## ****背景描述****

这是英国交通部 2016 年至 2020 年的公开数据。

它是从[https://data.gov.uk/dataset/cb7ae6f0-4be6-4935-9277-47e5ce24a11f/road-safety-data下载的，可在](https://data.gov.uk/dataset/cb7ae6f0-4be6-4935-9277-47e5ce24a11f/road-safety-data%E4%B8%8B%E8%BD%BD%E7%9A%84%EF%BC%8C%E5%8F%AF%E5%9C%A8) UK OGL 下使用。

## ****数据说明****

在讨论部分是将这些表导入 PostgreSQL 和 R 的脚本。笔记本中是我创建的快速集群脚本。

1. **Road-Safety-Open-Dataset-Data-Guide 20212310 v\_02.xlsx**是我根据现有数据字典创建的查找表。它的用途是使数据更容易缩小范围（例如天气状况、事故是否发生在城市、涉及多少行人等）。  
   在此文件中，标记为“事故查找”、“伤亡查找”和“车辆查找”的选项卡应单独保留，因为它们只是对“查询查找”的引用。

它是在谷歌表格中创建的，但也应该在 Excel 中打开。这是谷歌表格中的[链接。](https://docs.google.com/spreadsheets/d/1rwlW6zbYyNNhbHGi_nqK2CRJxC41s3g2fDX-DqOQSIA/edit?usp=sharing)

1. **dft-road-casualty-statistics-accident-last-5-years.csv、dft-road-casualty-statistics-casualty-last-5-years.csv 和 dft-road-casualty-statistics-vehicle-last-5- years.csv**是感兴趣的主要文件。

### 如果您希望将它们连接在一起，每个共享一个****accident\_index****的公共键。

## ****数据来源****

<https://www.kaggle.com/datasets/benjaminfox/uk-road-traffic-data-2016-2020>

## ****问题描述****

天气状况、事故是否发生在城市、涉及多少行人等预测