013), traditional methods that follow a linear sequence of stages: Needs Analysis: At this stage, all system requirements are collected and documented. The result is a requirements specification; Design: Once requirements are gathered, a design phase is carried out to determine the system architecture, including hardware and software specifications to be used; Implementation: In this stage, the system design is converted into program code. Each system component that has been designed is implemented as a software unit; Integration and Testing: Once implementation is complete, all software units are integrated and tested to ensure that the system works according to predetermined specifications; Application: Once the system passes the testing phase, it is then deployed to a production environment where it will be used by end-users; Maintenance: At this stage, the system is updated and repaired if any errors are found.

3 Result and Discussion

The results of this programming are the implementation of the PHP Triad program package operation which consists of the PHP programming language, Apache Server, and MySQL database (Radinschi et al., 2008; Gangwar et al., 2014; Nixon, 2014; Grippa & Kuzmichev, 2021).

3.1 Starting to Use the Application

Use of the implementation begins by displaying the identity of this application which provides 3 (three) types of main menus available in Figure 1.

Figure 1. The initial appearance of the application (in Indonesia language)
Photograph and permission by I Gusti Agung Sadnyana Putra

3.2 Analysis of Data Centering Measures

To obtain data centralization analysis, you can select the first menu, namely "Centralization Size" which is available in Figure 1. After that Figure 2 will appear, which contains menus for the required data centralization analysis.

Figure 2. Main data center analysis form (in Indonesia language)
Photograph and permission by I Gusti Agung Sadnyana Putra

Each menu is explained as follows.

Entering New Statistical Data. The "Input Data" facility on the main menu can be used to enter new data to be analyzed as shown in Figure 3.

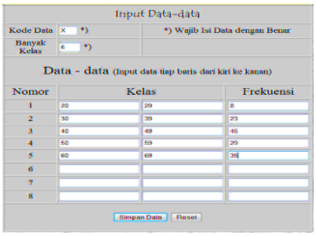


Figure 3. New statistical data input (in Indonesia language)
Photograph and permission by I Gusti Agung Sadnyana Putra

View Statistical Data. The "View Data" facility can be used to view previously input data as shown in Figure 4.

6	0	0	0
7	0	0	0
8	0	0	0
Kode Data	X		
Banyak Kelas	6		
Nomor	Kelas		Frekuensi
	Batas Bawah	Batas Atas	
1	20	29	8
2	30	39	23
3	40	49	45
4	50	59	29
5	60	69	35
6	0	0	0
7	0	0	0
8	0	0	0
Kode Data	Z		
Banyak Kelas	5		
Nomor	Kelas		Frekuensi
	Batas Bawah	Batas Atas	
1	60	62	5
2	63	65	18

Figure 4. View statistical data (in Indonesia language)
Photograph and permission by I Gusti Agung Sadnyana Putra

Using the Data Center Analysis Facility. The available facilities are arithmetic mean, median value, mode, 1-2-3 Quartiles, and Deciles, each of which can be used in data-centering analysis according to user needs.

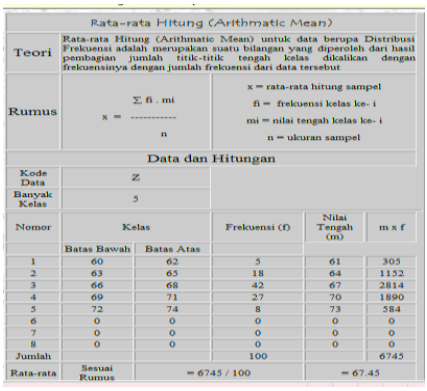
How to use. From the main data center menu, select one of the menus available in Figure 2. After selecting the existing data that will be analyzed as shown in Figure 5.

Daftar Data Tersimpan			
Kode Data	Z		Pilih
Banyak Kelas	5		
Nomor	Kelas		Frekuensi
	Batas Bawah	Batas Atas	
1	60	62	5
2	63	65	18
3	66	68	42
4	69	71	27
5	72	74	8
6	0	0	0
7	0	0	0
8	0	0	0
Kembali			

Figure 5. Selecting data for analysis (in Indonesia language)
Photograph and permission by I Gusti Agung Sadnyana Putra

In the next section, you can see the results of the analysis obtained, consisting of the following: Title: is a description of the type of analysis carried out; Theory: the theory that is the basis of analysis; Formula: approach formula for applying theory; Data and calculations: the main part of the analysis which contains calculations from applying formulas to obtain final results; "Back" button to return to the main centering size menu; The above steps can be repeated for other analyses.

Data Center Analysis Results. Arithmetic Mean: In essence, the arithmetic mean for data is in the form of a frequency distribution is the value obtained from dividing the number of class median values multiplied by the data frequency by the amount of data. The results obtained are shown in Figure 6.



Rata-rata Hitung (Arithmatic Mean)

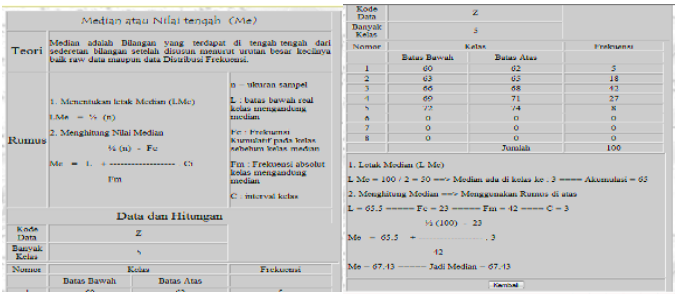
Teori	Rata-rata Hitung (Arithmatic Mean) untuk data berupa Distribusi Frekuensi adalah merupakan suatu bilangan yang diperoleh dari hasil pembagian jumlah titik-titik tengah kelas dikalikan dengan frekuensinya dengan jumlah frekuensi dari data tersebut	
Rumus	$x = \frac{\Sigma fi \cdot mi}{n}$	x = rata-rata hitung sampel fi = frekuensi kelas ke- i mi = nilai tengah kelas ke- i n = ukuran sampel

Data dan Hitungan

Kode Data	Z				
Banyak Kelas	5				
Nomor	Kelas		Frekuensi (f)	Nilai Tengah (m)	m x f
	Batas Bawah	Batas Atas			
1	60	62	5	61	305
2	63	65	18	64	1152
3	66	68	42	67	2814
4	69	71	27	70	1890
5	72	74	8	73	584
6	0	0	0	0	0
7	0	0	0	0	0
8	0	0	0	0	0
Jumlah			100		6745
Rata-rata	Sesuai Rumus		= 6745 / 100		= 67.45

Figure 6. Arithmetic mean analysis results (in Indonesia language)
Photograph and permission by I Gusti Agung Sadnyana Putra

Middle Value (Median). The basic definition of the median is the value in the middle of the frequency distribution table after sorting it according to class. Determine the location of the median by dividing the amount of data by 2, then use the approximate formula as shown in the results. The analysis results obtained are shown in Figure 7.



Median atau Nilai tengah (Me)

Teori	Median adalah Bilangan yang terdapat di tengah tengah dari sederetan bilangan setelah diurutkan menurut urutan besar kecilnya baik raw data maupun data Distribusi Frekuensi.
Rumus	1. Menentukan letak Median (LMe) LMe = ½ (n) 2. Menghitung Nilai Median $Me = L + \frac{½(n) - Fc}{Fm} \cdot C$ n = ukuran sampel L : batas bawah real kelas mengandung median Fc : Frekuensi Kumulatif pada kelas sebelum kelas median Fm : Frekuensi absolut kelas mengandung median C : interval kelas

Data dan Hitungan

Kode Data	Z		
Banyak Kelas	5		
Nomor	Kelas		Frekuensi
	Batas Bawah	Batas Atas	
1	60	62	5
2	63	65	18
3	66	68	42
4	69	71	27
5	72	74	8
6	0	0	0
7	0	0	0
8	0	0	0
		Jumlah	100

1. Letak Median (L Me)
L Me = 100 / 2 = 50 ⟶ Median ada di kelas ke . 3 ——— Akumulasi = 65
2. Menghitung Median ⟶ Menggunakan Rumus di atas
L = 65.5 ——— Fc = 23 ——— Fm = 42 ——— C = 3

$Me = 65.5 + \frac{½(100) - 23}{42} \cdot 3$

Me = 67.43 ——— Jadi Median = 67.43

Kembali

Figure 7. Median analysis results (in Indonesia language)
Photograph and permission by I Gusti Agung Sadnyana Putra

Modus (Mode). The basic definition of mode is the value in the frequency distribution data classes that have the highest occurrence or frequency. The approximate formula is used as shown in the results. The results obtained are shown in Figure 8.

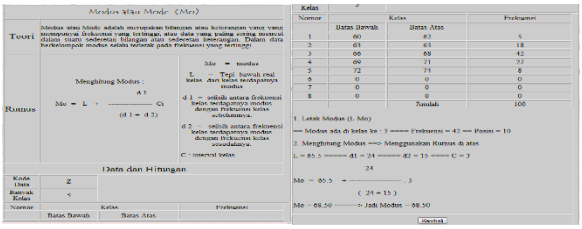


Figure 8. Mode analysis results (in Indonesia language)
Photograph and permission by I Gusti Agung Sadnyana Putra

Quartiles. The term quartile is defined as a value that is ranked 1st, 2nd, or 3rd in four (4) parts of the frequency distribution table of the same magnitude. Determine the position of the quartile by dividing the amount of data by the calculated quartile position. To get the results, an approximation formula is used as shown in Figure 9.

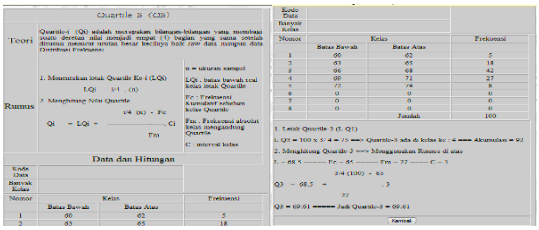


Figure 9. Results of quartile analysis (in Indonesia language)
Photograph and permission by I Gusti Agung Sadnyana Putra

Decile: Deciles divide the data into 10 equal parts, to obtain deciles 1 to 9. To determine the decile position is calculated by dividing the amount of data by the position of the decile to be calculated. The approximate formula is used as shown in the results. The analysis results obtained are shown in Figure 10

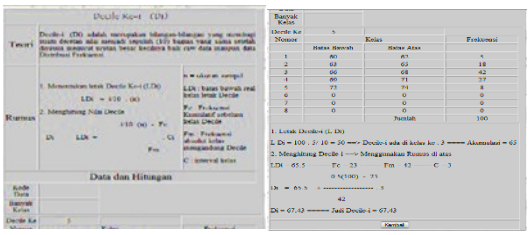


Figure 10. Results of decile analysis (in Indonesia language)
Photograph and permission by I Gusti Agung Sadnyana Putra

4 Conclusion

The descriptive statistics application provides 3 main facilities, namely data centralization measures, data dispersion measures, and additions. This paper only displays data centralization analysis. The specialty of this application which is not found in other statistical applications is that each analysis element contains a sequence of titles, theories, formulas, and data along with calculation steps to obtain the final results. Quartiles 1-2-3, arithmetic mean, mode, decile, and median are the facilities available in centering analysis.

References

Adinugroho, H. I., & Wahyono, T. (2022). Development of mobile-based statistics learning applications for students to support distance learning. *IT-Explore: Journal of Application of Information and Communication Technology, 1*(1), 49-61.

Andjarwati, T., Budiarti, E., Susilo, K. E., Yasin, M., & Soemadijo, P. S. (2021). *Descriptive Statistics*. Zifatama Jawara.

Arrifiyah, R. (2020). *Translating an E-Book Entitled "Software Engineering a Practitioners Approach Seventh Edition" by Roger S. Pressman a Report of Final Project*. Doctoral Dissertation, Jember State Polytechnic.

Badrul, M. (2021). Application of the waterfall method for designing an inventory information system at the Bintang Terang Ceramics Store. *PROSISKO: Journal of Research Development and Computer Systems Observation, 8*(2), 57-52.

Dzakwan, N., Murtinugraha, R. E., & Arthur, R. (2021). Effectiveness of using e-modules in statistics courses in the building Engineering Education Study Program, Faculty of Engineering, Jakarta State University. *Ricenology, 6*(1b), 70-77.

Gangwar, P., Kumar, S., & Rastogi, N. (2014). Web Solution using More Secure Apache HTTP Server with the Concept of Full Virtualization. *International Journal of Computer Applications, 98*(22).

Ghozi, S., & Sunindyo, A. (2015). *Descriptive Statistics For Economics*. Deepublish.

Grippa, V. M., & Kuzmichev, S. (2021). *Learning MySQL*. O'Reilly Media, Inc.

Hakim, A., & Kumadji, S. (1997). *Introduction to statistics. Surabaya*. Media Image.

Nengsih, T. A., Arisha, B., & Safitri, Y. (2022). *Descriptive Statistics Using the R Program*. Sonpedia Publishing Indonesia.

Nixon, R. (2014). *Learning PHP, MySQL, JavaScript, CSS & HTML5: A step-by-step guide to creating dynamic websites*. O'Reilly Media, Inc.

Nur, S. M., Yusandi, S., Indonesia, P. C., & Kav, J. J. S. (2009). *Aplikasi Program Minitab 15: Statistics for Experiment Design*. Calprint Indonesia.

Radinschi, I., Scripcariu, L., Ciobanu, B., & Frunza, M. (2008). Online quizzes, an application of PHP Triad and Mysql. *Romanian Journal of Physics, 53*(1-2), 423-427.

Rizki, L. M., & Fauziddin, M. (2021). Case study of students who have difficulty studying statistics courses. *Tambusai Education Journal, 5*(3), 11304-11314.

Saputra, R. (2013). *Applied Statistics*. West Sumatra Pioneer Stikes, Padang.

Shafa, N., Tamba, J., Auliyah, P. S., Lalistya, S. A., & Syahputra, S. A. (2024). Development of learning media on statistical materials. *Merdeka: Multidisciplinary Scientific Journal, 1*(3), 72-78.

Syahfitri, N., Rahmadani, N., & Nasution, N. S. (2023). Use of descriptive statistics methods in analyzing election results. *Journal of Social Service, 2*(1), 65–76. https://jurnal.asrypersadaquality.com/index.php/baktisosial/article/view/ 109.

Windarti, T. (2013). Production optimization modeling to maximize profits using linear programming methods. *Spectrum Industry, 11*(2), 148.