

Analyzing road safety indicator data across Europe: describing, explaining and comparing

Authors: B. Wilmots, E. Hermans, T. Brijs & G. Wets Presenter: Dr. E. Hermans

Session 8

4th IRTAD CONFERENCE *Road safety data: collection and analysis for target setting and monitoring performances and progress*

Seoul, 16-17 September 2009



Overview

- 1. Introduction
- 2. Data description
- 3. Method and results:
 - 3.1. Regression
 - 3.2. Clustering Detailed
 - 3.3 Footprint methodology

look at Belgium

4. Conclusion

universiteit ▶▶hasselt



1. Introduction

- The road safety problem (WHO, 2009): 1.3 million people killed 20 to 50 million injured Introduction How to improve the level of road safety in a • Data country? description Studying various influencing indicators: Method and results Understand the current road safety situation Regression Compare the road safety situation with other •Clustering Footprint countries Formulate policy recommendations Conclusion
 - Here, we formulate indicators, collect indicator data and perform analyses to describe, explain and compare countries





2. Data description





Introduction

description

Method and

•Regression

Clustering

Footprint

Conclusion

Data

results

2. Data description (2)

- 60 indicators
- Data collected from publications & online databases
- 2003

lacksquare

- 21 European countries
- Missing values: mainly for Switzerland, Hungary, Cyprus and Estonia
- For some analyses indicator data were standardized





Introduction

Data description

Method and results

•<u>Regression</u> Goal Methodology Results Conclusion

•<u>Clustering</u> •<u>Footprint</u>

Conclusion

universiteit hasselt

3.1. Regression

- Goal: Determine the indicators that explain the number of road fatalities per million inhabitants
- Methodology:
 - 1: Standardizing the indicator data set
 - 2: Determining approaches (3) to enter variables in the model (=> generating 3 models) and performing the linear multiple regression analyses
 - 3: Testing each model on the assumptions concerning the error term
 - 4: Studying the statistical significance of each model and the degree of multicollinearity
 - 5: Assessing the explanatory power of each model
 - 6: Interpreting the regression results



universiteit

3.1. Regression (2) Results

| <u>Introduction</u> | Model 1: | | | | | | | |
|-------------------------------------------|--------------------------------------------------------|---------------------------------------------|---------------|--|--|--|--|--|
| <u>Data</u> | FATAL = 0.183 - 0.774 E1 - 0.4 | R ^{2′} : 78 % | | | | | | |
| <u>description</u> | Model 2: | | | | | | | |
| Method and | FATAL = 0.046 - 0.378 E1 - 0.376 PS + 0 | R ^{2′} : 83.9 % | | | | | | |
| <u>results</u> | 0.264 T12 + ε | | | | | | | |
| • <u>Regression</u> Goal | Model 3: | | | | | | | |
| Methodology | FATAL = 0.109 – 0.336 PS – 0.369 VE | R ^{2′} : 71 % | | | | | | |
| Results Conclusion | Legend: | | | | | | | |
| • <u>Clustering</u> • <u>Footprint</u> | FATAL: number of road fatalities per million | • D4 : Avg. # people / hor | usehold | | | | | |
| | inhabitants | • T1: # passenger cars /1 | rs /1000 inh. | | | | | |
| | • E1: GDP | • T12 : Avg. # kms driven per driver | | | | | | |
| <u>Conclusion</u> | • ALC: % drivers < alcohol limit | | | | | | | |
| | • PS : % seat belt use in front of cars or vans | | | | | | | |
| | • VEH: % cars < 6 years | | | | | | | |
| | | | | | | | | |



3.1. Regression (3) Conclusion

- The final models contain:
 - 3 SPI's
 - share of drivers respecting the legal alcohol limit
 - seat belt wearing rate in front of cars or vans
 - share of cars < 6 years
 - 4 background and context indicators
 - gross domestic product (E)
 - average number of persons within a family (D)
 - number of passenger cars per 1000 inhabitants (T)
 - average number of kilometers travelled by a driver (T)
 - These indicators affect the road safety output level
- Background and context indicators are less controllable
- Improving the scores on the SPI's will have a direct effect on the number of road fatalities per million inhabitants

4th IRTAD Conference, Seoul, 16-17 September 2009

Introduction

Data description

Method and results

```
•<u>Regression</u>
Goal
Methodology
Results
Conclusion
```

- •<u>Clustering</u>
- Footprint

Conclusion

universiteit



3.2. Clustering

Introduction

Data description

•

Method and results •Regression

•<u>Clustering</u> Goal Methodology Groups based on SPO's Groups based on SPI's Conclusion

•<u>Footprint</u>

Conclusion

universiteit ▶▶hasselt Goal: Identify similarly performing countries

- Methodology:
 - 1: Standardizing the indicator data set
 - 2: Determining the algorithm(s)
 - Here, first Ward's algorithm to determine the ideal number of groups; then, k-means algorithm
 - 3: Performing the cluster analyses
 - Hierarchical clustering based on Ward's algorithm
 - Non-hierarchical k-means clustering
 - 4: Interpreting the clustering result

3.2. Clustering (2) Identifying groups based on SOI's





Data description Method and results Regression Clustering Goal Methodology Groups based on SOI's Groups based on SPI's Conclusion

imob

INSTITUUT VOOR MOBILITEIT

Conclusion



3.2. Clustering (3) Interpreting groups based on SOI's

Introduction

VOOR MOBILITEIT

imop

Data description

Method and results

•<u>Regression</u>

•<u>Clustering</u> Goal Methodology **Groups** based on SOI's Groups based on SPI's Conclusion

•<u>Footprint</u>

Conclusion

universiteit ▶▶hasselt

- Group 1: SUN-countries, Denmark, Germany, Finland, France and Ireland
 - + Best performing countries
- Group 2: Belgium, Austria, Spain, Italy, Portugal, Slovenia and Estonia
 - +/- Average performing countries
- Group 3: Czech Republic, Greece, Hungary and Poland
 - Worst performing countries





3.2. Clustering (5) Interpreting groups based on SPI's

Introduction

Data description

Method and results

•<u>Regression</u>

•<u>Clustering</u> Goal Methodology Groups based on SOI's **Groups** based on SPI's Conclusion •Footprint

<u>Conclusion</u>

universiteit

- Group 1: Czech Republic, Hungary, Poland and Estonia + 'alcohol and drugs'
 - 'vehicle' and 'protective systems'
- Group 2: Belgium, Spain, Italy and Portugal
 - +/- 'roads', 'trauma management' and 'vehicle'
 - 'alcohol and drugs' and 'speed'
 - -/0 'protective systems'
- Group 3: Switzerland, Germany, France and Netherlands + 'trauma management', 'roads' and 'protective systems'
- Group 4: Austria, Denmark, Finland, Ireland, Sweden, Slovenia and United Kingdom
 - + all risk domains except 'trauma management' and 'roads'



3.2. Clustering (6) Conclusion

Introduction

<u>Data</u> description

Method and results

•<u>Regression</u>

•<u>Clustering</u> Goal Methodology Groups based on SOI's Groups based on SPI's **Conclusion**

• Footprint

Conclusion

- Countries were also grouped with respect to background and context indicators
- Belgium resembles the Netherlands on some background and context indicators but has a lower road safety level due to differences in SPI's
- The SUN-countries perform best and show some similarities concerning SPI's, economic and transport background





3.3. Footprint methodology

Introduction

<u>Data</u> description

Method and results

•<u>Regression</u> •Clustering

•<u>Footprint</u> *Description General footprint Detailed footprint Conclusion*

<u>Conclusion</u>



- Footprint (Morsink et al., 2005):
 - A schematic overview, valuable for describing and comparing countries
 - Provides a quick summary view on the road safety situation in a country
- 2 summary footprints for Belgium:
 - Comparing Belgium to the European average ('general footprint')
 - Comparing Belgium to the average of the SUNcountries ('detailed footprint')
- How?
 - 1: Determine boundaries for the colouring
 - 2: Visualize relative scores by red, orange and green
 - 4th IRTAD Conference, Seoul, 16-17 September 2009



3.3. Footprint methodology (2) General footprint

Road Safety Output Indicators

% fatalities per age group

0-14y

Road Safety Performance Indicators

Per risk domain

Alcohol & drugs

Introduction

Organizational Background

recent road safety initiatives

Level of enforcement

Data description

Method and results •Regression

•<u>Clustering</u>

•<u>Footprint</u> Description **General footprint** Detailed footprint Conclusion

| Max. speed limit | | % drivers < max. alcohol limit | · 15-24y |
|---------------------------------------------------|---|---------------------------------------------------------------|--------------------------------------|
| inside urban area | | Speed | · 25-34y |
| · outside urban area | | · % drivers < max. speed limit inside urban area | · 35-44y |
| motorways | | Protective system | · 45-54y |
| Context | | % seat belt use in front of cars and vans | · 55-64y |
| Demographic background | | Daytime running lights (DRL) | · 65+ |
| % 15-24y | | Regulation on DRL | Age unknown |
| % 65+ | | Vehicle | % fatalities per person class |
| Economic background | _ | % cars < 6 years | Vehicle occupant |
| Gross Domestic Product | | Roads | Motorized two-wheeler |
| Unemployment % | | Density of motorways | · Pedestrian |
| Geographic background | | Trauma management | Cyclist |
| Geographic position in Europe | W | Gross Domestic Product % spent on health care | 🛛 🕖 Unknown |
| # wintry months | | | % fatalities per road type |
| Transport background | | | Inside urban area |
| % pass. kms | | | Outside urban area |
| · Car | | | Motorway |
| Motorized two-wheeler | | | No motorway |
| · Bus | | | 🛛 🕔 Unknown |
| · Train | | | |
| Pers. kms per person per year | | | |
| · Bicycle | | | |
| · Walk | | | |
| % road length | | | |
| Motorways | | | |
| · National | | | |
| Secondary | | | |
| · Other | | | |

Conclusion





3.3. Footprint methodology (3) Detailed footprint

Road Safety Performance Indicators

Per risk domain

Alcohol & druas

Road

Indicators

Safety

% fatalities per age group
 0-14y

Output

Introduction

Organizational Background

recent road safety initiatives

Level of enforcement

Data description

Method and results •Regression

•Clustering

•<u>Footprint</u> Description General footprint **Detailed** footprint Conclusion

| Max. speed limit | | % drivers < max. alcohol limit | • 15-24y |
|---------------------------------------------------|---|-----------------------------------------------------------------------|-------------------------------------------|
| inside urban area | | Speed | • 25-34y |
| outside urban area | | % drivers < max. speed limit inside urban area | 35-44y |
| motorways | | Protective system | 45-54y |
| Context | | % seat belt use in front of cars and vans | • 55-64y |
| Demographic background | | Daytime running lights (DRL) | • 65+ |
| % 15-24y | | Regulation DRL | Age unknown |
| % 65+ | | Vehicle | % fatalities per person class |
| Economic background | | % cars < 6 years | Vehicle occupant |
| Gross Domestic Product | | Roads | Motorized two-wheeler |
| Unemployment % | | Density of motorways | Pedestrian |
| Geographic background | | Trauma management | Cyclist |
| Geographic position in Europe | / | % Gross Domestic Product spent on health care | Unknown |
| # wintry months | | | % fatalities per road type |
| Transport background | | | Inside urban area |
| % pass. kms | | | Outside urban area |
| • Car | | | Motorway |
| Motorized two-wheeler | | | No motorway |
| • Bus | | | Unknown |
| Train | | | |
| Pers, kms per person per year | | | |
| Bicycle | | | |
| • Walk | | | |
| % road length | | | |
| Motorways | | | |
| National | | | |
| Secondary | | | |
| Other | | | |

Conclusion

universiteit



3.3. Footprint methodology (4) Conclusion

Introduction

۲

Data description

Method and results

- •<u>Regression</u>
- •<u>Clustering</u>

•<u>Footprint</u> Description General footprint Detailed footprint **Conclusion**

<u>Conclusion</u>



- Belgium underperforms in many areas
 - <-> the EU average:
 - Concerning the SPI's: 'protective systems', 'alcohol and drugs' and 'speed'
 - Concerning the SOI's: 'share of fatalities between 15 and 44 years' and 'share of fatalities on motorways'
- <-> the SUN-average:
 - Concerning the SPI's: on all risk domains except 'roads'
 - Concerning the SOI's: 'share of fatalities between 25 and 44 years', 'share of fatalities inside urban areas' and 'share of fatalities on motorways'
- The footprint methodology can be applied to every country revealing its best and worst characteristics



4. Conclusion

Introduction

<u>Data</u> description

Method and results

- <u>Regression</u>
- •<u>Clustering</u> •Footprint

Conclusion

universiteit

- A data set of RS indicators was analysed using:
 - Regression analysis
 - Cluster analysis
 - Footprint methodology
- Indicators explaining the number of road fatalities per million inhabitants:
 - 'alcohol and drugs' (% drivers < alcohol limit)</p>
 - 'protective systems' (% seat belt use in front of cars & vans)
 - ➤ 'vehicle' (% cars < 6 years)</p>
 - 'average number of persons within a family'
 - 'gross domestic product'
 - 'average number of kilometers traveled by a driver'
 - 'number of passenger cars per 1000 inhabitants'
 - 4th IRTAD Conference, Seoul, 16-17 September 2009



4. Conclusion (2)

- Clustering:
 - Belgium is often grouped with South-European countries
 - The SUN-countries show apart from their similar road safety level some resemblances concerning their economic and transport background
- The footprints revealed Belgium's most important problem areas and are a valuable instrument for policy makers and other users
- Future research:
 - Take more indicators into account
 - Gathering data for more countries
 - Impute missing data
 - 4th IRTAD Conference, Seoul, 16-17 September 2009

Introduction

Data description

Method and results

- <u>Regression</u>
- •<u>Clustering</u>
- •<u>Footprint</u>

Conclusion

universiteit



Thank you for your attention! Questions?

