

PT 分析(パーソントリップ分析)

交通研 B4 前田歩美

約半世紀前に始まり現在まで続いている PT 調査は、一定の調査対象地域内における「人の移動」を調べる交通実態調査としては最も基本的な調査である。ゾーン単位の移動を集計するこの調査は、21 世紀に登場した PP 調査と比べると精度が低く分析しづらいように感じられるかもしれないが、広域な人口流動を観察する際にゾーンを利用した適度な抽象化はデータ量が節約でき、逆に扱いやすい場合もある。また交通調査の基本を押さえるという意味でも、PT 分析の手法は理解しておくべきである。

目次

1. 交通計画の手法
 - 1.1 データ収集の手法
 - 1.2 PT 調査とは
 - 1.3 PT 調査の基本概念
 - 1.4 交通計画の手順
 - 1.5 PT 分析のステップ
2. データの抽出
 - 2.1 準備
 - 2.2 抽出の手順
3. 集計
 - 3.1 サンプル属性に基づく集計
 - 3.2 トリップ出発地/到着地に基づく集計
4. QGIS での可視化

1. 交通計画の手法

1.1 データ収集の手法

人口流動の分析を行うには、まずデータを収集する必要がある。データ収集のために様々な調査が考案・実施されてきた。

- 1928 道路交通センサス
- 1958 OD 調査(自動車起終点調査)
 - 戦災復興→(環状)ネットワーク整備
- 1960 大都市交通センサス(鉄道局)
 - 郊外化への対応
- 1967 PT 調査(都市局)@広島

地域の総合計画立案が可能に

1972 物資流動調査

2002 世界初長期間 PP 調査@松山

1.2 PT 調査とは

PT 調査とは、どのような人が「いつ」「何の目的で」「どの移動手段を使って」「どこから」「どこへ」移動したのかを調べるアンケート調査であり、被験者の平均的な一日の移動を捉えることができる。

首都圏では過去に 6 回(1968/1978/1988/1998/2008/2018)されており、約 88 万人分のデータが収集されている(抽出率 3~7%)。

The form is titled "個人票" (Individual Ticket) and is for the "世帯票" (Household Ticket) survey. It is organized into several main sections:

- はじめに (Introduction):** Includes a header with the survey name and date, a section for recording the starting location (before 9 AM), and a list of 15 possible purposes (e.g., school, work, shopping, leisure).
- 1日のはじめにいた場所 (Starting location):** A table for recording the starting location and mode of transport.
- 1番目に行った場所 (1st Trip):** A detailed section for the first trip, including departure/arrival times, purpose, mode of transport, transfer points, and parking.
- 2番目に行った場所 (2nd Trip):** A detailed section for the second trip, including departure/arrival times, purpose, mode of transport, transfer points, and parking.
- 3番目に行った場所 (3rd Trip):** A detailed section for the third trip, including departure/arrival times, purpose, mode of transport, transfer points, and parking.
- 4番目に行った場所 (4th Trip):** A detailed section for the fourth trip, including departure/arrival times, purpose, mode of transport, transfer points, and parking.
- Legend (左):** Explains various symbols and codes used in the form, such as "表1" for school types, "表2" for purposes, "表3" for transport modes, and "表4" for parking types.

図 1 PT 調査 個人票

1.3 PT 調査の基本概念

(1) トリップ

人または車両がある目的を持って起点から終点へ移動する場合に、その移動を表す概念。移動を定量的に表す際の基本的単位。

出発地から目的地までの一連の移動を「リンクトトリップ」、リンクトトリップを移動手段が変わるご

とに分割したものを「アンリンクトトリップ」と言う。
 代表交通機関は、鉄道>バス>自動車>二輪車>徒歩 の順。

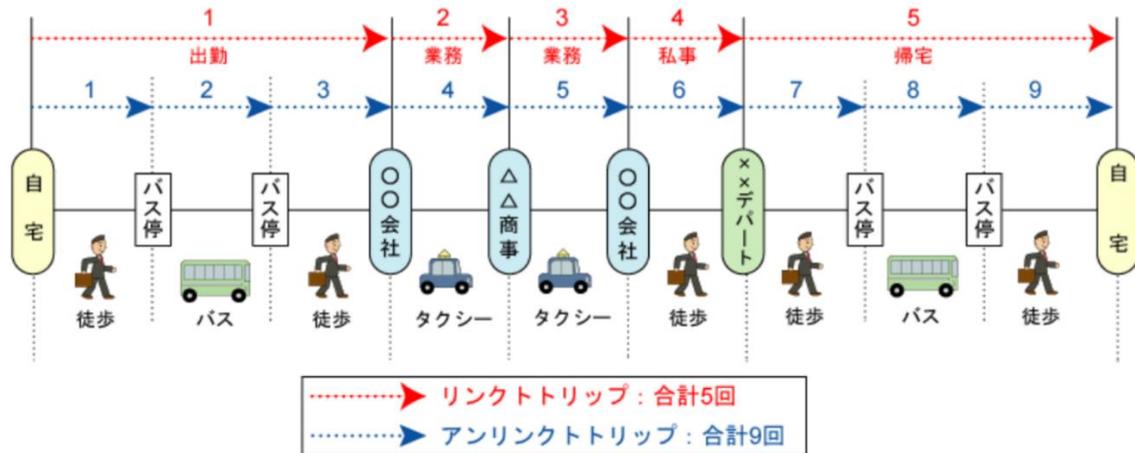


図 2 トリップ例

(2)ゾーン

交通計画立案の単位となる区域。地理的・歴史的な地域のまとまりを考慮して定められている。
 PTにおいて、位置情報(出発地・到着地等)は全て小ゾーン単位で表され、小ゾーン内での交通の様相は分からない。

各小ゾーンには 5 桁のゾーンコードが割り当てられている。毎回の調査でゾーンコードは基本的に変わらないが、合併等の影響で変わっている場合もあるので注意すること。

ゾーンコードの体系は下表の通り。

表 ゾーンコードの体系

大		中		計画基本	小	地区	
0	0	1	1	1	0	1	

各地域の小ゾーンコードは右のページから確認できる。<https://www.tokyo-pt.jp/data/map>



図 3 江東区周辺の小ゾーン

(3)小ゾーンとトリップ

ゾーン間トリップは、各ゾーンの中心点(セントロイド)間の移動と考える。またセントロイドと実際の

ノードとは仮想リンクで結ばれていると考える。

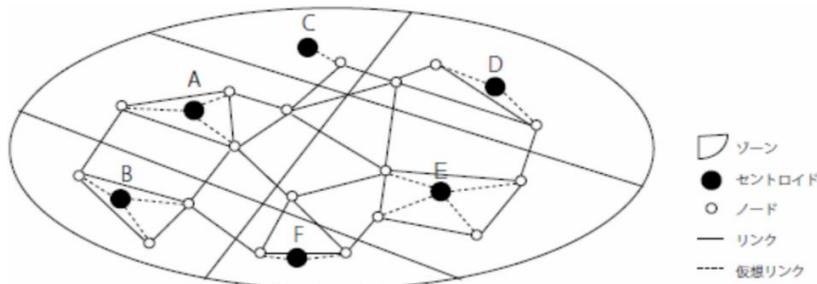


図 4 ゾーン間トリップのイメージ

(4)OD 表

ゾーン間のトリップを表にまとめたもの。

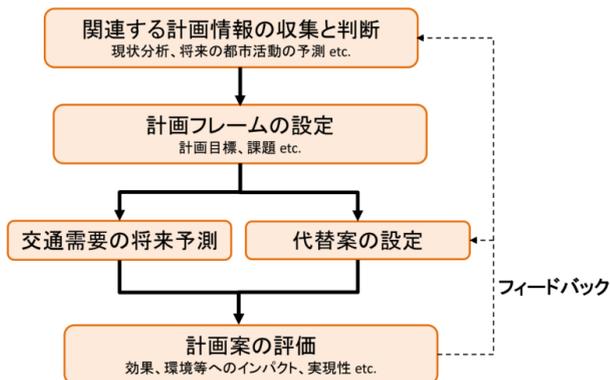
		終点(Destination)					
		1	...	j	...	N	発生量
起点(Origin)	1	?	?	?	?	?	G_1
	...	?	?	?	?	?	...
	i	?	?	T_{ij}	?	?	G_i
	...	?	?	?	?	?	...
	N	?	?	?	?	?	G_N
集中量		A_1	...	A_j	...	A_N	T

Annotations for the table:

- ゾーンiから発生するトリップ数 (Trips generated from zone i) - points to the G_i column.
- ゾーンiからゾーンjに向かうトリップ数 (Trips from zone i to zone j) - points to the T_{ij} cell.
- ゾーン全体で発生(集中)するトリップ数 (Total trips generated/concentrated in the zone) - points to the T cell.
- ゾーンjに集中するトリップ数 (Trips concentrated in zone j) - points to the A_j cell.

図 5 OD 表のイメージ

1.4 交通計画の手順



交通計画の立案は一般的には左の手順で行われる。PT調査・分析は「関連する計画情報の収集と判断」に当たる。

計画立案の際には交通需要の予測が非常に重要になるが、収集したPTデータやPP(プローブパーソン)データを基に予測が行われる。

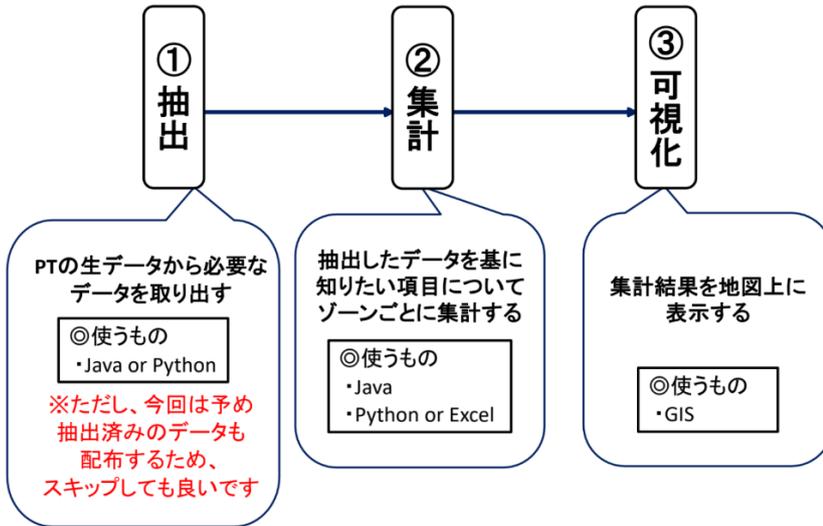
(1)四段階推定法

現状のPTデータと将来人口予測等から交通需要を推定する方法。長期の幹線交通網計画に用いる。ゾーン単位での分析・ゾーン平均での評価が行われる。

(2)非集計分析法

PPデータを用いて交通需要を予測する方法。細かいサービス変更(例：道路有料化)の影響の分析等、集計分析が適当でないスケールでの分析の際に利用される。個人単位での分析が可能。

1.5 PT分析のステップ



2. データの抽出

※今回は抽出済みのデータも配布するので、このステップは飛ばしても良いです。抽出済みデータに含まれない情報(性別、旅行時間等)を用いた分析が行いたい人は是非挑戦してください。以下では Python の Jupyter Notebook を使った抽出方法を説明しますが、他のツールが使える人はそちらを使っても構いません。

抽出済みデータ：

i) PT_shutoken2018.csv：東京都市圏 PT 調査(2018 年)の生データから下の項目を抽出したもの。

	A	B	C	D	E	F	G	H	I
1	出発地：\	到着地：\	拡大係数	移動の目的	代表交通手段	年齢	就業（形態）	dep_szone	arr_szone
2	347004	345008	139	3	1	84	9	3470	3450
3	345008	347004	139	19	1	84	9	3450	3470

ii) PT_tokyo2018.csv：i)の内、出発地または到着地が東京都内のトリップを抽出したもの。

iii) PT_tokyo2008.csv：ii)の 2008 年版。

※データ数：i)…554907 行 ii)…244749 行 iii)…869693 行

2.1 準備

Python の環境構築を行う。下のページを参考に Anaconda をインストールすること。

<https://techfun.cc/python/anaconda-install.html>

Anaconda とは、「Python 本体」 + 「各種データサイエンスライブラリ」などをまとめた Python パッケージである。

また、江東区>06_PT データ>PT 配布用.ipynb もダウンロードしておく。

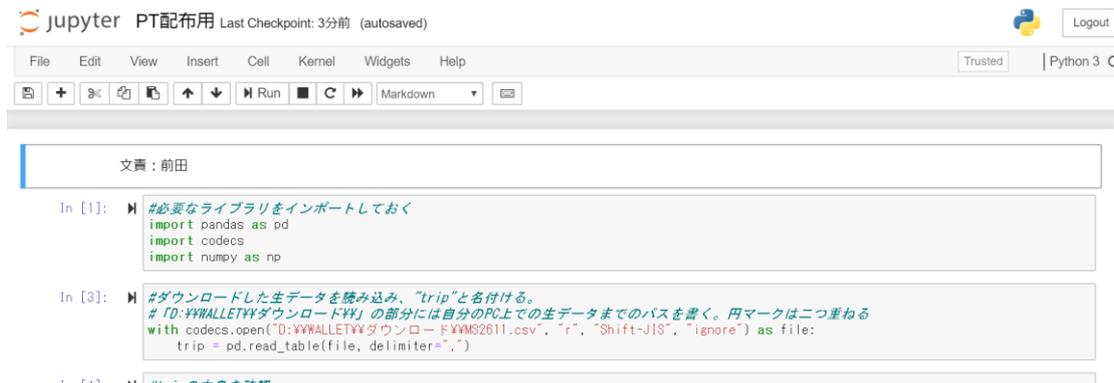
2.2 抽出の手順

①2018年東京都市圏PT調査の生データがGoogle driveの 江東区>06_PTデータ>東京都市圏PT2018>01 マスターデータ>MS2611.csvにあるので、ダウンロードしておく。

②PCのスタートメニュー>Anaconda3(64-bit)からJupyter Notebook(anaconda3)を開く。ブラウザが自動的に起動し、下のような画面が表示される。



③PT配布用.ipynbが入っているフォルダ（普通はDownloads）を開き、PT配布用.ipynbを開くと下の画面が開く。

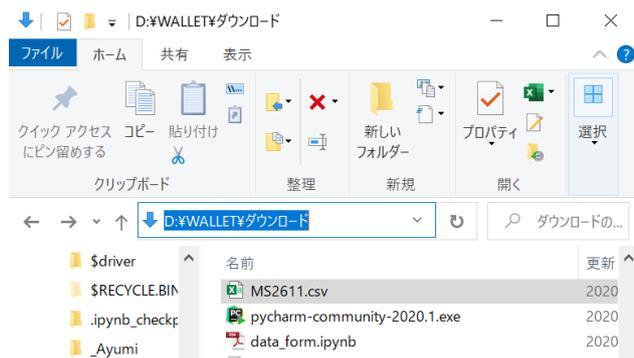


④書いてある指示に従って、上のセルから順番に実行していく。実行するには上の「Run」をクリックするか、「shift + Enter」を押す。

※③で生データファイルが見つからない場合

PC内にドライブが複数あり、生データが入っていない方のドライブが開かれている可能性があります。以下の手順で正しいドライブを開いてください。

①ダウンロードした生データがあるPC内の場所を確認する。エクスプローラーを開き、生データがある場所へ移動する。下の画像の場合、生データはDドライブ内にある。



②コマンドプロンプト(Command prompt)を開く。PC のタスクバーの検索機能を使うか、Windows ユーザーなら スタートメニュー>Windows システムツール からでも開ける。

起動すると黒い画面に白い文字が表示されている。一番下の行に書いてあるのがコマンドプロンプト上で現在いる場所である。例えば下画像の場合「C:¥User¥m>」と書いてあり、「Cドライブ内の User フォルダ内の m フォルダにいる」という意味である。



```
コマンド プロンプト
Microsoft Windows [Version 10.0.18363.836]
(c) 2019 Microsoft Corporation. All rights reserved.
C:¥User¥m>
```

③コマンドプロンプト上で、生データがある方のドライブへ移動する。D ドライブへ行きたい場合、下のように「D:」と打ち込んで Enter を押す。

④D に移動出来たら、jupyter notebook と打ち込んで Enter を押すとブラウザが開く。



```
コマンド プロンプト - jupyter notebook
Microsoft Windows [Version 10.0.18363.836]
(c) 2019 Microsoft Corporation. All rights reserved.
C:¥User¥m>D:
D:¥>jupyter notebook
```

3. 集計

※以下では Excel を使った集計方法を説明しますが、Python の Jupyter Notebook を使用することもでき、環境構築に多少時間はかかりますが慣れると Excel より速く集計できます。やり方は PT 配布用.ipynb の後半に載せているので、挑戦したい人は 2 章を参考に Jupyter Notebook でこのファイルを開いてみてください。

3.1 サンプル属性に基づく集計

例として PT_shutoken2018.csv のデータから、門前仲町を含む小ゾーン(コード：3470)を到着地とするトリップを、年齢階層別に集計してみる。年齢階層は下のように区分する。

1: 5~9, 2: 10~14, 3: 15~19, 4: 20~24, 5: 25~29, 6: 30~34, 7: 35~39, 8: 40~44, 9: 45~49, 10: 50~54, 11: 55~59, 12: 60~64, 13: 65~69, 14: 70~74, 15: 75~79, 16: 80~84, 17: 85~

①PT_shutoken2018.csv を Excel で開く。

②表に新しい列「年齢階層」を追加する。年齢階層は「年齢÷5 (小数点以下切り捨て)」なので、下のように ROUNDDOWN 関数を使ってセルの関数を定義する。但し、この定義の場合 18、19 といった年齢階層のデータも出てきてしまうことに注意。

オートフィル機能を使って、表の一番下まで各データの年齢階層を求める。(参考：<http://office-qa.com/Excel/ex206.htm>)



	A	B	C	D	E	F	G	H	I	J	K
1	出発地：\	到着地：\	拡大係数	移動の目的	代表交通	年齢	就業 (形態)	dep_szone	arr_szone		年齢階層
2	347004	345008	139	3	1	84	9	3470	3450		=2/5, 0)
3	345008	347004	139	19	1	84	9	3450	3470		
4	347004	347008	139	3	1	82	8	3470	3470		

③右側に下の M~P 列のような表を作る。一旦 P 列に、各年齢階層(1~20)におけるトリップ数を記入しておく。トリップ数のカウントには SUMIFS 関数を用いる。SUMIFS 関数の引数は順に、合計対象範囲、条件範囲 1、条件 1、条件範囲 2、条件 2、である。

※拡大係数…サンプルの属性から推定される、母集団(現実)においてそのトリップがなされる数。

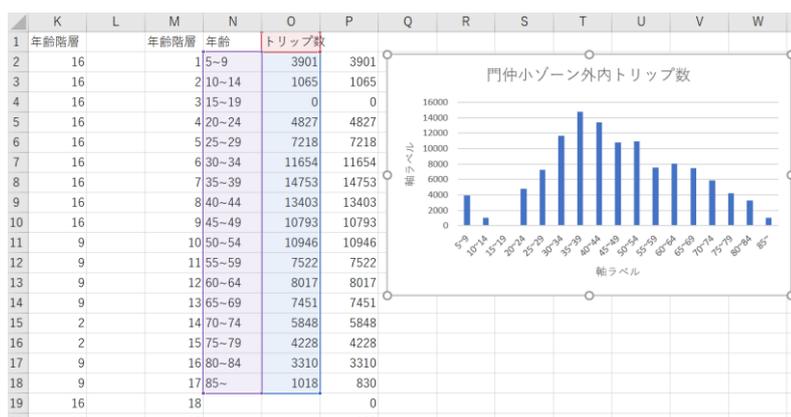
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	出発地：\	到着地：\	拡大係数	移動の目的	代表交通手段	年齢	就業（形態）	dep_szone	arr_szone		年齢階層		年齢階層	年齢	トリップ数	
2	347004	345008	139	3	1	84	9	3470	3450		16		1	5~9	3901	3901
3	345008	347004	139	19	1	84	9	3450	3470		16		2	10~14	1065	1065
4	347004	347008	139	3	1	82	8	3470	3470		16		3	15~19	0	0
5	347008	347004	139	19	1	82	8	3470	3470		16		4	20~24	4827	4827
6	345008	346005	172	7	7	80	9	3450	3460		16		5	25~29	7218	7218
7	346005	347012	172	7	7	80	9	3460	3470		16		6	30~34	11654	11654
8	347012	345008	172	3	1	80	9	3470	3450		16		7	35~39	14753	14753
9	345008	345008	172	12	1	80	9	3450	3450		16		8	40~44	13403	13403
10	345008	345008	172	19	1	80	9	3450	3450		16		9	45~49	10793	10793
11	345012	2121101	194	1	5	46	2	3450	21211		9		10	50~54	10946	10946
12	2121101	345012	194	19	5	46	2	21211	3450		9		11	55~59	7522	7522
13	345012	311011	168	1	5	45	4	3450	3110		9		12	60~64	8017	8017
14	311011	345012	168	19	5	45	4	3110	3450		9		13	65~69	7451	7451
15	345012	346003	213	2	1	11	7	3450	3460		2		14	70~74	5848	5848
16	346003	345012	213	19	1	11	7	3460	3450		2		15	75~79	4228	4228
17	348005	347005	167	1	1	48	2	3480	3470		9		16	80~84	3310	3310
18	347005	348005	167	19	1	48	2	3470	3480		9		17	85~	1018	830
19	341008	340005	139	12	7	82	9	3410	3400		16		18			0
20	340005	341008	139	19	7	82	9	3400	3410		16		19			188
21	341008	341008	134	6	1	74	9	3410	3410		14		20			0
22	341008	341008	134	19	1	74	9	3410	3410		14			合計	115954	

④ ②と同じくオートフィル機能を使って、全ての年齢階層におけるトリップ数を数える。

年齢階層 18~20 は存在しないことを考慮して、O 列に正しい年齢階層別トリップ数を記入する。

合計を求めたい場合は SUM 関数を使う。

⑤上のツールバーの挿入>グラフから好きな形式のグラフを選んで生成する。このとき、合計は含めないように注意。



※PT_tokyo2018.csv は CSV 形式のファイルなので、グラフを保存しておくことができず、ファイルを閉じるとグラフは消えてしまう。グラフを保存しておきたい場合は、新しく XLSX 形式のファイルを作成し、上画像の M~O 列の表をコピー→新ファイルに数値で貼り付けして、新ファイル上でグラフを生成する。

※SUMIFS 関数は集計するデータの条件をいくつでもつけられるので、例えば下画像のように「門仲小ゾーンを到着地とするトリップを、年齢階層別かつ目的別に集計する」といったことも可能である。

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1																							
2	5-9	0	1659	0	0	0	0	0	0	0	0	426	370	0	0	0	0	0	0	1872	0	0	4327
3	10-14	0	213	0	0	0	0	0	0	0	0	426	0	0	0	0	0	0	0	852	0	0	1491
4	15-19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	20-24	1816	0	0	0	257	0	0	0	0	0	349	0	0	0	0	0	349	0	2056	0	0	4827
6	25-29	4416	0	0	0	468	0	0	0	0	0	0	0	0	0	0	602	0	0	2510	0	0	7996
7	30-34	4252	0	857	0	284	0	0	0	0	789	0	396	0	0	413	439	0	239	4473	0	0	12142
8	35-39	6894	0	671	91	512	0	0	0	0	792	0	329	0	0	640	369	0	188	4729	0	0	15215
9	40-44	6270	0	570	0	363	0	97	0	0	885	0	822	0	0	142	340	0	147	4029	0	0	13665
10	45-49	6556	0	725	131	447	0	168	0	0	0	160	153	0	0	390	193	167	357	1809	0	0	11256
11	50-54	6316	0	200	0	911	71	77	0	0	77	134	0	0	0	167	758	534	168	1906	0	0	11319
12	55-59	4703	0	734	0	413	0	0	0	0	130	0	127	80	0	39	315	0	277	1120	0	0	7938
13	60-64	3861	0	548	0	240	0	43	0	0	183	0	903	0	0	249	120	0	73	2095	0	0	8315
14	65-69	2337	0	1539	95	161	0	119	0	0	0	462	91	0	703	63	0	343	1963	0	0	7876	
15	70-74	1283	0	1006	0	236	0	740	0	0	0	820	0	0	565	0	0	135	1195	0	0	5980	
16	75-79	80	0	1211	0	585	0	176	0	0	0	606	0	164	544	84	0	0	778	0	0	4228	
17	80-84	0	0	348	0	361	0	336	267	0	0	487	0	0	109	0	0	0	1293	109	0	3310	
18	85-	0	0	328	0	0	0	188	188	0	0	157	0	0	0	0	0	0	157	0	0	1018	
19	合計	48784	1872	8737	317	5238	71	1944	455	0	2856	1146	5981	171	164	3961	3283	1050	1927	32837	109	0	120903

3.2 トリップ出発地/到着地に基づく集計

例として PT_shutoken2018.csv のデータから、門前仲町を含む小ゾーン(コード:3470)を到着地とするトリップを、出発地小ゾーン別に集計してみる。小ゾーン毎の数値を集計する際は、後で QGIS 等で集計データを可視化できるので、必ず新しい CSV ファイルを作成してそこに表を張り付けること。

- ①Google drive の 東京都圏 PT2018>szoncode.csv をダウンロードし、開く。
- ②「KEY_CODE」列を列ごとコピーし、PT_shutoken2018.csv の表の右側に張り付ける。
- ③SUMIFS 関数を使って、到着地小ゾーンが 3470、出発地小ゾーンが各小ゾーンのトリップ数を集計する。因みに KEY_CODE の行数(小ゾーン数)は 1706 である。

COUNTIFS =SUMIFS(\$C\$2:\$C\$554908,\$I\$2:\$I\$554908,3470,\$H\$2:\$H\$554908,M2)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	出発地:)	到着地:)	拡大係数	移動の目的	代表交通	年齢	就業(形態)	dep_szone	arr_szone		年齢階層		KEY_CODE	トリップ数
2	347004	345008	139	3	1	84	9	3470	3450		16		100	08, M2)
3	345008	347004	139	19	1	84	9	3450	3470		16		101	0
4	347004	347008	139	3	1	82	8	3470	3470		16		102	137
5	347008	347004	139	19	1	82	8	3470	3470		16		103	558

- ④集計できたら、作成した表(上の画像でいう M-N 列)をコピーし、新しい白紙の CSV ファイルに数値で貼り付ける。後で QGIS での可視化に使用するのできちんと保存しておく。(arr_monaka.csv と名付けた。)

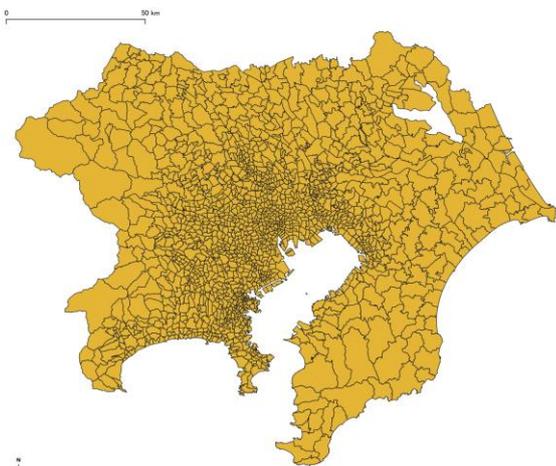
	A	B	C
1	KEY_COD	トリップ数	
2	100	0	
3	101	0	
4	102	137	

4. QGIS での可視化

例として、3.2 で作成した CSV ファイルを QGIS 上で可視化する手順を説明する。他のツール(ArcGIS 等)を使っている人はそちらで可視化しても構わない。

- ①Google drive の 東京都圏 PT2018>03 ゾーン図>H30_gis.zip をダウンロードし、展開する。
- ②QGIS を立ち上げ、任意のプロジェクトを開く。
- ③画面左の「ブラウザ」から、たった今展開した小ゾーンのシェイプファイル(H30_szone.shp)をダブル

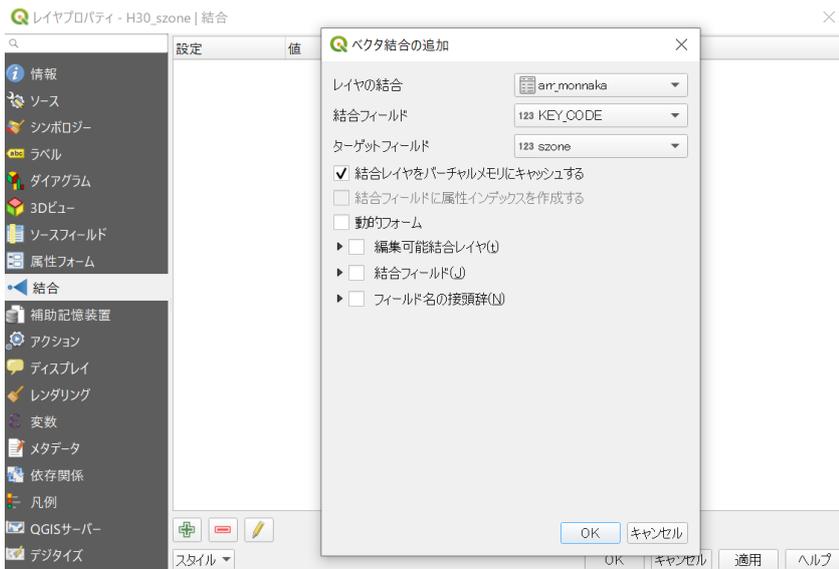
クリックで開く。下のような画面が表示される。



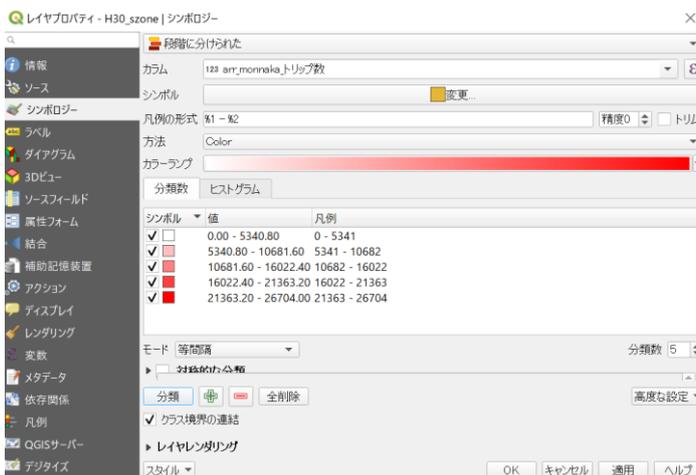
③3.2 で作成した CSV ファイルを読み込む。上のツールバーのレイヤ>レイヤの追加>デリミテッドテキストレイヤの追加 を選択すると、「データソースマネージャー」ウィンドウが開く。読み込む CSV ファイルを右上の四角(…)から選択し、下画像のように指定する。「追加」を押すとこのファイルが新しいレイヤとして追加されるが、「ジオメトリなし」なのでそのままでは地図上には表示できない。



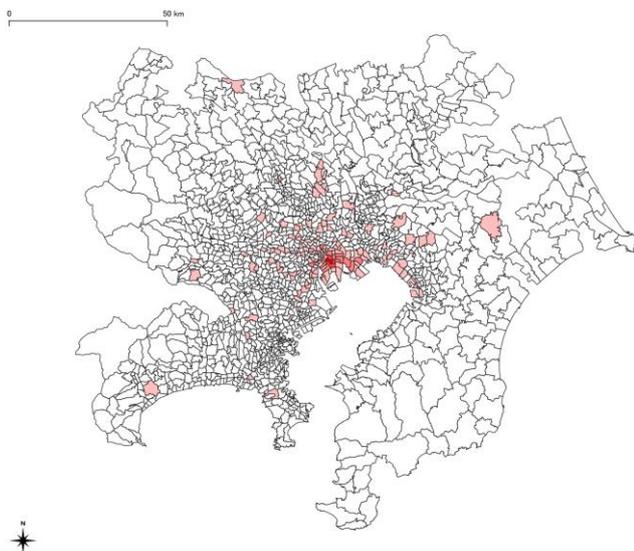
④追加した CSV を小ゾーンレイヤ(H30_szone)と結合する。画面左「レイヤ」内の H30_szone を右クリック>プロパティを選択すると、「レイヤプロパティ」ウィンドウが開く。その中の「結合」ページで、左下の✚を押して、下画像のようにレイヤの結合・結合フィールド・ターゲットフィールドを指定する。これを適用すると、レイヤ H30_szone において各小ゾーンに「その小ゾーンを出発地、門仲小ゾーンを到着地とするトリップ数」の数値が付与される。



⑤小ゾーンを、付与した数値の大きさと色分けする。引き続き H30_szzone のレイヤプロパティで、「シンポロジー」ページを開く。「単一シンボル」となっているところを「段階に分けられた」、カラムを「arr_monnaka(CSV ファイル名)_トリップ数」として「分類」を押す。モードは適宜変えてみる。



⑥分類を適用し、下のようになれば完成。



参考

- ・ 東京都市圏交通計画協議会「東京都市圏パーソントリップ調査 データ利用の手引き」
<https://www.tokyo-pt.jp/static/hp/file/data/tebiki.pdf>
- ・ 東京都市圏交通計画協議会「小ゾーン確認マップ」<https://www.tokyo-pt.jp/data/map/tokyo#>