Mathematics and Statistics 10(4): 747-753, 2022 http://www.hrpub.org
DOI: 10.13189/ms.2022.100405

Cluster Analysis on Various Cluster Validity Indexes with Average Linkage Method and Euclidean Distance (Study on Compliant Paying Behavior of Bank X Customers in Indonesia 2021)

Solimun Solimun, Adji Achmad Rinaldo Fernades*

Faculty of Mathematics and Natural Science, Brawijaya University, Malang, East Java, Indonesia

Received October 6, 2021; Revised January 18, 2022; Accepted April 15, 2022

Cite This Paper in the following Citation Styles

(a): *[1] Solimun Solimun, Adji Achmad Rinaldo Fernades , "Cluster Analysis on Various Cluster Validity Indexes with Average Linkage Method and Euclidean Distance (Study on Compliant Paying Behavior of Bank X Customers in Indonesia 2021)," Mathematics and Statistics, Vol. 10, No. 4, pp. 747 - 753, 2022. DOI: 10.13189/ms.2022.100405.*

(b): *Solimun Solimun, Adji Achmad Rinaldo Fernades (2022). Cluster Analysis on Various Cluster Validity Indexes with Average Linkage Method and Euclidean Distance (Study on Compliant Paying Behavior of Bank X Customers in Indonesia 2021). Mathematics and Statistics, 10(4), 747 - 753. DOI: 10.13189/ms.2022.100405.*

Abstract This study aims to examine the differences in various cluster validity indexes in the grouping of credit customers at Bank X Malang City, Indonesia using the average linkage and Euclidean distance methods. This study uses primary data with the variables used are service quality, environment, mode, willingness to pay, and obedient paying behavior obtained through a questionnaire with a Likert scale through purposive sampling distributed to 100 respondents. The data are then analyzed by clusters using the ward linkage and Euclidean distance methods on various validity cluster indexes, including the Silhouette Index, Krzanowski-Lai, Dunn, Gap, Davies-Bouldin, Index C, Global Sillhouette, Goodman-Kruskal in this study used as a tool analysis. This study uses R software. The results show that the Krzanowski-Lai, Dunn, Gap, Global Sillhouette, and Goodman-Kruskal indices have the same cluster members, as well as the Silhouette and Davies-Bouldin indices. The best cluster indexes are the Silhouette and Davies-Bouldin indexes. All validity indices produce variance between and within the same cluster. The novelty of this study is to compare 8 validity indices, namely Sillhouette Index, Krzanowski-Lai, Dunn, Gap, Davies-Bouldin, Index C, Global Sillhouette, and Goodman-Kruskal simultaneously.

Keywords Cluster Analysis, Cluster Validity, Euclidean Distance

1. Introduction

It is now impossible to isolate data from practically every aspect of life. This information can be used to assess risk, predict events, and make decisions. Statistics is a science that is well-known for its ability to handle data and alter information. It is used in a variety of fields. The banking industry is the only one that still requires data. Banks, according to Kashmir [1], are financial institutions that engage in activities such as collecting monies from the general public, paying these funds to the general public, and providing other banking services.

Credit is one of the services provided by the bank. Credit, according to Law No. 10 of 1998, is a payment or debt based on a contract or credit arrangement between a bank and a third party, in which the borrower is required to pay interest and repay the loan within a specified time frame. Before granting a loan to a debtor, the Bank must have a good track record of determining the debtor's ability to repay its debts. One issue with loans is that some customers' repayment behaviour is inconsistent, neutral, or even inappropriate. Clearly, this is a credit issue that has to be addressed. Cluster analysis is one statistical analysis that can be used to solve this challenge.

Cluster analysis is a technique for dividing a sample into discrete subgroups known as clusters. Subgroups within each cluster are frequently the same while also being highly distinct (non-identical) from elements in other groups. Many associations can be employed to build a cluster in cluster analysis. The linking method, according to Supranto [3], consists of single, complete, and average linkage. The average linkage approach is tested on seven different types of cluster validity indices in this study. In this study, the distance between each joint was measured using Euclidean distance. Different cluster identification findings and cluster accuracy scores will almost likely produce different outcomes. We wish to see how the cluster validity index affects the difficulty of grouping bank X clients in this study.

This study refers to research conducted by Petrofic [4], entitled "A Comparison between the Silhouette Index and the Davies-Bouldin Index in Labeling IDS Clusters" which aims to compare the silhouette index and Davies-Bouldin. This study gives the results that the Silhouette Index and Davies Bouldin give the same results but there are a few insignificant differences.

2. Literature Review

2.1. Score Interpretation Criteria

In each indicator or questionnaire, the variables are measured on the basis of design criteria developed on the basis of theory and research results. The purpose of identifying these research variables is to identify the frequency distribution of respondents' responses to the distributed questions in descriptive statistical analysis and to deeply identify the variables in the study. The frequency distribution was obtained from the table of respondents' responses. The interpretation of the scores is shown in table 1.

Table 1. Average Score Criteria

No.	Average Score Criteria	Criteria
1	1.00 – 1.5	Very Low/Very Weak
2	1.5 > - 2.5	Low/Weak
3	2.5 > - 3.5	Moderate
4	3.5 > - 4.5	High/Good
5	4.5 >	Very High/Very Good

Source: Solimun et al. [5]

2.2. Cluster Analysis

Cluster analysis is an analytical method for classifying items into groups in which some members of the group are homogenous (similar) and others are heterogeneous (different) [5].

The hierarchical and non-hierarchical methods of cluster analysis are used to partition the results. When information on the number of clusters is unavailable, hierarchical clustering is applied. The hierarchical approach's general premise is to group things that share a trait. When the number of stars is known or discovered, non-basic processing becomes necessary [6].

2.3. Hierarchical Method

Hierarchical Method is the process of grouping that begins with the combination of two or more entities that share a common characteristic. Then go on to the next closest object and repeat the process. As a result, a tree is created with a hierarchy or level ranging from the most similar to the most dissimilar [6]. More information on the implantation procedure or the popularity of the dendrogram can be found in this tree.

According to Johnson and Wichern [6], there are two approaches to group formation in the hierarchical method: agglomerative hierarchical techniques and divisive hierarchical methods. The Agglomerative Method starts with the notion that everything is interconnected. After that, combine the two adjacent objects. The procedure is repeated until a particular unit is formed. Hierarchical agglomerative is a popular approach. It's one of the most advanced collection aggregation algorithms. The average linkage method is one of the agglomerative hierarchy method algorithms employed in group building in this study.

2.4. Average Linkage Method

In the average linkage approach, the distance between two clusters is determined using the average distance between all members of one cluster and all other cluster members. The following is the formula for determining distance: (1).

$$_{(ij)k} = \frac{\sum_i \sum_j d_{(ij)}}{N_{ij} N_k} \quad (1)$$

Where:

$d_{(ij)k}$: the separation between the subsample (ij) and the cluster k

d_{ik}: the separation between sub-sample i and cluster k

d_{jk}: the separation between sub-sample j and cluster k

2.5. Distance in Cluster Analysis

The Euclidean distance, which is used to calculate the distance between two points, is the number of clusters used in this study.

$$d(x_i, x_j) = \sqrt{\sum_{z=1}^{p} (x_{ki} - x_{kj})^2} \quad (2)$$

where:

$d(x_i, x_j)$: the Euclidean separation between the i object and the j object
x_{ki}: the value of the i object in the variable k
x_{kj}: the value of the j object in the variable k
z: the variable to z, $z = 1,2,3, \dots, p$

2.6. Validity Index in Cluster Analysis

(a) Silhouette Index

The silhouette validity index is a statistical metric that can be used to address the challenge of establishing the optimal number of K clusters by providing a brief graphic depiction of how effectively each object is situated within the cluster. The silhouette index compares the average distance between objects in one cluster to the distance between objects in other clusters to determine the placement of each object in each cluster [7]. The term "silhouette" relates to an interpretive method and the validity of consistency in a group of data. This method generates a clear graphical depiction of how successfully each object was classified. By looking at the maximum average value of silhouette S, the best number of clusters may be determined (i). The optimal number of K clusters is a pricing estimate that maximizes the silhouette validity index S's average value (i).

Assume that the data has been clustered. Let an I be the average distance between objects in the same cluster, and b I be the average minimum distance between object I and all objects in a cluster that is not a cluster member for each object. The silhouette validity index can be written with the following equation based on the explanation provided:

$$S(i) = \frac{b(i) - a(i)}{Max\,\{a(i), b(i)\}} \quad (3)$$

Where:

$a(i)$: the average difference between I objects and the rest of the group
$b(i)$: the smallest mean difference between I objects and all other objects in other groups (in the closest group)

The silhouette validity index equation can be written as follows:

$$S_{(i)} = \begin{cases} 1 - \frac{a(i)}{b(i)}, if\ a(i) < b(i) \\ 0, \quad\quad if\ a(i) = b(i) \\ \frac{b(i)}{a(i)} - 1, if\ a(i) > b(i) \end{cases} \quad (4)$$

The average $S_{(i)}$ of all objects in a cluster shows how closely the objects in a cluster are similar, which also indicates how precisely the objects have been grouped. The closer $S_{(i)}$ to 1, the better the grouping of objects. Conversely, the closer $S_{(i)}$ it is to -1, the worse the grouping of objects will be. The optimal number of clusters is an estimate of the price that most maximizes the average value of $S_{(i)}$ and if there is one cluster whose members consist of one object, the average value $S_{(i)}$ is 0.

(b) *Krzanowski-Lai* Index

The optimal number of clusters is then defined as the k value that maximizes CH (k). The *Krzanowski-Lai* index is determined by:

$$KL(k) = \left| \frac{DIFF(k)}{DIFF(k+1)} \right|, \quad (5)$$

Where:

$$DIFF(k) = (k-1)^{\frac{2}{p}} W(k-1) - (k)^{\frac{2}{p}} W(k) \quad (6)$$

The number of features in the data collection is denoted by p. If KL (k) is maximized, the value of k will be optimal. The next section compares the performance of the three indices BB, CH, and KL using data from two different datasets: the fake dataset and the Iris dataset.

(c) *Dunn* Index

J. C. Dunn [8] proposed the Dunn index as a statistic for cluster validity. The smallest distance between observations in different clusters divided by the biggest distance in each data cluster is Dunn's index. The Dunn index can be calculated by dividing d_{min} by d_{max}.

Where:

$$C = \frac{d_{min}}{d_{max}} \quad (7)$$

$_{min}$= the smallest distance between observations in different clusters
$_{max}$= the largest distance in each data cluster

(d) *Gap* Index

A gap statistic can be used to estimate the ideal number of clusters [9]. Assume that X_{ij} is a j-th variable observation on the ith object. Then divide the data into k clusters, with $C_1, C_2, \dots, C_3$ 3 to C r representing the observations in the *r*-cluster and n_r representing the number of objects the *r*-cluster, as follows:

$$D_r = \sum_{(ik,jk)} \ _{ik} \quad (8)$$

where D_r is the total distance of all points in the cluster *r* and $_{ik}$ is the distance between the *i* object and the *k* object.

$$W_k = \sum_{r=1}^{k} \frac{1}{2_{n_r}} D_r \quad (9)$$

where W_k is the sum of the squares combined in the cluster.

(e) *Davies-Bouldin* Index

The Davied-Bouldin index formula can be written as:

$$DB = \frac{1}{n}\sum_{i=1}^{n} \max_{i \neq j} \left[\frac{d`(c_i)+d`(c_j)}{d(c_i,c_j)}\right] \quad (10)$$

where

n= number of groups

d(ci,cj)= the distance between groups ci and cj

d'(ck)= the distance in groups c_k

The small Davies-Bouldin index value indicates a good group.

(f) C-Index

This index can be explained as follows:

$$C=\frac{S-S_{min}}{Smin_{max}} \quad (11)$$

where:

S: the number of distances in all pairs of observed objects from the same group, with ℓ is the number of pairs.

$S_{min:}$ the sum of the ℓ smallest distances if all sample pairs are in different groups.

S_{max}: the sum of the ℓ greatest distances of all pairs.

The smaller C value indicates that a good group will be [10].

(g) *Global Sillhouette* Index

To get the Silhouette S (i) index the following formula is used:

$$S(i) = \frac{(b(i) - a(i))}{max\{a(i), b(i)\}} \quad (12)$$

where

a(i): the average difference of the i-object with all other objects in the same group.

b(i): the minimum value of the mean difference of i-objects with all objects in other groups (in the closest group).

The number of best groups is determined by the highest value from the Global Silhouette Index, which is then used to determine the ideal group. The formula for the Global Silhouette is as follows:

$$GS_u = \frac{1}{n}\sum_{i=1}^{n} S(i) \quad (13)$$

where

S(i)= *Silhouette* of *i* group

n= number of groups

(h) *Goodman-Kruskal* (GK) Index

Suppose that the four pairs of all observed objects are (q, r, s, t), where *d(x,y)* is the distance between the object x and y. The four pairs of objects are said to be concordant if they meet the conditions *d(q,r) < d(s,t),* where q and r are in the same group and s and t are in different groups. Conversely, four pairs of objects are said to be discordant if they meet the following conditions: *d(q,r) < d(s,t)* where q and r are in different groups and s and t are in the same group.

The GK index is calculated from the calculation of the value of the concordant and discordant pairs with the formula:

$$GK = \frac{S_c - S_d}{S_c + S_d} \quad (14)$$

di mana

S_c = number of concordant pairs

S_d = number of discordant pairs

Large values indicate the optimum group (Bolshakova, 2003).

All calculated indices can give the optimum number of groups, but each index can give different results. Pouwels [11] provides an alternative to choosing the optimum number of groups by combining the group validity index, which can then be selected for the optimum number of groups when the index is the most combined. The step is to calculate the five validity indices, then rank each possible number of groups on each index. The optimum number of groups is obtained at the highest average ranking.

2.7. Operational Definition of Service Quality Variables

Service quality is a way of working for a firm that strives to enhance the processes, products, and services it produces on a constant basis. In order to balance consumer expectations, service quality also includes an effort to meet the requirements and wishes of customers, as well as the precision with which they are delivered. According to Parasuraman et al. [12]'s research, five factors can be used to assess service quality: 1) Reliability; 2) Responsiveness; 3) Assurance; 4) Empathy; 5) Tangibles.

(1) Reliability

Reliability is the ability to provide the promised service appropriately and the ability to be trusted, especially in providing services.

(2) Responsiveness

Responsiveness is the ability to help what consumers need quickly, precisely, and responsively.

(3) Assurance

Assurance is the ability to eliminate customer doubts and make them feel exogenous from dangers and risks.

(4) Empathy

Empathy is the company's ability to understand consumer needs and ease of communication or relationships.

(5) Tangibles

Tangibles are the availability of physical facilities, equipment, and others in a company.

2.8. Operational Definition of Environmental Variables

According to Simamora [13], the work environment is the internal/psychological environment of the company and the human resource policies accepted by company employees. According to Carr's research [14], three dimensions can measure the work environment, namely: 1) Physical Work Environment; 2) Temporary Work Environment; 3) Psychological Work Environment.

1) Physical Work Environment

Measured through six indicators, namely: lighting, use of color, air circulation, noise, cleanliness, and safety.

2) Temporary Work Environment

Measured by two indicators, namely: working hours and rest periods.

3) Psychological Work Environment

Measured through three indicators, namely: boredom, fatigue, and work relations.

2.9. Operational Definition of Fashion Variables

Fashion is a model, method, style, or form of habit. Fashion is not only related to clothing styles, but there are also relationships with cosmetic styles, accessories, hairstyles, etc. to support one's appearance.

According to Karlyle; "Fashion is a symbol of the soul. Clothing cannot be separated from the development of human life history and culture. In other words, clothing can be interpreted as a social skin that contains messages and also the way of human life".

The benefits of fashion in everyday life include providing self-confidence for women where psychologically, every woman who looks attractive and comfortable has more confidence than women who look unattractive. Besides, fashion can give its charm, especially in connection with politeness and friendliness there will be an attractive charisma. Fashion can also make you happy because there is a feeling of satisfaction with fashion that is a concern. The indicator of the fashion variable:

• Activities

According to KBBI, an activity is an activity. Activities in fashion relate to clothing design work, footwear design and other fashion accessories design, making patterned clothes and accessories, mid-fashion dialogues, and fashion product distributors.

• Interests

Understanding interest in language (etymology) Interest is the desire and desire to learn (learn) and get something. In the sense of (the word), interest is the desire, enthusiasm and desire to do something. Attention is a feeling of love and desire for something.

• Opinions

In general, the notion of opinion is an opinion, response, view, or the result of a person's thoughts in explaining or addressing a matter but its nature is not objective and the truth is uncertain. Opinions are subjective and everyone may have different opinions about an event or object.

2.10. Operational Definition of Willingness to Pay Variable

According to Zhao and Kling [15], willingness to pay is the maximum price of goods that consumers are willing to buy at a certain time. Another thing, willingness to pay can be interpreted as the willingness of the community to accept payments at a certain level of payment. Willingness to pay is important to protect consumers from the dangers of corporate monopoly related to price and product supply [16].

According to research by Priambodo and Najib [17], 4 indicators can measure a person's willingness to buy, namely:

1) Options

Indicators that aim to choose to pay credit on the Bank BNI people's housing credit assessment (KPR).

2) Benefits

The added value of an expected outcome in using Bank BNI KPR.

3) Sacrifice

An act of giving up something sincere and sincere moral awareness in using Bank BNI KPR.

4) Be consistent

Payments are made on public housing loans repeatedly from time to time, to be fair and accurate.

2.11. Operational Definition of Compliant Paying Behavior Variable

Customer compliance is the customer has the willingness to fulfill his debt obligations following applicable regulations without any investigation, joint investigation, warning, and application of sanctions both legally and administratively [18], customer actions in fulfilling their debt obligations following regulatory provisions between customers and the leasing party or bank [19]. Based on this theory, it can be concluded that customer compliance is the customer's action in fulfilling their debt obligations according to the previously agreed regulations and is willing to accept sanctions if they do not comply. According to Law No. 6 of 1983, customer compliance can be measured through: timeliness, data accuracy, and sanctions.

3. Methodology

The variables used in this study are service quality, fashion, environment, willingness to pay, and obedient behavior to pay at Bank X. The data used in this study is primary data, and the variables used in this study are service quality, fashion, environment, willingness to pay, and obedient behavior to pay at Bank X. Questionnaires with a Likert scale level yielded data. The average score of each item is used to measure variables in primary data. Purposive sampling was utilized to collect data. A total of 100 people are expected to answer the survey. Because the central limit theory states that the sampling distribution curve (for a sample size of 30 or more) will center on the value of the population parameter and will have all the properties of a normal distribution, a sample of 100 customers was chosen.

The silhouette index, Karzanowski-Lai, Dunn, Gap, Davies-Bouldin, C-Index, Global Sillhouette, and Goodman-Kruskal were used as analytical tools in this study, followed by descriptive analysis and cluster analysis using the ward linkage and Euclidean distance methods for various cluster compatibility indices, including the silhouette index, Karzanowski-Lai, Dunn, Gap, Davies-Bouldin, C-Index, Global Sillhouette This study also made use of the R programming language.

4. Results

According to the findings of this study, the number of members in each cluster is the same for all indices, i.e., Cluster 1 has 42 members per index, whereas Cluster 2 has 58 members for each index. Table 2 shows the findings of the number of cluster members on each index.

Table 2. Number of Clusters for Each Index

Index	Cluster 1	Cluster 2
1. Silhouette	42	58
2. Krzanowski-Lai	42	58
3. Dunn	42	58
4. GAP	42	58
5. Davies-Bouldin	42	58
6. C-index	42	58
7. Global Sillhouette	42	58
8. Goodman-Kruskal	42	58

After getting the cluster and its members, then the average is sought to find out the differences in the members of each index and variable. The following are the average results obtained which can be seen in table 3.

Table 3 shows that the majority of Cluster 1 consumers perceive Bank X to have the greatest service quality, environment, fashion payment availability, and customer compliance behavior in Indonesia. Index of Validity in Cluster 2, the majority of Bank X customers agree that service quality, surroundings, fashion, willingness to pay, and behavioral consistency are all appropriate in Indonesia. The validity index C, on the other hand, demonstrates that the average findings for all static variables differ by 3.5 points for each cluster.

The Krzanovsky-Lai, Dune, Gap, Global Silhouette, Goodman-Kirskl, Silhouette, and Davis-Boldin indices have the same mean yield as the cluster mean, and the C index has the mean yield. They are different from the others. This means that the members of the same cluster in the Karzanovsky-Lai, Dunn, Gap, Global Silhouette, and Goodman-Kriskel indexes are the same as the Silhouette and Davis-Bold. From table 3, the best cluster indexes are Silhouette and Davis-Baldin.

Table 3. The Average Cluster Members for Each Index

Index	Average									
	X1		X2		X3		Y1		Y2	
	C1	C2	C1	C2	C1	C2	C1	C2	C1	C2
Silhouette	3.94	3.34	4.06	3.30	3.92	3.39	3.98	3.30	3.97	3.32
Krzanowski-Lai	3.94	3.17	4.05	3.10	3.93	3.23	3.97	3.13	4.00	3.11
Dunn	3.94	3.17	4.05	3.10	3.93	3.23	3.97	3.13	4.00	3.12
GAP	3.94	3.17	4.05	3.10	3.93	3.23	3.97	3.13	4.00	3.12
Davies-Bouldin	3.94	3.34	4.06	3.30	3.92	3.39	3.98	3.30	3.97	3.32
C-index	3.52	3.48	3.54	3.47	3.54	3.52	3.58	3.41	3.61	3.40
Global Sillhouette	3.94	3.17	4.05	3.10	3.93	3.23	3.97	3.13	4.00	3.12
Goodman-kruskal	3.94	3.17	4.05	3.10	3.93	3.23	3.97	3.13	4.00	3.12

The results of this study are consistent with research conducted by Petrovic [4] entitled "A Comparison Between the Silhouette Index and the Davies-Bouldin Index in Labeling IDS Clusters", indicating that the Silhouette Index and Davies Bouldin give the same results but there is a slight difference, which is meaningless.

The results of the comparison of the variance within the cluster and the variance between clusters on each validity index can be seen in table 4 below.

Table 4. The Variance Within and Between Groups of Each Index and Linkage Method

Index	Variance Within *Cluster*		Variance Between *Cluster*
	***Cluster* 1**	***Cluster* 2**	
Silhouette	10.83	6.54	4.29
Krzanowski-Lai	10.83	6.54	4.29
Dunn	10.83	6.54	4.29
GAP	10.83	6.54	4.29
Davies-Bouldin	10.83	6.54	4.29
C-index	10.83	6.54	4.29
Global Sillhouette	10.83	6.54	4.29
Goodman-kruskal	10.83	6.54	4.29

The different validity indices in each linkage method give the same results, so it can be concluded in this study that the difference in the validity index does not make a difference to the variance within and between clusters.

5. Conclusions

Based on the results of the analysis that has been done, the following conclusions can be drawn.

1. The cluster analysis with various validity indices can be applied to classify the compliance behavior of bank X customers in Indonesia.
2. The application of cluster analysis with different cluster validity indices results in the same number of clusters and variance between and within clusters.
3. The difference in the cluster validity index results in different cluster members, which can be seen from the average of each cluster for each index.

REFERENCES

[1] Kasmir, 2011, “Analisis Laporan Keuangan”, Raja Grafindo Persada: Jakarta.

[2] Bank Indonesia. 1998. UU No.10 tahun 1998, Tentang Perubahan Terhadap UU No. 7 tahun 1992, Jakarta.

[3] Supranto, J. (2004). Analisis Multivariat Arti dan Interpretasi. Bandung: PT. Rineka Cipta

[4] Petrovic, S. (2006, October). A comparison between the silhouette index and the davies-bouldin index in labeling ids clusters. In Proceedings of the 11th Nordic Workshop of Secure IT Systems (Vol. 2006, pp. 53-64). sn.

[5] Solimun, Fernandes, A. A. R., dan Nurjannah. 2017. Metode Statistika Multivariat Pemodelan Persamaan Struktural (SEM) Pendekatan WarPLS. Malang. UB Press

[6] Johnson, R. A. dan Wichern, D. W. 1992. Applied Multivariate. Analysis. Upper Saddle River, NJ: Prentice Hall.

[7] Aini, F. N. (2014). Clustering Business Process Model Petri Net Dengan Complete Linkage.

[8] Dunn, J. C. (1973). A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters.

[9] Friedman, J., Hastie, T. and Tibshirani, R., 2001. The elements of statistical learning (Vol. 1). Springer, Berlin: Springer series in statistics

[10] Bolshakova, N., F. Azuaje. 2001. Improving Expression Data Mining Through Cluster Validation. Department of Computer Science. Trinity College Dublin. Ireland.

[11] Pouwels, E.1998. Selecting An Optimum Clustering by Combining Cluster Validity Indices. http://www.estat.kuleuven.ac.be/`pauwels/pocv-98. Tanggal akses: 24 Februari 2006. Santoso, S. 2002. Statistika Multivariat dengan SPSS. PT. Elex Media Komputindo. Jakarta.

[12] Parasuraman, A., Valarie A. Zeithaml, and Leonard L. Berry. (1990). Delivering Quality Service: Balancing Customer Perceptions and Expectations. New York: THE TREE PRESS.

[13] Simamora, Henry. 1997. “Manajemen Sumber Daya Manusia”. Yogyakarta. STIE YKPN.

[14] Carr, A. (2004). Positive Psychology; The Science of Happiness and Human Strengs. New York: Brunner Routledge.

[15] Zhao, & Kling. (2005). Willingness To Pay, Compensating Variation, and the Cost of Commitment. Economic Inquiry, Vol. 42, N, 503–517

[16] Grace Laumahina & Njo Anastasia, 2014, Kesediaan untuk Membayar pada Green Residential, FINESTA vol.2 no.1 page 82-86, Universitas Kristen Petra

[17] Priambodo, L. H., & Najib, M. (2014). Analisis Kesediaan Membayar (Willingness to Pay) Sayuran Organik dan Faktor-Faktor yang Mempengaruhinya. Jurnal Manajemen dan Organisasi, 5(1), 1-14.

[18] Gunadi. 2013. Panduan Komprehensif Pajak Penghasilan. Jakarta: Bee Media Indonesia.

[19] Siti Kurnia Rahayu, 2010 .PERPAJAKAN INDONESIA: Konsep dan Aspek Formal, Yogyakarta: Graha Ilmu..