

Analysis of the Accidents Dataset

Information on over 150,000 road accidents with personal injury in 2010 in Great Britain was obtained from the Government of the UK. It contains information on the date, time, and other variables about the accident.

We investigated the distribution of the number of accidents by the reported minute of the hour (Figure 1 below). Notice the clear evidence of data heaping where the reported time of the accident is much larger at the hour and half hour, and at five minute intervals. This heaping is strictly an artifact of people recording the time of the accident to convenient boundaries and has no relevance for policing. What is surprising, is the number of accidents with reported times that are NOT rounded to convenient boundaries, e.g. at 23 past the hour! It is unclear why these accident times are not rounded and this requires further investigation.

