# Car accidents in the US between 2016 and 2020

Catchy subtitle

Written by Eva Verkade on Month ##, 2023.

# Car accidents in the US between 2016 and 2020

## introduction

text

## description of dataset

* This dataset contains data on car accidents in the United States of America that occurred between … 2016 and … 2020.
* The size of the dataset is sizeable: … GB.
* Description of main sets of variables. Link to the pdf?

**Variables:**

## Preparation

The dataset was imported into Power BI.

* Remove columns: Because of the size of the dataset it was important to remove columns that are not relevant for the analysis. This will improve performance, an important aspect because of the dataset size.
* Twilight
* Separate datetime + date table: Datetime columns are a known to use a lot of memory. The datetime column was split into a date and a time column. After that, a date table was created and all date columns were changed into an integer date key. An active relationship was established between the Start Date Key and DateKey, and an inactive relationship was established between the End Date Key and DateKey. DateKey is used in the report and a role-playing dimension to filter the results.
* No star schema
* Deal with missing values/nulls
* Group columns/variables (weather, road obstructions, location)

## Dashboard

[embedded report from power BI service]

Explanation of the dashboard, give an overview of the included pages.

## Page 1 explanation

In-depth into page 1

## Page 2 explanation

In-depth into page 2

## Page # explanation

In-depth into page #

## Page # explanation

In-depth into page #

## Conclusion (overall)

Conclusion of the accidents. Any advice for insurance companies?