



# Shopping Behavior At Certain Times And Weeks Using Data Mining Methods Analysis Of Consumer

Luky Fabrianto<sup>1</sup>, Novianti Madhona Faizah<sup>2</sup>, Widyat Nurcahyo<sup>3</sup>, Sakius Tarigan<sup>4</sup>

<sup>1</sup>Program Studi Bisnis Digital, Universitas Nusa Mandiri

<sup>2,3,4</sup>Program Studi Ilmu Komputer, Universitas Tama Jagakarsa

E-mail: [luky.lfb@nusamandiri.ac.id](mailto:luky.lfb@nusamandiri.ac.id)

## ARTICLE INFO

Article history:  
Received: May 30, 2022  
Revised : Jun 15, 2022  
Accepted: Aug 01, 2022

**Keywords:**  
Datamining, Association, Rules, Support, Confidence.

## ABSTRACT

Many interesting things could be obtained from sales data, sales data can be viewed from various angles, such as sales side, shopping habits of customers: general and members, distributors of goods, shopping time, transactions by type of goods, association of goods and market basket analysis This study analyze buyer habits such as: shopping time, shopping week, type of goods that sells best at a certain time and finds out pattern of association rules for goods sold. The research stages start from collecting raw data, adding new attributes for the purpose of pre-processing to support data mining methods implementation, data transformation, running data mining algorithms and conclusions. The best result obtained is in the afternoon of the first week, if a consumer buys Beverage, probability that consumer buying a snack is 85% (Beverage => Snack, confidence: 85%), the minimum support value in this study is 50%.

Copyright © 2022 Jurnal Mantik.  
All rights reserved.

## 1. Introduction

Analysis of the sales data of a trading business seems endless to discuss, the discussion can be viewed from various perspectives such as from the sales side [1], shopping habits of customers: general and members, distributors of goods, shopping time, transactions by type of goods, transaction value to the arrangement of display racks, as well as profit forecasting [2].

Sales data generally has attributes such as date, transaction number, time, product code, product description, quantity, price and so on. From the existing sales data, additional attributes can be made such as clustering, shopping time, shopping week and so on, additional attributes are made to support the application of data mining methods [3].

The data mining method that most often used in research is market basket analysis using the association rule method with aim that buyers feel comfortable when reaching items to be purchased, also can use to make offers for related goods or arrange shopping shelves [4]. This study analyzes buyer habits such as: shopping time, shopping week, type of goods that sells best at a certain time and finds out pattern of association rules for goods sold.

The object of this research is sales data belonging to minimarkets around Sunter, Central Jakarta. There are 9,886 transaction receipts consist of 88,991 transactions data in 2021, in other words there are about 240 sales data lines or around 27 transactions that occur every day, the purpose of this study is to find out pattern of association rules that are formed specifically [5] based on what time and week used by customers to shopping. Although sales data of each market has different characteristics depending on several things such as location, type of customer, market segmentation etc, the larger data analyzed are better the analysis results will be, so that the insights obtained are better and more accurate. for the progress of the business.

Conventional stores should turn themselves into modern stores by utilizing information technology [6]. In his research, which uses the association rule method and is carried out per event or certain period related to Muslim holidays, the highest rule is Snacks => Groceries with 46% confidence and 16% support which occurs in the month of Ramadan [7]. The paper [8] discusses association rules for data mining extraction knowledge



from databases and new supermarket store layouts based on associations between categories. This approach allows managers to group products around related buying opportunities. Data mining techniques, association rule mining and provide new algorithms that can help to check customer habits to increase sales [9]. Extensive new data and information resources make scientists, engineers and business people need efficient analytical techniques to extract useful knowledge and effectively uncover valuable new knowledge patterns [10].

## 2. Method

The research stage begins with collecting sales data for one year, 2021. The raw data obtained has attributes as shown in Table 1. below.

TABEL 1  
SALES DATA ATTRIBUTES

Attribute	Description
Date	Transaction date
No_trans	Transaction number
Anggota	Membership type
Code	Item code
Keterangan	Item detail description
Kuantiti	Number of item items
Harga	Item unit price
Total	Total item price
Dist	Goods distributor

Furthermore, after raw data is obtained we add several attributes, namely: Time, Week and Cluster. These new attributes are added for purpose of pre-processing to support data mining methods implementation. Table 2. Represents additional attributes and descriptions.

TABEL 2  
ADDITIONAL ATTRIBUTES

Attribute	Description
Time	Taken from shopping hours on a transaction
Week	Taken from the date of a transaction
Cluster	Item Type Grouping

Cluster: The grouping of types goods follows the pattern of placement goods [11]

Waktu: Time division based on transaction hours (8<sup>am</sup> – 10<sup>am</sup> = Morning, 10:01<sup>am</sup> – 2<sup>pm</sup> = Midday, 2:01<sup>pm</sup> – 6<sup>pm</sup> = Afternoon, 6:01<sup>pm</sup> – 10<sup>pm</sup> = Night) [12]

Pekan: Week Division based on date (1 – 8 = First, 9 – 16 = Second, 17 – 24 = Third, 24 – end = Fourth)

The grouping of types goods (Cluster) is labeling of goods according to their categorization, there are eleven categories, as follows [11]:

1. Dry food is dry food in packaging that has a longer shelf life and usually requires further handling for consumption, for example: instant noodles, sardines, fried onions and so on.
2. Frozen is frozen food that is stored in the freezer and usually requires further handling for consumption such as the first category, for example: frozen chicken, ice cream, yogurt and so on.
3. Snacks are foods that are not considered as the main menu, for example: biscuits, wafers, sweets and so on.
4. AMDK (bottled drinking water) contains clear water, sold in several different sizes, for example: cup or glass, bottles of 600 ml, 330 ml, 1 liter and other sizes.
5. Beverage are liquids that can be directly consumed and packaged in certain sizes and flavors, for example: tea, coffee, milk, juice and so on.
6. Sachet Beverage are materials with advanced handling that can be consumed orally, generally packaged in sachets or pouches, such as: coffee, tea, milk, ginger and so on.

7. Baby Needs are a combination of all baby needs, such as: soap, shampoo, powder, telon oil, food and special milk for babies.
8. Non-consumption is a combination of personal care and oral hygiene such as soap, shampoo, toothpaste and others. As well as house hold, house ware and paper good such as plastic bags, carbolic acid, camphor, tissue, brush, detergent and so on.
9. Fruit that is sold in ordinary shelves (not refrigerator or freezer)
10. Groceries are 9 types of basic needs of the community according to the Decree of the Minister of Industry and Trade No. 115/MPP/Kep/2/1998 dated February 27, 1998, for example: cooking oil, rice, eggs, sugar and so on.
11. Medicine is a substance used to prevent, reduce, eliminate, cure disease or symptoms of disease.

After obtaining new attributes, the most frequently used data mining method is implemented to determine shopping patterns of customers, in this case the shopping basket analysis using A Priori algorithm, which is run with WEKA software version 3.8.4 and Microsoft Excel as pre-processing tools.

### 2.1 Association Rules

Association Rules mostly referred to as Market Basket Analysis (MBA) [13]. Association rule mining is a technique to find associative rules in the form of a combination of hidden items in one transaction contained in a large dataset, for example in a transaction is to find out how likely it is that one type of item is purchased together with another type of item [9], Table 3. is an example of a list of transactions.

TABEL 3  
EXAMPLE OF ASSOCIATION RULES

No Trans	Item
1	Snack, Oral Hygiene, House Hold
2	Personal Care, Oral Hygiene, Snack
3	Snack, Personal Care

Interesting relationships can be represented in the form of the following associative rules:

Personal Care → Snack

The basic method of association analysis is divided into two stages:

#### a. High Frequency Pattern Analysis

This stage is looking for a combination of goods that match the minimum requirements of support value in sales data. Support value of an item is obtained by following formula:

$$Support(A) = \frac{Transaction\ Amount\ contains\ A}{Total\ Transaction} \tag{1}$$

Support value of the 2 items is obtained from following formula:

$$Support(A \cap B) = \frac{Transaction\ Amount\ contains\ A\ and\ B}{Total\ Transaction} \tag{2}$$

#### b. Formation of Association Rules

After all high-frequency patterns are found, then look for the associative rule that match minimum requirements for confidence by calculating the confidence of the associative rule  $A \Rightarrow B$ . The confidence value of the rule  $A \Rightarrow B$  is obtained from the following formula:

$$Confidence = P(A / B) = \frac{Transaction\ Amount\ contains\ A\ and\ B}{Transaction\ Amount\ contains\ A} \tag{3}$$

## 3. Result and Discussion

### 3.1 Attribute Selection



In this study, not all attributes are used, to find customer shopping patterns using data mining method, only a few attributes are needed as shown in table 4. below.

TABEL 4  
ATTRIBUTES USED

Attribute	Description
No_trans	Transaction number
Kuantiti	Number of items purchased
Waktu	Shopping Time
Pekan	Shopping Week
Cluster	Item Type Grouping

### 3.2 Data Exploration

Table 5. is a statistical description of sales data based on number of clusters, statistical descriptions are useful for obtaining a real form of data to make it easier to understand for other people who are interested in the results of the research conducted [14].

TABEL 5  
DESCRIPTION OF GOODS CLUSTER STATISTICS

Cluster	Mean	Median	Std Dev	Min	Max	Sum	Count
Beverage	1.771	1	2.088	1	101	23,637	13,345
Snack	2.044	1	2.127	1	50	25,479	12,465
Groceries	86.545	2	283.818	1	4,031	913,240	10,607
Non-consumption	27.541	1	420.288	1	9,999	267,143	10,193
Frozen	1.836	1	1.604	1	45	6,920	3,770
Dry Food	3.913	2	10.785	1	400	13,090	3,345
Sachet Beverage	1.984	1	3.186	1	70	3,983	2,008
Medicine	1.532	1	1.287	1	16	2,188	1,428
Baby Needs	1.543	1	1.555	1	19	872	565
AMDK	14.107	2	32.117	1	384	6,701	475
Fruit	1.275	1	0.694	1	5	463	363

Sales data in a year consists of 9,886 transaction receipts, 88,991 transaction data and 1,281,529 items sold based on shopping receipts, Table 6. and Figure 1. is a visualization of the goods sold by cluster.

TABEL 6  
SALES RECAPITULATION BASED ON CLUSTER

Cluster	Number of Item Sold	Percentage
AMDK	6,701	0.53%
Sachet Beverage	3,983	0.32%
Fruit	463	0.04%
Dry Food	13,090	1.04%
Frozen	6,920	0.55%
Baby Needs	872	0.07%
Snack	25,479	2.02%
Beverage	23,637	1.87%
Non-consumption	267,143	21.14%
Medicine	2,188	0.17%
Groceries	913,240	72.27%

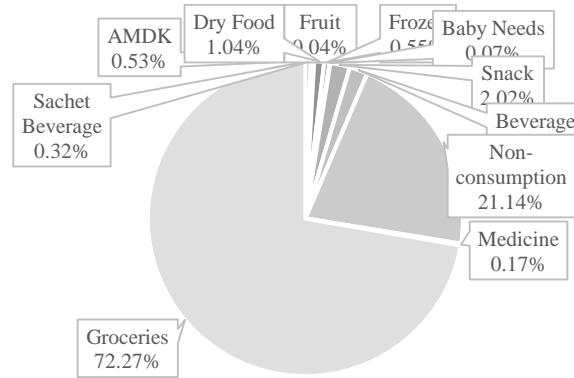


Figure 1. Cluster Distribution

Apart from cluster distribution, this study also analyzes customer spending based on shopping times and shopping weeks that are usually done by customers, some tables and figures below visualize the condition of sales data based on shopping times and weeks.

TABEL 7 SALES BASED ON SHOPPING TIME

Time	Percentage	Number of Transaction Receipts	Total (Rp)
Morning	13%	1,319	208,376,141
Midday	25%	2,508	385,025,203
Afternoon	35%	3,488	510,119,790
Night	26%	2,571	461,554,349
Total	100%	9,886	1,565,075,483

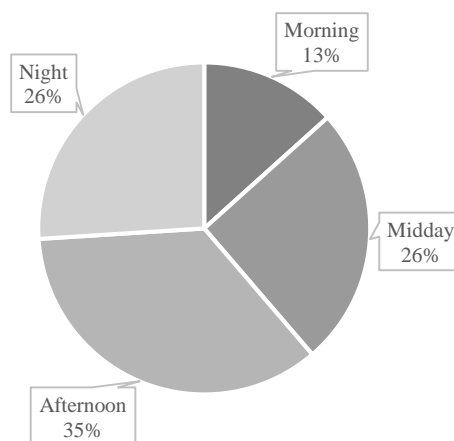


Figure 2. Distribution Based on Shopping Time

TABEL 8 SALES BASED ON SHOPPING WEEK

Week	Percentage	Number of Transaction Receipts	Total (Rp)
------	------------	--------------------------------	------------



First	23%	2,271	346,074,480
Second	23%	2,318	368,990,929
Third	26%	2,619	416,761,532
Fourth	27%	2,678	433,248,542
<b>Total</b>	<b>100%</b>	<b>9,886</b>	<b>1,565,075,483</b>

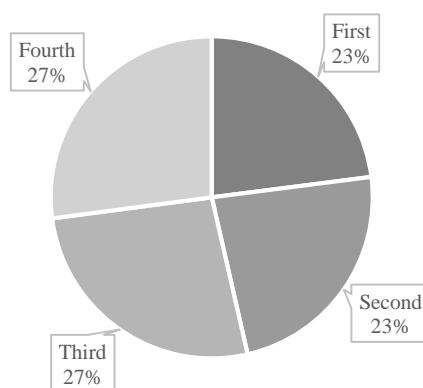


Figure 3. Distribution Based on Shopping Week

### 3.3 Data Transformation

In association rules or shopping cart analysis, attributes needed is sales data based on type of goods (count), for example in one transaction there are purchases of five pouches of cooking oil, so only one pouch is counted, because what will be seen is association or relation between goods purchased. Therefore, a data transformation or binarization process is needed, in other words the data type is converted into binary form, it can be 1 or 0 or YES or NO, in this case YES or 1 is considered to buy [15] as in the example in Table 9. and Table 10. an example of tabulation before and after the transformation is presented.

TABEL 9  
EXAMPLE BEFORE DATA TRANSFORMATION

No Transaction	Number of Items Sold							
	AMDK	Sachet Beverage	Dry Food	Frozen	Snack	Beverage	Non-Consumption	Groceries
1	0	0	0	0	5	7	0	0
2	0	0	0	0	0	0	0	0
3	0	0	0	3	0	0	0	0
4	0	0	0	0	7	7	3	9
5	0	4	0	0	0	0	3	8

TABEL 10  
EXAMPLE AFTER DATA TRANSFORMATION

No Transaction	Number of Items Sold							
	AMDK	Sachet Beverage	Dry Food	Frozen	Snack	Beverage	Non-Consumption	Groceries
1	0	0	0	0	1	1	0	0
2	0	0	0	0	0	0	0	0
3	0	0	0	1	0	0	0	0

4	0	0	0	0	1	1	1	1
5	0	1	0	0	0	0	1	1

To find the association rules, it only takes a count of goods sold, for example in Transaction No. 1 at Table. 10. sold: Snacks and Beverages, A Priori algorithm calculation is not affected by the number of goods sold as in first transaction in Table 9. after the binary transformation is carried out, composition of distribution of types of goods sold changes as shown in Table 11. and Figure 4. below.

TABEL 11  
CLUSTER DISTRIBUTION AFTER DATA TRANSFORMATION

Cluster	Number of Items Sold	Percentage
AMDK	442	1.10%
Sachet Beverage	1,915	4.76%
Fruit	291	0.72%
Dry Food	3,067	7.63%
Frozen	3,214	7.99%
Baby Needs	573	1.42%
Snack	7,793	19.38%
Beverage	7,235	17.99%
Non-consumption	7,027	17.47%
Medicine	1,375	3.42%
Groceries	7,284	18.11%

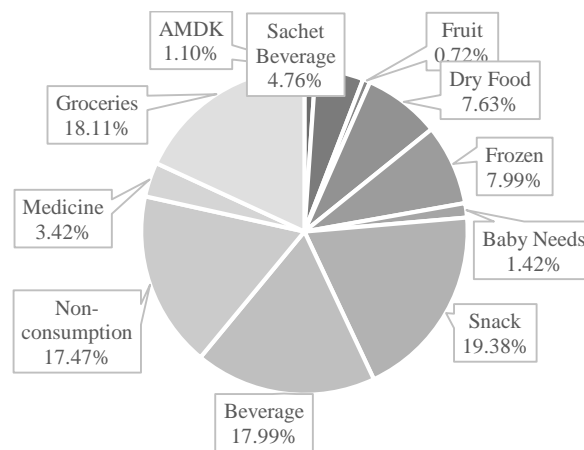


Figure 4. Cluster Distribution After Data Transformation

### 3.4 Association Rules

Based on total number of transactions and number of clusters, as shown in Table 11., threshold ( $\theta$ ) used in this study is 50%, which means that only rules that have minimum support value  $\geq 50\%$  are taken. The study used Weka software version 3.8.4, results obtained are summarized in Table 12. below.

TABEL 12 BEST RULE

		Best Rules	Confidence
All Trans	Beverage	==> Snack	81%
	Groceries	==> Snack	79%



		Non-consumption	==>	Snack	76%
		Non-consumption	==>	Groceries	74%
<b>First Week</b>		Beverage	==>	Snack	82%
		Groceries	==>	Snack	79%
		Non-consumption	==>	Snack	77%
		Non-consumption	==>	Groceries	75%
<b>Second Week</b>		Beverage	==>	Snack	80%
		Groceries	==>	Snack	78%
		Non-consumption	==>	Snack	75%
		Beverage	==>	Groceries	74%
<b>Third Week</b>		Beverage	==>	Snack	82%
		Groceries	==>	Snack	79%
		Non-consumption	==>	Snack	77%
		Groceries	==>	Beverage	75%
<b>Fourth Week</b>		Beverage	==>	Snack	81%
		Groceries	==>	Snack	79%
		Non-consumption	==>	Snack	77%
		Non-consumption	==>	Groceries	76%
<b>Morning</b>		Snack	==>	Groceries	78%
		Non-consumption	==>	Groceries	78%
		Beverage	==>	Groceries	78%
		Beverage	==>	Snack	77%
<b>Midday</b>		Beverage	==>	Snack	82%
		Groceries	==>	Snack	77%
		Groceries	==>	Beverage	76%
		Non-consumption	==>	Snack	75%
<b>Afternoon</b>		Beverage	==>	Snack	83%
		Groceries	==>	Snack	80%
		Non-consumption	==>	Snack	78%
		Non-consumption	==>	Groceries	74%
<b>Night</b>		Beverage	==>	Snack	81%
		Groceries	==>	Snack	80%
		Non-consumption	==>	Snack	77%
		Groceries	==>	Snack	75%

Table 12. above is the best rules obtained as whole transactions and based on consumer shopping week and consumer shopping time, obtained several items that are included in these best rules, such as: Beverage, Groceries, Non-consumption and Snack. The rule with the highest confidence is Beverage ==> Snack = 83% in the afternoon. Table 13. below is the best rule in form of a correlation metric between time and week of consumer shopping.

TABEL 13  
BEST RULE IN FORM CORRELATION METRIC

		Week							
		First	Conf.	Second	Conf.	Third	Conf.	Fourth	Conf.
Time	Morning	Beverage		Beverage		Non-consumption		Snack	
		↓	81%	↓	81%	↓	80%	↓	76%
		Groceries		Snack		Groceries		Groceries	
	Midday	Beverage		Beverage		Beverage		Beverage	
		↓	82%	↓	79%	↓	81%	↓	82%
		Snack		Snack		Snack		Snack	
	Afternoon	Beverage		Beverage		Beverage		Beverage	
		↓	85%	↓	81%	↓	84%	↓	82%
		Snack		Snack		Snack		Snack	
	Night	Beverage		Beverage		Beverage		Groceries	
		↓	81%	↓	78%	↓	82%	↓	82%
		Snack		Snack		Snack		Snack	

#### 4. Conclusions

Data mining using association rule method and A Priori algorithm can provide an overview of pattern of relationships between items sold in a trading business, this study describe association rules in more detail in consumer shopping time and consumer shopping week. The research results obtained can provide new insights that can be followed up by store managers, minimarkets, supermarkets, etc.

The sales data used in this study are minimarkets located in fairly densely populated settlements of households and individuals, items in beverage, groceries, non-consumption and snack clusters dominate sales transactions, but with same method and with different location and different sales data, it is more possible that variety number of clusters will frequently appear and will be more attractive.

#### References

- [1] V. L. Jenneson, F. Pontin, D. C. Greenwood, G. P. Clarke, and M. A. Morris, "A systematic review of supermarket automated electronic sales data for population dietary surveillance," *Nutr. Rev.*, vol. 80, no. 6, pp. 1711–1722, May 2022, doi: 10.1093/NUTRIT/NUAB089.
- [2] A. Mansur, F. Dwi, and J. Sofalina, "Sales Prediction Based on Sales Data Analysis Using the Moving Average Algorithm (Case in hawada refill drinking water)," *IJISTECH (International J. Inf. Syst. Technol.)*, vol. 6, no. 1, pp. 127–133, Jun. 2022, doi: 10.30645/IJISTECH.V6I1.220.
- [3] J. M. Chae and E. H. Kim, "Sales Pattern and Related Product Attributes of T-shirts," *J. Korean Soc. Cloth. Text.*, vol. 44, no. 6, pp. 1053–1069, Jul. 2020, doi: 10.5850/JKSCT.2020.44.6.1053.
- [4] F. Kurniawan, B. Umayah, J. Hammad, S. M. S. Nugroho, and M. Hariadi, "Market Basket Analysis to Identify Customer Behaviours by Way of Transaction Data," *Knowl. Eng. Data Sci.*, vol. 1, no. 1, p. 20, Dec. 2017, doi: 10.17977/UM018V1I12018P20-25.
- [5] P. Patel, B. Sivaiah, and R. Patel, "Relevance of Frequent Pattern (FP)-Growth-Based Association Rules on Liver Diseases," *Lect. Notes Networks Syst.*, vol. 431, pp. 665–676, 2022, doi: 10.1007/978-981-19-0901-6\_58/COVER.
- [6] N. Edward, R. Intan, and A. T. Infra, "Analisis Consumer Behaviour Pada Toko Retail Dengan Metode APRIORI-SD," *J. Infra*, vol. 7, no. 2, pp. 185–189, 2019, [Online]. Available: <http://publication.petra.ac.id/index.php/teknik-informatika/article/download/8829/7972>.
- [7] L. Fabrianto, N. M. Faizah, J. H. Prasetyo, B. S. Prakoso, and G. Wiharso, "Aturan Asosiasi Antar Item Terjual pada Data Penjualan Minimarket Milik Komunitas di Hari Besar Tertentu Menggunakan Algoritma Apriori," *J. AKUNTANSI, Ekon. dan Manaj. BISNIS*, vol. 9, no. 2, pp. 208–215, Dec. 2021,



- doi: 10.30871/JAEMB.V9I2.3621.
- [8] A. Madan Mirajkar, A. Prafulla Sankpal, P. S. Koli, R. A. Patil, A. R. Pradnyavant, and A. Professor, “Data Mining Based Store Layout Architecture for Supermarket IRJET Journal Data Mining Based Store Layout Architecture for Supermarket,” *Int. Res. J. Eng. Technol.*, 2016, Accessed: Jul. 24, 2022. [Online]. Available: [www.irjet.net](http://www.irjet.net).
- [9] M. Kaur and S. Kang, “Market Basket Analysis: Identify the Changing Trends of Market Data Using Association Rule Mining,” *Procedia Comput. Sci.*, vol. 85, pp. 78–85, 2016, doi: 10.1016/J.PROCS.2016.05.180.
- [10] R. R. Shelke HVPM COET, A. R. V Dharaskar Former Director, and V. M. Thakare, “Data Mining For Supermarket Sale Analysis Using Association Rule,” *Int. J. Trend Sci. Res. Dev.*, vol. 1, no. 4, pp. 2456–6470, Accessed: Jul. 24, 2022. [Online]. Available: [www.ijtsrd.com](http://www.ijtsrd.com).
- [11] A. St Sujana, “PARADIGMA BARU DALAM MANAJEMEN RETAIL MODERN,” Accessed: Jul. 24, 2022. [Online]. Available: [www.grahailmu.com](http://www.grahailmu.com).
- [12] “Beranda | Badan Pengembangan dan Pembinaan Bahasa - Kemendikbudristek.” <https://badanbahasa.kemdikbud.go.id/> (accessed Jul. 24, 2022).
- [13] K. N. Wijaya, “Analisa Pola Frekuensi Keranjang Belanja Dengan Algoritma Apriori (Studi kasus: Minimarket Adi),” *Annu. Res. Semin.*, vol. 3, no. 1, pp. 9–12, 2017, [Online]. Available: <http://seminar.ilkom.unsri.ac.id/index.php/ars/article/view/1707>.
- [14] H. A. Irianto, *Statistik Konsep Dasar; Aplikasi, dan Pengembangannya*. Kencana, 2016.
- [15] “Data Mining Menggunakan Android, Weka, dan SPSS - Indah Werdiningsih, S.Si., M.Kom., Barry Nugoba, S.Si., M.Kom., Muhammadun, S.Si., M.Si. - Google Books.” [https://books.google.co.id/books?id=xEwCEAAQBAJ&pg=PA182&lpg=PA182&dq=Data+Mining+Menggunakan+Android,+Weka,+dan+SPSS&source=bl&ots=z3T\\_p2D1jO&sig=ACfU3U3ttL8RwIh50mvxS1o3G4laXf8i0w&hl=en&sa=X&ved=2ahUKEwjzMvPwJD5AhVZT2wGHZ12AWsQ6AF6BAGEEAM#v=onepage&q=Data Mining Menggunakan Android%2C Weka%2C dan SPSS&f=false](https://books.google.co.id/books?id=xEwCEAAQBAJ&pg=PA182&lpg=PA182&dq=Data+Mining+Menggunakan+Android,+Weka,+dan+SPSS&source=bl&ots=z3T_p2D1jO&sig=ACfU3U3ttL8RwIh50mvxS1o3G4laXf8i0w&hl=en&sa=X&ved=2ahUKEwjzMvPwJD5AhVZT2wGHZ12AWsQ6AF6BAGEEAM#v=onepage&q=Data%20Mining%20Menggunakan%20Android%20Weka%20dan%20SPSS&f=false) (accessed Jul. 24, 2022).