NOVA IMS Information Management School

Accidents' point of incidence in Lisbon

Beatriz Macedo
m20201719@novaims.unl.pt

Frederico Ferreira
m20201723@novaims.unl.pt

Ricardo Martins
m20201443@novaims.unl.pt

Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

Context

Objectives

Urban mobility plays an important role in addressing urban livability

Traffic accidents have an important impact in urban mobility, affecting the livability of the cities

External factors such as the condition of the roads or the meteorological conditions, for example, can affect or increase accidents

Identify the points with the highest incidence of road accidents and their correlation with exogenous factors

Workflow

1

Received data from LxDataLab

- Road Accident Occurrences*
- Accident data from the National Road Safety Agency
- Slopes
- Crossroads
- Signage
- Traffic lights position
- Altimetry
- Traffic data*

Obtain geo data from ArcGIS Lisboa

- Lisbon Parish polygon
- Metro Stations

Get Lisbon parish data from INE

- Parish demographic data
- Parish mean house pricing

2

Trim weekdays

Format date time columns

Time of accident binned into “morning”, “afternoon” “evening” and “night”

Parish organization from 2011 converted to organization since 2013

Accidents with missing GPS coordinates were dropped (for visual representations)

Accidents intersected with Lisbon parish polygons for accurate representation

Number of passengers, vehicles and pawns merged with accidents

Slope, Altimetry, crossroads and traffic lights were merged with the accidents data

3

Graphical representation of accidents

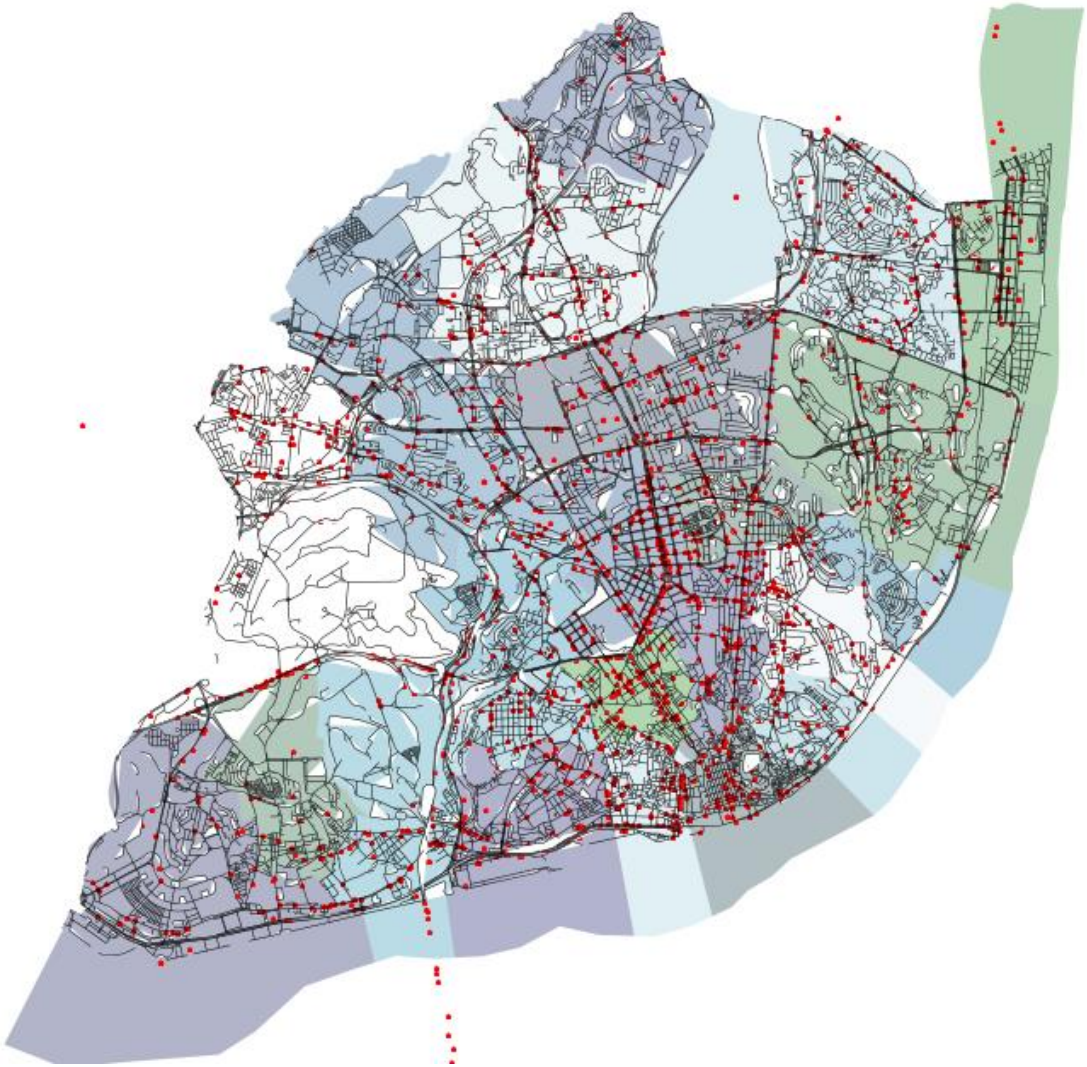
Exploratory analysis

Correlation analysis

Linear regression

* Although available this data was not considered for analysis

The big picture...



Strategy

Strategy

When…

do accidents happen?

1

- Distribution of the accidents during 2019
- Distribution of the accidents per weekday
- Distribution of the accidents per hour of the day

Why…

do accidents happen?

2

- Road conditions and types of road
- Road signaling
- Meteorological and lighting conditions

No conclusions reached

Where…

do accidents happen?

3

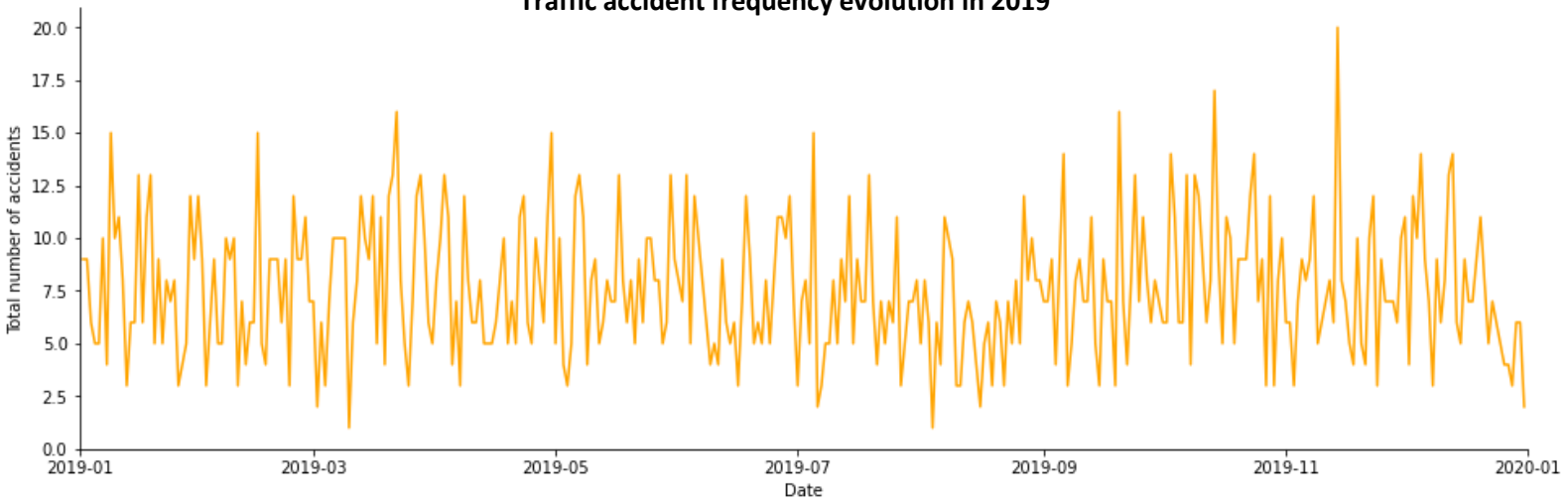
- Frequency of accidents per parish
- Frequency of the types of accident per parish
- Evaluate the impact of the parish's endogenous characteristics in the number of accidents

Project Focus

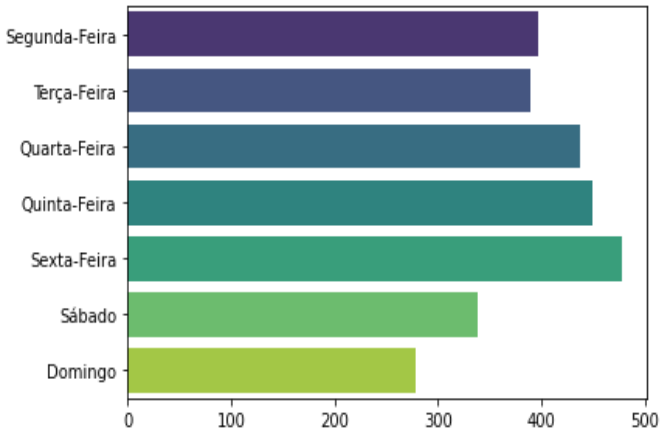
Results

Temporal analysis of accidents

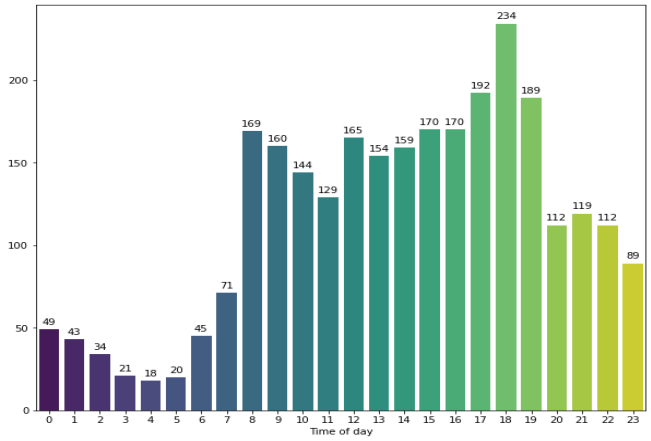
Traffic accident frequency evolution in 2019



Number of traffic accidents registered per weekday



Number of traffic accidents registered per hour



The typical accident in Lisbon

Ambient
- day
- dry
- good weather

Roadside
- well paved

Road direction / inclination
- Straight
- flat

Road
- outside highways
- in a good state
- asphalt
- clearly marked

Lane
- Right

Intersection
- Roads with no intersection

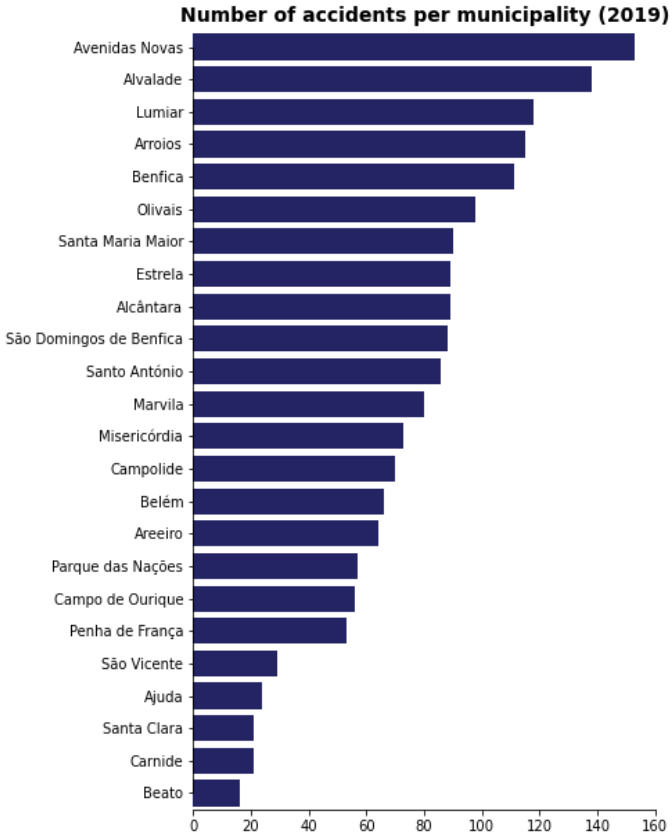
Context

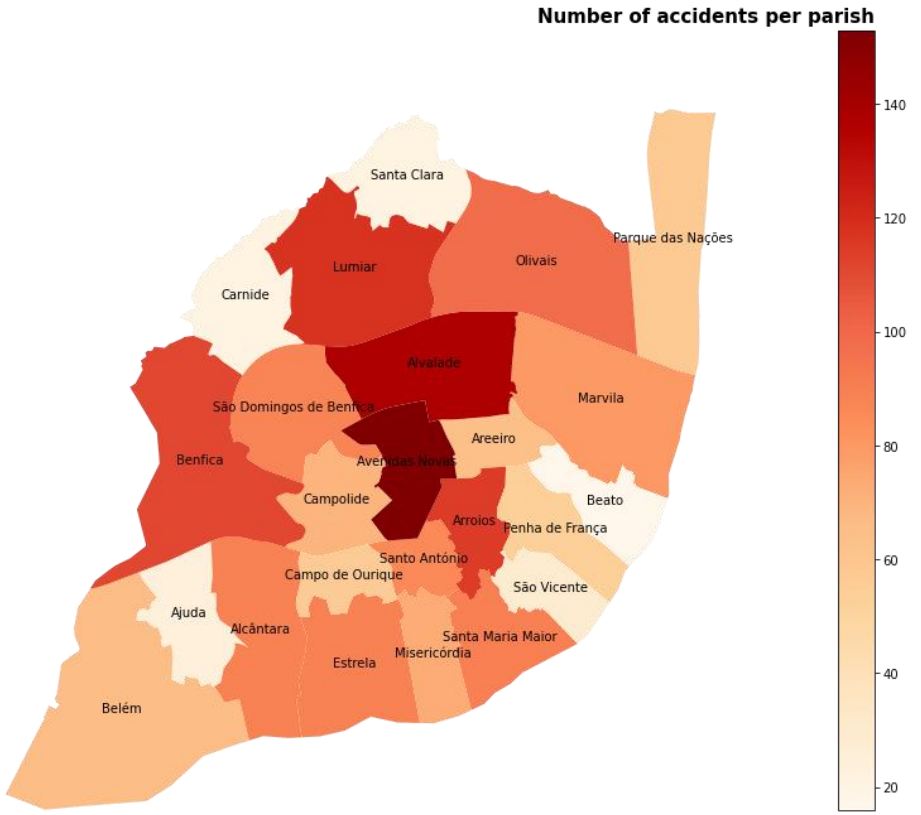
Most accidents occurred in the presence of favorable external factors

- Exogenous characteristics as the road condition or weather do not seem to influence the outcome
- Human error may have a greater impact on the occurrence of accidents than external factors

Focus on the parish characteristics

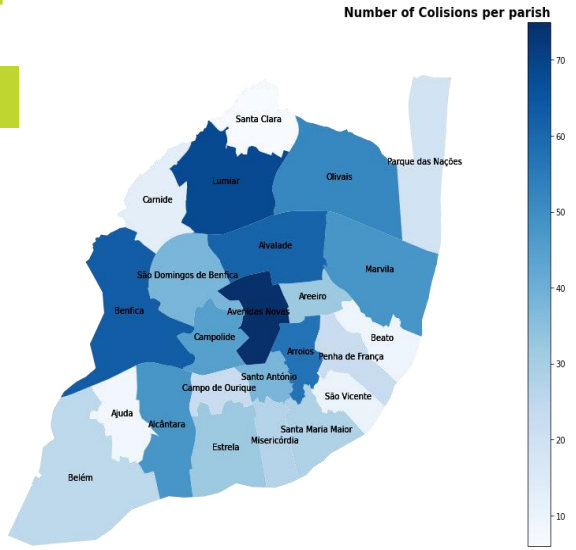
Frequency of accidents in Lisbon

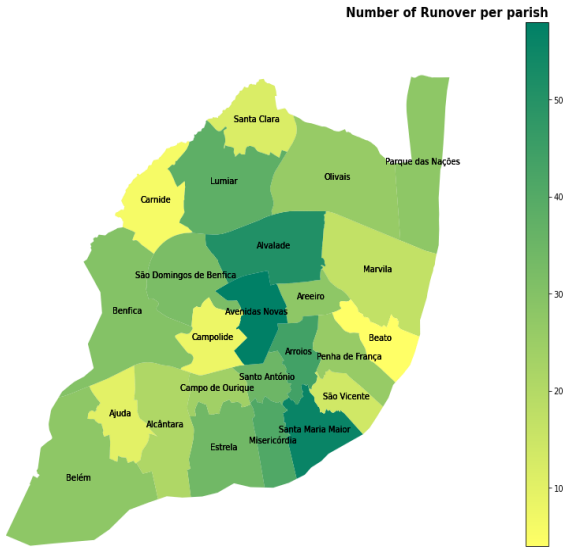


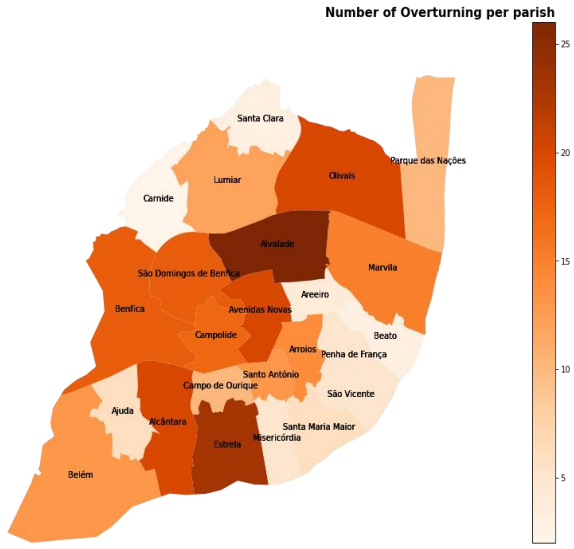


- The parishes with higher incidence of accidents are **Avenidas Novas** and **Alvalade**
- However, this may be related with endogenous characteristics of these parishes (e.g. resident population, number of people commuting, complexity of the roads, area, etc.)
- In order to corroborate this analysis, the number of accidents per parish were compared taking into consideration the endogenous characteristics of the parishes

Frequency of each type of accident in Lisbon







- When considering only collisions, Avenidas Novas, Lumiar and Benfica are the top 3 parishes, followed by Alvalade and Arroios
- Considering runover accidents by parish, besides Avenidas Novas and Alvalade, Santa Maria Maior also stands out
- Plotting the overturn accidents, it is possible to observe that Alvalade and Estrela are the parishes with higher number of accidents
- **Conclusion:** Parishes that have a higher number of accidents, generally, also register a higher number of accidents of each type (Alvalade and Avenidas Novas stand out in all the categories)

Regression model

OLS Regression Results

Dep. Variable:	N_Accidents	R-squared:	0.870
Model:	OLS	Adj. R-squared:	0.834
Method:	Least Squares	F-statistic:	24.17
Date:	Sat, 08 May 2021	Prob (F-statistic):	2.12e-07
Time:	16:13:39	Log-Likelihood:	-9.5387
No. Observations:	24	AIC:	31.08
Df Residuals:	18	BIC:	38.15
Df Model:	5		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	1.527e-16	0.085	1.8e-15	1.000	-0.178	0.178
Commute_Automovel_Pass	-0.6934	0.215	-3.222	0.005	-1.146	-0.241
Commute_Moto	0.6585	0.178	3.706	0.002	0.285	1.032
AREA_M2	0.4683	0.115	4.070	0.001	0.227	0.710
N_Metro_Stations	0.3292	0.185	1.782	0.092	-0.059	0.717
N_Crossroads	0.4726	0.179	2.645	0.016	0.097	0.848

Omnibus:	0.170	Durbin-Watson:	1.742
Prob(Omnibus):	0.919	Jarque-Bera (JB):	0.119
Skew:	0.132	Prob(JB):	0.942
Kurtosis:	2.779	Cond. No.	6.30

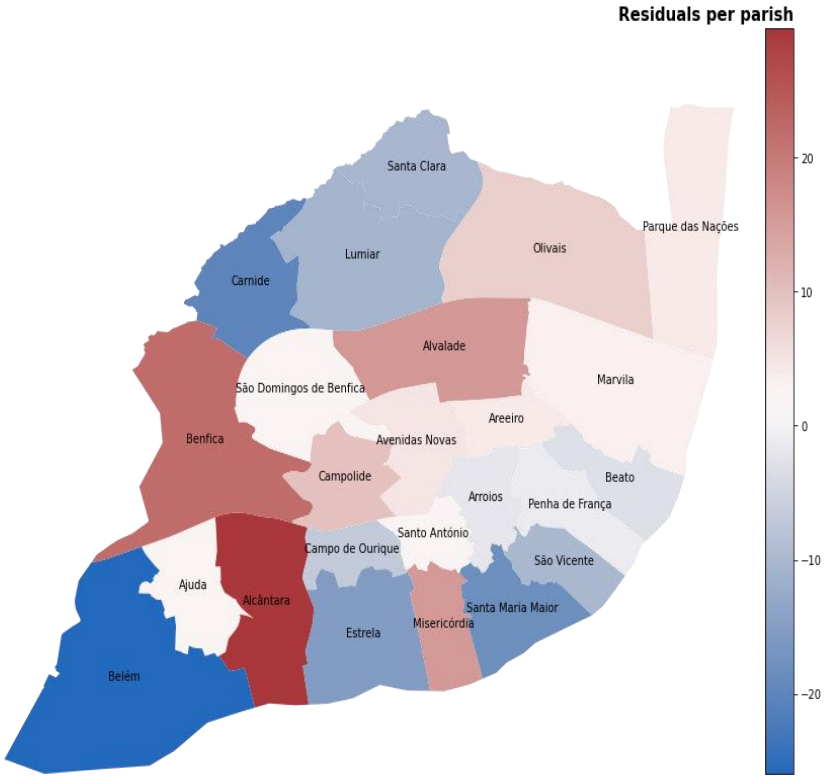
Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Commute_Automovel_Pass	Number of passenger commutes in cars
Commute_Moto	Number of commutes in motorcycles
Area_M2	Parish area
N_Metro_Stations	Number of metro stations
N_Crossroads	Number of crossroads

$$\boldsymbol{N\_Accidents}_i = 1.527.10^{-16} - 0{,}69\ \text{Commute\_Automovel\_Pass}_i + 0.66\ \text{Commute\_Moto}_i + 0.47\ \text{AREA\_M2}_i + 0.33\ \text{N\_Metro\_Stations}_i + 0.47\ \text{N\_Crossroads}_i$$

- All the regression coefficients are significant at least at the 10% level
- 87% of the dependent variable variability is explained by the regressors
- Most features have a positive marginal impact on the target, except for Commute_Automovel_Pass

Predicted accidents VS Registered accidents



- The model was used to predict the of number of accidents for each parish
- The predicted values were compared with the observed values in order to determine if the endogenous characteristics influence the number of accidents observed
- For most parishes, the deviations between the predicted and the observed value is small
- For **Belém** the model is overestimating the number of accidents and for **Alcântara** is underestimating the number of accidents

Final Thoughts

Conclusions + Critiques + Challenges

Challenges

- The dataset did not comprise all the accidents registered during the year (is the sample representative?)
- Waze data did not have timestamps - did not allow to match traffic data with accidents
- INE data was not from the same period as the accidents

- Unable to identify the external factors that contribute to the occurrence of accidents – No way of comparing the characteristics of the accidents with the characteristics of "non-accidents"

Critiques

Conclusions

- Endogenous factors of the parishes influence the occurrence of accidents
- The model was not accurate at predicting the number of accidents for Belém and Alcântara, indicating there may be some characteristics explaining the occurrence of accidents other than the ones contemplated in the model
- Avenidas Novas and Alvalade registered the highest number of accidents, however it cannot be implied that these parishes are more dangerous than others

NOVA IMS Information Management School

Thank you

Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa