



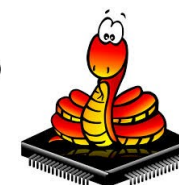
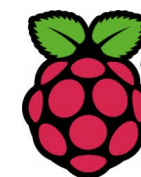
福岡県の交通事故を 見える化してみた

- ・ 危険な時間帯は、いつよ？
- ・ 事故が多発する危険なエリアは、どこよ？
- ・ 福岡県で最も危険な交差点は、どこよ？

自己紹介



- 田中丸 祐治 malo21st
- 某電力会社で技術系の営業をしているPython好き



1 ~ 2 年前

ここ最近

いまからやること（ざっくり）



CKAN

自治体オープンデータ

<https://ckan.open-governmentdata.org>

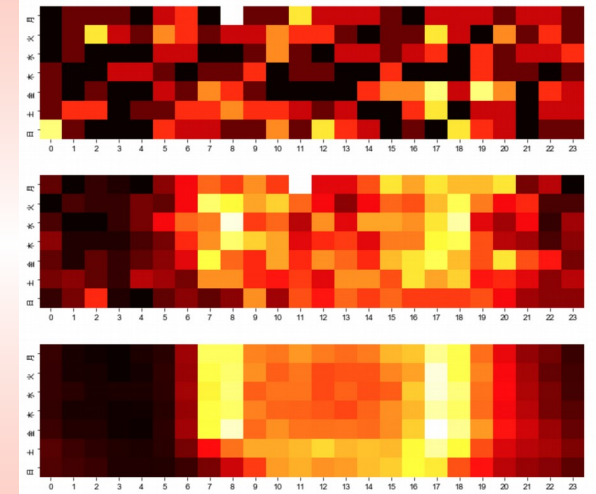
福岡県内で発生した
人身事故データ

・平成28年，平成29年

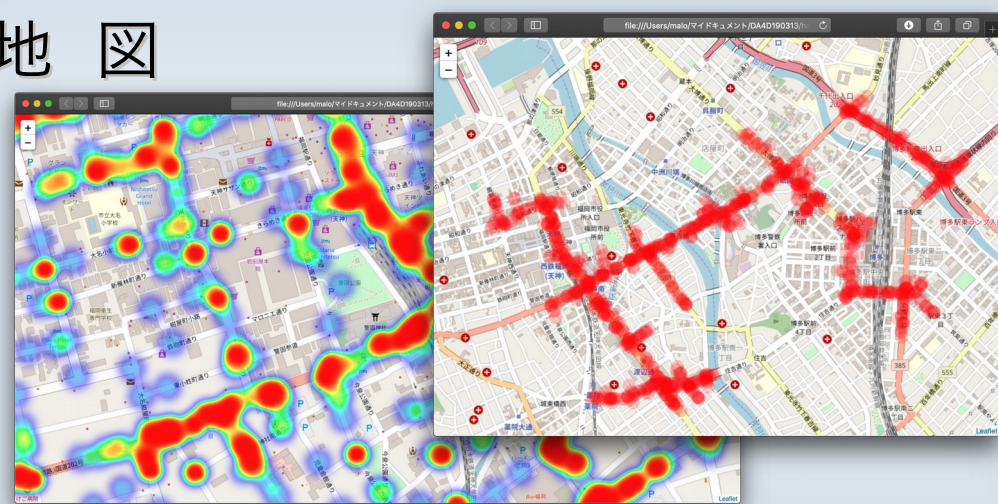
データ

- ・データ形式：エクセル
- ・項目数：25
- ・レコード数：72,170

ヒートマップ



地図



ヒートマップとは



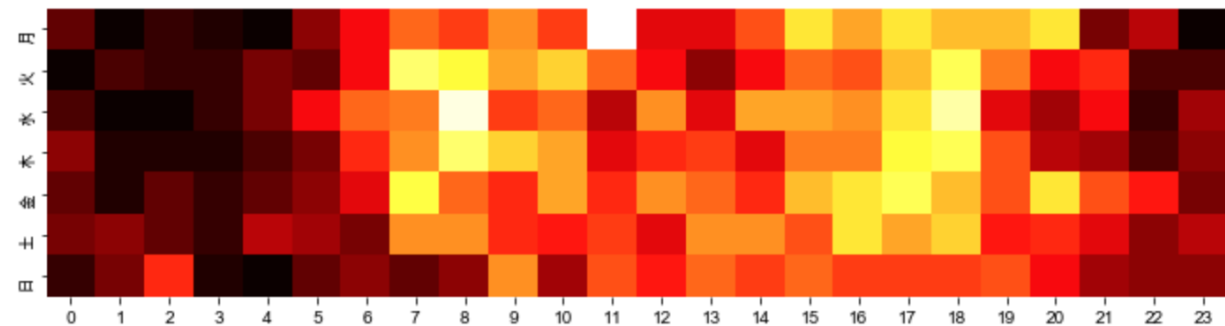
数値データの大小を
色で視覚化する方法



ヒートマップのデモ

- ・ 数値データ
- ・ 地図（位置データ）

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
月	5	1	3	2	1	7	12	17	15	19	15	32	11	11	16	23	20	23	21	21	23	6	9	1
火	1	4	3	3	6	5	12	27	24	20	22	17	12	7	12	17	16	21	26	18	12	14	4	4
水	4	1	1	3	6	12	17	18	31	15	17	9	19	11	20	20	19	23	29	11	8	12	3	8
木	7	2	2	2	4	6	14	19	27	22	20	11	14	15	11	18	18	24	26	16	9	8	4	1
金	5	2	5	3	5	7	11	25	17	14	20	14	19	17	14	21	23	26	21	16	23	16	13	6
土	6	7	5	3	9	8	6	19	19	14	13	15	11	19	19	16	23	20	22	13	14	11	7	9
日	3	6	14	2	1	5	7	5	7	19	8	16	13	17	15	17	15	15	15	16	12	8	7	1
計	31	23	33	18	32	50	79	130	140	123	115	114	99	97	107	132	134	152	160	111	101	75	47	42





デモンストレーション



交通事故が多発する危険なエリアとは？

👉 交通事故が集中して、よく発生しているエリア

👉 なら、発生場所は近いんじゃないか

👉 発生場所が近いなら同類とみなして同じグループ

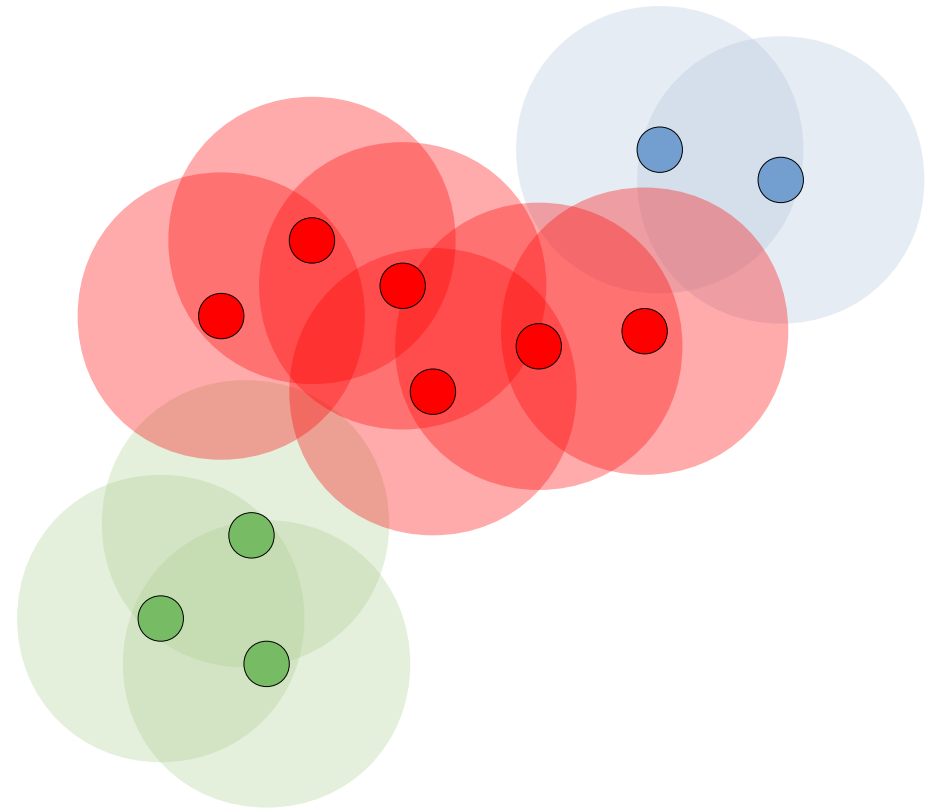


- ・発生場所が近いなら同類とみなして同じグループ

⇒ レコード総当たりで データ相互の距離 を計算

同じグループにする条件

- ・危険なエリア
条件：50m以内
- ・危険な交差点
条件：10m以内





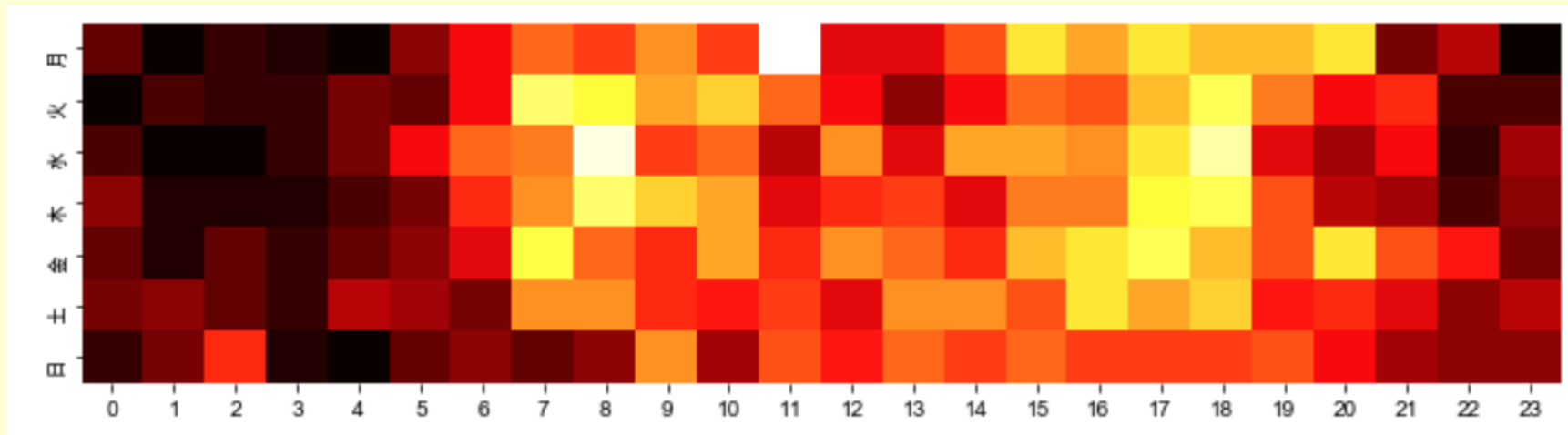
デモンストレーション



ビッグ(大量の)データの見せ方

👉 情報量を上手に減らして、見える化！

・ 曜日 と 時間帯 で集計 72,170 👉 168 (= 7 * 24)





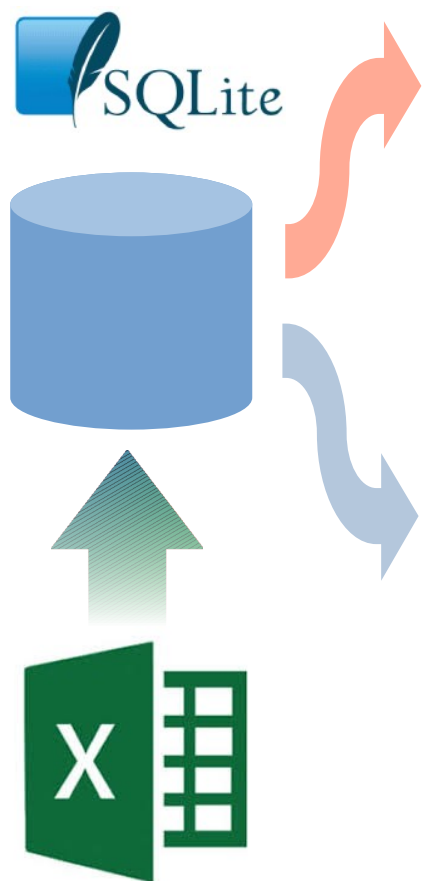
ご清聴ありがとうございました

本日の資料：

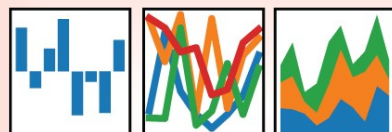
<https://github.com/malo21st/DA4D190313>



(補足) 処理の流れ と Pythonライブラリ



ヒートマップ

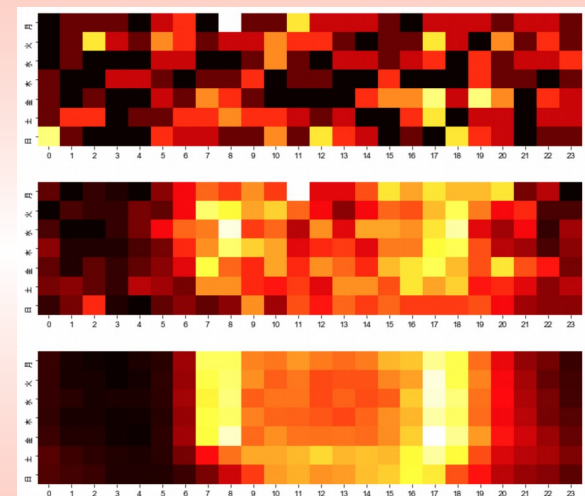


$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$

pandas

ipywidgets

seaborn



地図



NetworkX



folium

