Analysis of Multiple Correspondence Against Crimes in Sleman Regency*

E Widodo¹‡ and R Maggandari²

¹,²Department of Statistics, Islamic University of Indonesia, Indonesia
‡corresponding author: edywidodo@uui.ac.id

Abstract
Crime is a bad behavior, from social and religious norms and its make psychology and economics harm. Stealing, ill-treatment, embezzlement, deception, deception/embezzlement, and adultery are the most crime in the last 9 months. Therefore, for identify the type of crime in the community we need a method to see the tendency of a category using multiple correspondence analysis methods. Analysis of multiple correspondence is one of the descriptive statistics that use to describe a pattern of relationships from contingency’s table with the aim of finding liability between categories. The results of the correspondence analysis is that the tendency of criminal suspect to be related to this types of crime of stealing and ill-treatment to be done by students or students less than 25 years old and were male, suspect of deception and adultery tends to be done by women over 40 years old and does not work, and suspect of embezzlement tends by workers and their ages around 25 to 40 years. The liability of the relation between criminal incidents and the types of crime is the types of crime of ill-treatment and adultery that are most prone to occur in shops with vulnerable hours 00:00-05:59 and 18:00-23:59

Keywords: crime, multiple correspondence, Sleman regency.

* Received: Jan 2021; Reviewed: Jan 2021; Published: Jun 2021
1. Introduction

In this modern era, it is undeniable that times are getting most sophisticated, technology is most blooming until the social life is change. Because the social life getting complicated it make the criminal act can be done anybody without know the age, religion, race, and gender (Kartono, 2011).

Criminal act continues to grow, especially in quantity or the types. Criminal act getting blooming especially in the developing country ex Indonesia. Urban is a central of crime, its because in the urban always occur high competition and not health. One of which is in Sleman Regency.

Sleman Regency be composed from 17 sub-district areas, 86 villages, and 1212 backwoods. Sleman Regency is famous as a central from big university. Based on the result of survey from Environmental Survey Agency published Daerah Istimewa Yogyakarta 2015th-2016th in CNN Indonesia states that Sleman Regency is the region that have a high crime in Daerah Istimewa Yogyakarta Province. Its cause Sleman regency is a central of big university in this province and the regency always increased from the economic activity until business trade (CNN, 2016).

Based on observations made by (Mutijo, 2016) in his book with title “Analisis Informasi Statistik Pembangunan Daerah Istimewa Yogyakarta” got the result is the reported crime in Daerah Istimewa Yogyakarta Province in 2012nd- 2014th is below:

<table>
<thead>
<tr>
<th>Kabupaten/Kota</th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kulonprogo</td>
<td>629</td>
<td>306</td>
<td>859</td>
<td>1794</td>
</tr>
<tr>
<td>Bantul</td>
<td>580</td>
<td>617</td>
<td>2.570</td>
<td>3767</td>
</tr>
<tr>
<td>Gunung Kidul</td>
<td>402</td>
<td>344</td>
<td>303</td>
<td>1049</td>
</tr>
<tr>
<td>Sleman</td>
<td>2.009</td>
<td>2002</td>
<td>1843</td>
<td>5854</td>
</tr>
<tr>
<td>Yogyakarta</td>
<td>1913</td>
<td>1911</td>
<td>1778</td>
<td>5602</td>
</tr>
<tr>
<td>MAPOLDA DIY</td>
<td>1194</td>
<td>1181</td>
<td>11347</td>
<td>13722</td>
</tr>
<tr>
<td>D.I.Yogyakarta</td>
<td>6727</td>
<td>6361</td>
<td>18700</td>
<td>31788</td>
</tr>
</tbody>
</table>

Source: regional police/resort police D.I. Yogyakarta

From Table 1 we can see total of reported crime in 2012nd until 2014th from regional police’s database. In this case, Sleman Regency is a regency with high crime case that 5854 cases, although in 2014th the big crime case occur in Bantul Regency.

From police’s database in January until September 2019th it had more than 10 types crime in occurs. There are had 6 types crime case that most occur, among them, embezzlement, fraud, stealing, persecution, and ill-treatment.

Figure 1 showed total of criminal case in Sleman Regency that we know if total of suspect of crime each one crime case more than compared with total of the case. this is influenced by total of incident that evenly involve more than one suspect of crime in every their case.

The high crime rate needed prevention. Crime information more needed to use give an overview or data about incident of crime, other than that it can to use be a crime gauge and the conclusion. Statistical analysis can be applied in criminology or we know with criminal statistical. Criminal anatomy is one of from criminal statistical
that explained about type of crime that comprise crime scene, time of incident, victim, and etc.

Figure 1: Crime case based on incident of crime and suspect of crime in Sleman Regency from January-September 2019th

The prevention of crime be a obligation of all parties, especially the police. One part from criminology is criminal anatomy. Anatomy criminal is study of types crime include the crime scene, time of incident, chronology, victim and etc (Nurfitasari A, 2017). One of statistic methods that can applied in criminal anatomy is correspondence analysis. Data of anatomy criminal’s occur from many inefficient variables if we use simple correspondence analysis, therefore we used multiple correspondence analysis. Multiple correspondence analysis is selected because can combine reduction technic and data mapping with more than two variables.

Multiple correspondence analysis can be applied for looking the accountability from crime case in Sleman Regency with use variable of suspect with some of characteristics namely age, gender and job, and variable of incident of crime with characteristic namely crime scene, time and place.

Nurfitasari A (2017) in Crime Mapping in Madiun Territory with Multiple Correspondence Analysis showed if characteristic of place and time of incident of crime like stealing, persecution, and fraud have the same characteritics in public places, settlement in Madiun City and SSWP 2 territory at 06.00AM-11.59AM, gambling case and rape have same characteristic in settlement in SSWP 3 territory at 18.00PM-23.59PM, and illegal logging case tend to be different with the other crime that happened in SSWP 1 and SSWP 4 forest at 00.59PM-05.59PM, while the suspect of crime, ill-treatment, and rape have similar characteristic with age less than 25th year old with level of education more than senior high school, suspect of gambling with age more than 25th year old with level of education just until primary school, suspect of fraud tend to be different with the other crime case namely to do from womankind, and suspect of illegal logging have characteristic with level of education until junior high school.

According to Alia Lesari and Muhammad Hajarul Aswad A (2016th) in “Mapping of
Criminal Case in Palopo City in 2015th showed if high criminal case in law territory of Palopo’s Police is vehicle theft and the smallest is ill-treatment with time more than occur at 24.00PM-07.00AM that located in Wara Regency (Lestari & Aswad, 2018).

Refers to previous research that have done there some similarity namely researching with use object of criminal case and use correspondence analysis’s methods with use suspect of crime’s data with characteristic from age, gender, education and data of incident criminal case with characteristic of crime scene, time and place. Things that differentiate this previous research is in use of attributes. This research use the criminal database that occur in Sleman Regency from January-September 2019, and in the suspect of crime’s database that to use is age, gender and job. Characteristic of education changed into be job that caused by insufficient data completeness in Sleman’s Police, Yogyakarta.

This finding and analyzes are expected give an overview and gauge for police for knowing the mapping and characteristic of criminal incident that often occur in Sleman Regency if we see from suspect’ viewpoint and criminal incidency. Other than that, this research expected to use be device in prevention and making decisions on criminal acts.

2. Materials and Methods

2.1 Criminality

Criminalization is considered a criminal act by the community (Soekanto et al., 1981). In the labeling perspective, criminalization is the decision of the criminal law-making body to label human behavior as a crime or a criminal act (Barlow & Kauzlarich, 1984). In his book (Kartono, 1998), crime is a behavior that violates the law and violates social norms, so that people oppose it (Kartono, 2011).

According to (Abdussalam, 2007) crime is divided into two, that is according to law (juridis) and non-law or crime according to sociology:

1. Crime according to the law (juridis) Crime is an act that shouldn’t be done, and is determined by the state in criminal law and is punishable by a penalty or punishment.
2. Crime according to non-law or according to sociology. Crime is a human behavior created by society, although society has different behaviors, but has the same patterns

2.2 Chi-Square Test

Chi-Square Test is one of the association tests to find out whether there is a relationship between two variables. This association test uses data from a contingency table between two variables that will be seen it’s relation [9]. The hypothesis from the Chi-Square test is:

\( H_0 \): there is no relationship between variable A and variable B, and
\( H_1 \): there is a relationship between variable A and variable B.

\[
\chi^2_{hitung} = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}
\]

(1)

Test statistics from the Chi-Square test are as follows:
where $O_{ij}$ = the observed frequency in the $i$-th row and $j$th column contingency table, $E_{ij}$ = the expected frequency of the contingency table in the $i$-th row and $j$-th column. The expected frequency of the contingency table between the two variables is calculated by the formula:

$$E_{ij} = \frac{n_i n_j}{n}$$

(2)

where $E_{ij}$ = the expected frequency from the contingency table in the $i$-th row and $j$-th column, $n_i$ = the number of the $i$-th row frequency, $n_j$ is the number of the $j$-th column frequency, $n$ = the sum of all frequencies from the contingent block.

The results of the test statistics are compared with the value of the chi-square distribution with free degrees (db) of $(r-1)(c-1)$. The rejection criterion is $H_0: \chi^2 \text{counts} > \chi^2(\text{db}, \alpha)$ or if the $p$-value < $\alpha$. If $H_0$ is rejected this means that there is a relationship between the two variables tested.

2.3 Correspondence Analysis

Correspondence analysis is a part of multivariate analysis that studies the relationship between two or more variables by demonstrating rows and columns simultaneously from two-way contingency tables in a low dimensional vector space (two). Correspondence analysis is used to reduce variable dimensions and describe row and column vector profiles of a data matrix from contingency tables [10]. The first stage of correspondence analysis is to change the contingency table into a correspondence matrix by using a proportion value.

$$P = \begin{bmatrix}
\frac{O_{11}}{N} & \frac{O_{12}}{N} & \ldots & \frac{O_{1c}}{N} \\
\frac{O_{21}}{N} & \frac{O_{22}}{N} & \ldots & \frac{O_{2c}}{N} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{O_{r1}}{N} & \frac{O_{r2}}{N} & \ldots & \frac{O_{rc}}{N}
\end{bmatrix}$$

(3)

$P$ is a correspondence matrix that is defined as a matrix whose elements are elements of matrix $N$ divided by the total number of elements of matrix $N$ and $O_{ij}$ is the marginal probability of the $i$-th row and $j$-th column. The diagonal matrix of the vector number of $r$ lines is the matrix $D_r$ with size $(i \times i)$, whereas the vector of column number $c$ is the matrix $D_c$ with size $(j \times j)$.

$$D_r = \text{diag}(r) = \begin{bmatrix}
p_{1r} & 0 & \ldots & 0 \\
0 & p_{2r} & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & p_{ir}
\end{bmatrix}$$

$$D_c = \text{diag}(c) = \begin{bmatrix}
p_{1c} & 0 & \ldots & 0 \\
0 & p_{2c} & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & p_{jc}
\end{bmatrix}$$

(4)

Next determine the row and column profile. Profile is the proportion of rows and columns of the correspondence matrix where each frequency of observation of the $i$-th row of the $j$-th column is divided by the sum of each total row and column respectively (Mattjik et al., 2011).
the matrix $R$ is called the row profile in dimension $b$ space where the inverse diagonal matrix of rows is multiplied by the matrix $P$, whereas the matrix $C$ is called the column profile in dimension a space where the inverse diagonal matrix column is multiplied by the matrix $P$ transpose, $P_i^r = \text{relative frequency of the } i\text{-th row}$, and $P_j^c = \text{relative frequency of the } j\text{-th column}$. The value of inertia shows the contribution of the $i$-th row of column $j$-th to total inertia, while the total inertia is the sum of the squares of the weighted (Greenacre & Michael, 2007).

The distance between the $i$-th row to the $j^{th}$ row cell

$$d_{ij} = \sum_{p_{ij}} \frac{1}{\sqrt{p_{ij}}} \left( \frac{p_{ij}}{p_{i^*}} - \frac{p_{ij}}{p_{j^*}} \right)$$

where $p_{ij} = \text{relative frequency of the } i\text{-th row cell}$.
relative frequency of the $i$-th line, $p_{ri} = \text{relative frequency of the } j\text{-th column}$. Furthermore, to determine the accuracy of the relationship between one variable with another variable contingency coefficient is used by using the following equation (Greenacre & Michael, 1984):

$$ C = \sqrt{\frac{x^2}{N + x^2}} $$

(11)

where $C$ is a contingency coefficient, $x^2$ is a chi-square statistic, and $N$ is the number of sample population. Relative contribution or relative contribution (KR) is the range of a point that can be explained by its main axis.

Correlation of $k$-axis and $i$-th row $= \frac{(\text{total baris ke}-i)(Q_{ik})}{\text{inertia baris ke}-i}$

(12)

Correlation of $k$-axis and $j$-th column $= \frac{(\text{total kolom ke}-j)(Q_{jk})}{\text{inertia kolom ke}-j}$

(13)

where $Q_{ik} = \text{line profile coordinates to-i on the } k\text{-axis}$, and $Q_{jk} = \text{column profile coordinates to-j on the } k\text{-axis}$. Relatively contribution (KR) is the proportion of diversity explained by each point towards its main axis. Relatively contribution or correlation of line to-$I$ or column to-$j$ with component $k$ is axis contribution to inertia line to-$i$ or column to-$j$, stated in inertia percent to-$I$ or column to-$j$ in $k$-axis (Kartono, 1998).

Correlation of $k$-axis and $i$-th row $= \frac{(\text{total baris ke}-i)(Q_{ik})}{\text{inertia sumbu ke}-k}$

(14)

Absolute contribution (KM) is a diversity proportion that explained each of point against its main axis. Result of absolute contribution used for determine a point that into a factor or dimension with criteria that the point in to a factor is had a value or biggest proportion. Where $Q_{ik}$= profile's coordinat to-$i$ to axis-$k$ and $Q_{jk}$= coordinate profile column to-$j$ to axis (Anggraini, 2011).

Correlation of $k$-axis and $j$-th column $= \frac{(\text{total kolom ke}-j)(Q_{jk})}{\text{inertia sumbu ke}-k}$

(15)

3. Results and Discussion

The data used in this study are secondary data about types of crime that happened in Sleman Regency that which is reported live by public in Sleman Police. Types of crime that used in this research are stealing, ill-treatment, embezzlement, fraud, and aldutery. These six acts are the most happening in Sleman Regency during 9 months ago. The other type of crime is ignored in this research because its rarely occur in 9 months ago.

Criminal incident’s data taken based on record from police that reported live from public. Data that use in this research as much 334 suspects of crime and 255 incidents of crime that record in Sleman Police.
3.1 Mapping and Trends of Criminal Actors with Characteristics of Age, Gender, and Occupation

The data used in this study is secondary data so that there are no assumptions that support for the data, the data analysis process is to test chi-square. This test is used to see whether there is a relationship between the variables of the suspect of crime and their characteristics, namely age, gender, and occupation.

<table>
<thead>
<tr>
<th>Criminal offenders</th>
<th>Value</th>
<th>df</th>
<th>P-value</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age * type of crime</td>
<td>47.116</td>
<td>10</td>
<td>0.000</td>
<td>Reject H₀</td>
</tr>
<tr>
<td>Gender * Types of Criminal Acts</td>
<td>7.198</td>
<td>5</td>
<td>0.206</td>
<td>Failed to reject H₀</td>
</tr>
<tr>
<td>Occupation * Type of Crime</td>
<td>36.185</td>
<td>10</td>
<td>0.000</td>
<td>Reject H₀</td>
</tr>
</tbody>
</table>

Based on Table 2 it can be seen that the value \( p \)-value less than \( \alpha = 0.05 \) The available data does not support the null hypothesis that there is a relationship between the variable age (\( X_1 \)) against the variable type of crime (\( Y \)) and employment (\( X_3 \)) against the variable type of crime (\( Y \)). while for the gender variable (\( X_2 \)) with type of crime (\( Y \)) support the null hypothesis namely there is no relationship.

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Primary Inertia</th>
<th>Proportion of Inertia</th>
<th>Cumulative Inertia</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.449</td>
<td>17.677</td>
<td>17.677</td>
</tr>
<tr>
<td>2</td>
<td>0.352</td>
<td>14.079</td>
<td>32.046</td>
</tr>
</tbody>
</table>

Based on Table 2, it can be seen that the size of the line (dimension 1) on the variables that affect the other variables that will be used as a coordinate point is able to contribute 0.449 and the distribution of the coordinate points is 17.677\%, and the size of space (dimension 2) in the variables that affect other variables that will be used as a coordinate point are able to contribute 0.352 and the distribution of these coordinate points is 14.079\%
Figure 2 and Table 4 explains that the red circle shows that the criminal’s suspect against the types crime of stealing (Y₁) and ill-treatment (Y₂) tend to be done by students (P₃) of male sex (JK₁) with a fairly young age that is less than 25 years (U₁). Type of stealing is the most crime that often committed by children (Sudarsono, 1991). Delinquency is an expression from tension, frustration (Darajat, 1969). There are some etiologic of reason about juvenile delinquency among other a) the child lacks of attention and affection and parental education guidance especially from father, b) unmet physical and psychological need and c) the child never gets physical and mental exercise (Kartono, 1998).

Blue circles indicate that the types of crime is deception (Y₄) and adultery (Y₆) and the characteristics is female (JK₂) with an age of more than 40 years (U₃) and the status is not working (P₂). Basically women are just made slaves of biology (Docherty, 1988). A women tend to commit crime act influenced generative a pregnant, menstruation, and menopause, that happened because when a women be a pregnant mom, menstruation, and into a menopause periode there are hormonal change that
demand the fulfillment of certain needs and satisfaction (Abbott et al., 2006).

The type of embezzlement ($Y_3$) is related to the age of the offender between 25 to 40 years ($U_2$) and they’re work ($P_1$) circled in green indicating that embezzlers tend to be carried out by people who work with an age range of 25 to 40 years. Someone that have certain position tend to be by officials. The form of deviation leads to the embezzlement that began from interested that go out of their way (Saripa, 2015). Type of fraud ($Y_5$) is a criminal type that anyone can do without see the age, gender and their occupation.

3.2 Mapping and Incidence of Criminal Events with Characteristics of Crime Scene, Time, and Region

The data used in this study are secondary data so that there are no assumptions that support for the data, the data analysis process is to do a chi-square test. This test is used to see whether there is a relation between the variable crime events with their characteristics are the crime scene, time, and region.

<table>
<thead>
<tr>
<th>Table 5: Test Chi-Square Crime Incident</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suspect</td>
</tr>
<tr>
<td>Crime Scene</td>
</tr>
<tr>
<td>Time * Types of Crimes</td>
</tr>
<tr>
<td>Region * Type of Crime</td>
</tr>
</tbody>
</table>

Based on Table 5 it can be seen that the p-value is less than $\alpha = 0.05$ The available data does not support the null hypothesis that there is a relationship between the crime scene variable ($X_1$) and the crime type variable ($Y$) and time ($X_2$) with the crime type variable ($Y$), whereas for region variable ($X_3$) with type of crime variable ($Y$) support the null hypothesis namely it doesn’t have relationship between both of them.

<table>
<thead>
<tr>
<th>Table 6: Inertia Value for Criminal Incident</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dimension</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 7: Table of Incident of Crime Against Types of Crime</th>
</tr>
</thead>
<tbody>
<tr>
<td>Circle</td>
</tr>
<tr>
<td>Blue</td>
</tr>
<tr>
<td>Green</td>
</tr>
<tr>
<td>Red</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
Based on Table 6, it can be seen that the size of the line (dimension 1) on the variables that influence the other variables that will be used as a coordinate point is able to contribute 0.354 and the distribution of these coordinate points is 10.906% and the size of the space (dimension 2) on the variable variables that affect other variables that will be used as a coordinate point can contribute 0.344 and the distribution of these coordinate points is 10.596%.

Figure 3: Mapping Insidents of Crime Against Types of Crime

Figure 2 and Table 7 explain that the green circle shows the types of crime of ill-treatment ($Y_2$) and adultery ($Y_6$) is to the right of the origin, so that the ill-treatment ($Y_2$) and adultery ($Y_6$) have similar characteristics from the crime scene and time. The crime scene from ill-treatment ($Y_2$) and adultery ($Y_6$) is shop (TKP$_3$). Meanwhile the prone time is occur at 00:00PM (T$_1$) and at 18:00PM-23:59PM (T$_4$). The shops and supermarket are the prone place apart from automated teller machine (Polda, 2019). The crime tend to do from 22.00PM when many public will take a rest (Putra, 2019).

The blue circle has showed the type of crime ($Y_1$) and fraud ($Y_4$) that have relation with time of incident, crime scene, and region. The crime scene of this act at 06.00AM until 11:59AM (T$_2$) and 12:00AM-17:59PM (T$_3$) with the region is hierarchy III (W$_2$) and hierarchy IV (W$_3$). One of the public place that often occur the crime is highway. According to a article that was upload by 100KPJ.com in November 2019 said if highway is a public place in Indonesia that not save and not comfortable from crime aspect (Soedibyo, 2019). Yogyakarta City is one of a big city that part of the area is full residential and susceptible of crime. In 2011 the crime is dominated by big fraud as much 58%, while in 2012 is dominated by stealing as much 55% and fraud as much 37%. It showed if the majority of crimes occur in settlements that where is a
central of economy and gathering of people (Yuliandarmaji, 2014).

The red circle is a type of crime by embezzlement ($Y_3$) and fraud ($Y_5$) located close together, while this type has similar characteristic. The crime scene ($Y_3$) and the embezzlement ($Y_5$) is institution. There are several etiology among them are economic factor, desire factor, chance factor, and weak faith factor (Lubis, 2017). This factors influenced the embezzlement in a institution, the suspect usually have a position while can utilize and abuse the position for personal interests.

4. Conclusion

The conclusions obtained from the study conducted based on the results of the analysis and discussion are as follows:

1. The tendency of suspect of crime with some characteristic namely age, gender and occupation that often occur in Sleman Regency during January- September 2019th are:
   a. The tendency from stealing and ill-treatment performed by student with male gender and their age less than 25 year old
   b. The tendency from fraud and adultery tend to do by women that more than 40 year old and unemployment
   c. The tendency from embezzlement tend to do by a official with age range between 25-40 year old

2. The tendency of criminal incident with characteristic by crime scene, time and region that occur in Sleman Regency from January- September 2019th are:
   a. The tendency from ill-treatment and adultery tend to be in shop with time range between 00:00PM-05:59AM and 18:00PM-23:59PM
   b. The tendency from stealing and fraud tend to be in settlement in hierarchy III and IV with time range between 06.00AM-11:59AM and 12.00AM-17:59PM
   c. The tendency from embezzlement and fraud tend to be in a institution

References


