Etapas do processo de limpeza dos dados:

* Decidimos utilizar como padrão, os dados referentes ao ano de 2016;
* Analisamos as colunas que não seriam úteis e as removemos de todos os anos, com isso as seguintes colunas foram removidas:
  + ID
  + PREDIAL1
  + LOCAL
  + DATA
  + HORA
  + FONTE
  + BOLETIM
  + CONT\_ACID
  + CONT\_VIT
  + CONSORCIO
  + CORREDOR
* Trocamos a posição das colunas "LONGITUDE" e "LATITUDE" entre elas;
* Depois de organizar as colunas dos dados de 2016, começamos a organizar e verificar a dos anos anteriores;
* Os dados dos anos 2000 até 2012 não tinham as seguintes colunas "QUEDA\_ARR", "FERIDOS\_GR" e "ONIBUS\_MET", acrescentamos essas colunas e definimos elas com o valor 0 (zero);
* Padronizamos o valor da coluna "DATA\_HORA" com o seguinte formato yyyy-mm-dd'T'hh:mm:ss.

Para poder padronizar a data e hora neste formato utilizamos a seguinte função:

*value.substring(0,4)+"-"+value.substring(4,6)+"-"+value.substring(6,8)+"T"+value.substring(8)+":00";*

*value.replace(" ", "");*

* Padronizamos as colunas "LATITUDE" e "LONGITUDE" com o carácter "." ao invés do carácter "," nos dados de 2000 a 2014;
* Constatamos que os dados referentes a latitude e longitude não estavam mostrando a localização exata no mapa, então subtraímos 800 da latitude para mostrar uma localização exata no mapa;
* Definimos que se algum campo das colunas "LATITUDE", "LONGITUDE" e "UPS" estivessem vazias a linha inteira foi removida.