Accidents in the city of Porto Alegre

Felipe Paula

11/8/2018

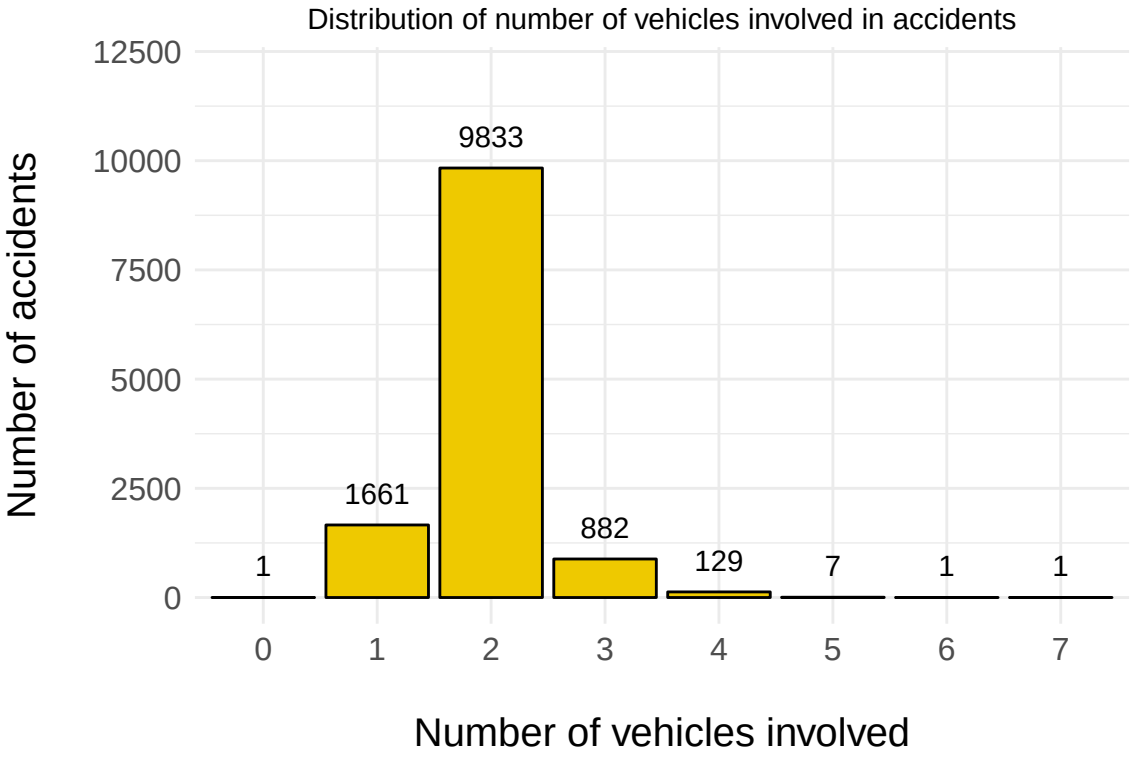
Data

In this document, we analyze the data regarding the accidents that occured in the city of POA (Porto Alegre). The data is provided by the city's center of open data #DataPOA. In our analisys, we focus in the year of 2016. There's also a dictionary of the semantics of the data available in the website.

The data presents 12515 rows and 44 columns. Including, information about date and time of the accident, vehicles involved, number of people injuried, location, and weather. Since 2016 is a leap year, the dataset spans 366 days, with accidents in everyday.

How many cars are involved in accidents?

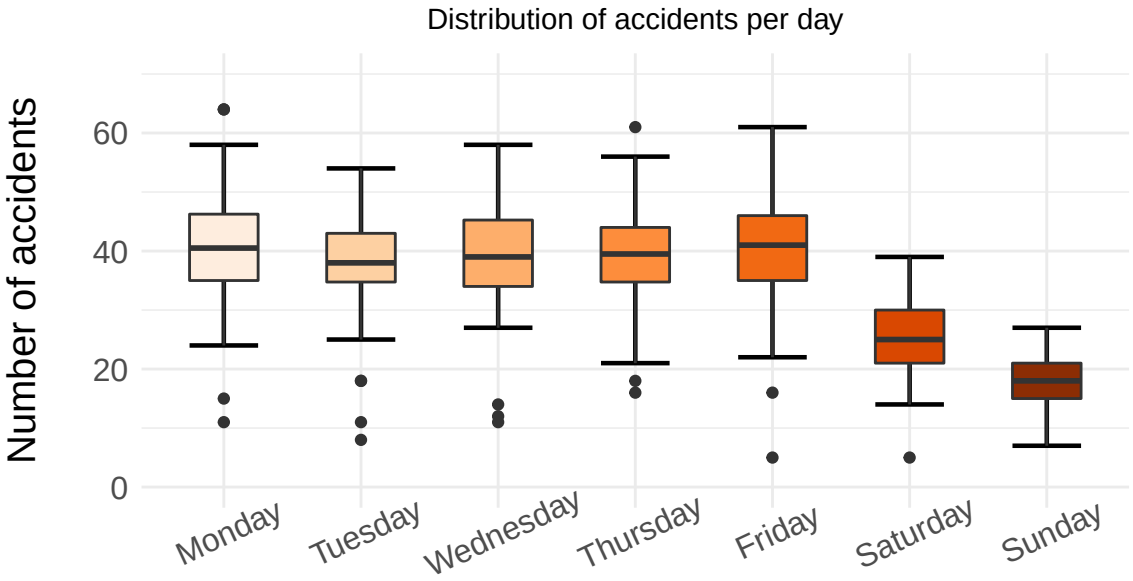
For each accident, the dataset indicates how many cars, taxis, buses, trucks, motobikes, bikes and others are involved. To answer the question, we plot the distribution of the sum of the vehicles involved. As expected, the majority of the accidents involve 2 cars. The accident which involved 0 vehicles was a run over with no vehicle registred.



Is there a specific weekday with more accidents?

To access if there is a day of a week which more accidents occur, we also plot the distribution of the number of accidents in each week day. An one way one factor anova test reveals differences between the means of the

days. We also used Tukey's test to find significant pairwise differences. There was not a weekday with the mean greater than all others, so the answer of the question is no. Sunday present the smaller mean number of accidents ($p < 0.01$), followed by Saturday.



Is there a time of the year with more accidents?

Since during the weekend there are less accidents, we aggregate the data by weeks. The figure below shows the number of accidents evolution through the weeks of the year. Also, we mark interesting months in the plot, with their mean number of accidents over the days. The mean accident per week in 242 (sd = 31.5). One way anova did not show difference between the number of accidents among the months. So we can't say that there is a time of the year with more accidents.