

## **STATISTICAL INFERENCE-BASED DECISION MAKING FOR SMALL AND MEDIUM-SIZED ENTERPRISE**

**Suparman,**

*Student, Department of Statistics,  
Universitas Terbuka, (INDONESIA)  
alosuparman074@gmail.com*

### **Abstract**

More than 22 millions Indonesians running Small and Medium-sized enterprise (UKM stands for Usaha Kecil dan Menengah), absorbed more than 43.9 millions workers and contributed more than Rp 1.6 trillions to Indonesia's Product Domestic Brutto (PDB)" (a citation from Sri Wahyuningsih, Lecturer, Universitas Wahid Hasyim, "PERANAN UKM DALAM PEREKONOMIAN INDONESIA", 2009). "The PKL" stands for Pedagang Kaki Lima (street vendor) selling foods, garment or other small sized commodities is the most popular example of the UKM and it has been chosen as the subject of this study. Survival and development of the UKM are indeed important to Indonesia's economy. One important issue among UKM is in decision making process about whether to stop the business or continue with/without needs for capex. Scientific approach making use of riel data and statistical inference can well be utilized for this purpose and produce reliable basis for a decision making. When cost of goods and selling price have been established and ensure profitability, the remaining question is about the plausible sales projection. Trial running the subject shop of this experiment for 35 days, generates sufficient variable random data of sales, visitors and other relevant information. Further, by making use of statistical inference theory, a reliable estimate of sales potential and other related data will be available for decision making. When the scientific-approached concept is developed among UKM individuals and becomes a culture, it will rectify the path toward enjoying the benefit of the forth coming Society 5.

Keywords: Street vendor, Scientific Approach, Variable random data, Statistical Inference, Society 5.0

## 1 INTRODUCTION

More than 22 million Indonesians running Small and Medium-sized enterprise, the UKM (Usaha Kecil dan Menengah), absorbed more than 43.9 million workers and contributed more than Rp 1.6 trillion to Indonesia's Product Domestic Brutto (PDB)", (a citation from S. Wahyuningsih, Lecturer Universitas Wahid Hasyim, "Peranan UKM Dalam Perekonomian Indonesia", (2009)) [1]. "PKL=Pedagang Kaki Lima" (street vendor) selling foods, garment or other small sized commodities is a popular example of this business entity and it has been chosen as the subject of this study

PKL is a label attached to those small-sized enterprises operating on roadsides or sidewalk/ pedestrian. They are one, yet most popular example of The UKM group, the Small and Medium-Sized Enterprise.

The purpose of this paper is to collect and analyze sales, customer's data of a PKL and precipitation forecast; additionally obtain also other related data about its operation. Eventually, based upon the statistics of data sample and its inference, a conclusion and basic argument for further steps of decision making related to survival or development if any of the PKL can be formulated.

## 2 METHODOLOGY

### 2.1 Subject of The Study

The PKL chosen for this study, located in an area at a busy main street in South Jakarta. It bears the name "Sate Maranggi 29", selling beef, chicken meat and skin 'sate', a typical Indonesian meat barbecue. This relatively new PKL, in operation for less than a month, planned to open 7 days a week starting 13.00 to 09.00 pm, employing two helping boys. It is equipped with visitors' facilities i.e. a semi-permanent plastic covered hood/tent, 2 two meters long tables and 10 chairs. I counted more than a dozen of PKL selling food nearby. Despite the PKL selling similar food quite limited, competition among them is tough. The owner of 'Sate Maranggi 29' explained that buyers prefer delicious taste, low price and acceptable size. Buyers are

mostly those travelling on the way home from their jobs, driving motor bikes. Take away order as well as on-line also offered to buyers.

The owner manages to formulate and produce the foods at a competitive price yet yielding the targeted maximum gross profit of 100 percent/profit margin of 50%.

He has been successful in finding the source of good quality raw materials especially the low price beef and finding a unique tasty recipe of the sauce for the barbecue. The remaining issue to be answered is its business prospect: Its survival or growth or else the other-way-round, bankrupt?

The entrepreneurs running UKM, according to the data published by BPS in 2006, 73% of them graduated from junior high school (SLTP) or lower, 21% from senior high school (SLTA), and only 6 % has a high education background (vocational diploma, bachelor or other university degree) [1].

In view of that, it is very probably that most UKM is managed in a traditional way, particularly in decision making process. I am tempted to take action by offering optional solution, especially the scientific-based one.

The subject being studied, 'Sate Maranggi 29' at the preliminary stage, employed 2 sale boys and paid allowance of Rp50,000/day/person plus free Rp25,000 lunch meal every day. Although this allowance is lower than the minimum regional remuneration requirement of Jakarta (UMR Jakarta 2021): Rp 4.416.186 [2], it is still acceptable because difficulty to find job.

According to BPS Publication (2006) in UKM sectors, 31.7 million out of 43.9 million workers (60%) are not paid, because they are families or the owner himself [1].

The success, failure or in another words the life of any trading business depends basically on whether in a certain period of time, a day, a week, a month or a year, it is capable of getting gain from the exchange between their goods or services they sold and the value in terms of money they collected.

Cash money in an UKM circulates very fast as owners will directly consume gain of their income to finance their daily living expenses. If they can get a picture about their business prospect early enough, they can make decision in time, of continuing the enterprise and make available any additional facilities or cut expenses if needed, or else to stop the current one to avoid any further loss and look for new option.

## 2.2 Design of Experiment

Pictures of the long period run (example 1 year) of an enterprise, which indicates among others it prospects, will be assumed as population. We can get the important and fairly accurate data and information about the population by running the enterprise in the intended duration. However, this method suffers from two disadvantages – long waiting time and costly. Both are not acceptable.

Thanks to the statistical sampling method and inference theory, the needs to collect all data from population which will be numerous in number can be avoided. To do the inference we take a representative sample, namely a sample which well represents the population. A sample which conforms to the requirement is the random sample fitting the homogenous population and be drawn randomly (not bias).[3]

### 2.2.1 *Data, Information and Sampling*

Scientific-based analysis will need two tools, the riel random data and the theoretical science: mathematics, statistics and the other ones which are neither statistics nor mathematics.

Our analysis concerned with inference—that is, reaching valid conclusion concerning the population on the basis of information from the sample. [3]

Sale in rupiah, number of buyers, climatic condition-precipitation/rain, and traffics (vehicle also people) are very probably the variables which correlated each other and constitute the variable vectors of the analyzed population. This study will cover only 3 variable vectors, namely:

Sale in rupiah ( $X_1$ ),  
 Number of buyer ( $X_2$ ), and  
 Precipitation data ( $X_3$ ).

Sale data and number of guest are provided by the entrepreneur and precipitation data is taken from those published by <https://climate-data.org> [4].

Sample size between 30 and 50 is considered as large, and in multivariate case sample is large if  $(n - p)$ , wherein  $n$ =number of observations,  $p$ =number of variables) is large (30 - 50) [5]. When the sample is large, tests of hypotheses and confidence region for  $\mu$ , the population mean, can be constructed without the assumption of normal population and inferences about  $\mu$  are based on  $\chi^2$  [5]. This is the reason of drawing sample size 35. Thus, test for normal distribution of data is no needed.

### 2.2.2 Statistical Multivariate

The goal of our analysis is inference of the population means, namely to make a valid conclusion about the means of a population based on a sample. As we have  $p \geq 1$  variables, multivariate statistical approach shall be used instead of univariate one. We analyze data jointly. Simultaneous analysis yields stronger test, with better error control. [6]

The values of the variables are all recorded for each distinct item, individual, or experimental unit. Accordingly, as this study sample size,  $n = 35$  and variable,  $p=3$ , array of data can be displayed as follow [6]:

Table 1. Observed variable data

Variable, p	Sample, n					
1	$X_{11}$	$X_{21}$	$X_{31}$	.....	....	$X_{351}$
2	$X_{12}$	$X_{22}$	$X_{32}$	.....	....	$X_{352}$
3	$X_{13}$	$X_{23}$	$X_{33}$	.....	....	$X_{353}$

And we can display these data in form of a matrix  $X_{jk}$  of  $p$  rows and  $n$  columns:

$$X_{jk} = \begin{bmatrix} X_{11} & X_{21} & \dots & X_{351} \\ X_{12} & X_{22} & \dots & X_{352} \\ X_{13} & X_{23} & \dots & X_{353} \end{bmatrix} \quad \begin{array}{l} j=1,2, \dots j=n=35 \\ k=1,2, k=p=3 \end{array}$$

The statistical methods designed to elicit information from those types of data sets and the main objective of this scientific investigation is *data/structural reduction, dependence among variables, prediction, hypothesis construction, and testing* [5]

A large data set is bulky and its very mass poses a serious obstacle to an attempt to visually extract pertinent information. By calculating certain summary numbers in accordance with descriptive statistics analysis, similar information can be obtained. Arithmetic average or sample mean is a central value, and the average of squares distances of all data from means provide measure of spread/variation. [7]

Mean of a sample can be computed from  $n$  data on each  $p$  variables, so that there will be  $p$  sample means [6]:

$$\bar{x}_k = \frac{1}{n} \sum_{j=1}^n x_{jk} \quad k=1,2,\dots,p \quad (1)$$

In our particular experiment,  $n=35$ ,  $p=3$ . the vector mean of  $X$ :

$$\bar{X} = \begin{bmatrix} \bar{x}_1 = \frac{1}{35} \sum_{j=1}^n x_{j1} \\ \bar{x}_2 = \frac{1}{35} \sum_{j=1}^n x_{j2} \\ \bar{x}_3 = \frac{1}{35} \sum_{j=1}^n x_{j3} \end{bmatrix} \quad \begin{matrix} j=1,2, \dots, 35 \\ k=1,2,3 \end{matrix}$$

A measure of spread is provided by sample mean variance-covariance, for  $p$  variables and  $n$  observations it can be calculated by the following formulas [6]

$$s_k^2 = \frac{1}{n} \sum_{j=1}^n (x_{jk} - \bar{x}_k)^2 \quad k=1,2,\dots,p \quad (2)$$

$$s_{ik} = \frac{1}{n} \sum_{j=1}^n (x_{ji} - \bar{x}_i)(x_{jk} - \bar{x}_k) \quad i=1,2,\dots,p, \quad k=1,2,\dots,p \quad (3)$$

In our case,  $n = 35$  and  $p=3$

The variance covariance:

$$\begin{aligned} s_{11} &= 1/35 \sum (x_{j1} - \bar{x}_1)^2 \\ s_{22} &= 1/35 \sum (x_{j2} - \bar{x}_2)^2 \quad ; j = 1, 2, 3 \dots\dots 35 \\ s_{33} &= 1/35 \sum (x_{j3} - \bar{x}_3)^2 \\ s_{12} &= 1/35 \sum [(x_{j1} - \bar{x}_1)((x_{j2} - \bar{x}_2))] \\ s_{13} &= 1/35 \sum [(x_{j1} - \bar{x}_1)((x_{j3} - \bar{x}_3))] \\ s_{23} &= 1/35 \sum [(x_{j2} - \bar{x}_2)((x_{j3} - \bar{x}_3))] \end{aligned} \quad \Sigma = \begin{bmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{bmatrix}$$

In fact  $s_{12} = s_{21}$ ;  $s_{13} = s_{31}$ ;  $s_{23} = s_{32}$

$s_{11}$ ,  $s_{22}$ ,  $s_{33}$  are variants

$s_{12}$ ,  $s_{13}$ ,  $s_{21}$ ,  $s_{23}$ ,  $s_{31}$ ,  $s_{32}$  are covariants

If there is no particular association between two variables, covariants are approximately zero.

Final descriptive statistic is  $r$ , the sample correlation coefficient (Pearson's product-moment correlation coefficient), defined as follows [6]:

$$r_{ik} = \frac{S_{ik}}{\sqrt{S_{ii}} \sqrt{S_{kk}}} = \frac{\sum_{j=1}^n (X_{ji} - \bar{X}_i)(X_{jk} - \bar{X}_k)}{\sqrt{\sum_{j=1}^n (X_{ji} - \bar{X}_i)^2} \sqrt{\sum_{j=1}^n (X_{jk} - \bar{X}_k)^2}} \quad (4)$$

Value of  $r$  ranges from -1, 0, and +1 and it shows degree of correlation between two variables. 1 is the highest, - and + signs meaning increase of one variable will decrease the other (case -), and increase of one variable increasing also others (case +), and 0 means no correlation.

The summary of matrixes of the vector means, variance and correlation are as follow:

$$\bar{X} = \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \\ \bar{X}_3 \end{bmatrix} \quad \Sigma = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix} \quad R = \begin{bmatrix} 1 & r_{12} & r_{13} \\ r_{21} & 1 & r_{23} \\ r_{31} & r_{32} & 1 \end{bmatrix}$$

sample mean                      sample variance covariance                      sample correlation

### Inferences about A Vector Mean of the Population.

Descriptive statistics provided us with 3 summary statistics, namely vector sample mean, matrix variance covariance, and matrix correlation.

Inference about vectors means of population based on sample data, is carried out by testing the null hypotheses:

$$H_0 : \mu = \mu_0 \quad \text{and} \quad H_1 : \mu \neq \mu_0$$



This test will affirm the plausability of  $\mu_0$  as a value for the population mean.

Since we are working with multivariate data, the proper test formula is the Hotteling T<sup>2</sup> formula [5]:

$$T^2 = (\bar{X} - \mu_0)' \left( \frac{1}{n} S \right)^{-1} (\bar{X} - \mu_0) = n(\bar{X} - \mu_0)' S^{-1} (\bar{X} - \mu_0) \quad (5)$$

Where [5]

$$\bar{x}_{(p \times 1)} = \frac{1}{n} \sum_{j=1}^n x_j, \quad S_{(\rho \times \rho)} = \frac{1}{n-1} \sum_{j=1}^n (x_j - \bar{x})(x_j - \bar{x})', \quad \text{and } \mu_0 (p \times 1) = \begin{bmatrix} \mu_{10} \\ \mu_{20} \\ \vdots \\ \mu_{p0} \end{bmatrix} \quad (6)$$

X = vector mean of respective variable; n= sample size; p= number of variable; S = variance –covariance;  $\mu_0$  = population means of respective variable.

Although we do not assume that data is a normal distribution [X<sub>j</sub> not similar to Np ( $\mu, \Sigma$ )]; we can still make inferences (hypotheses & make confidence statements) about population data since our sample size is large namely  $n - (35 - p) = 32$  ( $32 > 30$ ). However, instead of F distribution, the test should be based on Chi-square distribution,  $\chi^2_p$ . [5]

$$T^2 \approx \chi^2_p \quad (\text{as } n \rightarrow \infty, (n-1)p/(n-p) F_p, n-p \rightarrow \chi^2_p) \quad (7)$$

Let  $x_1, x_2, \dots, x_n$  be random sample from a population mean  $\mu$  and positive definitive covariance matrix  $\Sigma$ . When  $(n - p)$  is large ( $\geq 30$ ), the hypotheses  $H_0: \mu = \mu_0$  is rejected in favour of  $H_1: \mu \neq \mu_0$ , at level of significance approximately  $\alpha$ , if the observed. [5]

$$n(\bar{x} - \mu_0)' S^{-1} (\bar{x} - \mu_0) > \chi^2_p(\alpha) \quad (8)$$

Where in for  $n=35$ ,  $p=3$ ,  $\alpha = 0.05$  from table Chi Square:

$$\chi^2_3 (0.05) = 7.81 [3]$$

Then, in our case, rejection criteria is:

$$35(\bar{x}_1 - \mu_1 \quad \bar{x}_2 - \mu_2 \quad \bar{x}_3 - \mu_3) \begin{bmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{bmatrix}^{-1} \begin{bmatrix} \bar{x}_1 - \mu_1 \\ \bar{x}_2 - \mu_2 \\ \bar{x}_3 - \mu_3 \end{bmatrix} > \chi^2(0.05) = 7.81 \quad (9)$$

The plausible means region is [5]:

$$\begin{aligned} \bar{x}_1 \pm \sqrt{\chi_p^2(\alpha)} \sqrt{\frac{s_{11}}{n}} & \text{ contains } \mu_1 \\ & \vdots \\ \bar{x}_p \pm \sqrt{\chi_p^2(\alpha)} \sqrt{\frac{s_{pp}}{n}} & \text{ contains } \mu_p \end{aligned} \quad (10)$$

And for  $n=35$ ,  $p=3$ ,  $\alpha = 0.05$ ,  $\chi^2_3 (0.05) = 7.81$ , the region of

$$\begin{aligned} \mu_1 & \text{ is } \bar{x}_1 \pm \sqrt{(7.81)} \sqrt{[(s_{11})/35]} ; \mu_2 & \text{ is } \bar{x}_2 \pm \sqrt{(7.81)} \sqrt{[(s_{22})/35]} ; \mu_3 & \text{ is } \\ \bar{x}_3 & \pm \sqrt{[(7.81)]} \sqrt{[(s_{33})/35]} \end{aligned} \quad (11)$$

### 3 RESULTS

#### 3.1 Observed Sample Data and Data Matrix

*Table 2. Sale Data in Thousand Rupiah*

411	735	332	299	659	261	170
74	185	379	222	165	318	254
338	182	92	133	171	286	173
181	437	177	660	572	260	324
199	150	240	256	679	739	517

*Table 3. Number of Customer*

9	15	11	7	18	3	6
3	7	8	6	3	10	7
9	8	5	7	7	5	6
6	15	8	8	11	8	8
6	5	5	4	6	8	11

*Table 4. Precipitation Data/Rainfall, in mm*

1.4	2.1	1.9	1.4	1.3	1.2	0.7
0.9	1.4	1.7	2.1	2	2.3	2
1.6	2	1.5	0.9	1.2	1.2	1.9
1.6	2.2	2.5	1.7	2	2.5	1.7
2	1.9	1.5	2	1.9	2.3	1.9

Vectors of Observed Data Matrix\_X

$$X = \begin{bmatrix} 411 & 735 & 332 & 299 & 659 & 261 & 170 & 74 & 185 & 379 & 222 & 165 & 318 & 254 & 338 \\ 182 & 92 & 133 & \underline{171} & 286 & 173 & 181 & 437 & 177 & 660 & 572 & 260 & 324 & 199 & 150 \\ 240 & 256 & 679 & \underline{739} & 517 & & & & & & & & & & \\ \\ 9 & 15 & 11 & 7 & 18 & 3 & 6 & 3 & 7 & 8 & 6 & 3 & 10 & 7 & 9 \\ 8 & 5 & 7 & \underline{7} & \underline{5} & 6 & 6 & 15 & 8 & 8 & 11 & 8 & 8 & 6 & 5 \\ 5 & 4 & 6 & 8 & 11 & & & & & & & & & & \\ \\ 1.4 & 2.1 & 1.9 & 1.4 & 1.3 & 1.2 & 0.7 & 0.9 & \underline{1.4} & 1.7 & 2. & 2. & 2.3 & 2 & 1.6 \\ 2 & 1.5 & 0.9 & 1.2 & 1.2 & 1.9 & 1.6 & 2.2 & \underline{2.5} & 1.7 & 2 & 2.5 & 1.7 & 2 & 1.9 \\ 1.5 & 2 & 1.9 & 2.3 & 1.9 & & & & & & & & & & \end{bmatrix}$$

### 3.2 Summary of Descriptive Statistics and Profit Forecast

The aforementioned formulas (1) to (10) are used to compute the followings:

Matrixes of the vector means, variance and correlation are:

$$\bar{X} = \begin{bmatrix} 409 \\ 10 \\ 1.7 \end{bmatrix} \quad \Sigma = \begin{bmatrix} 42824.4 & 599.8 & 20.9 \\ 599.8 & 16.3 & 0.3 \\ 20.9 & 0.3 & 0.2 \end{bmatrix} \quad R = \begin{bmatrix} 1 & 0.72 & 0.23 \\ 0.72 & 1 & 0.18 \\ 0.23 & 0.18 & 1 \end{bmatrix}$$

sample mean      sample variance covariance      sample correlation

To check the region of each variables mean we use the Formula (10). The result are summarized in *Table 5*, and it shows also the correlation between two variables.

*Table 5. Statistics Descriptive*

No.	item	Variable, p					
		X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>			
1	Vector Mean	509	10	1.7			
2	Vector Mean, min	311	8	1.5			
3	Vector Mean,max	507	12	1.9			
4	Correlation				X <sub>1</sub> - X <sub>2</sub>	X <sub>1</sub> - X <sub>3</sub>	X <sub>2</sub> - X <sub>3</sub>
					0.72	0.23	0.18

*Table 6. Profit Forecast, (in Rupiah)*

No	Item	Daily	Monthly (25 days)	Annually (12months)
1	Sale, average	409,000	10.225.000	122.700.000
2	Gross profit margin, 50%	204,500	5.112.500	61.350.000
3	Sales cost (*)	To be decided		
4	Utilities			
4.1	Investment Rp 2.400.000	4.000	100,000	1,200,000
4.2	(Depreciation, 24 months) Water, electricity, etc.	1.200	30,000	360,000
5	Gross Profit (2 – 4)	199.300	4.982.500	59.790.000

\* Sale cost needs be decided later by the owner which will depend on recruitment of the sales force. The cost will reduce the profit.

### 3.3 Discussion

1. At running operation of 25 days/month, this business offers a potency monthly profit of Rp 4.982.500 (average), which is slightly above the standard Jakarta regional living cost (UMR Jakarta 2021) of Rp 4.416.186. per month [2] and therefore worth to be continued.
2. The owner shall decide how many sales force to be recruited and at what cost be paid. The cost will diminish his earning.
3. Numbers of current available chairs is the same as mean of customer and together with all other facilities considered to be kept or else added two more chairs as anticipation.
4. A relatively high correlation ( $R = 0.72$ ) between sale and number of customers; on-line marketing and circulating brochures can help inviting the interest of more customers.
5. Precipitation variable correlates very low with both sale and number of customer. The reason is probably due to the fact that this study covered only the dry season period. The study in wet season may give different result.
6. This study is only the beginning. In the future other experiment can be conducted studying effect of other variables such as traffics, differences between week day and week end, new recipe or new foods and drinks. Since multiplying earnings for this enterprise will only be effectively done by opening new outlets, the same new experiment will work in providing basis for decision.

Big Data with its 3Vs feature – Big Volume, very high Variability, and immense Velocity, capable of processing data at any format in in addition to storage and retrieval, in real time. [8]

Society 5.0 according to one definition: "A human-centered society that balances economic advancement with the resolution of social problems by a system that highly integrates cyberspace and physical space" and Big Data is a part of it. [9].

We expect therefore when Society 5.0 is around, it will provide services to anybody, at affordable cost, real time, to process and analyze data and information to generate the required answer for solution.

That is the situation wherein PKLs can enjoy the benefit of the coming Society 5.0 and Big Data, as the latter can provide inputs for PKLs decision making in real time. By merely upload the required data any time and in any format, Big data will take care the manipulation of data and generate the solution for decision making.

#### **4 CONCLUSIONS**

This study has provided its subject – Sate Maranggi 29 with a set of arguments for decision making based on a scientific approached. This PKL according to this statistical analysis, in the long run offers an estimated mean income Rp199.300 daily/Rp 4.982.500 monthly/Rp59.790.000 annually, provided all of the sale-related variables remains similar and it operates 25 days a month, 12 months a year. The monthly average net profit namely the monthly income of the owner, is well above minimum living cost limit, thus the business worth to be run. The owner has to decide himself concerning number and salary of sale person in view of the tolerated reduction of his monthly earning. Only minor adjustment needed in view of equipment to entertain the customers.

If the owner intend to develop the business still further to achieve higher gain, multiple to the maximum existing mean sale, the proper strategy is to open other similar enterprise at other location (outlet).

This study has to be considered as the beginning/preliminary introduction of the technic. Some others such as difference between week day and week end, recipe changes, new food and drink menu, can be run in accordance with the needs but similar in process to this experiment.

Introduction and developing familiarity about scientific-based decision making among small and medium-sized enterprises will provide them with a plausible technic in making policy, strategy and

forecasting their businesses. The success of this process will rectify a pavement for small and medium-sized enterprises to enjoy the benefit from the coming Society 5.0 and Big Data.

## ACKNOWLEDGEMENTS

Special thanks is due to Mr Rizky Jatnika, owner of the 'Sate Maranggi 29' for his collaboration and supplying data/information.

## REFERENCES

- [1] Sri Wahyuningsih, "Peranan UKM dalam Perekonomian Indonesia", Semarang, Indonesia, Fakultas Pertanian, Wahid Hasyim University, 2009, Retrieval 2021.
- [2] Rincian UMR Jakarta 2021 dan Daerah Sekitarnya, Kompas.com - 29/03/2021, Retrieval 2021.
- [3] Sri H. Kartiko and Suryo Guritno, "Metode Statistika Multivariat", Edisi 1, Jakarta, UT University, 2008.
- [4] PONDOK PINANG WEATHER IN AUGUST and SEPTEMBER// WEATHER AVERAGES, <https://climate-data.org>, Retrieval 2021.
- [5] Carolyn J. Anderson, "Inferences About a Mean Vector", Illinois, USA, Department of Educational Psychology, University of Illinois, 2017, Retrieval 2021.
- [6] T. W. Anderson, "An Introduction To Multivariate Statistical Analysis", Third Edition, Stanford CA, Stanford University, Department of Statistics, John Wiley & Sons, Inc., 2003, Retrieval 2021.
- [7] Richard A. Johnson and Dean W Wichern, "Applied Multivariate Statistical Analysis, Sixth Edition, New Jersey, USA, Pearson Prentice Hall, 2007, Retrieval 2021.
- [8] Big data, Wikipedia, the free encyclopedia, Retrieval 2021.
- [9] Achieving Society 5.0, <https://www8.cao.go.jp> > cstp > english > society5\_0, Retrieval 2021.