Analysis of Traffic Accident Correspondence in Central Java Province

Regita Nurul Fitriani¹, Nur Khasanah²

¹² Mathematics Department, Faculty of Science and Technology, UIN Walisongo Semarang Jl. Prof. Dr. Hamka Nomor 3-5, Tambakaji, Ngaliyan Districts, Semarang City 50185

Korespondensi: Nur Khasanah, Email: khasanah.nur@walisongo.ac.id

Abstract

The transportation sector can facilitate and accelerate human activities. However, the development of the transportation sector has a negative impact that causes traffic jams and causes traffic accidents. Traffic accidents can cause fatalities which are divided into 3 categories, namely minor injuries, serious injuries, and death. This study aims to analyze the relationship between district or city variables in the province of Central Java with the status of the category of accident victims that occurred in 2019 – 2020 as the main source data on this study. This study uses R Studio software as a tool to facilitate data analysis. The method used is correspondence analysis of secondary of total amount accident data mentioned before by representing the results obtained through the correspondence graph that is formed. In two consecutive years, accidents were dominated by victims of the minor injury status category in the city of Surakarta with a percentage of 94,7% in 2019 and 94,6% in 2020.

Keywords: transportation, minor injuries, serious injuries, death, R software

Introduction

The field of transportation is a technology that supports the movement of goods or people from one place to another. The development of the transportation sector can facilitate and accelerate human activities in terms of moving goods and traveling. However, behind its advantages, this field also has other impacts, such as traffic jams due to too many vehicles and causing traffic accidents [1].

As we know, human populations continue to rise over time. This has certainly created an increasing need for transportation in an area. Such increases can trigger a traffic accident that can cause injury or even death. The rise in cases of accidents is comparable to the increase in deaths. According to a law on the traffic and transport of Indonesian traffic accidents. Article 1 clause 24 act number 24 can be construed as an unpredictable and accidental road event involving vehicles and causing damage to
property [2]. In addition to property losses, traffic accidents can also cause trauma, disability, injury and even death to accident victims. Traffic accidents are certainly not easy to reduce but will continue to increase along with the increase in road construction and the movement of a vehicle [3].

Central Java Province has a fairly high traffic accident rate. Based on traffic accident data obtained from the Central Java Province BPS in each district or city there were a total of 34,792 accidents in 2019 and 27,633 accidents in 2020 [4]. The accident caused fatalities which were divided into three categories, namely death, serious injury, and minor injury. From the traffic accident, several accident victims applied for compensation to an agency known as PT Jasa Raharja.

PT Jasa Raharja provides basic protection through two social insurance programs, namely passenger accident insurance on public transportation which is implemented based on Law number 34 of 1964 concerning the Road Traffic Accident Fund. PT Jasa Raharja as one of the insurance businesses of non-bank financial institutions is becoming increasingly important. This is because, in addition to its business activities that protect the public, insurance is also an institution that collects funds sourced from the receipt of insurance premiums from the public and distributes them with claims [5].

Based on previous research [6] regarding traffic accidents conducting a clustering analysis using the K-Means method. The clustering with the K-Means method has the aim of finding the best quality cluster results. The results of his research can provide information regarding traffic accidents that often occur in the Tanah Laut Police area. Further research in [7] regarding the comparative analysis of the K-Means Clustering method and the a priori method of the factors causing traffic accidents in the Medan area. This aims to produce association and clustering patterns contained in the accident data in the area.

In research [8] on correspondence analysis to evaluate the level of consumer satisfaction. This study aims to evaluate product quality based on customer ratings. The more satisfied customers are with the product, the better the quality and reputation of the product. Another study in [9] regarding the use of simple correspondence analysis in mapping potential disaster areas in the province of Central Sulawesi. It aims to map areas with potential disasters in the province of Central Sulawesi. So that in the analysis several regions will be grouped based on the row profile value that is greater than the average. The results of the grouping can be the basis for the focus of the government and the community in tackling disasters that are dominant in their respective areas so as to minimize the impact when a disaster occurs.

This research uses the correspondence analysis method. Correspondence analysis is an exploratory method that allows researchers to analyze data and develop perceptual maps to improve interpretation and communication findings [10,11]. The purpose of those studies are to see whether there is a relationship between variables visually, as well as to see the closeness of a category on one variable to another category of variables [12].

Based on the previous studies on the analysis correspondence on some areas, then this study continued by focusing on the evaluation of analysis correspondence of traffic accident in Central Java Province as the main goal of this study.

Research Methods
The research method used is case studies in accidents that occur in The Central Java district or provincial cities in 2019-2020 as obtained from the official page of the https://jateng.bps.go.id. It was treated with corresponding analysis and with the assistance of R studio software. The measures taken by researchers are as follows.

**Categorical data in a two-directional contingency table**
A two-way conference table is a chart that contains observational data that involves two variables that are row and column variables. As for the appropriate tests to see if there is a link between the two category variables that the contingency table is Chi-Square Test [13]
\[ \chi^2 = \sum \frac{(O-E)^2}{E} \]  

(1)

where,

\( O \) : Frequency of observation

\( E \) : Expected frequency, with \( E = \frac{\text{Number of lines} \times \text{Number of columns}}{\text{amount of data}} \)

Moreover, the next observation and calculation is on about the test statistics used by the Chi-Square test, that are as the following hypothesis:

\( H_0 \) : There is no significant relationship between district or city variables and the accident victim status category variable

\( H_1 \) : There is a significant relationship between district or city variables and the accident victim status category variable

Hypothetical withdrawal can be made by looking at the test criteria set in Chi-Square Test. The test criteria on this part are described as if \( \chi^2 \) count \( \leq \chi^2 \) table \( H_0 \) accepted, and on the other hand if \( \chi^2 \) count \( > \chi^2 \) table \( H_0 \) reject it. This also could be represented by using test criteria with if Sig. \( \chi^2 \) count \( > \alpha \) \( H_0 \) accepted, and for the converse if Sig. \( \chi^2 \) count \( < \alpha \) \( H_0 \) rejected.

If significant, value \( \chi^2 \) count \((p-value) > \alpha \) table then \( H_0 \) accepted. Meaning there is no significant link between the variables of the district or city and the category of crash victim status. Where, as if a significant value. \( \chi^2 \) count \((p-value) > \alpha \) table then \( H_0 \) reject it. This suggests that there is a significant link between the variables of the district or city and those of the category of crash victims.

**Determine the probability table**

Correspondence matrix is a matrix that contains sized \( a \times b \) with elements \( x_{ij} \) as the frequency. First, a matrix was built \( P(a \times b) \) which aims to obtain a visualization of the original data rows and columns in a low dimension. The correspondence matrix \( P(a \times b) \) is defined as the relative frequency matrix of \( x \), then:

\[ p_{ij} = \frac{n_{ij}}{n} \]  

(2)

where,

\( p_{ij} \) : Correspondence matrix entries

\( n_{ij} \) : Marginal probability of row \( i \) and column \( j \)

\( n \) : The total number of frequencies of the correspondence matrix

**Compute a set of weights for the columns and the rows (masses)**

This step is constructing the properties of the mass or vector of the number of rows \( (r) \), the mass or vector of the number of columns \( (c) \) and the diagonal matrix of the elements of the vector number of rows \( r \) is \((D_r)\) temporary \((D_c)\) is a diagonal matrix of the elements vector of the number of columns \( c \) with the following detail.

\[ r = \sum_{j=1}^{b} p_{ij}, i = 1,2,\ldots,a \]  

(3)

\[ c = \sum_{i=1}^{a} p_{ij}, j = 1,2,\ldots,b \]  

(4)
\[
D_r = \text{diag}(r) = \begin{bmatrix}
p_1 & 0 & \Lambda & 0 \\
0 & p_2 & 0 & M \\
M & 0 & 0 & 0 \\
0 & 0 & \Lambda & p_a \\
\end{bmatrix},
D_c = \text{diag}(c) = \begin{bmatrix}
p_1 & 0 & \Lambda & 0 \\
0 & p_2 & 0 & M \\
M & 0 & 0 & 0 \\
0 & 0 & \Lambda & p_b \\
\end{bmatrix}
\]

where,
- \( r \) : Mass or vector number of rows
- \( c \) : Mass or vector number of columns
- \( D_r \) : Diagonal matrix of the elements of the vector number of rows \( r \)
- \( D_c \) : Diagonal matrix of the elements of the vector number of columns \( c \).

**Define row and column profile matrix**

The row and column profile matrices can be calculated by dividing the row and column vectors by their respective masses. Therefore, the row profile matrix and column profile matrix can be expressed by the formula:

\[
R = D_r^{-1}P
\]

\[
C = PD_c^{-1}
\]

where,
- \( R \) : Row profile matrix
- \( C \) : Column profile matrix
- \( D_r^{-1} \) : Inverse matrix \( D_r \)
- \( D_c^{-1} \) : Inverse matrix \( D_c \)
- \( P \) : Correspondence matrix

**Determine the coordinates of the row and column points**

Singular Value Decomposition (SVD) is one of the concepts of matrix algebra and the concept of eigen decomposition which consists of eigenvalues and eigenvectors [14].

\[
D_r^{-\frac{1}{2}}(P - rc')D_c^{-\frac{1}{2}} = UAV'
\]

\[
P - rc' = D_r^{\frac{1}{2}}UAV'D_c^{\frac{1}{2}}
= A\Lambda B'
\]

where,  
- \( U \) : Eigenvectors matrix  
- \( \Lambda \) : Diagonal \((\lambda_1, \lambda_2, \Lambda, \lambda_n)\)
- \( V' \) : Eigen value

whereas,
\[
A = D_r^{\frac{1}{2}}U
\]
\[
B = D_c^{\frac{1}{2}}V.
\]

**Determining the value of Inertial decomposition**

Determining the value of inertial decomposition can be done by calculating the sum of the squares of the singular values indicating the contribution of the row \( a \) and column \( j \) to the total inertia. Total inertia itself is the variance of the data obtained by the sum of the weighted squares of the distance to the center and mass [15].
The total inertia is written by the following equation:

$$\hat{\chi}^2 \over n = \sum_{i=1}^{b} p_j (c_j - r) \mathbf{D}^{-1} (c_j - r) = \sum_{i=1}^{k} \hat{\lambda}_{ij}^2$$  \hspace{1cm} (9)

where,

$\hat{\chi}^2$ : Chi-Square test value

$c_{ij}$ : column $-j$

$\hat{\lambda}_{ij}$ : row diagonal $-i$

**Specifies the Chi-Square distance to create a two-dimensional image**

The following formula is given to determine the Chi-Square distance that will be used to create a correspondence graph.

$$d_{ij}^2 = (r_i - r_j) \mathbf{D}^{-1} (r_i - r_j)$$  \hspace{1cm} (10)

where,

$d_{ij}^2$ : Chi-Square distance

$r_i, r_j$ : Row to $-i$, row to $-j$

**Results and Discussion**

In this section, the results and discussion of the research of this study is discussed with the following details.

**Chi-Square Test**

The Chi-Square test is used to determine whether there is a significant relationship between row variables and column variables [16].

*Table 1. Chi-Square Test Results*

<table>
<thead>
<tr>
<th>Year</th>
<th>Person Chi-Square Test</th>
<th>x-squared</th>
<th>Df</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>2019</td>
<td>1.507,5</td>
<td>68</td>
<td></td>
<td>&lt; 2.2e-16</td>
</tr>
<tr>
<td>2020</td>
<td>627,84</td>
<td>68</td>
<td></td>
<td>&lt; 2.2e-16</td>
</tr>
</tbody>
</table>

Source: Chi-Square Test Output Results with R Studio software

The level of significance used in this study is 5%. Based on the results of calculations with the help of R Studio software, it can be seen that the p-value in 2019 and 2020 < 0,05 i.e. 2.2e-16 < 0,05 then $H_0$ rejected. So it can be concluded that there is a relationship between row variables and column variables. Furthermore, to find out the pattern of variable relationships, it can be continued by doing a correspondence analysis.

**Row Profile Matrix**

The formation of a row profile matrix in this case provides information on the level of traffic accidents in each district or city based on the number of accident victims in each district or city.

*Table 2. Row Profile Matrix*

<table>
<thead>
<tr>
<th>District/City</th>
<th>2019</th>
<th>2020</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Accident Victim Status</td>
<td>Accident Victim Status</td>
</tr>
<tr>
<td></td>
<td>LB</td>
<td>LR</td>
</tr>
<tr>
<td>Banjarnegara</td>
<td>0.007</td>
<td>0.853</td>
</tr>
</tbody>
</table>
Based on the row profile matrix, it shows that the largest proportion is in the category of victims of minor injuries with a proportion of 0.946 in 2019 and 0.944 in 2020 in the city of Surakarta. Therefore, it can be concluded that of all districts/cities in Central Java province, most cases of accidents cause victims to suffer minor injuries.

### Column Profile Matrix

The formation of a column profile matrix in this case provides information on the level of traffic accidents in each district or city based on the number of accident victims in each accident victim status category. 

**Table 3. Column Profile Matrix and Mass**

<table>
<thead>
<tr>
<th>District/City</th>
<th>2019</th>
<th>2020</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LB</td>
<td>LR</td>
</tr>
<tr>
<td>Banjarnegara</td>
<td>0.052</td>
<td>0.018</td>
</tr>
<tr>
<td>Banyumas</td>
<td>0.072</td>
<td>0.037</td>
</tr>
<tr>
<td>Batang</td>
<td>0.135</td>
<td>0.021</td>
</tr>
<tr>
<td>Blora</td>
<td>0.020</td>
<td>0.018</td>
</tr>
</tbody>
</table>

www.fourier.or.id JURNAL FOURIER (2022) 11 78-87
Accidents that caused the most serious injuries were found in the Wonosobo district with the largest proportion of 0.260 in 2019 and the Klaten district with the largest proportion of 0.142 in 2020. The most minor injuries category was in Klaten district with the largest proportion of 0.075 in 2019 and in Banyumas district with the largest proportion of 0.073 in 2020. Furthermore, accidents that cause the most deaths are in the city of Banyumas with the largest proportion of 0.054 in 2019 and 0.062 in 2020.

**Row Profile Mass Value**

The mass value of the row profile provides information about the number of accident victims in each district or city as a whole regardless of the accident victim status category. From the profile row above, it can be seen that the largest mass value in 2019 was in Klaten District at 0.071, which means that in Klaten District there were the most traffic accidents in 2019. This shows the number of accidents in Klaten District in 2019 was 7, 11% (2,474 accidents) of the total accidents that occurred in 2019 numbered 34,792 accidents. While in 2020 there were in Banyumas District (0.071) which means that in Banyumas District there were the most traffic accidents in 2020. This shows that the number of accidents in Banyumas District in 2020 was 7.176% (1983 accidents) of the total accidents that occurred in 2020 as many as 27,633 cases.

**Column Profile Mass Value**
The mass value of the column profile provides information about the number of accident victims in each
district or city as a whole based on the category of accident victim status. From the column profile above,
it can be seen that the largest mass value for two consecutive years was traffic accidents that caused the
victims to suffer minor injuries. The percentage of accidents in the category of victims experiencing
minor injuries was 87.821% (30,555 victims) of the total accidents that occurred in 2019 as many as
34,792 cases. While in 2020 with a percentage of 87.236% (24,107 victims) of the total accidents that
occurred in 2020 as many as 27,633 cases.

### Total Inertia

Total inertia is the variance of the data obtained by the sum of the squares weighted to the center of mass.
The following is the value of the proportion of inertia in Table 4.

<table>
<thead>
<tr>
<th></th>
<th>2019</th>
<th></th>
<th>2020</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Value</td>
<td>%</td>
<td>Cum%</td>
<td>Value</td>
</tr>
<tr>
<td>1</td>
<td>0.024813</td>
<td>57.3</td>
<td>57.3</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0.018515</td>
<td>42.7</td>
<td>100</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>0.043328</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

We can see The Principal Inertia Cumulative in 2019 or the total inertia of dimension 1 of 57.3%.
This means that with 1 dimension is able to explain 57.3% of the diversity of the data so that later the
resulting image will be very representative of the actual configuration. While The Principal Inertia
Cumulative in 2020 dimension 1 is 85.2%. This means that dimension 1 contains 85.2% of information.

### Correspondence Graph

The correspondence graph produced in this mini-research explains how the relationship between district
or city variables and the accident victim status variable.

![Traffic Accident Correspondence Graph in Central Java Province in 2019](image)

Figure 1 shows a graph of the correspondence between districts or city variables and the status variable
of traffic accident victims in Central Java province. Dimension 1 and dimension 2 explain the diversity
of data. The information obtained from Figure 1 is that dimension 1 is 85.3% of the total singular value,
this means that dimension 1 contains information of 85.3% of the total data while dimension 2 contains
the remaining information.

The red triangle shows the accident victim status category which includes minor injuries, serious
injuries and death. While the gray circle indicates the category of Districts or City in Central Java
Province. It can be noticed that almost all of the gray circles for the Districts or City variable are close to
the red triangle for the status variable for the accident victim, died and suffered minor injuries.
Furthermore, in Figure 1 it can be seen that there are 13 districts or cities that are located very close to
the red triangle status of victims of minor injuries, namely Cilacap, Banyumas, Magelang, Boyolali,
Klaten, Sukoharjo, Karanganyar, Sragen, Pati, Kudus, Brebes Districts, Surakarta City and Semarang.
City. This shows that the 13 districts or cities in Central Java province have similarities in the number of accidents that result in minor injuries compared to other districts or cities. Furthermore, it can also be noted that there are 3 districts that are located quite close to the red triangle status of the accident victim who died, namely Kendal, Pemalang and Tegal District. This indicates that the 3 districts have similarities in the number of accident victims and fatalities compared to other districts in Central Java province.

Figure 2. Traffic Accident Correspondence Graph in Central Java Province in 2020

Figure 2 shows a graph of the correspondence between Districts or City variables and the status variable of traffic accident victims in Central Java province. Dimension 1 and dimension 2 explain the diversity of data. Figure 2 informs that dimension 1 is 85.2% of the total singular value, this means that dimension 1 contains information of 82.3% of the total data while dimension 2 contains the remaining information.

The red triangle shows the accident victim status category which includes minor injuries, serious injuries, and death. While the gray circle indicates the category of districts or cities in Central Java Province. It can be seen that almost all the blue circles for the districts or cities variables are close to the red circles for the status variable for the accident victim, who died and suffered minor injuries. Furthermore, in Figure 2 it can be seen that there are 11 districts or cities that are located very close to the red circle with the status of victims of minor injuries, namely Wonogiri, Sukoharjo, Banyumas, Purworejo, Temanggung, Tegal, Kebumen, Jepara, Purbalingga, Cilacap, and Surakarta City. This shows that the 11 districts or cities in Central Java Province have similarities in the number of accidents that result in minor injuries compared to other districts or cities. Furthermore, it can also be noted that 2 districts that are located quite close to the red circle with the status of a deceased accident victim, namely Pekalongan districts and Cilacap Districts. This indicates that in the 2 districts there are similarities in accidents with fatalities compared to other districts in Central Java Province.

Conclusions

In 2019, the highest number of accident victims was in Klaten district with a percentage of 7.11%, the category of minor injuries in Surakarta district was 94.7%, serious injuries in Wonosobo district was 26.04% and the category of death in Banyumas district was 5.4%. While in 2020 the highest total accident victims were in Banyumas district with a percentage of 7.17%, the category of minor injuries in Surakarta district was 94.6%, serious injuries in Klaten district was 14.2% and the category of death in Banyumas district was 6.2%. The highest number of accident victims was in the category of minor injuries with a percentage of 87.83% in 2019 and 87.23% in 2020.

Dimensions 1 and 2 in the correspondence graph illustrate the diversity of the data. The red triangle represents the category of accident victims which includes serious injuries, minor injuries and death. While the gray circles represent districts or cities in the province of Central Java. The correspondence graph for 2019 contains information that there are 13 districts or cities that have similarities in the number of minor injuries, namely Cilacap, Banyumas, Magelang, Boyolali, Klaten, Sukoharjo, Karanganyar, Sragen, Pati, Kudus, Brebes, Surakarta and Semarang cities. Meanwhile, for accident victims who died
there were 3 districts that had similarities, namely Kendal, Pemalang and Tegal districts. While the accident correspondence chart in 2020 contains information that there are 11 districts or cities that have similarities in the number of minor injuries, namely Wonogiri, Sukoharjo, Banyumas, Purworejo, Temanggung, Tegal, Kebumen, Jepara, Purwalingga, Cilacap and Surakarta districts. In the category of dead victims, there are 2 districts that have similarities, namely Pekalongan and Cilacap districts. The results of the interpretation of this correspondence graph can be used as material for decision making or policy on government efforts in overcoming the number of accidents that exist in various districts or cities in the province of Central Java. This can be done by focusing on areas that have a fairly high number of accident victims.

Referensi