International Journal of Mathematics and Statistics Studies, 11 (3), 27-46, 2023
Print ISSN: 2053-2229 (Print),
Online ISSN: 2053-2210 (Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

# Optimal Scaling Categorical Principal Components Analysis: Road Traffic KSI Car Accidents in England (STATS19)

**Mohammad M R Sheikh**
(Statistical Researcher, School of Mathematics and Computing, Kingston University London;
Phone: +4407723382812, email: m.sheikh@kingston.ac.uk)

**Citation**: Sheikh MMR (2023) Optimal Scaling Categorical Principal Components Analysis: Road Traffic KSI Car Accidents in England (STATS19), *International Journal of Mathematics and Statistics Studies*, 11 (3), 27-46

**ABSTRACT:** *Categorical principal component analysis (CATPCA) technique was applied in the road killed or seriously injured (KSI) car accidents in England based on STATS19 data so that the categorical variables of KSI car accidents can be transferred into few components with reduction of dimensionality. Finally selected 20 variables in KSI car accident database were divided to create four principal components by applying "optimal scaling CATPCA" procedure in SPSS. The statistically significant KSI car accident variables, particularly the most accountable categorical variables, were identified and quantified for developing models as well as leading to aims to reduce as well as to prevent the car accidents, particularly the KSI car accidents. It also leads to map out the possible safety improvement strategies as well as to inform the policymakers on how best to reduce the number and severity of car crashes*.

**KEYWORDS:** KSI, CATPCA, optimal scaling, dimensionality, principal components, quantification

## INTRODUCTION

This research is to model road traffic KSI car accidents in England based on STATS19 database by applying 'categorical principal component analysis (CATPCA)' method, so that the variables in KSI car accidents can be transferred into few components with reduction of dimensionality. CATPCA performs principal components analysis on a set of variables that can be given mixed optimal scaling levels, and the relationships among observed variables are not assumed to be linear. This technique quantifies categorical variables alongside while reducing the dimensionality of the data. The goal of CATPCA is to reduce an original set of categorical variables into a smaller set of uncorrelated components representing most of the information observed in the original variables. The technique is most useful when many categorical variables forbid effective interpretation of the relationships between objects (i.e., subjects and units), and a few components rather than many variables are interpreted by reducing the dimensionality. The optimal-scaling approach allows categorical variables to be scaled at different levels so that categorical variables are optimally quantified in the specified dimensionality, and nonlinear relationships between categorical variables can be modelled.

The aims/ objectives of the research study are as follows:

a) To reduce the dimensionality of categorical variables in killed or seriously injured (KSI) car accidents.
b) To quantify the most accountable categorical variables involved in KSI car accidents.
c) To develop models by applying categorical principal component analysis method.
d) To inform policymakers on how best to reduce the number and severity of car crashes.

The key research questions area as follows:

A) What are the most significant variables in road KSI car accidents?
B) How can these KSI car accidents be reduced?

The structure of the study is detailed as **Section 2** is research methods including data source, CATPCA method, defining scale and weight of variables, discretisation, missing vale option, output/ loading plots etc. **Section 3** is a brief of KSI car accident database construction and data manipulation extracted from STATS19 database. **Section 4** is CATPCA application in KSI car accidents, including design, variable selection, iteration, quantification, variance accounted for (VAF), rotated variance accounted for (RVAF), correlation matrix of original variables as well as transformed variables, and component loadings as well as rotated component loadings. **Section 5** is study findings, discussions, and conclusion.

## 2. RESEARCH METHODS

### 2.0. Data Source

The study is to use the secondary data from DfT-STATS19 database developed by UK Police and Department of Transport, UK. The DfT has undertaken work to link data from STATS19. Data for KSI car accidents in England are to be extracted from the DfT-STATS19 database. This secondary data covers the period from 1979 through 2015.

### 2.1. CATPCA Method

The CATPCA simultaneously not only quantifies categorical variables, but also decreases the dimensionality of the data. It is to bring down an original set of variables into a smaller set of uncorrelated components, representing most of the information found in the original variables. The technique is most useful in STATS19 data as many variables exclude effective interpretation of the relationships between subjects and units. The optimal-scaling approach allows variables to be scaled at different levels so that categorical variables are optimally quantified in the specified dimensionality as well as nonlinear relationships between variables can be modelled.

In CATPCA data considerations, string variable values are always converted into positive integers by ascending alphanumeric order, where user-defined missing values or system-missing values or any values less than 1 are considered missing. The data contain at least three valid cases mandatorily and it must be positive integer. The discretization option is to automatically categorise a fractional-valued variable by grouping its values into categories with a close to "normal" distribution and it is to automatically convert values of string variables into positive integers.

### 2.2. Defining Scale and Weight of Variables in CATPCA

Optimal scaling level is set for analysis variables and supplementary variables, as well as the weight for analysis variables can be set specifying a positive integer. The scaling level to be used to quantify each variable can also be selected so that a) the order of the categories of the observed ordinal variable is preserved in the optimally scaled variable; b) the order of the categories of the observed nominal variable is not preserved; c) the order of the categories and the equal distances between category numbers of the observed discrete variable are preserved in the optimally scaled variable.

### 2.3. Discretisation in CATPCA

The variables are to be discretised by applying 'grouping method', where a specified number of categories is recoded for categorical variables as well as discrete variables are recoded into categories defined by equally sized intervals.

### 2.4. Missing Values in CATPCA

Missing values with the quantification of an extra category can be replaced so that objects with a missing value on the variable are considered to belong to the same (extra) category.

### 2.5. Options in CATPCA

The Options provide the controls to select the initial configuration, specify iteration and convergence criteria, as well as to select a normalisation method, to choose the method for labelling plots, and to specify supplementary objects. The case number of the object, or the first and last case numbers of a range of objects to make supplementary, is specified if it is needed. If an object is specified as supplementary, then case weights are ignored for that object. One of [1]five options (i.e., Variable Principal, Object Principal, Symmetrical, Independent and Custom) for normalising the object scores and the variables, is specified. It is noted that only one normalisation method can be used in each analysis.

The maximum number of iterations can be specified so that the procedure can go through in its computations. A convergence criterion value can also be selected. The algorithm stops iterating when the difference in total fit between the last two iterations is less than the convergence value or if the maximum number of iterations is reached. Variables and value labels or variable names

[1] **Normalising Methods**:

a) **Variable Principal** option is applied to optimise the association between variables so that the coordinates of the variables in the object space are the component loadings (correlations with principal components, such as dimensions and object scores).
b) **Object Principal** option is applied to optimise the distances between objects.
c) **Symmetrical** option is used for the relationship between objects and variables.
d) **Independent** option is used to examine distances between objects and correlations between variables separately.
e) **Custom** option specifies any real value in the closed interval $[-1, 1]$ where value of '1' is for the Object Principal option; '0' is for the Symmetrical option; '–1' is for the Variable Principal option. Any value greater than '–1' or less than '1' is specified so that the eigenvalue over both objects and variables are spreaded.

and values can be specified to be used in the plots. All dimensions in the solution are displayed in a scatterplot matrix. One of [2]five options for rotation method (i.e., Varimax, Quartimax, Equamax, Oblimin and Promax) to obtain rotated results, can be selected.

### 2.6. Output and Loading Plots in CATPCA

1) Displays the component loadings for all variables that were not given multiple nominal scaling levels. The component loadings by size can be sorted. The component loadings can be displayed in Scatter Matrix.
2) Shows for each iteration, the variance accounted for, loss, and increase in variance accounted for (VAF).
3) Shows the correlation matrix of the original variables and the eigenvalues of that matrix.
4) Shows the correlation matrix of the transformed (optimally scaled) variables and the eigenvalues of that matrix.
5) Displays the amount of variance accounted for (VAF) by centroid coordinates, vector coordinates, and total (centroid and vector coordinates combined) per variable and per dimension. The VAF can be shown in Bar Charts.
6) Gives the category quantifications and coordinates for each dimension of the variables that are selected.

## 3. Road KSI Car Accident Data Manipulation and Database Construction

The reported/ recorded data of road traffic KSI car accidents extracted from the datasets of DfT-STATS19 under Department for Transport for the period of 1979-2015, had been used for database construction and then, the data in the databases were manipulated individually using 'data function' and 'transform function' of SPSS 26.0.1 version and onwards following research methodology as briefed in the previous section (i.e., Section 2).

Two accident databases based on DfT-STATS19 database, were constructed containing the existing variables and the new variables computed from existing variable(s) following the database design such as determining the purpose of the database, findings, and organising the information required, dividing the information items into tables, turning information items into

**[2] Rotated Methods:**

f) **Varimax** orthogonal rotation method is to minimise the number of variables to have high loadings on each component.
g) **Quartimax** rotation method is to minimise the number of components that are needed to explain each variable.
h) **Equamax** rotation method is a combination of the Varimax method, and the Quartimax method so that the number of variables loading highly on a component and the number of components needing to explain a variable, are minimised.
i) **Oblimin** is a method for oblique (non-orthogonal) rotation. The value of Delta must be less than or equal to 0.8.
j) **Promax** is an oblique (non-orthogonal) rotation, to allow the components to be correlated. The amount of correlation (obliqueness) that is allowed is controlled by the kappa parameter. The value of **Kappa** must be greater than or equal to 1 and less 10,000.

columns, specifying the primary keys, applying the normalisation rules, refining the design and setting up the table relationship.

Total 49 variables, containing 10 discrete variables as well as 13 dichotomous, 14 nominal and 12 ordinals, were included in road KSI car accident database shown in Table 3.

Table 3. Valid/ Missing KSI cases in Road Car Accident Data based in Accident Index in England

| Valid and Missing KSI Cases in Road Traffic Car Accident Database based on Accident Index in England | | | | | | | |
|---|---|---|---|---|---|---|---|
| # | Variable/ Factor | Type | Valid Case | Valid Case % | Missing Case | Missing Case % | Other Case | Other Case % |
| 1 | Accident Year | Continuous | 1089560 | 100.00 | 0 | £0.00 | 0 | 0.00 |
| 2 | First Road Number | Continuous | 815907 | 74.90 | 273653 | £25.10 | 0 | 0.00 |
| 3 | Latitude | Continuous | 332388 | 30.51 | 757172 | £69.49 | 0 | 0.00 |
| 4 | Longitude | Continuous | 332388 | 30.51 | 757172 | £69.49 | 0 | 0.00 |
| 5 | Number of Casualty per KSI Accident | Continuous | 1089560 | 0.00 | 0 | £0.00 | 0 | 0.00 |
| 6 | Number of Vehicle per KSI Accident | Continuous | 1089560 | 0.00 | 0 | £0.00 | 0 | 0.00 |
| 7 | OSGR Easting | Continuous | 1086224 | 99.69 | 3336 | £0.31 | 0 | 0.00 |
| 8 | OSGR Northing | Continuous | 1087796 | 99.84 | 1764 | £0.16 | 0 | 0.00 |
| 9 | Second Road Number | Continuous | 1069885 | 98.19 | 19675 | £1.81 | 0 | 0.00 |
| 10 | Speed Limit | Continuous | 1089560 | 100.00 | 0 | £0.00 | 74 | 0.00 |
| 11 | Built-non-Built-up Speed Area | Dichotomous | 1089557 | 100.00 | 3 | £0.00 | 0 | 0.00 |
| 12 | Carriageway Hazards-non-Hazards | Dichotomous | 1082347 | 99.30 | 7213 | £0.70 | 0 | 0.00 |
| 13 | First Classified-non-Classified Road Class | Dichotomous | 1089560 | 100.00 | 0 | £0.00 | 0 | 0.00 |
| 14 | First Classified Trunk-non-Trunk Road Class | Dichotomous | 797819 | 73.20 | 291741 | £26.80 | 0 | 0.00 |
| 15 | First Numbered-non-Numbered Road | Dichotomous | 1089505 | 100.00 | 55 | £0.00 | 0 | 0.00 |
| 16 | Junction Control-non-Control | Dichotomous | 607766 | 55.80 | 481794 | £44.20 | 0 | 0.00 |
| 17 | Junction-non-Junction Details | Dichotomous | 1089471 | 100.00 | 89 | £0.00 | 0 | 0.00 |
| 18 | Pedestrian Crossing at Human Control | Dichotomous | 3450 | 0.30 | 100 | £0.00 | 0 | 0.00 |
| 19 | Pedestrian Crossing at Human Control-non-Control | Dichotomous | 1089882 | 98.20 | 100 | £0.00 | 0 | 0.00 |
| 20 | Pedestrian Crossing at Physical-non-Physical Facilities | Dichotomous | 1089890 | 98.20 | 19670 | £1.80 | 0 | 0.00 |
| 21 | Police Officer's Attandance-non-Attendance at Accident Sc | Dichotomous | 321525 | 29.50 | 768035 | £70.50 | 0 | 0.00 |
| 22 | Road Environment Urban-non-Urban | Dichotomous | 479862 | 44.00 | 609698 | £56.00 | 0 | 0.00 |
| 23 | Special Conditions-non-Conditions at Site | Dichotomous | 1067904 | 98.00 | 21656 | £2.00 | 0 | 0.00 |
| 24 | Carriageway Hazards (5-Level) | Nominal | 18983 | 1.70 | 7213 | £0.70 | 1063364 | 97.60 |
| 25 | First Classified Road Class (5-Level) | Nomial | 797819 | 73.20 | 100 | £0.00 | 291741 | 26.80 |
| 26 | England Region (9-Level) | Nomial | 1089560 | 100.00 | 100 | £0.00 | 0 | 0.00 |
| 27 | Junction Control (4-Level) | Nomial | 596831 | 54.80 | 481794 | £44.20 | 10935 | 1.00 |
| 28 | Junction Details (8-Level) | Nomial | 597571 | 54.80 | 89 | £0.00 | 491900 | 45.10 |
| 29 | Lights Conditions (5-Level) | Nomial | 1089413 | 100.00 | 147 | £0.00 | 0 | 0.00 |
| 30 | Pedestrian Crossing at Physical Facilities (5-Level) | Nomial | 134016 | 12.30 | 19670 | £1.80 | 935874 | 85.90 |
| 31 | Police Officer's Attandance at Accident Scene (3-Level) | Nominal | 321525 | 29.50 | 768035 | £70.50 | 0 | 0.00 |
| 32 | Road Environment (3-Level) | Nomial | 479862 | 44.00 | 609698 | £56.00 | 0 | 0.00 |
| 33 | Road Surface Conditions (7-Level) | Nomial | 1088290 | 99.90 | 1270 | £0.10 | 0 | 0.00 |
| 34 | Road Type (5-Level) | Nomial | 1070965 | 98.30 | 18595 | £1.70 | 0 | 0.00 |
| 35 | Second Road Class (5-Level) | Nomial | 578347 | 53.10 | 511213 | £46.90 | 0 | 0.00 |
| 36 | Special Conditions at Site (7-Level) | Nomial | 19691 | 1.80 | 21656 | £2.00 | 1048213 | 96.20 |
| 37 | Weather Conditions (8-Level) | Nomial | 1089460 | 100.00 | 100 | £0.00 | 0 | 0.00 |
| 38 | Accident Day (7-Level) | Ordinal | 1089560 | 100.00 | 100 | £0.00 | 0 | 0.00 |
| 39 | Accident Month (12-Level) | Ordinal | 1089560 | 100.00 | 100 | £0.00 | 0 | 0.00 |
| 40 | Accident Time (12-Level) | Ordinal | 1089472 | 100.00 | 88 | £0.00 | 0 | 0.00 |
| 41 | Car per KSI Accident (3-Level) | Ordinal | 1089560 | 100.00 | 100 | £0.00 | 0 | 0.00 |
| 42 | Casualty per KSI Car Accident (3-Level) | Ordinal | 1089560 | 100.00 | 100 | £0.00 | 0 | 0.00 |
| 43 | First Numbered Road Digit (6-Level) | Ordinal | 815907 | 74.90 | 55 | £0.00 | 273598 | 25.10 |
| 44 | First Numbered Road Zone (4-Level) | Ordinal | 793467 | 72.80 | 22495 | £2.10 | 273598 | 25.10 |
| 45 | Latitude Band (5-Level) | Ordinal | 332384 | 30.50 | 757176 | £69.50 | 0 | 0.00 |
| 46 | Longitude Band (5-Level) | Ordinal | 332388 | 30.50 | 757172 | £69.50 | 0 | 0.00 |
| 47 | OSGR Easting Band (5-Level) | Ordinal | 1086224 | 99.70 | 3336 | £0.30 | 0 | 0.00 |
| 48 | OSGR Northing Band (7-Level) | Ordinal | 1087796 | 99.80 | 1764 | £0.20 | 0 | 0.00 |
| 49 | Speed Limit Zone (8-Level) | Ordinal | 1089557 | 100.00 | 3 | £0.00 | 0 | 0.00 |

## 4. CATPCA in KSI Car Accidents

By following El-Basyouny and Sayed (2009), Harrel (2015), Kleinbaum and Klein (2010), Saukani and Ismail (2019), and Hung and Tai-Jin. (2018), categorical principal component analysis (CATPCA) method was applied in KSI car accidents.

### 4.0. CATPCA Design

The CATPCA based on "optimal scaling level for one set of variables that some factor(s) are not niminal" was run on a 20-factor/ factor dataset that measured desired 'road traffic KSI car

accident' characteristics on 1089560 cases/ observations. The suitability of CATPCA was assessed prior to analysis. Inspection of the correlation matrix showed that all variables had at least one correlation coefficient greater than 0.3. Grouping method and normal distribution were used for discretisation. The methods such as 'factor principal' and 'Varimax for rotation with Kaiser nomalisation' were applied. Each factor was defined/ scaled as '1' for its weight that was analysed through optimal scaling level. Preliminary assumption checking revealed: variables were measured in ordinal/ nominal/ dichotomous formats; there were linear relationships between all variables, as assessed by scatterplots; the data had sampling adequacy (large sample size, $n = 1089560$ ); the dataset were suitable for data reduction; there should be no significant outlier(s), as assessed by boxplot.

### 4.1. Variables in CATPCA of Car Accidents

1089560 cases of 20 variables were used in analysis, where the common cases that were valid were 579496.

a) Only one discrete variable such as 'accident year' was used.
b) Seven dichotomous variables such as 'built-non-built up speed area', 'carriageway hazards-non hazards', 'junction-non-junction details', 'junction control-non-control', 'pedestrian crossing at human control-non-control', 'pedestrian crossing at physical-non-physical facilities', and 'special conditions-non-conditions', were used.
c) Four nominal variables such as 'England region', 'lighting conditions', 'road type', and 'weather conditions', were used.
d) Eight ordinal variables such as 'OSGR easting band', 'OSGR northing band', 'speed limit zone', 'casualty per accident', 'car per accident', 'accident time', 'accident day', and 'accident month', are used.

### 4.2. Iterations in CATPCA of Car Accidents

The process stopped at the $34^{th}$ iteration because the convergence test value was reached. The iteration history has been displayed in Table 4.2. The $0^{th}$ iteration displays the solutions of the statistics with all variables. It displays the eigenvalues for each iteration of the analysis. These are used to determine the percentage of variance accounted for (a type of effect size) and therefore, larger eigenvalues are preferred over smaller ones to get a better solution (higher eigenvalue).

Table 4.2. Iteration History in CATPCA of KSI Car Accidents in England

**Iteration History in CATPCA of KSI Car Accidents in England**

| Iteration Number | Variance Accounted For | | Loss | | |
|---|---|---|---|---|---|
| | Total | Increase | Total | Centroid Coordinates | Restriction of Centroid to Vector Coordinates |
| 0 | 6.33255 | 0.00151 | 73.667 | 72.134 | 1.5329 |
| 1 | 7.09878 | 0.76623 | 72.901 | 72.134 | 0.7667 |
| 2 | 7.37783 | 0.27905 | 72.622 | 71.793 | 0.8291 |
| 3 | 7.47787 | 0.10004 | 72.522 | 71.718 | 0.8044 |
| 4 | 7.53631 | 0.05844 | 72.464 | 71.683 | 0.7810 |
| 5 | 7.57372 | 0.03741 | 72.426 | 71.662 | 0.7639 |
| 6 | 7.60111 | 0.02739 | 72.399 | 71.646 | 0.7524 |
| 7 | 7.62449 | 0.02338 | 72.376 | 71.629 | 0.7461 |
| 8 | 7.64729 | 0.02280 | 72.353 | 71.608 | 0.7446 |
| 9 | 7.67190 | 0.02460 | 72.328 | 71.581 | 0.7476 |
| 10 | 7.70020 | 0.02830 | 72.300 | 71.545 | 0.7546 |
| 11 | 7.73325 | 0.03306 | 72.267 | 71.502 | 0.7650 |
| 12 | 7.77106 | 0.03781 | 72.229 | 71.452 | 0.7774 |
| 13 | 7.81209 | 0.04103 | 72.188 | 71.398 | 0.7900 |
| 14 | 7.85318 | 0.04109 | 72.147 | 71.345 | 0.8013 |
| 15 | 7.89036 | 0.03718 | 72.110 | 71.299 | 0.8108 |
| 16 | 7.92060 | 0.03024 | 72.079 | 71.261 | 0.8185 |
| 17 | 7.94300 | 0.02240 | 72.057 | 71.232 | 0.8245 |
| 18 | 7.95850 | 0.01550 | 72.042 | 71.212 | 0.8293 |
| 19 | 7.96873 | 0.01023 | 72.031 | 71.198 | 0.8329 |
| 20 | 7.97518 | 0.00646 | 72.025 | 71.189 | 0.8358 |
| 21 | 7.97916 | 0.00398 | 72.021 | 71.183 | 0.8381 |
| 22 | 7.98161 | 0.00245 | 72.018 | 71.179 | 0.8399 |
| 23 | 7.98313 | 0.00152 | 72.017 | 71.176 | 0.8413 |
| 24 | 7.98408 | 0.00095 | 72.016 | 71.174 | 0.8424 |
| 25 | 7.98467 | 0.00060 | 72.015 | 71.172 | 0.8432 |
| 26 | 7.98506 | 0.00038 | 72.015 | 71.171 | 0.8439 |
| 27 | 7.98530 | 0.00024 | 72.015 | 71.170 | 0.8444 |
| 28 | 7.98546 | 0.00016 | 72.015 | 71.170 | 0.8449 |
| 29 | 7.98556 | 0.00010 | 72.014 | 71.169 | 0.8452 |
| 30 | 7.985624 | 0.000067 | 72.014 | 71.169 | 0.8455 |
| 31 | 7.98567 | 0.000044 | 72.014 | 71.169 | 0.8457 |
| 32 | 7.98570 | 0.000028 | 72.014 | 71.168 | 0.8458 |
| 33 | 7.98571 | 0.000014 | 72.014 | 71.168 | 0.8459 |
| 34 | 7.98572 | 0.000008 | 72.014 | 71.168 | 0.8459 |

Iteration 0 displays the statistics of the solution with all variables, except variables with optimal scaling level Multiple Nominal, treated as numerical.

The iteration process stopped at 34th because the convergence test value was reached.

### 4.3. Quantification of Variables in CATPCA of Car Accidents

The quantification for dichotomous variables followed by nominal and ordinal variables, is detailed and displayed in Table 4.3 that display the frequency, the quantification value assigned, the centroid coordinates, and the vector coordinates of each response category for each item.

Table 4.3. Quantification in CATPCA of KSI Car Accidents in England

**Quantification in CATPCA of KSI Car Accidents in England**

| Variable Type | Variable | Category | Frequency | Quantification | Centroid Coordinates Dimension 1 | 2 | 3 | 4 | Vector Coordinates Dimension 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Dichotomous | Built-non-Built up Speed Area | non-Built up Speed Area | 355803 | -1.448 | -1.252 | -0.052 | -0.287 | -0.129 | -1.254 | -0.054 | -0.269 | -0.146 |
| | | Built up Speed Area | 733954 | 0.686 | 0.597 | 0.027 | 0.109 | 0.066 | 0.595 | 0.026 | 0.127 | 0.069 |
| | | Missing | 3 | | | | | | | | | |
| | Carriageway Hazards-non-Hazards | No Hazards | 1083364 | -0.025 | -0.002 | 0.008 | 0.001 | 0.016 | <0.001 | 0.001 | 0.003 | <0.001 |
| | | Hazards | 18863 | 7.574 | -0.136 | -0.359 | -0.924 | 0.040 | -0.137 | -0.360 | -0.924 | 0.037 |
| | | Missing | 7213 | | | | | | | | | |
| | Junction Control-non-Control | Not at Junction or Within 20 Metres | 10935 | -3.530 | -0.522 | 0.436 | -1.993 | -0.551 | -0.703 | 0.032 | -1.971 | -0.202 |
| | | At Junction or Within 20 Metres | 596831 | 1.264 | 0.261 | 0.009 | 0.704 | 0.054 | 0.252 | -0.011 | 0.706 | 0.072 |
| | | Missing | 481794 | | | | | | | | | |
| | Junction-non Junction Details | Not at Junction or within 20 Metres | 491900 | -1.126 | -0.331 | -0.009 | -0.000 | -0.031 | -0.331 | -0.010 | -0.899 | -0.047 |
| | | At Junction or within 20 Metres | 597571 | 0.883 | 0.269 | 0.010 | 0.704 | 0.054 | 0.269 | 0.008 | 0.706 | 0.037 |
| | | Missing | 99 | | | | | | | | | |
| | Pedestrian Crossing at Human Control-non-Control | Not at Crossing or within 50 Metres | 1066432 | 1.011 | -0.011 | -0.008 | -0.017 | -0.089 | -0.011 | -0.008 | -0.017 | -0.089 |
| | | At Crossing or within 50 Metres | 3450 | -0.292 | 0.730 | -0.068 | 0.030 | -0.063 | 0.003 | 0.002 | 0.005 | 0.026 |
| | | Missing | 19678 | | | | | | | | | |
| | Pedestrian Crossing at Physical-non-Physical Facilities | Not at Physical Crossing Facilities or within 50 Metres | 935874 | -0.392 | -0.164 | -0.050 | -0.047 | -0.073 | -0.160 | -0.042 | -0.031 | 0.016 |
| | | At Physical Crossing Facilities or within 50 Metres | 134016 | 2.656 | 1.077 | 0.279 | 0.191 | -0.200 | 1.081 | 0.287 | 0.208 | -0.108 |
| | | Missing | 19670 | | | | | | | | | |
| | Special Conditions-non-Conditions at Site | No Conditions | 1048213 | 0.988 | -0.005 | -0.009 | -0.015 | -0.097 | -0.014 | -0.010 | -0.017 | -0.095 |
| | | Special Conditions | 19891 | 1.891 | -0.287 | -0.053 | -0.063 | -0.135 | -0.028 | -0.019 | -0.032 | -0.183 |
| | | Missing | 21656 | | | | | | | | | |
| Nominal | England Region | North-East Region | 43070 | -1.452 | 0.142 | -1.400 | -0.069 | -0.114 | -0.152 | -1.358 | -0.053 | -0.184 |
| | | North-West Region | 131651 | -0.060 | 0.203 | -0.078 | -0.032 | -0.010 | -0.006 | -0.056 | -0.002 | -0.008 |
| | | Yorkshire & the Humber Region | 113204 | -1.255 | 0.057 | -1.202 | -0.010 | -0.113 | -0.131 | -1.173 | -0.040 | -0.159 |
| | | East Midlands Region | 99882 | -1.369 | -0.186 | -1.277 | -0.027 | -0.097 | -0.142 | -1.271 | -0.050 | -0.173 |
| | | West Midlands Region | 117765 | -0.829 | 0.044 | -0.813 | -0.052 | 0.072 | -0.087 | -0.775 | -0.030 | -0.105 |
| | | Eastern Region | 127453 | 0.091 | -0.360 | 0.133 | -0.078 | -0.013 | 0.009 | 0.085 | 0.003 | 0.012 |
| | | London Region | 170391 | 1.684 | 0.563 | 1.478 | 0.120 | 0.480 | 0.174 | 1.556 | 0.061 | 0.211 |
| | | South-West Region | 107590 | 0.510 | -0.316 | 0.542 | -0.083 | -0.079 | 0.053 | 0.477 | 0.019 | 0.065 |
| | | South-East Region | 178554 | 0.548 | -0.273 | 0.585 | -0.035 | -0.172 | 0.057 | 0.513 | 0.020 | 0.070 |
| | Lighting Conditions | Darkness with Lights Lit | 278734 | -0.819 | 0.555 | 0.007 | 0.001 | 0.137 | 0.432 | -0.006 | 0.240 | 0.047 |
| | | Darkness with Lights Unlit | 6525 | 0.376 | -0.058 | 0.052 | -0.350 | -0.072 | -0.198 | 0.003 | -0.110 | -0.021 |
| | | Darkness without Lighting | 100015 | 2.997 | -1.496 | 0.030 | -1.043 | -0.106 | -1.579 | 0.022 | -0.880 | -0.171 |
| | | Darkness with Lighting Unknown | 8761 | 0.057 | 0.085 | -0.098 | -0.299 | 0.202 | -0.030 | <0.001 | -0.017 | -0.003 |
| | | Daylight | 695378 | -0.071 | -0.018 | -0.004 | 0.125 | -0.017 | 0.038 | -0.001 | 0.021 | 0.004 |
| | | Missing | 147 | | | | | | | | | |
| | Road Type | Roundabout | 31355 | -0.246 | 0.189 | -0.046 | 0.713 | -0.184 | -0.006 | -0.014 | 0.011 | -0.216 |
| | | Dual Carriageway | 140983 | -0.213 | -0.403 | 0.082 | 0.177 | -0.182 | -0.005 | -0.012 | 0.009 | -0.186 |
| | | One Way Street/ Slip Road | 4333 | -1.311 | 0.542 | 0.115 | -0.119 | -1.185 | -0.030 | -0.075 | 0.057 | -1.150 |
| | | Unknown RT | 13604 | 8.868 | 0.155 | 0.530 | -0.404 | 7.775 | 0.200 | 0.510 | -0.386 | 7.776 |
| | | Single Carriageway | 880090 | -0.068 | 0.029 | -0.026 | -0.014 | -0.059 | -0.002 | -0.004 | 0.003 | -0.059 |
| | | Missing | 18595 | | | | | | | | | |
| | Weather Conditions | Raining without High Winds | 138819 | 0.143 | -0.033 | 0.021 | -0.015 | 0.112 | -0.005 | 0.006 | -0.010 | 0.115 |
| | | Snowing without High Winds | 5808 | 0.414 | -0.237 | -0.237 | -0.285 | 0.317 | -0.013 | 0.017 | -0.028 | 0.333 |
| | | Fine with High Winds | 18209 | -0.054 | -0.160 | -0.177 | -0.118 | -0.051 | 0.002 | -0.002 | 0.004 | -0.044 |
| | | Raining with High Winds | 15156 | -0.096 | -0.175 | -0.052 | -0.161 | -0.088 | 0.003 | -0.004 | 0.008 | -0.069 |
| | | Snowing with High Winds | 1629 | 0.209 | -0.560 | -0.378 | -0.380 | 0.135 | -0.007 | 0.009 | -0.014 | 0.168 |
| | | Fog or Mist | 10051 | 0.074 | -0.634 | -0.161 | -0.319 | 0.016 | -0.002 | 0.003 | -0.005 | 0.060 |
| | | Other/ Unknown WC | 46003 | 4.735 | -0.134 | 0.206 | -0.330 | 3.809 | -0.153 | 0.198 | -0.323 | 3.809 |
| | | Fine without High Winds | 853095 | -0.252 | 0.021 | -0.004 | 0.006 | -0.203 | 0.008 | -0.011 | 0.017 | -0.203 |
| | | Missing | 100 | | | | | | | | | |
| Ordinal | OSGR Easting Band | OSE 0-300 km | 44900 | -0.812 | -0.185 | 0.425 | -0.183 | 1.357 | 0.031 | 0.581 | -0.002 | 0.051 |
| | | OSE 300-400 km | 248730 | -0.812 | -0.002 | 0.074 | -0.052 | 0.103 | 0.031 | 0.581 | -0.002 | 0.051 |
| | | OSE 500-600 km | 345648 | -0.812 | 0.094 | 0.927 | 0.016 | -0.132 | 0.031 | 0.581 | -0.002 | 0.051 |
| | | OSE 600-700 km | 39062 | -0.196 | -0.263 | 0.153 | -0.011 | -0.010 | 0.008 | 0.137 | <0.001 | 0.012 |
| | | OSE 400-500 km | 407864 | 1.278 | -0.046 | -0.885 | -0.017 | -0.068 | -0.049 | -0.883 | 0.003 | -0.061 |
| | | Missing | 3336 | | | | | | | | | |
| | OSGR Northing Band | OSN 0-100 km | 53026 | -0.737 | -0.118 | 0.018 | -0.051 | 0.282 | -0.041 | -0.607 | -0.028 | 0.099 |
| | | OSN 200-300 km | 224099 | -0.737 | -0.209 | -0.495 | -0.075 | 0.038 | -0.041 | -0.607 | -0.028 | 0.099 |
| | | OSN 300-400 km | 203295 | -0.737 | -0.006 | -0.679 | -0.019 | 0.002 | -0.041 | -0.607 | -0.028 | 0.099 |
| | | OSN 400-500 km | 158274 | -0.737 | 0.120 | -0.738 | -0.016 | -0.005 | -0.041 | -0.607 | -0.028 | 0.099 |
| | | OSN 500-600 km | 54495 | -0.737 | 0.023 | -1.115 | -0.107 | 0.156 | -0.041 | -0.607 | -0.028 | 0.099 |
| | | OSN 600+ km | 13645 | -0.737 | 0.055 | -0.056 | -0.158 | 3.598 | -0.041 | -0.607 | -0.028 | 0.099 |
| | | OSN 100-200 km | 380962 | 1.361 | 0.071 | 1.125 | 0.031 | -0.167 | 0.077 | 1.121 | 0.051 | 0.162 |
| | | Missing | 1764 | | | | | | | | | |
| | Speed Limit Zone | 20 mph (32 km/h) Zone | 3029 | -1.193 | 0.114 | -0.048 | -0.186 | -0.357 | -1.009 | -0.062 | -0.182 | -0.136 |
| | | 40 mph (64 km/h) Zone | 94099 | -1.193 | -0.126 | -0.088 | 0.122 | -0.075 | -1.009 | -0.062 | -0.182 | -0.136 |
| | | 50 mph (80 km/h) Zone | 19878 | -1.193 | -1.079 | 0.212 | -0.087 | -0.280 | -1.009 | -0.062 | -0.182 | -0.136 |
| | | 60 mph (96 km/h) Zone | 257059 | -1.193 | -1.270 | -0.086 | -0.266 | -0.115 | -1.009 | -0.062 | -0.182 | -0.136 |
| | | 70 mph (112 km/h) Zone | 78868 | -1.193 | -1.216 | -0.010 | -0.406 | -0.137 | -1.009 | -0.062 | -0.182 | -0.136 |
| | | 30 mph (48 km/h) Zone | 636228 | 0.835 | 0.707 | 0.045 | 0.108 | 0.112 | 0.706 | 0.043 | 0.127 | 0.095 |
| | | Missing | 3 | | | | | | | | | |
| | Casualty per KSI Car Accident | Single Casualty | 729823 | -0.682 | 0.368 | 0.005 | -0.123 | 0.011 | 0.374 | 0.006 | -0.100 | -0.004 |
| | | Double Casualties | 204861 | 1.045 | -0.574 | 0.010 | 0.153 | 0.025 | -0.574 | -0.009 | 0.154 | 0.007 |
| | | Multiple Casualties | 154876 | 1.844 | -1.022 | -0.029 | 0.233 | 0.024 | -1.012 | -0.015 | 0.271 | 0.012 |
| | Car per KSI Accident | Single Car | 415843 | -1.287 | 0.523 | -0.064 | -0.815 | -0.003 | 0.546 | -0.065 | -0.799 | -0.018 |
| | | Double Cars | 557518 | 0.729 | -0.250 | 0.045 | 0.492 | 0.000 | -0.309 | 0.037 | 0.453 | 0.010 |
| | | Multiple Cars | 116199 | 0.949 | -0.730 | 0.025 | 0.367 | 0.013 | -0.402 | 0.048 | 0.589 | 0.013 |
| | Accident Time | AT(00:01-02:00) | 48285 | -3.521 | 0.159 | 0.129 | -1.331 | 0.040 | 0.183 | 0.136 | -1.359 | -0.005 |
| | | AT(02:01-04:00) | 25357 | -3.521 | 0.227 | 0.121 | -1.409 | -0.013 | 0.183 | 0.136 | -1.359 | -0.005 |
| | | AT(04:01-06:00) | 54296 | -2.818 | -0.014 | 0.142 | -1.101 | -0.071 | 0.131 | 0.109 | -1.088 | -0.004 |
| | | AT(06:01-08:00) | 55148 | 0.222 | -0.219 | 0.057 | 0.152 | 0.001 | -0.010 | -0.009 | 0.086 | <0.001 |
| | | AT(08:01-10:00) | 98235 | 0.222 | -0.044 | 0.053 | 0.158 | 0.009 | -0.010 | -0.009 | 0.086 | <0.001 |
| | | AT(10:01-12:00) | 86294 | 0.222 | -0.058 | 0.018 | 0.148 | -0.028 | -0.010 | -0.009 | 0.086 | <0.001 |
| | | AT(12:01-14:00) | 109958 | 0.222 | -0.011 | -0.012 | 0.157 | -0.014 | -0.010 | -0.009 | 0.086 | <0.001 |
| | | AT(14:01-16:00) | 132191 | 0.222 | 0.007 | -0.028 | 0.070 | -0.014 | -0.010 | -0.009 | 0.086 | <0.001 |
| | | AT(16:01-18:00) | 136823 | 0.222 | 0.028 | -0.013 | 0.088 | 0.007 | -0.010 | -0.009 | 0.086 | <0.001 |
| | | AT(18:01-20:00) | 97240 | 0.222 | -0.037 | 0.006 | 0.044 | 0.040 | -0.010 | -0.009 | 0.086 | <0.001 |
| | | AT(20:01-22:00) | 102938 | 0.222 | -0.039 | -0.015 | -0.103 | 0.127 | -0.010 | -0.009 | 0.086 | <0.001 |
| | | AT(22:01-00:00) | 101097 | 0.352 | 0.030 | -0.067 | 0.136 | 0.019 | -0.016 | -0.014 | 0.136 | <0.001 |
| | | Missing | 88 | | | | | | | | | |
| | Accident Day | Sunday | 143338 | -2.095 | -0.126 | 0.028 | -0.392 | 0.028 | -0.126 | 0.028 | -0.393 | 0.019 |
| | | Monday | 143900 | 0.243 | 0.009 | 0.001 | 0.068 | 0.012 | 0.011 | -0.003 | 0.036 | -0.002 |
| | | Tuesday | 143945 | 0.243 | 0.033 | 0.009 | 0.085 | 0.009 | 0.011 | -0.003 | 0.036 | -0.002 |
| | | Wednesday | 147130 | 0.243 | 0.011 | 0.008 | 0.091 | 0.007 | 0.011 | -0.003 | 0.036 | -0.002 |
| | | Thursday | 155489 | 0.243 | 0.021 | -0.003 | 0.076 | 0.015 | 0.011 | -0.003 | 0.036 | -0.002 |
| | | Saturday | 171857 | 0.243 | -0.012 | -0.011 | -0.111 | 0.013 | 0.011 | -0.003 | 0.036 | -0.002 |
| | | Friday | 184101 | 0.243 | 0.012 | -0.016 | 0.033 | 0.025 | 0.011 | -0.003 | 0.036 | -0.002 |
| | Accident Month | January | 87008 | -0.221 | 0.011 | 0.012 | -0.056 | 0.047 | -0.002 | 0.004 | 0.009 | -0.015 |
| | | February | 77658 | -0.221 | 0.020 | 0.005 | -0.048 | 0.009 | -0.002 | 0.004 | 0.009 | -0.015 |
| | | March | 85749 | -0.221 | 0.038 | 0.010 | -0.014 | 0.003 | -0.002 | 0.004 | 0.009 | -0.015 |
| | | April | 83574 | -0.221 | 0.009 | 0.018 | -0.006 | -0.028 | -0.002 | 0.004 | 0.009 | -0.015 |
| | | May | 91104 | -0.221 | -0.002 | 0.016 | 0.007 | -0.033 | -0.002 | 0.004 | 0.009 | -0.015 |
| | | June | 88922 | -0.221 | -0.029 | 0.016 | 0.023 | -0.044 | -0.002 | 0.004 | 0.009 | -0.015 |
| | | July | 91308 | -0.221 | -0.055 | 0.022 | 0.022 | -0.051 | -0.002 | 0.004 | 0.009 | -0.015 |
| | | August | 90287 | -0.221 | -0.076 | 0.007 | 0.009 | -0.043 | -0.002 | 0.004 | 0.009 | -0.015 |
| | | September | 93049 | -0.221 | -0.026 | 0.005 | 0.021 | -0.022 | -0.002 | 0.004 | 0.009 | -0.015 |
| | | October | 101330 | 0.099 | <0.001 | 0.002 | 0.008 | 0.016 | 0.001 | -0.002 | -0.004 | 0.007 |
| | | December | 96657 | 2.294 | 0.015 | -0.048 | -0.128 | 0.164 | 0.021 | -0.046 | -0.095 | 0.154 |
| | | November | 102916 | 2.294 | 0.019 | -0.039 | -0.074 | 0.143 | 0.021 | -0.046 | -0.095 | 0.154 |

#### 4.3.1. Dichotomous quantification

a) Built-up speed area manifested positive stance/ attribute, but non-built-up speed area contained negative stance.
b) Carriageway hazards manifested positive stance/ attribute, but non-hazards expressed negative stance.
c) 'At junction' manifested positive stance, but 'out of junction' had negative stance.
d) 'Junction control' had positive stance, but 'non-control' had negative stance.
e) Pedestrian crossing at human control quantificated negative stance, but non-control quantificated positive stance.
f) Pedestrian crossing at physical facilities quantificated positive stance, but non-physical quantificated negative stance.
g) Both special conditions and non-conditions quantificated positive stance.

#### 4.3.2. Nominal quantification

h) Eastern region as well as London, South-West, and South-East of England region, manifests positive stance. On the other hand, North-East region as well as North-West, Yorkshire/ the Humber, East Midlands, and West Midlands, contains negativity.
i) 'Day light' as well as 'darkness with light unlit', 'darkness without lighting', and 'darkness with lighting unknown' of lighting conditions quantificated positive stance. Only 'darkness with lights lit' quantificated negative stance.
j) All the categories of 'road type', except 'unknown RT', had negative stance.
k) 'Raining without high winds' as well as 'snowing without high winds', 'snowing with high winds', 'fog or mist', and 'other/ unknown' of weather conditions had positive quantification. On the other hand, 'fine without high winds', 'fine with high winds', and 'raining with high winds' had negative stance.

#### 4.3.3. Ordinal quantification

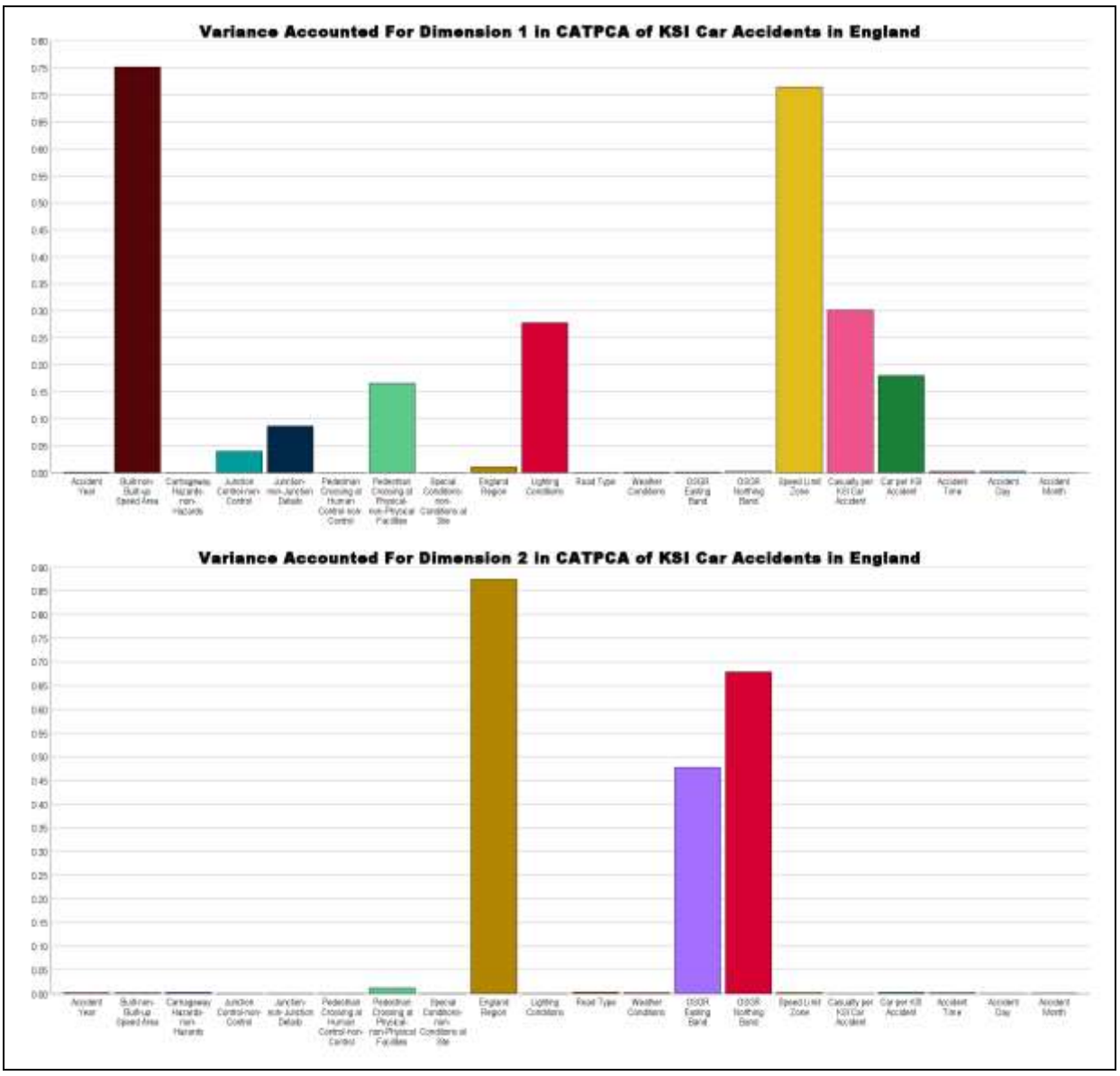
l) All categories of OSGR easting band, except OSE 400-500km, quantificated negative stance.
m) All categories of OSGR northing band, except OSN 100-200km, quantificated negative stance.
n) All categories of speed limit zone, except 30 mph, quantificated negative stance.
o) Single casualty per accident had negative stance, but double as well as multiple casualties per accident had positive stance.
p) Single car per accident had negative stance, but double as well as multiple cars per accident had positive stance.

### 4.4. Variance Accounted For (VAF) in Variables

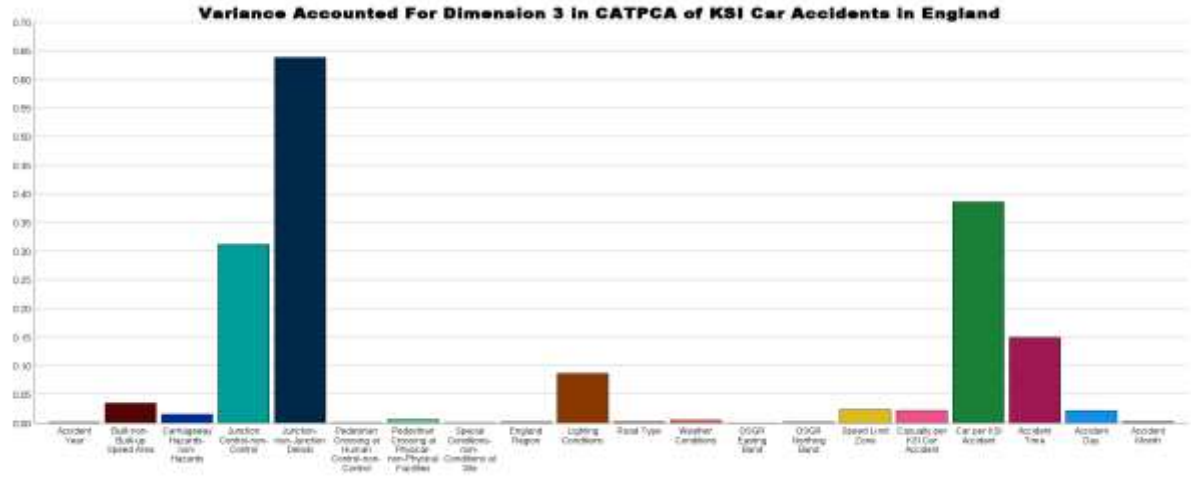
The highest VAF in vector coordinates of the first component was in 'built-non-built-up speed area (0.690)' followed by 'speed limi zone (0.652)', 'lighting conditions (0.314)', 'casualty per accident (0.168)', and 'pedestrian crossing at physical-non-physical facilities (0.161)'. England region (0.694) followed by OSGR northing band (0.502), and OSGR easting band (0.397) had the highest in the second component. On the other hand, 'road type (0.678)' had the highest in the third component followed by 'weather conditions (0.586)', 'accident year (0.152)', 'special-non-special conditions (0.007)', 'accident month (0.006)', and 'pedestrian crossing at

human control-non-control (0.006)'. Further, in the fourth component, 'car per accident (0.526)' had the highest followed by 'junction-non-junction details (0.329)', 'junction control-non-control (0.169)', 'accident time (0.126)', 'accident day (0.011)', and 'carriageway hazards-non-hazards (0.009)'. These are detailed in Table 4.4 and displayed in Figure 4.4. It displays the coordinates for each item on each dimension in relation to the centroid (0, 0). It is noted that displayed a very small mean coordinate (very close to or below 0.100) indicates these items not contributing substantially to the principal components.

CATPCA Loading plots of eigenvalue (VAF) are as follows:



International Journal of Mathematics and Statistics Studies, 11 (3), 27-46, 2023
Print ISSN: 2053-2229 (Print),
Online ISSN: 2053-2210 (Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK




next page

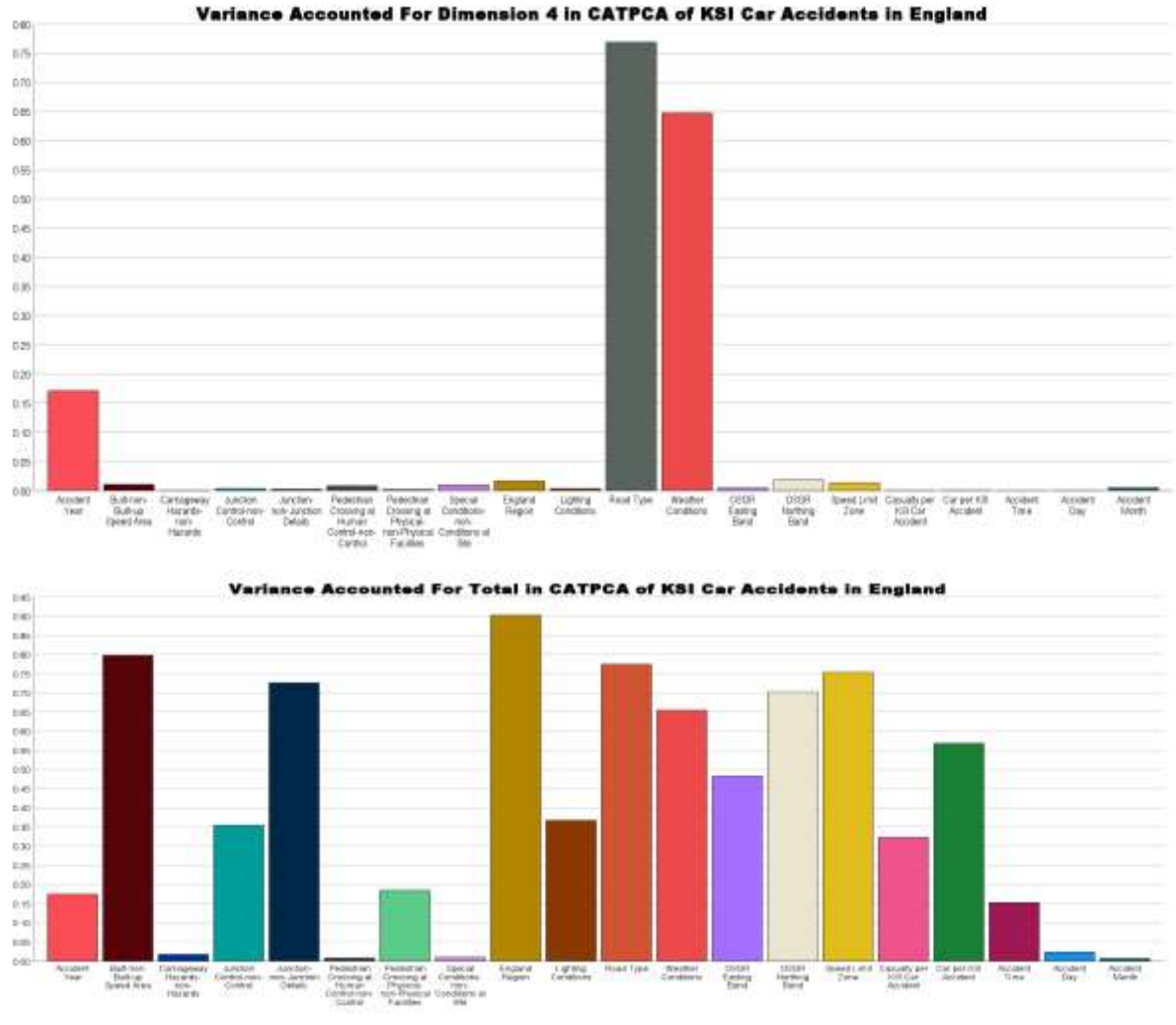



Figure 4.4: Loading Plots of Eigenvalue (VAF) in CATPCA Components

Table 4.4. Variance Accounted For (VAF) in CATPCA of KSI Car Accidents in England

| Variance Accounted For (VAF) in CATPCA of KSI Car Accidents in England | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Centroid Coordinates | | | | | Total (Vector Coordinates) | | | | |
| | Dimension | | | | Mean | Dimension | | | | Total |
| | 1 | 2 | 3 | 4 | | 1 | 2 | 3 | 4 | |
| Built-non-Built-up Speed Area | **0.690** | 0.070 | 0.002 | 0.037 | 0.200 | **0.690** | 0.070 | 0.001 | 0.037 | 0.798 |
| Speed Limit Zone | **0.739** | 0.071 | 0.005 | 0.046 | 0.215 | **0.652** | 0.055 | 0.003 | 0.043 | 0.753 |
| Lighting Conditions | **0.315** | 0.051 | 0.009 | 0.027 | 0.101 | **0.314** | 0.051 | 0.001 | 0.002 | 0.367 |
| Casualty per KSI Car Accident | **0.168** | 0.019 | 0.001 | 0.135 | 0.081 | **0.168** | 0.019 | 0.001 | 0.135 | 0.323 |
| Pedestrian Crossing at Physical-non-Physical Facilities | **0.162** | 0.003 | 0.013 | 0.015 | 0.048 | **0.161** | 0.003 | 0.007 | 0.014 | 0.185 |
| England Region | 0.273 | **0.712** | 0.030 | 0.005 | 0.255 | 0.200 | **0.694** | 0.008 | <0.001 | 0.903 |
| OSGR Northing Band | 0.119 | **0.557** | 0.229 | 0.009 | 0.229 | 0.103 | **0.502** | 0.096 | 0.001 | 0.701 |
| OSGR Easting Band | 0.098 | **0.463** | 0.104 | 0.005 | 0.167 | 0.079 | **0.397** | 0.007 | <0.001 | 0.483 |
| Road Type | 0.059 | 0.043 | **0.678** | 0.039 | 0.205 | 0.035 | 0.033 | **0.678** | 0.029 | 0.775 |
| Weather Conditions | 0.021 | 0.032 | **0.586** | 0.025 | 0.166 | 0.012 | 0.032 | **0.586** | 0.024 | 0.655 |
| Accident Year | 0.003 | 0.003 | **0.220** | 0.025 | 0.063 | 0.002 | 0.001 | **0.152** | 0.021 | 0.175 |
| Special Conditions-non-Conditions at Site | 0.003 | <0.001 | **0.007** | 0.001 | 0.003 | 0.002 | <0.001 | **0.007** | 0.001 | 0.010 |
| Accident Month | 0.001 | <0.001 | **0.007** | 0.001 | 0.002 | <0.001 | <0.001 | **0.006** | 0.001 | 0.007 |
| Pedestrian Crossing at Human Control-non-Control | 0.002 | <0.001 | **0.006** | 0.001 | 0.003 | 0.001 | <0.001 | **0.006** | 0.001 | 0.008 |
| Car per KSI Accident | 0.029 | 0.005 | 0.028 | **0.526** | 0.147 | 0.012 | 0.003 | 0.027 | **0.526** | 0.568 |
| Junction-non-Junction Details | 0.307 | 0.058 | 0.033 | **0.329** | 0.182 | 0.307 | 0.058 | 0.032 | **0.329** | 0.726 |
| Junction Control-non-Control | 0.146 | 0.032 | 0.011 | **0.169** | 0.090 | 0.146 | 0.031 | 0.009 | **0.169** | 0.355 |
| Accident Time | 0.010 | 0.012 | 0.013 | **0.130** | 0.041 | 0.008 | 0.010 | 0.009 | **0.126** | 0.153 |
| Accident Day | 0.009 | 0.003 | 0.003 | **0.014** | 0.007 | 0.008 | 0.003 | 0.002 | **0.011** | 0.024 |
| Carriageway Hazards-non-Hazards | 0.006 | <0.001 | 0.002 | **0.009** | 0.004 | 0.006 | <0.001 | 0.002 | **0.009** | 0.017 |
| **Active Total** | **3.159** | **2.136** | **1.988** | **1.549** | **2.208** | **2.905** | **1.962** | **1.640** | **1.478** | **7.986** |

### 4.5. Rotated Variance Accounted For (RVAF) in Variables

The rotated VAF had a little change and there was also a little change of variable in component to component. These are detailed in Table 4.5.

Table 4.5. Rotated Variance Accounted For (RVAF) in CATPCA of KSI Car Accidents in England

| Rotated Variance Accounted For (RVAF) in CATPCA of KSI Car Accidents in England | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Centroid Coordinates | | | | | Total (Vector Coordinates) | | | | |
| | Dimension | | | | Mean | Dimension | | | | Total |
| | 1 | 2 | 3 | 4 | | 1 | 2 | 3 | 4 | |
| Built-non-Built-up Speed Area | **0.752** | 0.001 | 0.035 | 0.010 | 0.200 | **0.752** | 0.001 | 0.035 | 0.010 | 0.798 |
| Speed Limit Zone | **0.806** | 0.004 | 0.037 | 0.014 | 0.215 | **0.714** | 0.003 | 0.023 | 0.013 | 0.753 |
| Casualty per KSI Car Accident | **0.301** | <0.001 | 0.022 | <0.001 | 0.081 | **0.301** | <0.001 | 0.022 | <0.001 | 0.323 |
| Lighting Conditions | **0.285** | <0.001 | 0.111 | 0.006 | 0.101 | **0.278** | <0.001 | 0.086 | 0.003 | 0.367 |
| Pedestrian Crossing at Physical-non-Physical Facilities | **0.169** | 0.012 | 0.007 | 0.010 | 0.049 | **0.166** | 0.012 | 0.006 | 0.002 | 0.185 |
| England Region | 0.096 | **0.877** | 0.005 | 0.042 | 0.255 | 0.011 | **0.874** | 0.001 | 0.016 | 0.903 |
| OSGR Northing Band | 0.014 | **0.722** | 0.003 | 0.178 | 0.229 | 0.003 | **0.679** | 0.001 | 0.018 | 0.701 |
| OSGR Easting Band | 0.008 | **0.577** | 0.002 | 0.086 | 0.168 | 0.001 | **0.477** | <0.001 | 0.004 | 0.483 |
| Junction-non-Junction Details | 0.087 | <0.001 | **0.638** | 0.002 | 0.182 | 0.087 | <0.001 | **0.638** | 0.002 | 0.726 |
| Car per KSI Accident | 0.193 | 0.003 | **0.392** | <0.001 | 0.147 | 0.180 | 0.003 | **0.386** | <0.001 | 0.568 |
| Junction Control-non-Control | 0.072 | 0.004 | **0.559** | 0.008 | 0.161 | 0.040 | <0.001 | **0.312** | 0.003 | 0.355 |
| Accident Time | 0.006 | 0.003 | **0.154** | 0.002 | 0.041 | 0.002 | 0.001 | **0.149** | <0.001 | 0.153 |
| Accident Day | 0.002 | <0.001 | **0.026** | <0.001 | 0.007 | 0.002 | <0.001 | **0.021** | <0.001 | 0.024 |
| Carriageway Hazards-non-Hazards | <0.001 | 0.002 | **0.015** | <0.001 | 0.004 | <0.001 | 0.002 | **0.015** | <0.001 | 0.017 |
| Road Type | 0.025 | 0.005 | 0.021 | **0.782** | 0.208 | 0.001 | 0.003 | 0.002 | **0.769** | 0.775 |
| Weather Conditions | 0.007 | 0.003 | 0.007 | **0.647** | 0.166 | 0.001 | 0.002 | 0.005 | **0.647** | 0.655 |
| Accident Year | 0.003 | 0.005 | 0.004 | **0.239** | 0.063 | 0.001 | 0.002 | 0.001 | **0.171** | 0.175 |
| Special Conditions-non-Conditions at Site | 0.002 | <0.001 | <0.001 | **0.010** | 0.003 | <0.001 | <0.001 | <0.001 | **0.009** | 0.010 |
| Pedestrian Crossing at Human Control-non-Control | 0.002 | <0.001 | <0.001 | **0.008** | 0.003 | <0.001 | <0.001 | <0.001 | **0.008** | 0.008 |
| Accident Month | 0.001 | <0.001 | 0.003 | **0.005** | 0.002 | <0.001 | <0.001 | 0.002 | **0.005** | 0.007 |
| **Active Total** | **2.828** | **2.219** | **2.040** | **2.051** | **2.284** | **2.540** | **2.060** | **1.705** | **1.680** | **7.986** |

### 4.6. Correlation in Original Variables

Out of 190 unique pairs of twenty original variables (i.e., $20_{c_2} = 190$) in the analysis, only three pairs had greater ($> 0.30$) or lesser ($< -0.30$) correlation and such pair of variables are 'junction-non-junction details **vs** Junction control-non-control ($\rho = -0.966$)', 'speed limit zone **vs** built-non-built-up speed area ($\rho = 0.464$)', and 'OSGR northing band **vs** England region ($\rho = 0.399$)'. The correlation matrix of original variables are displayed in Table 4.6.

Table 4.6. Correlation Matrix of Original Variables in CATPCA of KSI Car Accidents in England

**Correlation Matrix of Original Variables in CATPCA of KSI Car Accidents in England**

| | Accident Year | Built-non-Built-up Speed Area | Carriageway Hazards-non-Hazards | Junction Control-non-Control | Junction-non-Junction Details | Pedestrian Crossing at Human Control-non-Control | Pedestrian Crossing at Physical-non-Physical Facilities | Special Conditions-non-Conditions at Site | England Region | Lighting Conditions | Road Type | Weather Conditions | OSGR Easting Band | OSGR Northing Band | Speed Limit Zone | Casualty per KSI Car Accident | Car per KSI Accident | Accident Time | Accident Day | Accident Month |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Accident Year | 1.000 | | | | | | | | | | | | | | | | | | | |
| Built-non-Built-up Speed Area | -0.010 | 1.000 | | | | | | | | | | | | | | | | | | |
| Carriageway Hazards-non-Hazards | -0.020 | -0.044 | 1.000 | | | | | | | | | | | | | | | | | |
| Junction Control-non-Control | -0.031 | -0.285 | 0.093 | 1.000 | | | | | | | | | | | | | | | | |
| Junction-non-Junction Details | -0.005 | 0.298 | -0.060 | **-0.966** | 1.000 | | | | | | | | | | | | | | | |
| Pedestrian Crossing at Human Control-non-Control | 0.004 | 0.028 | 0.002 | -0.015 | 0.015 | 1.000 | | | | | | | | | | | | | | |
| Pedestrian Crossing at Physical-non-Physical Facilities | 0.074 | 0.234 | -0.022 | -0.176 | 0.181 | 0.060 | 1.000 | | | | | | | | | | | | | |
| Special Conditions-non-Conditions at Site | -0.002 | -0.035 | 0.028 | 0.017 | -0.018 | <0.001 | 0.001 | 1.000 | | | | | | | | | | | | |
| England Region | -0.030 | -0.054 | 0.015 | 0.008 | -0.015 | -0.009 | -0.002 | -0.007 | 1.000 | | | | | | | | | | | |
| Lighting Conditions | 0.038 | -0.138 | -0.010 | 0.020 | -0.022 | 0.017 | -0.067 | 0.012 | 0.031 | 1.000 | | | | | | | | | | |
| Road Type | -0.051 | 0.130 | -0.002 | 0.027 | -0.025 | 0.006 | -0.076 | -0.031 | -0.012 | 0.027 | 1.000 | | | | | | | | | |
| Weather Conditions | 0.052 | 0.022 | -0.002 | -0.017 | 0.017 | -0.002 | <0.001 | -0.014 | 0.005 | 0.088 | 0.009 | 1.000 | | | | | | | | |
| OSGR Easting Band | 0.025 | -0.041 | 0.005 | 0.034 | -0.022 | -0.001 | -0.017 | 0.003 | -0.052 | 0.004 | -0.014 | 0.021 | 1.000 | | | | | | | |
| OSGR Northing Band | 0.014 | 0.099 | -0.012 | -0.044 | 0.049 | 0.002 | 0.089 | -0.006 | **0.399** | -0.024 | -0.018 | 0.011 | -0.025 | 1.000 | | | | | | |
| Speed Limit Zone | -0.042 | **0.464** | -0.021 | -0.158 | 0.168 | 0.017 | 0.122 | -0.024 | -0.040 | -0.063 | 0.131 | 0.015 | -0.057 | 0.073 | 1.000 | | | | | |
| Casualty per KSI Car Accident | 0.004 | -0.260 | 0.003 | 0.095 | -0.100 | -0.015 | -0.100 | 0.024 | -0.012 | -0.028 | -0.028 | -0.029 | 0.015 | -0.048 | -0.191 | 1.000 | | | | |
| Car per KSI Accident | 0.018 | -0.151 | -0.033 | -0.080 | 0.080 | -0.023 | -0.125 | 0.029 | 0.069 | 0.111 | -0.023 | -0.006 | 0.014 | -0.007 | -0.114 | 0.259 | 1.000 | | | |
| Accident Time | -0.045 | 0.052 | -0.011 | -0.033 | 0.036 | -0.011 | 0.005 | -0.015 | -0.022 | -0.090 | 0.043 | -0.001 | 0.007 | -0.013 | 0.030 | -0.017 | 0.025 | 1.000 | | |
| Accident Day | -0.017 | 0.021 | -0.003 | -0.004 | 0.006 | 0.001 | 0.010 | -0.006 | -0.001 | -0.021 | 0.007 | -0.005 | 0.001 | -0.002 | 0.014 | -0.019 | -0.010 | 0.039 | 1.000 | |
| Accident Month | -0.017 | -0.005 | 0.002 | -0.004 | 0.004 | 0.002 | <0.001 | <0.001 | 0.002 | -0.057 | -0.001 | -0.027 | 0.003 | -0.004 | -0.007 | 0.005 | 0.013 | 0.009 | 0.002 | 1.000 |
| **Dimension** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** | **16** | **17** | **18** | **19** | **20** |
| **Eigenvalue** | **2.461** | **1.588** | **1.415** | **1.196** | **1.164** | **1.042** | **1.013** | **1.004** | **0.996** | **0.988** | **0.974** | **0.959** | **0.94** | **0.889** | **0.825** | **0.798** | **0.659** | **0.571** | **0.484** | **0.033** |

Missing values were imputed with variable mode

### 4.7. Correlation in Transformed Variables

Out of 190 unique pairs of twenty transformed variables (i.e., $20_{c_2} = 190$) in the analysis, only seven pairs had greater ($> 0.30$) or lesser ($< -0.30$) correlation and such pair of variables are 'junction-non-junction details **vs** Junction control-non-control ($\rho = 0.884$)', 'lighting conditions **vs** built-non-built-up speed area ($\rho = -0.434$)', 'weather conditions **vs** road type ($\rho = 0.354$)', 'OSGR easting band **vs** England region ($\rho = -0.535$)', 'OSGR northing band **vs** England region ($\rho = 0.700$)', 'speed limit zone **vs** built-non-built-up speed area ($\rho = 0.825$)', and 'speed limit zone **vs** lighting conditions ($\rho = -0.372$)'. The correlation matrix of transformed variables are displayed in Table 4.7.

Table 4.7. Correlation Matrix of Transformed Variables in CATPCA of KSI Car Accidents in England

**Correlation Matrix of Transformed Variables in CATPCA of KSI Car Accidents in England**

| | Accident Year | Built-non-Built-up Speed Area | Carriageway Hazards-non-Hazards | Junction Control-non-Control | Junction-non-Junction Details | Pedestrian Crossing at Human Control-non-Control | Pedestrian Crossing at Physical-non-Physical Facilities | Special Conditions-non-Conditions at Site | England Region | Lighting Conditions | Road Type | Weather Conditions | OSGR Easting Band | OSGR Northing Band | Speed Limit Zone | Casualty per KSI Car Accident | Car per KSI Accident | Accident Time | Accident Day | Accident Month |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Accident Year | 1.000 | | | | | | | | | | | | | | | | | | | |
| Built-non-Built-up Speed Area | -0.010 | 1.000 | | | | | | | | | | | | | | | | | | |
| Carriageway Hazards-non-Hazards | -0.021 | -0.026 | 1.000 | | | | | | | | | | | | | | | | | |
| Junction Control-non-Control | -0.072 | 0.272 | -0.048 | 1.000 | | | | | | | | | | | | | | | | |
| Junction-non-Junction Details | -0.005 | 0.296 | -0.055 | **0.884** | 1.000 | | | | | | | | | | | | | | | |
| Pedestrian Crossing at Human Control-non-Control | -0.004 | -0.028 | 0.001 | -0.014 | -0.015 | 1.000 | | | | | | | | | | | | | | |
| Pedestrian Crossing at Physical-non-Physical Facilities | 0.074 | 0.234 | -0.025 | 0.161 | 0.181 | -0.060 | 1.000 | | | | | | | | | | | | | |
| Special Conditions-non-Conditions at Site | -0.002 | -0.035 | 0.075 | -0.016 | -0.018 | <0.001 | 0.001 | 1.000 | | | | | | | | | | | | |
| England Region | -0.015 | 0.135 | -0.027 | 0.052 | 0.072 | 0.004 | 0.101 | -0.013 | 1.000 | | | | | | | | | | | |
| Lighting Conditions | 0.008 | **-0.434** | 0.029 | -0.194 | -0.212 | 0.010 | -0.136 | 0.009 | -0.057 | 1.000 | | | | | | | | | | |
| Road Type | -0.124 | 0.071 | -0.009 | 0.010 | 0.009 | 0.004 | -0.019 | -0.014 | 0.132 | -0.031 | 1.000 | | | | | | | | | |
| Weather Conditions | -0.111 | 0.024 | 0.001 | -0.003 | -0.007 | 0.004 | -0.017 | -0.006 | 0.098 | -0.007 | **0.354** | 1.000 | | | | | | | | |
| OSGR Easting Band | -0.009 | -0.059 | 0.021 | -0.008 | -0.034 | 0.001 | -0.047 | 0.004 | **-0.535** | 0.037 | -0.044 | -0.029 | 1.000 | | | | | | | |
| OSGR Northing Band | 0.025 | 0.076 | -0.021 | 0.033 | 0.041 | 0.001 | 0.086 | -0.005 | **0.700** | -0.030 | -0.046 | -0.036 | -0.237 | 1.000 | | | | | | |
| Speed Limit Zone | -0.023 | **0.825** | -0.018 | 0.250 | 0.271 | -0.026 | 0.203 | -0.035 | 0.145 | **-0.372** | 0.083 | 0.025 | -0.079 | 0.075 | 1.000 | | | | | |
| Casualty per KSI Car Accident | 0.004 | -0.283 | 0.004 | -0.089 | -0.099 | 0.016 | -0.101 | 0.024 | -0.068 | 0.128 | -0.019 | -0.005 | 0.024 | -0.046 | -0.278 | 1.000 | | | | |
| Car per KSI Accident | 0.013 | -0.124 | -0.007 | 0.121 | 0.143 | 0.024 | -0.124 | 0.023 | 0.014 | -0.023 | -0.013 | -0.015 | 0.002 | 0.024 | -0.130 | 0.223 | 1.000 | | | |
| Accident Time | -0.013 | 0.027 | -0.008 | 0.068 | 0.075 | -0.008 | 0.006 | <0.001 | -0.010 | -0.078 | 0.001 | -0.021 | 0.011 | -0.005 | 0.025 | -0.064 | 0.147 | 1.000 | | |
| Accident Day | -0.001 | 0.044 | -0.001 | 0.031 | 0.034 | -0.011 | 0.020 | -0.001 | 0.009 | -0.017 | -0.001 | <0.001 | -0.001 | 0.007 | 0.040 | -0.075 | 0.018 | 0.138 | 1.000 | |
| Accident Month | -0.015 | 0.004 | 0.001 | 0.001 | <0.001 | -0.005 | 0.015 | -0.001 | -0.008 | 0.015 | -0.004 | 0.050 | 0.006 | -0.007 | <0.001 | -0.004 | -0.023 | 0.003 | 0.021 | 1.000 |
| **Dimension** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** | **16** | **17** | **18** | **19** | **20** |
| **Eigenvalue** | **2.951** | **1.941** | **1.605** | **1.438** | **1.167** | **1.074** | **1.064** | **1.004** | **0.985** | **0.956** | **0.916** | **0.851** | **0.837** | **0.789** | **0.679** | **0.644** | **0.632** | **0.204** | **0.171** | **0.112** |

Missing values were imputed with the mode of the quantified variable

### 4.8. Component Loadings and Rotated Component Loadings

The reporting CATPCA for KSI car accidents is shown in Table 4.8. Also, the CATPCA component loadings plots as well as rotated loadings are displayed in Figure 4.8. The table for Component Loadings shows the coordinates for each item on each dimension and these are plotted in the next element of the output, the scatter plot displayed above-right.

#### 4.8.1. First Categorical Principal Component (CATPC-1)

First categorical principal component was good with internal consistency ([3]$Cronbach's\ \alpha = 0.7$). Also its eigenvalue was $\lambda_1 = 2.91$. It comprised five out of 20 original variables, but also five out of 20 variables in rotated model (see the report in Table 4.8). The variables in CATPC-1 reformed with increasing for 'built-non-built up speed area' as well as 'speed limit zone', and 'pedestrian crossing at physical-non-physical facilities', having positive associations with first component; but with decreasing for 'lighting conditions' as well as 'casualty per accident', having nagative associations with first component. These variables in CATPC-1 were in the 4-dimensional factor space to best approximate the data in the least square sense.

[3] **Internal Consistency for Cronbach's Alpha**: Excellent ($\alpha \geq 0.9$); Good ($0.7 \leq \alpha < 0.9$); Acceptable ($0.6 \leq \alpha < 0.7$); Poor ($0.5 \leq \alpha < 0.6$); Unacceptable ($\alpha < 0.5$).

### 4.8.2. Second Categorical Principal Component (CATPC-2)

Second categorical principal component was poor with internal consistency ($Cronbach's\ \alpha = 0.52$). Also its eigenvalue was $\lambda_2 = 1.962$. It comprised three out of 20 original variables, but as well as same number of original variables in rotated model (see the report). The variables in CATPC-2 reformed with increasing for 'England region' as well as 'OSGR northing band', having positive associations with second component; but with decreasing for 'OSGR easting band' having negative association with second component. These variables in CATPC-2 were in the 4-dimensional factor space for variance-maximising direction orthogonal to CATPC-1.

### 4.8.3. Third Categorical Principal Component (CATPC-3)

Third categorical principal component was under-poor/ unacceptable with internal consistency ($Cronbach's\ \alpha = 0.41$). Also its eigenvalue was $\lambda_2 = 1.640$. It comprised six out of 20 original variables but as well as same number of original variables in rotated model (see the report). The variables in CATPC-3 reformed with increasing for 'road type' as well as 'weather conditions', and 'accident month', having positive associations with third component; but with decreasing for 'accident year' as well as 'special condition-non-condition at site', and 'pedestrian crossing at human control-non-control', having negative associations with third component. These variables in CATPC-3 were in the 3-dimensional factor space for variance-maximising direction orthogonal to CATPC-1 and CATPC-2.

### 4.8.4. Fourth Categorical Principal Component (CATPC-4)

Fourth categorical principal component was under-poor/ unacceptable with internal consistency ($Cronbach's\ \alpha = 0.34$). Also its eigenvalue was $\lambda_2 = 1.48$. It comprised six out of 20 original variables but as well as same number of original variables in rotated model (see the report). The variables in CATPC-4 reformed with increasing for 'car per accident' as well as 'junction-non-junction details', 'junction controll-non-controll', 'accident time', 'accident day', having positive associations with fourth component; but with decreasing for 'carriageway hazards-non-hazards', having negative association with fourth component. These variables in CATPC-4 were in the 4-dimensional factor space for variance-maximising direction orthogonal to CATPC-1, CATPC-2, and CATPC-3.

Table 4.8. Component Loadings in CATPCA of KSI Car Accidents in England

**Reporting Component Loadings in CATPCA of KSI Car Accidents in England**

**Component Loadings**

| Variable | Dimension 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Built-non-Built-up Speed Area | **0.831** | -0.264 | 0.035 | -0.191 |
| Speed Limit Zone | **0.808** | -0.236 | 0.053 | -0.207 |
| Lighting Conditions | **-0.561** | 0.225 | 0.026 | -0.039 |
| Casualty per KSI Car Accident | **-0.410** | 0.138 | -0.031 | 0.368 |
| Pedestrian Crossing at Physical-non-Physical Facilities | **0.401** | -0.052 | -0.083 | -0.119 |
| England Region | 0.448 | **0.833** | -0.087 | 0.019 |
| OSGR Northing Band | 0.321 | **0.709** | -0.309 | -0.024 |
| OSGR Easting Band | -0.281 | **-0.630** | 0.085 | 0.002 |
| Road Type | 0.186 | 0.181 | **0.823** | 0.172 |
| Weather Conditions | 0.112 | 0.179 | **0.765** | 0.156 |
| Accident Year | -0.042 | -0.027 | **-0.389** | -0.144 |
| Special Conditions-non-Conditions at Site | -0.040 | -0.015 | **-0.084** | -0.032 |
| Accident Month | -0.002 | -0.003 | **0.079** | -0.023 |
| Pedestrian Crossing at Human Control-non-Control | -0.035 | -0.013 | **-0.077** | -0.032 |
| Car per KSI Accident | -0.108 | 0.057 | -0.166 | **0.725** |
| Junction-non-Junction Details | 0.554 | -0.241 | -0.180 | **0.574** |
| Junction Control-non-Control | 0.382 | -0.176 | -0.097 | **0.411** |
| Accident Time | 0.090 | -0.099 | -0.096 | **0.354** |
| Accident Day | 0.089 | -0.055 | -0.044 | **0.104** |
| Carriageway Hazards-non-Hazards | -0.076 | -0.012 | 0.048 | **-0.096** |

Variable Principal Normalization

**Rotated Component Loadings**

| Variable | Dimension 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Built-non-Built-up Speed Area | **0.867** | 0.037 | 0.186 | 0.101 |
| Speed Limit Zone | **0.845** | 0.052 | 0.153 | 0.114 |
| Casualty per KSI Car Accident | **-0.549** | -0.008 | 0.147 | 0.006 |
| Lighting Conditions | **-0.527** | 0.007 | -0.294 | -0.057 |
| Pedestrian Crossing at Physical-non-Physical Facilities | **0.407** | 0.108 | 0.078 | -0.041 |
| England Region | 0.104 | **0.935** | 0.037 | 0.127 |
| OSGR Northing Band | 0.056 | **0.824** | 0.038 | -0.134 |
| OSGR Easting Band | -0.038 | **-0.691** | 0.003 | -0.063 |
| Junction-non-Junction Details | 0.294 | 0.009 | **0.799** | 0.042 |
| Car per KSI Accident | -0.424 | 0.050 | **0.621** | 0.014 |
| Junction Control-non-Control | 0.199 | -0.009 | **0.558** | 0.057 |
| Accident Time | -0.046 | -0.039 | **0.386** | 0.001 |
| Accident Day | 0.047 | -0.011 | **0.146** | -0.007 |
| Carriageway Hazards-non-Hazards | -0.018 | -0.048 | **-0.122** | 0.005 |
| Road Type | 0.023 | 0.056 | -0.043 | **0.877** |
| Weather Conditions | -0.032 | 0.042 | -0.068 | **0.805** |
| Accident Year | 0.037 | 0.043 | -0.029 | **-0.413** |
| Special Conditions-non-Conditions at Site | -0.015 | -0.010 | -0.017 | **-0.097** |
| Pedestrian Crossing at Human Control-non-Control | -0.011 | -0.008 | -0.017 | **-0.088** |
| Accident Month | 0.009 | -0.020 | -0.041 | **0.067** |

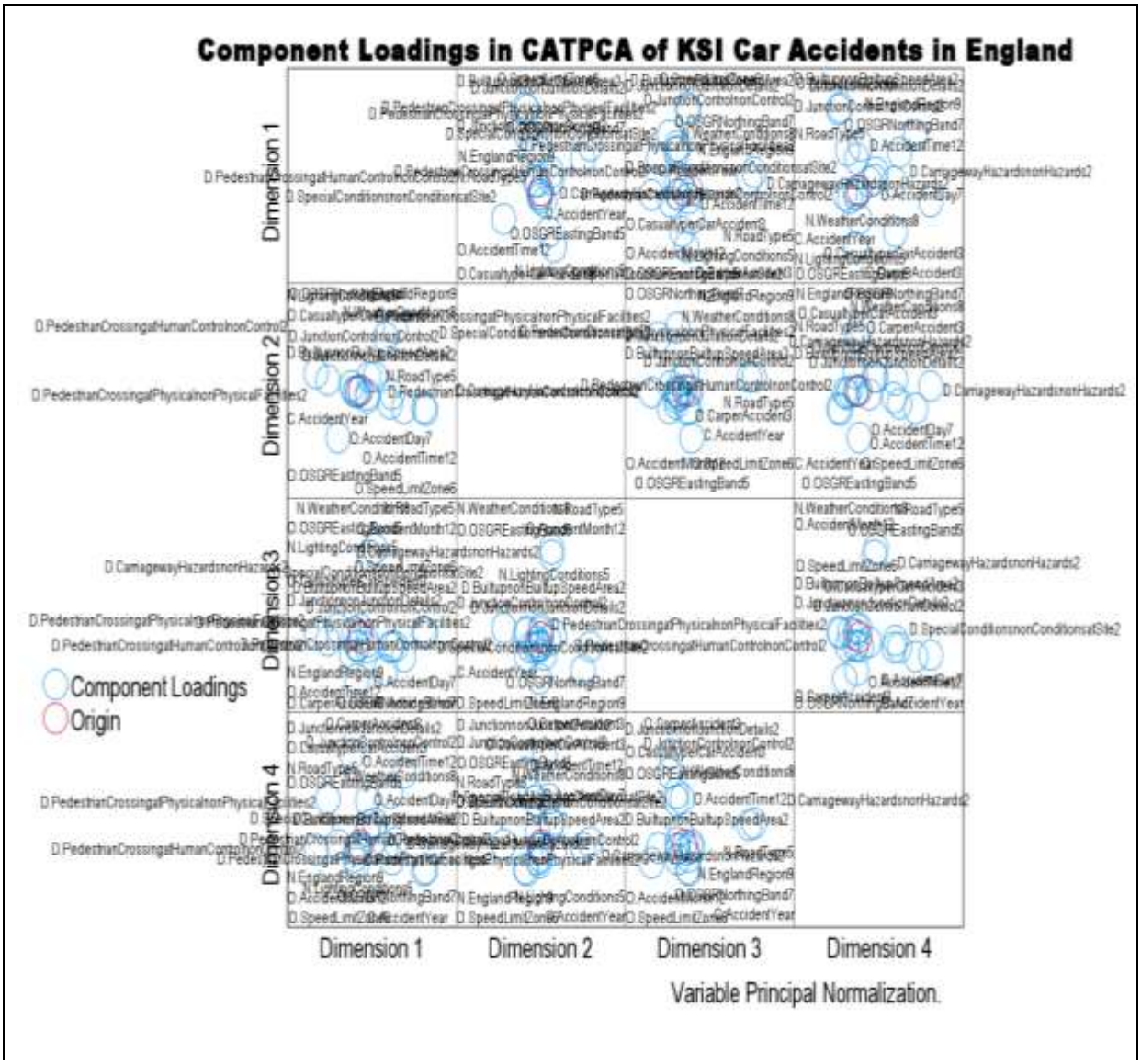
Rotation Method: Varimax with Kaiser Normalization. Rotation converged in 5 iterations.

| **Model Summary** | | | | |
|---|---|---|---|---|
| Cronbach's Alpha | 0.690 | 0.516 | 0.411 | 0.340 |
| Variance Accounted For (Total (Eigenvalue)) | 2.905 | 1.962 | 1.640 | 1.478 |
| Total Cronbach's Alpha in 4-dimensions | 0.921 | | | |
| Total Variance Accounted For (Total (Eigenvalue)) | 7.986 | | | |

| **Model Summary** | | | | |
|---|---|---|---|---|
| Cronbach's Alpha | 0.647 | 0.541 | 0.477 | 0.358 |
| Variance Accounted For (Total (Eigenvalue)) | 2.540 | 2.060 | 1.705 | 1.680 |
| Total Cronbach's Alpha in 4-dimensions | 0.921 | | | |
| Total Variance Accounted For (Total (Eigenvalue)) | 7.986 | | | |

| **Case Processing Summary** | |
|---|---|
| Valid Active Cases | 579496 |
| Active Cases with Missing Values | 510064 |
| Supplementary Cases | 0 |
| Total | 1089560 |
| Cases Used in Analysis | 1089560 |

**Credit**: CATPCA Version 2.0 by Leiden SPSS Group, Leiden University, The Netherlands


Component Loadings in CATPCA of KSI Car Accidents in England
Dimension 1
Dimension 2
Dimension 3
Dimension 4
Component Loadings
Origin
Variable Principal Normalization.

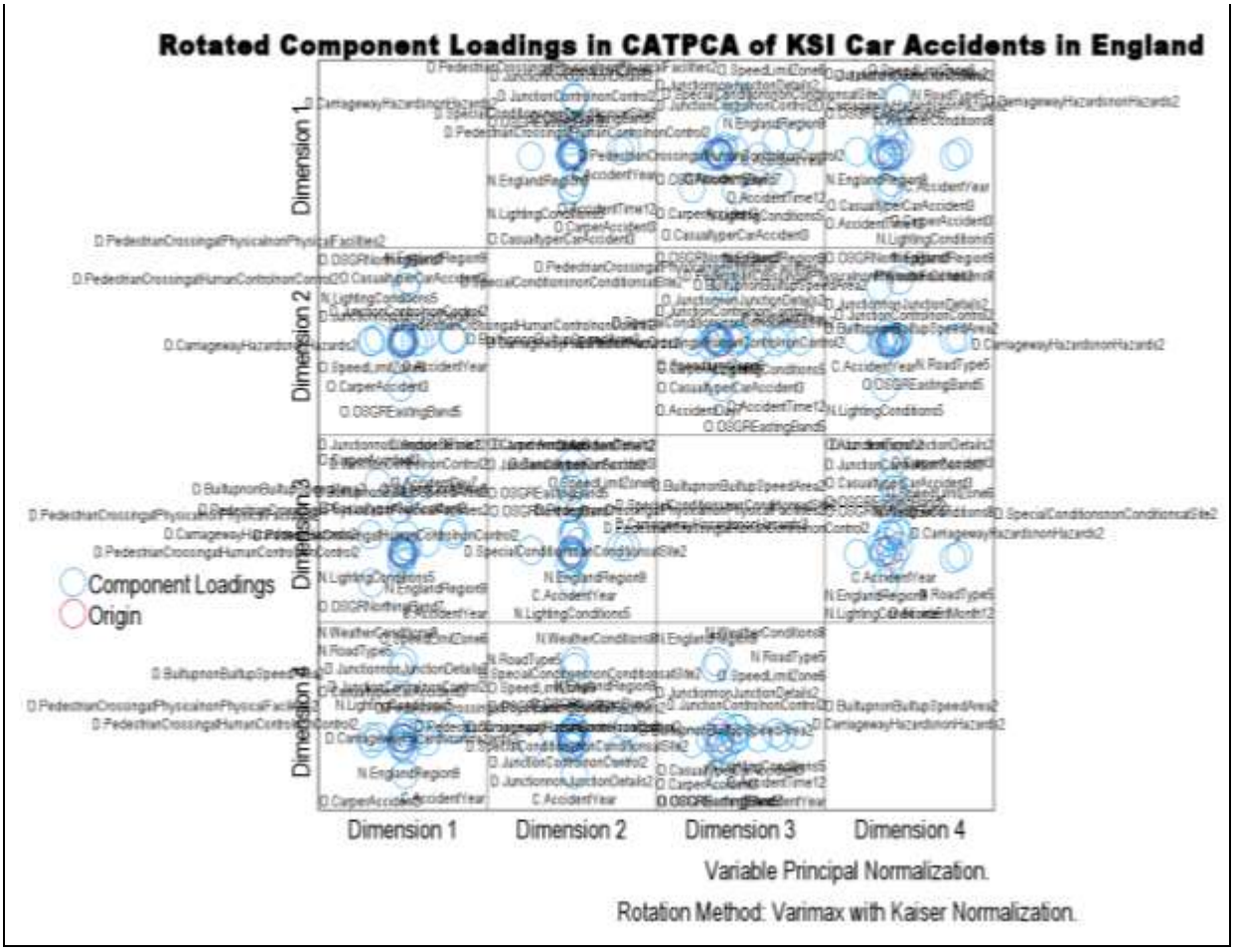



Figure 4.8: Scatter Matrix of Components Loadings and Rotated Component Loadings in CATPCA of KSI Car Accidents in England

### 4.9. Overview in CATPCA for Road KSI Accidents

CATPCA revealed that four-component solution met the interpretability criterion. The 20 variables were reduced to four components that explained 92.1% (total Cronbach's α based on grand total eigenvalue) of the total variance. A Varimax orthogonal rotation was applied to aid interpretability. The rotated solution exhibited 'simple structure' (Thurstone, 1947). The data was consistent with the accident attributes in the 20-variable dataset that was designed to measure with strong loadings of 5-variables on Component-1, 3-variables on Component-2, 6-variables on Component-3, and 6-variables on Component-4. Component loadings of the rotated solution had a little change.

## 5. Research Findings, Discussions, and Conclusion

The aim of this study was to identify the component structure of road KSI car accidents in England based on categorical variables in STATS19 database for the period of 1979-2015 by applying categorical principal component analysis. Based on 39 potential variables, this study used CATPCA as an alternative to reduce the large number of variables into the four dimensions of KSI car accidents that have been identified. This study proceeded with 33 fit variables after a series of analyses to determine the fit of variables to be selected. Then,

following the four steps of analysis in the CATPCA, this study reduced and grouped the 20 variables into four principal components. As the data consisted of mainly categorical types (i.e., dichotomous, nominal, and ordinal), the optimal scaling CATPCA allowed this study to jointly analyse and transform these categorical data into numerical values. The aggregated summation of the data of 1089560 cases (where only 579496 cases valid) on these 20 variables was later transformed into component/object scores. These scores are the manifestations of the surrogates of the four components of KSI car accidents identified in this study. These are the best composite scores to use.

1) First categorical principal component comprised five out of 20 original variables. It reformed with increasing for 'built-non-built up speed area' as well as 'speed limit zone', and 'pedestrian crossing at physical-non-physical facilities', having positive associations with first component; but with decreasing for 'lighting conditions' as well as 'casualty per accident', having nagative associations with first component.

2) Second categorical principal component consisted of three variables that reformed with increasing for 'England region' as well as 'OSGR northing band', having positive associations with second component; but with decreasing for 'OSGR easting band' having negative association with second component.

3) Third categorical principal component comprised six variables that reformed with increasing for 'road type' as well as 'weather conditions', and 'accident month', having positive associations with third component; but with decreasing for 'accident year' as well as 'special condition-non-condition at site', and 'pedestrian crossing at human control-non-control', having negative associations with third component.

4) Fourth categorical principal component comprised six variables that reformed with increasing for 'car per accident' as well as 'junction-non-junction details', 'junction controll-non-controll', 'accident time', 'accident day', having positive associations with fourth component; but with decreasing for 'carriageway hazards-non-hazards', having negative association with fourth component.

This research leads a road car safety practice/ policy as a key policy priority that should therefore be to plan for the long-term prevention of road car accidents. The identified significant variables leading KSI car accidents, can be taken special attention so that KSI car accidents can be reduced. This research can be used to develop targeted interventions aimed at road safety practice/ policy. The findings of this study have several important implications for future practice. These findings are dynamic if policymakers are to be able to identify those individuals at greatest threat of KSI car accident involvement. Policies to reduce the KSI car accidents can then be effectively implemented. Until such policies are implemented, KSI car accidents and associated casualties will remain unnecessarily high for certain individuals in the society, and as a society people will all suffer. Where there is an alert today, there is an alive tomorrow. Stay Alive – Think and Drive.

**REFERENCES**

El-Basyouny, K., and Sayed, T. (2009). Accident Prediction Models with Random Corridor Parameters. Accident Analysis & Prevention. 41(5), p.1118-1123.

Harrel, F. (2015). Regression Modelling Strategies. Springer Series in Statistics book series (SSS): p.219-274. Springer, Cham. https://doi.org/10.1007/978-3-319-19425-7_10.

Kleinbaum, D., and Klein, M. (2010). Ordinal Logistic Regression. In: Logistic Regression. Statistics for Biology and Health. Springer, New York, NY. https://doi.org/10.1007/978-1-4419-1742-3_13.

Saukani, N., and Ismail, N.A. (2019). Identifying the Components of Social Capital by Categorical Principal Component Analysis (CATPCA). *Soc Indic Res* **141**, 631–655. https://doi.org/10.1007/s11205-018-1842-2

Hung, Younshik, and Tai-Jin Song. (2018). Safety Analysis of Motorcycle Crashes in Seoul Metropolitan Area, South Korea: An Application of Nonlinear Optimal Scaling Methods. *International Journal of Environmental Research and Public Health* 15, no. 12: 2702. https://doi.org/10.3390/ijerph15122702