

事故多発地点解析ツール マニュアル

目次

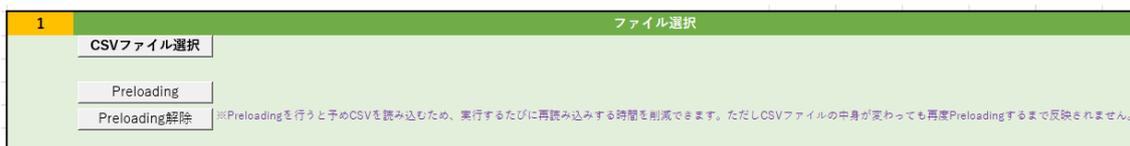
- 1 使い方
 - 1.1 ファイル選択
 - 1.2 集計
 - 1.2.1 経緯度グリッド集計
 - 1.2.2 距離グループ集計
 - 1.2.3 集計結果
 - 1.3 地図出力
 - 1.4 高度なフィルタ
- 2 集計アルゴリズムについて
 - 2.1 集計対象
 - 2.2 経緯度グリッド集計
 - 2.3 距離グループ集計
 - 2.4 集計出力
- 3 地図出力について
 - 3.1 地図の仕様

1 使い方

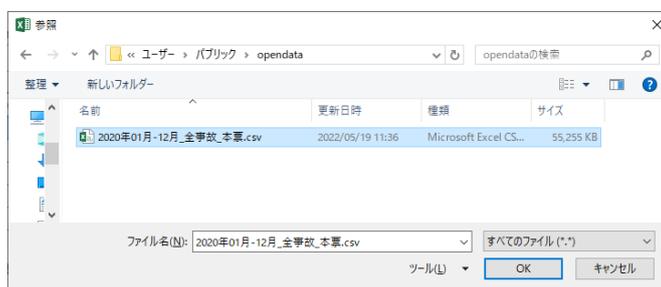
操作は「実行」シートから行います。

1.1 ファイル選択

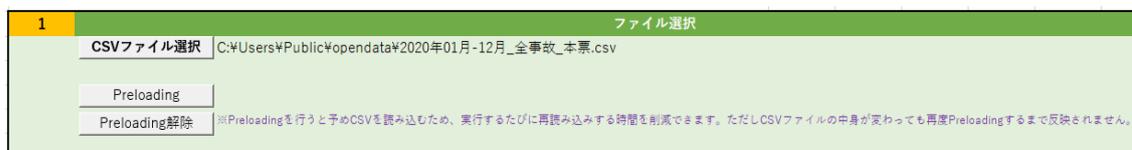
1 のファイル選択から警察庁の交通事故統計情報のオープンデータを選択します。



[CSV ファイル選択] ボタンをクリックすると、ファイル選択画面が開きます。



オープンデータの本票 CSV ファイルを選択して OK ボタンを押します。

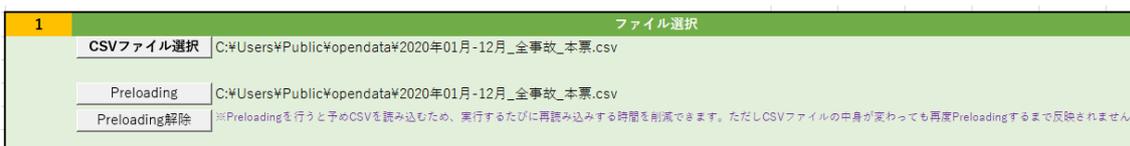


[CSV ファイル選択] ボタンの右に選択された CSV ファイルのパスが表示され、選択が完了します。

・ Preloading 機能

本票 CSV ファイルは 55MB 前後の大容量のため、集計時の読み込みに時間がかかります。連続して集計を行いたい時には[Preloading]ボタンから事前読み込みしておくことで、時間を短縮できます。(Preloading をしなくても集計は可能です。)

[Preloading] ボタンを押すと「* CSV ファイル、オープン中」の表示が始まり、「Preloading」成功のポップアップが出ると完了します。



[Preloading] ボタンの右に読み込んだファイルパスが表示されます。以後、集計はこの読み込んだデータを使い、CSV ファイルのオープンをスキップできます。

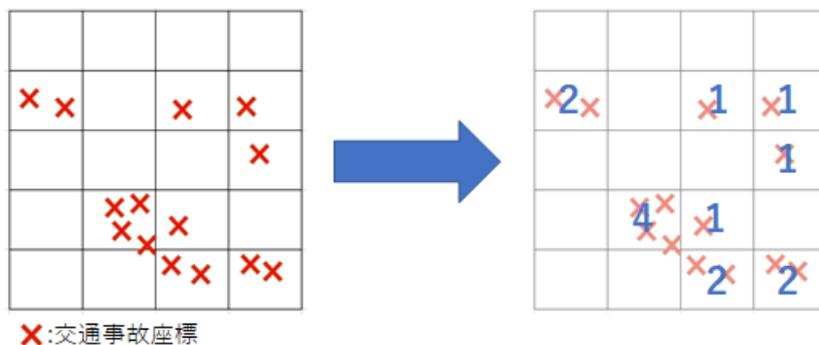
[Preloading 解除] ボタンで事前読み込みを解除できます。

1.2 集計

オープンデータの集計手法として、次の2種類を実行できます。

1.2.1 経緯度グリッド集計

経緯度グリッド集計では、地図を緯度幅と経度幅で指定した長方形のグリッドで地図を区切り、各グリッドに含まれる交通事故の件数を数え、件数が多い順にグリッドの中心の経緯度を出力します。



ファイル選択後、2の経緯度グリッド集計の項目から実行します。

2 経緯度グリッド集計			
緯度幅	1000	ミリ秒	
経度幅	1000	ミリ秒	
都道府県	全国		
上位出力件数	50	件	

件数集計 出力クリア

緯度幅と経度幅をミリ秒の単位で入力します。日本付近では1000ミリ秒=1度の長さは、緯度が約30メートル、経度が約25メートルになります。例えば共に1000ミリ秒を設定すると、南北30メートル・東西25メートルの長方形のグリッドで集計します。

都道府県を設定すると、対象地域の交通事故だけを集計します。この設定は「高度なフィルタ」を使うとさらに細かく定義することができます。

上位出力件数は集計結果の事故件数のうち上位何件を「多発地点出力」シートに出力するかを設定します。

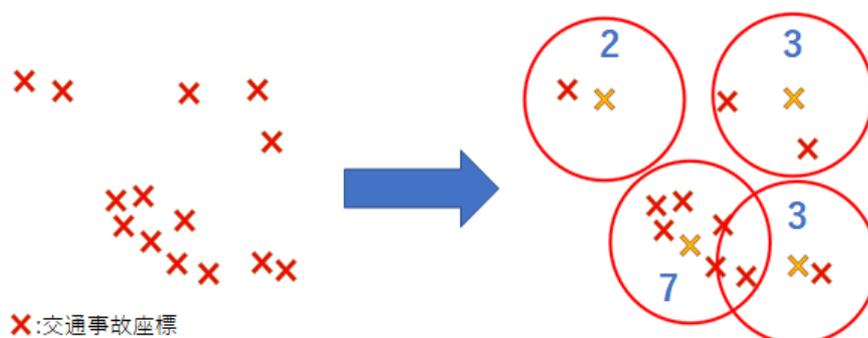
[件数集計]ボタンを押すと集計作業が開始され、「集計完了」のポップアップで完了します。事故件数の多い座標と含まれる交通事故統計情報のデータを「多発地点出力」シートに出力し、各座標の事故件数のリストを「プロット用出力」シートに出力します。

なお、すでに出力されている集計結果は上書き削除されます。

[出力クリア]ボタンを押すと「多発地点出力」と「プロット用出力」シートの内容が削除されます。

1.2.2 距離グループ集計

距離グループ集計では、各々の交通事故座標から一定半径以内の交通事故の件数を数え、件数が多い順に中心点の交通事故の座標を出力します。



ファイル選択後、2'の距離グループ集計の項目から実行します。

2'	距離グループ集計	
グループ化半径	30	メートル
都道府県(2と共通)	全国	
上位出力件数(2と共通)	50	件

件数集計 出力クリア

交通事故の件数を数える円の半径をグループ化半径(メートル)で指定します。
都道府県と上位出力件数は2の経緯度グリッド集計の項目で設定した値が入ります。
[件数集計]ボタンを押すと集計作業が開始され、「集計完了」のポップアップで完了します。経緯度グリッド集計よりも集計に時間を要します。
集計結果は経緯度グリッド集計と同様に出力されます。
この集計では円の中心点同士が重複する場合、件数が最も多い座標以外を除外しています。

1.2.3 集計結果

集計の結果は「多発地点出力」と「プロット用出力」シートに出力されます。「多発地点出力」は件数が上位(ただし最低でも2件以上)の座標について、個別事故の情報を列挙します。「プロット用出力」は集計を行った全座標と件数のみを出力します。

多発地点出力

集計方法とパラメータ								
A	B	C	D	E	F	G	H	
距離グループ	距離:	30						
北緯	東経	件数	資料区分	都道府県	警察署等	本票番号	事故内容	死
353936744	1394636246	14	1	30	107	48	2	0
			1	30	107	52	2	0
			1	30	107	90	2	0
			1	30	107	64	2	0
			1	30	107	108	2	0
			1	30	107	89	2	0
			1	30	107	28	2	0
			1	30	107	107	2	0
			1	30	107	126	2	0
			1	30	107	36	2	0
			1	30	107	75	2	0
			1	30	107	134	2	0
			1	30	107	73	2	0
			1	30	107	98	2	0
354428879	1394224652	11	1	30	555	169	2	0

プロット用出力

集計方法とパラメータ		
A	B	C
距離グループ	距離:	30
北緯	東経	件数
353936744	1394636246	14
354428879	1394224652	11
354020895	1393938508	11
354151982	1395337263	11
354234771	1393915407	10
353923573	1394624841	10
354518268	1394117067	10
354227714	1394634501	9
353843724	1393704778	9
354414593	1394248472	9
354103975	1392045782	9

1.3 地図出力

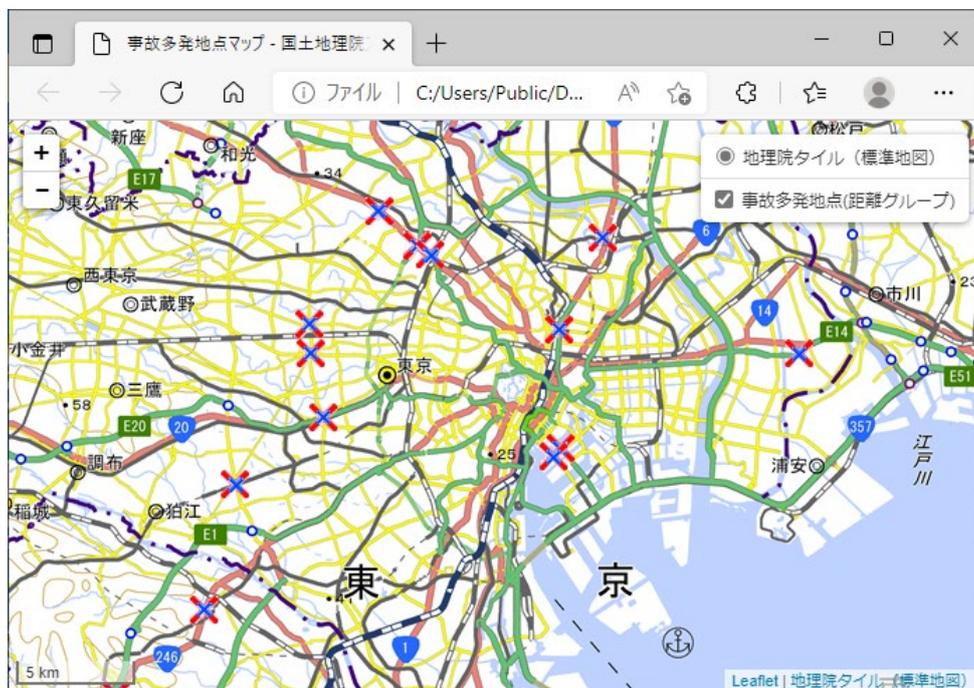
「多発地点出力」シートの内容を地図上に可視化することができます。表示にはインターネット接続が必要になります。

3 地図出力		
事故座標出力	有効	多発地点に含まれる個別の事故座標をプロットします
集計範囲出力	有効	多発地点の集計範囲を枠を表示します
地図出力		
多発地点出力のデータをGeoJSONに整形して国土地理院地図を利用したLeafletアプリケーションとして出力します。 出力されたHTMLファイルはインターネット環境に移して閲覧できます。 国土地理院情報普及課Github より https://github.com/gsi-cyberjapan/geojson-with-style-spec		

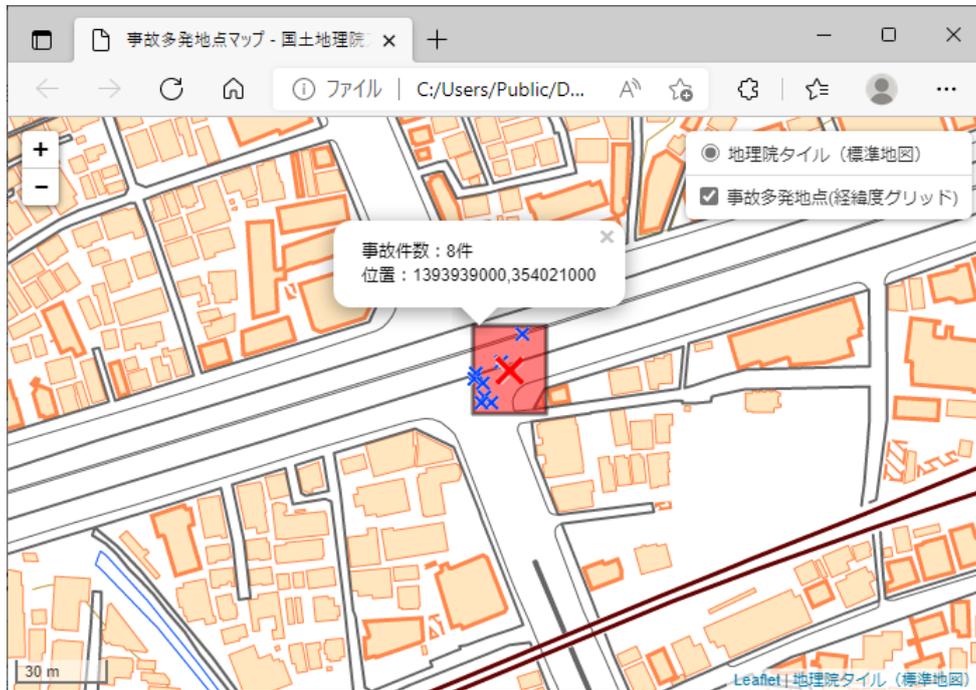
動作には 1.2 の集計で「多発地点出力」シートが出力されている必要があります。
事故座標出力を有効にすると、事故多発地点に含まれる個別事故の座標もプロットします。
集計範囲出力を有効にすると、集計を行った範囲を図形表示します。

「地図出力」ボタンを押すと、「事故多発地点マップ_(タイムスタンプ).html」の名前のファイルがツールと同じフォルダに出力されます。

このファイルはインターネットに接続できる環境で Web ブラウザを使って開くと、地図上に集計した事故多発地点が表示されます。

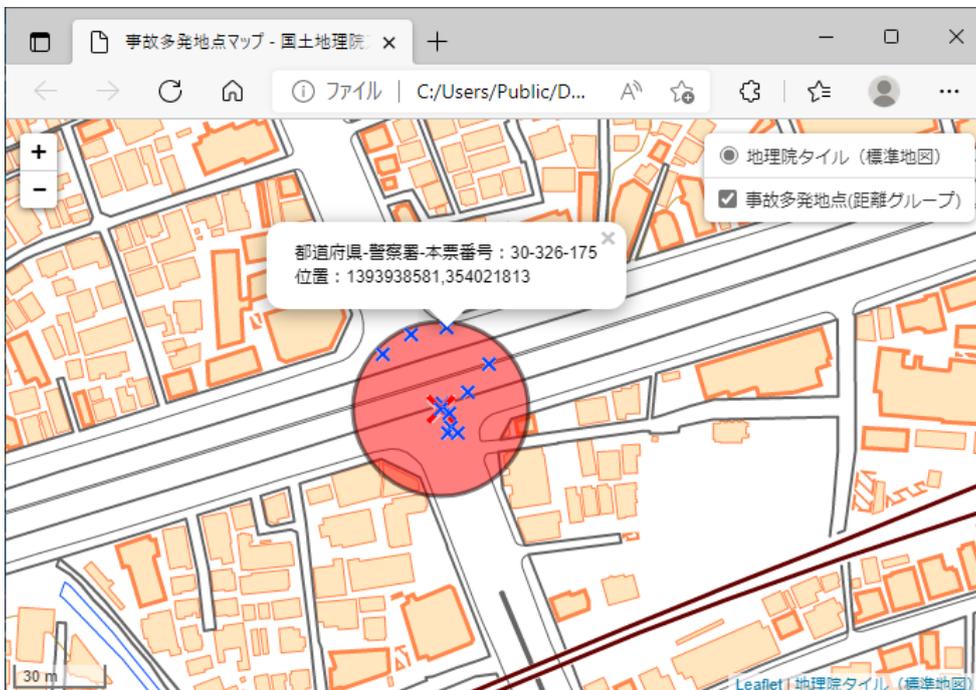


経緯度グリッド集計の表示



経緯度グリッド集計の集計範囲出力では、集計範囲は長方形になります。

距離グループ集計の表示



距離グループ集計では範囲は円形になります。

集計範囲をクリックすると事故件数と赤×印の座標を表示します。青×印は個別の事故の位置で、クリックすると座標と本票番号を表示できます。

1.4 高度なフィルタ

「高度なフィルタ」シートを使うと交通事故統計情報のデータを様々な条件で絞り込んで集計することができます。

フィルタ作成		C:\Users\Public\opendata\2020年01月-12月_全事故_本票.csv			
[フィルタ作成]ボタンを押すと、csvファイル内の列を取得して指定した値でデータを絞り込むシートを無指定でフィルタなし、値を指定すると指定した値を持つ列だけを集計対象とする。値のリストはコピーペースト使用可能：例1：1桁目が9の4桁の数字を指定したい場合は9###とする 例2：5000～5					
		フィルタ値 [OR条件]ー			
フィルタ列 [AND条件]ー	資料区分				
	都道府県コード				
	警察署等コード				
	本票番号				
	事故内容				
	死者数				
	負傷者数				
	路線コード				
	上下線				
地番コード					

1 のファイル選択後の状態で[フィルタ作成]ボタンを押すと、その CSV ファイルの 1 行目の列名をもとにしてフィルタ列を作成します。ただし、デフォルトで交通事故統計情報の本票データを使って作成済みとなっています。

2 経緯度グリッド集計		フィルタ値 [OR条件]ー			
緯度幅	1000	ミリ秒			
経度幅	1000	ミリ秒			
都道府県	東京		30		
上位出力件数	50	件			
件数集計		出力クリア			

「実行」シートの都道府県入力欄とは連動しており、片方を変更するともう片方も変更されます。高度なフィルタの都道府県コードに複数の都道府県を指定すると、都道府県入力欄は「カスタム」の表示になります。

高度なフィルタは、横列に値を並べていくと何れか一つ以上の横列の値に一致するデータだけを集計します(OR 条件)。縦行に値を並べていくとすべての縦行の値に一致するデータだけを集計します(AND 条件)。

例えば以下のように都道府県に 30,43,44,45、事故内容に 2 を入力すると、東京都・埼玉県・千葉県・神奈川県 の 4 都県で負傷者を出した交通事故だけを集計します。

		フィルタ値 [OR条件]ー			
フィルタ列 [AND条件]ー	資料区分				
	都道府県コード	30	43	44	45
	警察署等コード				
	本票番号				
	事故内容	2			
	死者数				

フィルタ値にはワイルドカードを使うことができます。

ワイルドカード文字	意味
?	任意の 1 文字
*	任意の数の文字
#	任意の 1 文字の数字(0~9)
[文字列]	括弧内の文字列のうち 1 文字
[!文字列]	括弧内の文字列に含まれない 1 文字

2 集計アルゴリズムについて

2.1 集計対象

読み込まれた CSV ファイルの各行は高度なフィルタで指定した値で Like 演算子を適用してフィルタリングされます。

データに極端に小さな経緯度の値があるとアフリカ付近にプロットされるため、北緯 20~46 度、東経 122~154 度の範囲以外を除外しています。離島を含めた日本全体がこの経緯度に収まります。

集計時に「地点 緯度 (北緯)」「地点 経度 (東経)」の列が必須になり、地図出力時には「都道府県コード」「警察署等コード」「本票番号」の列が必須になります。

2.2 経緯度グリッド集計

実際にグリッドを用意して件数をカウントしているわけではなく、各交通事故の経緯度について、経緯度をグリッド幅で丸め込んだ値でデータをソートして、連続する同一経緯度の個数を数えています。

2.3 距離グループ集計

緯度と経度でデータをソートしたあと、データ毎に付近の他のデータとの距離を計算してグループ化半径以下の件数をカウントしています。

そのまま出力するとたくさんの円が密集してしまうため、ある点をカウントした際、グループ化半径以内の点についてすでにカウント済みの点の件数を調べて、最も大きいものの以外の件数をマイナスにして除外する処理を行っています。



ここで除外された点は「プロット用出力」シートにも出力されません。

2.4 集計出力

集計した結果を「多発地点出力」シートと「プロット用出力」シートに書き出します。「多発地点出力」では上位出力件数に達するまで降順で多発地点を出力しますが、同一件数の座標が上位出力件数を超えて続く場合はすべて出力するため、上位出力件数の値を超えることがあります。また、件数が 1 件になると上位出力件数未満でも出力を停止します。これは同一件数を全て出力する条件と合わさって、集計を行った全データがシートに書き出されてしまうことを防ぐためです。

もしデータが極端に少なく 2 件以上が存在しない場合、「多発地点出力」シートには何も出力されません。

3 地図出力について

3.1 地図の仕様

インターネット環境で Web ブラウザから開くと Leaflet ライブラリと国土地理院地図を読み込んで動作します。

対応ブラウザは Leaflet に準拠し、この資料を記載した時点で Google Chrome、Microsoft Edge、Firefox など主要な Web ブラウザで表示できることを確認しています。