

# 神エクセルの「統計データ」を 一括オープンデータに自動変換するツール ～上田市統計データを丸ごとCSV化～

2019年1月25日

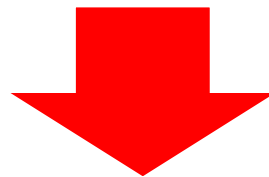
**長野大学・前川道博研究室 & ゼミ**



本資料は、クリエイティブ・コモンズ 表示4.0 国際 (CC BY 4.0) に従って利用できます。  
(<http://creativecommons.org/licenses/by/4.0/legalcode.ja>)

# 統計データの神エクセル問題

- お役所の紙文書文化の堅牢な慣習
- 都合よくスキマに入り込んだExcel統計表業務
- オープンデータに対応できない現場の戸惑い  
(手作業の場当たりの対応にならざるを得ない)



- 現行の「神エクセル」データに照準を合わせる
- 業務フローの根本改善を目指し「今」を変える
- Excel→CSV変換ツール開発で停滞を打破する
- 「統計」でできることは多業務でもできる！

# 地域課題

自治体の統計データはExcelで作られているため、それらはオープンデータとして再利用しにくい。Excelはとりわけ機械判読可能なデータ作成の障壁となる。自治体職員も統計業務の慣習で印刷物作成のためにExcelを使っているため、何が問題なのかを全く気付くことすらできない。全国に蔓延しているこういった状況から統計業務の現場を脱却させるための啓発的な課題解決が望まれている。

# 統計データのオープンデータ化課題

- ・「上田統計データ」の問題は遍く全国に共通の地域課題解決を図るモデルケースとなる
- ・既存統計データをCSVにするためには 極力手作業をなくし、明確な変換規則、処理のなり直しが容易なプログラムによる一括処理が欠かせない
- ・統計の担当者が人任せにしないチェック、改善が図れる支援が必要である
- ・大学等、課題解決に当たれる行政外部の力を活かす対策は有効である

# 行政の統計データの一括処理手順

プログラムを介在させ、極力手作業を排して省力化、質の確保を図る。

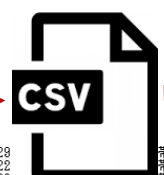
公開の統計サイト (CC-BY)



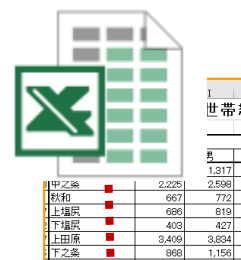
Excel  
リスト作成



ファイル名  
修正  
(手作業)



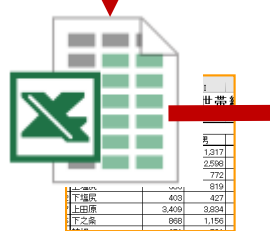
Excel  
一括転送



```
2011080715
2011080715
2011080715
20110807152003312.xls
20110807152003840.xls
20111201114818575.xls
20121203172048917.xls
90121909111004822.xls

20110807152000829
20110807152002822
20110807152002869
20110807152003312.xls
20110807152003840.xls
20111201114818575.xls
20121203172048917.xls
90121909111004822.xls

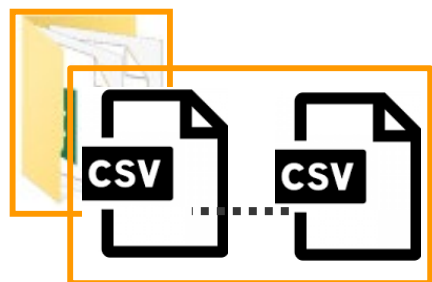
世帯数2006年 (H18)
世帯数2007年 (H19)
世帯数2008年 (H20)
世帯数2009年 (H21)
世帯数2010年 (H22)
世帯数2011年 (H23)
世帯数2012年 (H24)
全世帯人口
```



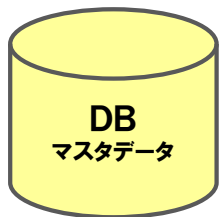
統計Excelファイル

Excel  
→CSV  
一括変換

データ  
修正  
(手作業)



CSVファイル (ワークシート別)



CSV公開

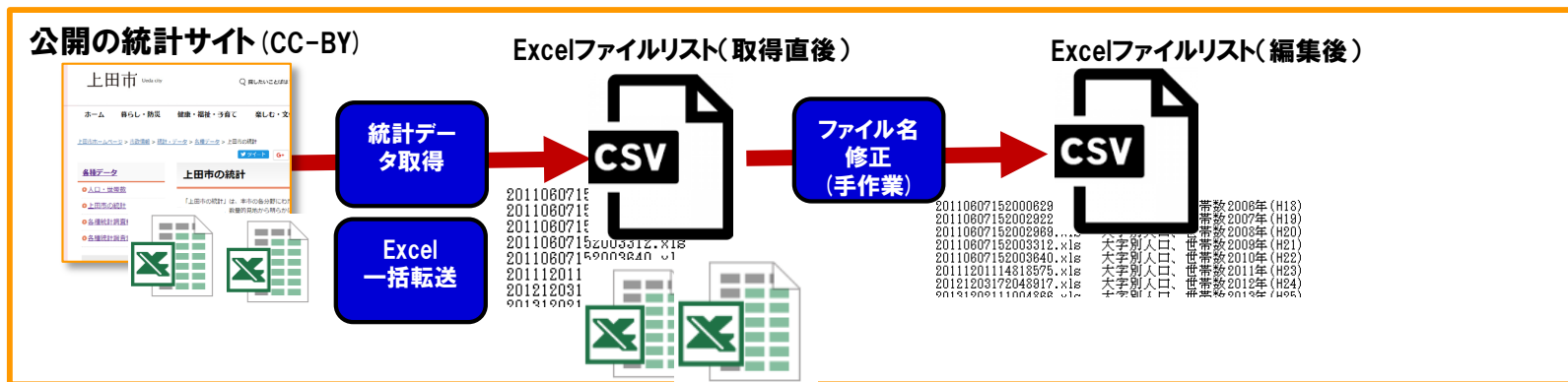
DB公開

オープンデータ  
公開サイト (CC-BY)



# ツール1: 統計データ取得

オープンデータ化 (CSV化)する統計Excelファイル情報とファイルを取得する



<https://mmdb.net/sdc/tokei/>

**統計データ取得**

統計サイトのURLを入力してください

統計Excelリストを作成する

Excelファイルを一括転送する

Excelの全ファイルのリストをCSV形式でGETする

```
https://www.city.ueda.na/2011120114818575.xls 大字別人口、世帯数2011年 (H23)
https://www.city.ueda.na/20121203172048917.xls 大字別人口、世帯数2012年 (H24)
https://www.city.ueda.na/20131202111004866.xls 大字別人口、世帯数2013年 (H25)
https://www.city.ueda.na/20141202095910842.xls 大字別人口、世帯数2014年 (H26)
https://www.city.ueda.na/jinko_jukijinko_2015.xls 大字別人口、世帯数2015年 (H27)
https://www.city.ueda.na/jinko_jukijinko_2016.xls 大字別人口、世帯数2016年 (H28)
https://www.city.ueda.na/jinko_jukijinko_2017.xls 大字別人口、世帯数2017年 (H29)
https://www.city.ueda.na/jinko_jukijinko_2018.xls 大字別人口、世帯数2018年 (H30)
https://www.city.ueda.na/jinko_jukijinko_2019.xls 大字別人口、世帯数2019年 (H31)
https://www.city.ueda.na/20110607152732826.xls 自治体別人口、世帯数2006年 (H18)
https://www.city.ueda.na/20110607152733590.xls 自治体別人口、世帯数2007年 (H19)
https://www.city.ueda.na/20110607152734261.xls 自治体別人口、世帯数2008年 (H20)
```



Excelファイルリスト(取得直後)

Excel等でファイルの日本語タイトルを正しく記述し保存する

Excelファイル一式 (ダウンロード)



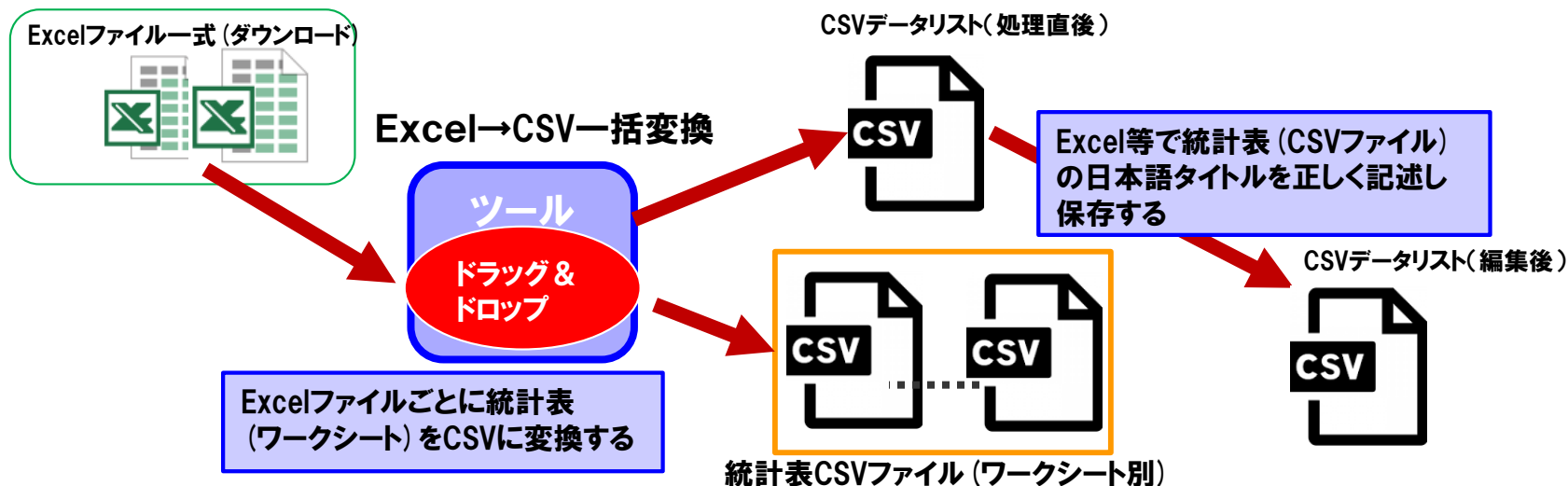
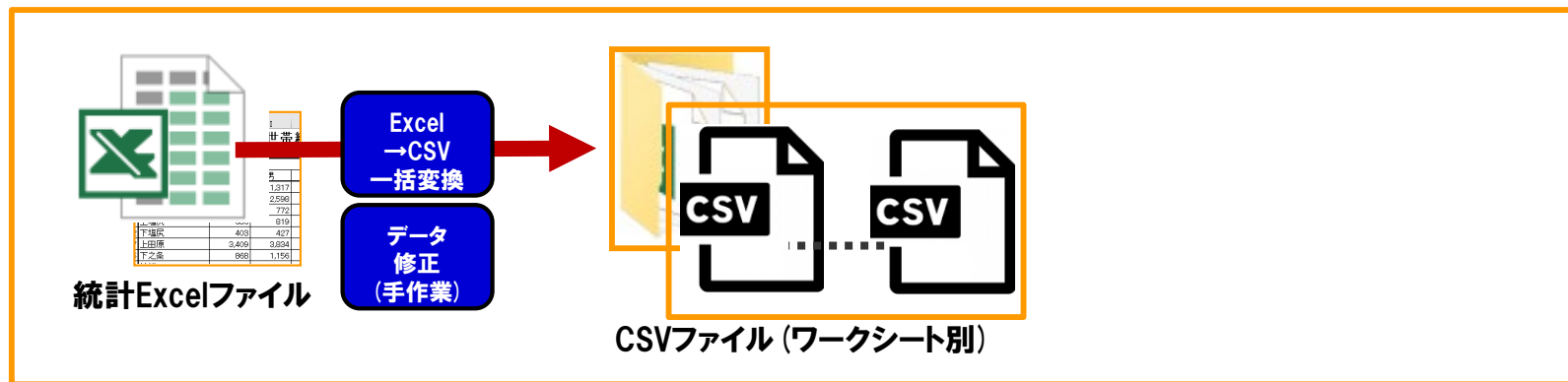
```
https://www.city.ueda.na/2011120114818575.xls 大字別人口、世帯数2011年 (H23)
https://www.city.ueda.na/20121203172048917.xls 大字別人口、世帯数2012年 (H24)
https://www.city.ueda.na/20131202111004866.xls 大字別人口、世帯数2013年 (H25)
https://www.city.ueda.na/20141202095910842.xls 大字別人口、世帯数2014年 (H26)
https://www.city.ueda.na/jinko_jukijinko_2015.xls 大字別人口、世帯数2015年 (H27)
https://www.city.ueda.na/jinko_jukijinko_2016.xls 大字別人口、世帯数2016年 (H28)
https://www.city.ueda.na/jinko_jukijinko_2017.xls 大字別人口、世帯数2017年 (H29)
https://www.city.ueda.na/jinko_jukijinko_2018.xls 大字別人口、世帯数2018年 (H30)
https://www.city.ueda.na/jinko_jukijinko_2019.xls 大字別人口、世帯数2019年 (H31)
https://www.city.ueda.na/20110607152732826.xls 自治体別人口、世帯数2006年 (H18)
https://www.city.ueda.na/20110607152733590.xls 自治体別人口、世帯数2007年 (H19)
https://www.city.ueda.na/20110607152734261.xls 自治体別人口、世帯数2008年 (H20)
```



Excelファイルリスト(編集後)

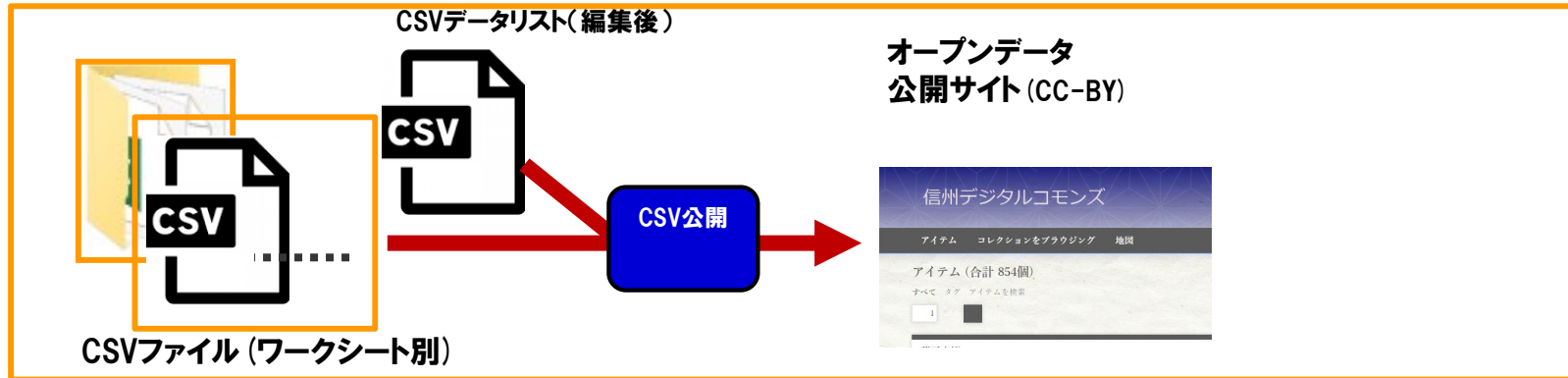
# ツール2: Excel→CSV変換

Excelファイル全ての統計データ(ワークシート)をCSVに一括変換する



# ツール3:オープンデータ公開

統計データ(CSVファイル)を取捨選択し統計データサイトに公開する



<https://mmdb.net/sdc/tokei/>

**統計データ公開**

CSVデータリストを選択してください

  
 CSVデータを一括転送する  
 Excelファイルも一括転送する  
CCライセンス 

<https://mmdb.net/sdc/tokei/>

**統計データ公開サイト**

アイテム コレクションをブラウジング 地図

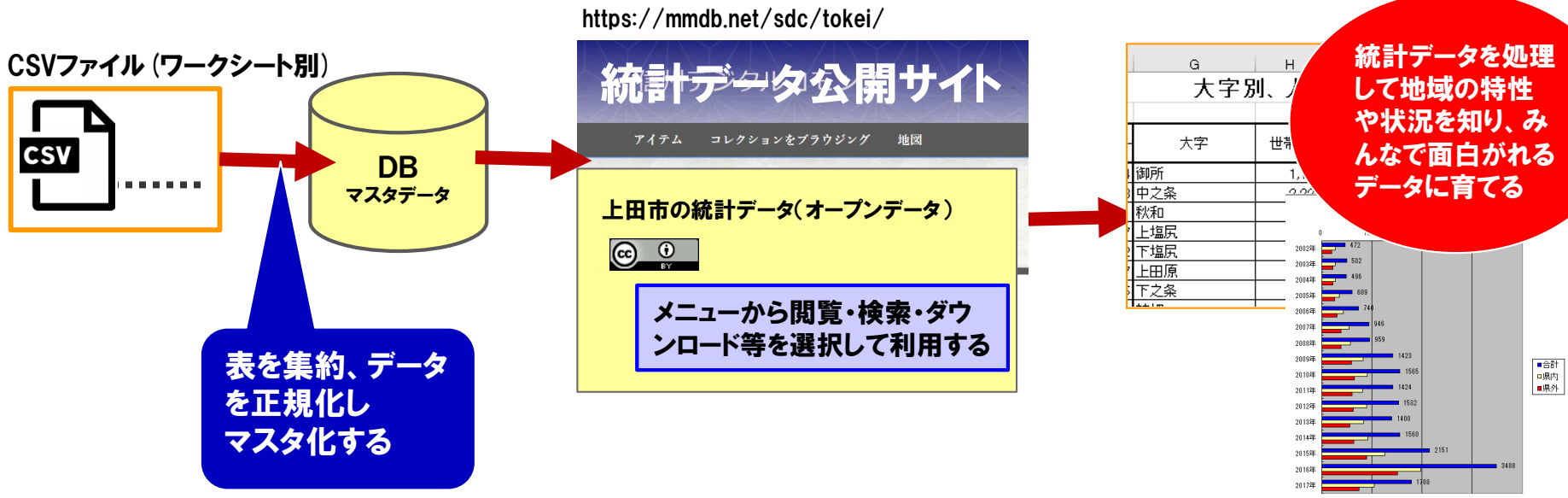
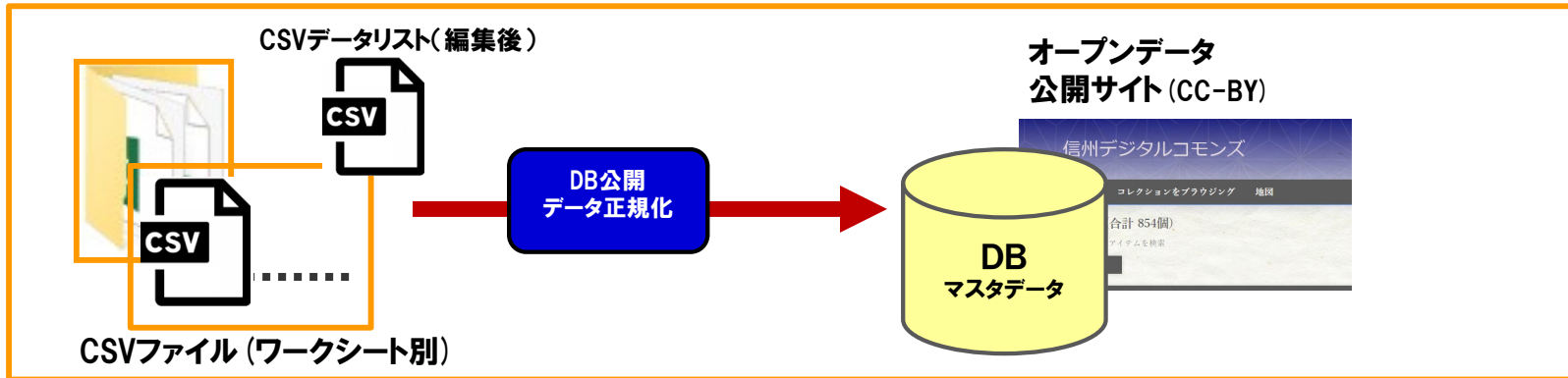
上田市の統計データ(オープンデータ)

メニューから閲覧・検索・ダウンロード等を選択して利用する



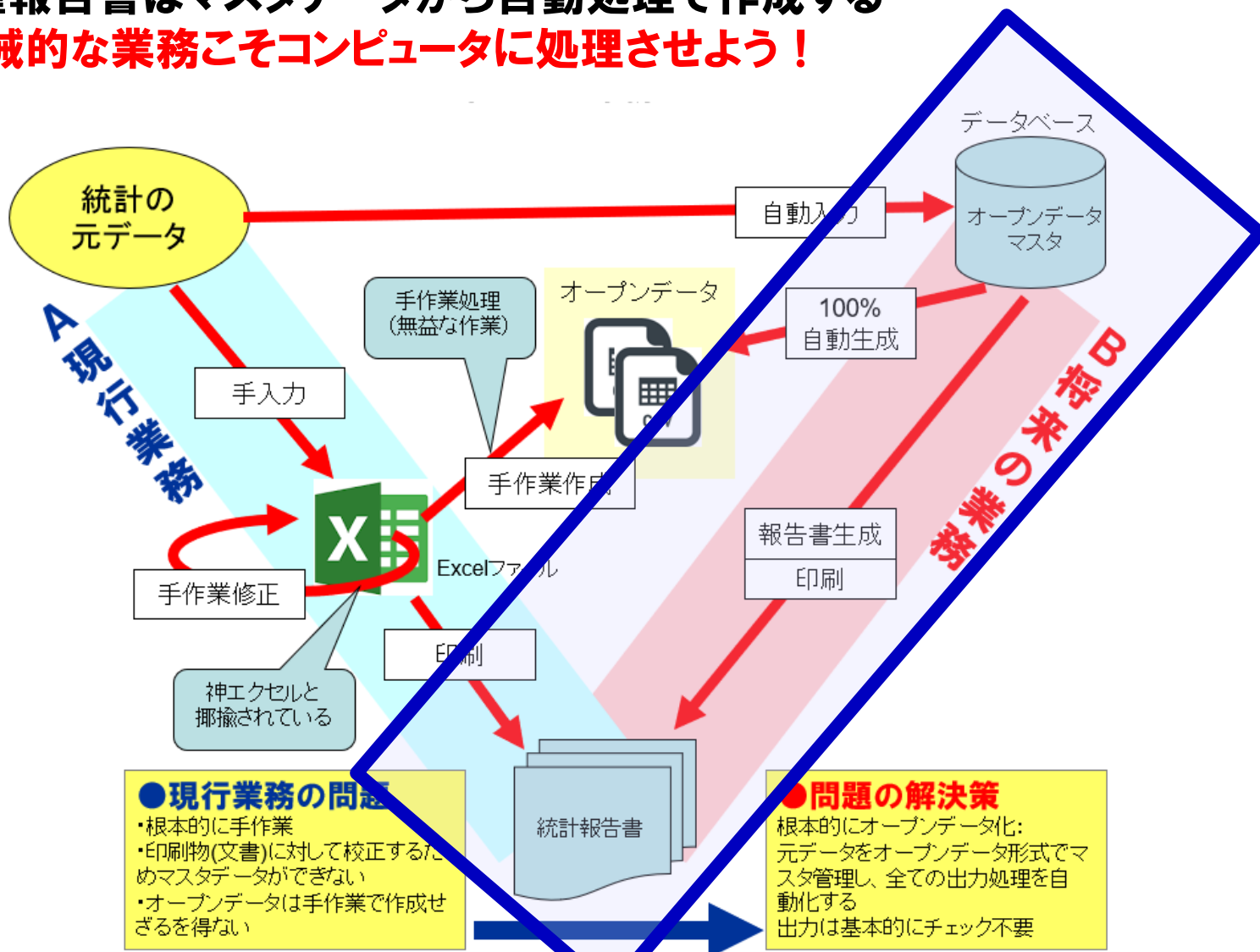
# ツール4: 今後の拡充

統計データをみんなで面白がれるデータに育て活用する



# ツール5:さらに進めて業務イノベーション

従来の各種報告書はマスターデータから自動処理で作成する  
こういう機械的な業務こそコンピュータに処理させよう！



# Excel→CSV一括変換

## 統計Excelファイル



✓ jinko\_jukijinko\_2015.xls

大字	世帯数	人口	世帯数	人口
総計	1,348	2,237	1,348	2,237
大字別	2,225	4,096	2,225	4,096
世帯別	867	722	911	1,723
世帯別	498	1,118	478	1,000
大字別	403	427	499	928
世帯別	3,810	6,554	4,051	7,923
大字別	866	1,426	1,417	2,573
世帯別	491	769	574	1,101
大字別	596	796	763	1,462
世帯別	793	981	881	1,242
大字別	289	386	371	723
世帯別	1,736	3,121	2,059	4,023
大字別	1,165	1,882	1,401	2,706
世帯別	299	444	379	683
大字別	911	982	947	1,222

複数ワークシート

Excel→CSV  
一括変換



✓ jinko\_jukijinko\_2015

## CSVファイル (ワークシート別)



✓ 大字別、人口、世帯統計\_2015.2.csv



✓ 大字別、人口、世帯統計\_2015.3.csv



✓ 大字別、人口、世帯統計\_2015.4.csv



✓ 大字別、人口、世帯統計\_2015.5.csv



✓ 大字別、人口、世帯統計\_2015.6.csv



✓ 大字別、人口、世帯統計\_2015.7.csv



✓ 大字別、人口、世帯統計\_2015.8.csv



✓ 大字別、人口、世帯統計\_2015.9.csv



✓ 大字別、人口、世帯統計\_2015.10.csv



✓ 大字別、人口、世帯統計\_2015.11.csv



✓ 大字別、人口、世帯統計\_2015.12.csv

# イタレート型 上田市大字別人口・世帯統計

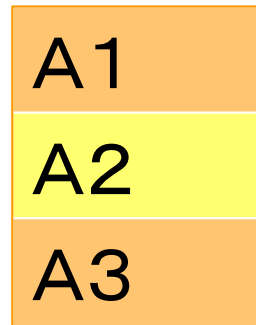
上田市自治会別人口・世帯統計

ファイル:年次 2015～ 表:月次

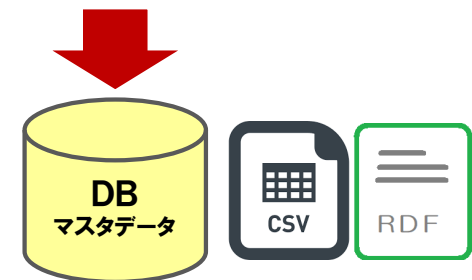
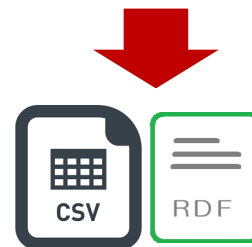
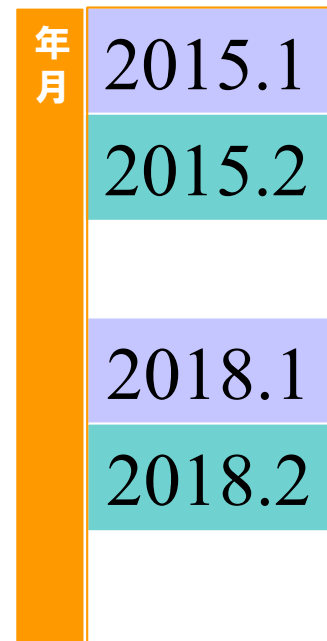
表1が上田地域、表2が丸子・真田・武石地域  
それぞれ1枚に収めるため、横3列に整列させている

平成18年12月1日現在															
大字別			人口			世帯			人口			世帯			
大字	世帯数	人口	大字	世帯数	人口	大字	世帯数	人口	大字	世帯数	人口	大字	世帯数	人口	
丸ノ下1丁目	167	155	302	328	308	東吉田	291	2,130	2,209	4,329	野倉	57	42	45	
丸ノ下2丁目	34	353	325	69	69	御所	1,325	1,356	2,682	別所湯原	25	713	526	1,631	
真田1丁目	22	132	147	27	27	中ノ郷	2,289	2,471	4,770	仁成田	522	509	1,031		
真田2丁目	46	411	454	61	61	武石	810	803	1,605	野	32	419	435	877	
武石1丁目	29	279	312	320	63	上ノ郷	188	800	810	1,610	野原	169	493	339	1,031
武石2丁目	281	330	335	68	68	下郷尻	397	464	942	1,000	郷	169	282	293	541
丸ノ下3丁目	171	160	158	33	33	上田原	3,095	3,740	3,647	7,893	下郷原	453	567	616	1,201
丸ノ下4丁目	817	737	673	1,411	1,411	下之郷	709	961	1,090	2,041	上郷原	290	315	352	667
丸ノ下5丁目	171	160	158	33	33	神原	509	750	764	1,514	野原	16	16	31	
丸ノ下6丁目	74	63	85	18	18	郷	917	691	718	1,409	【上田地域】	47,229	59,667	62,216	121,866
丸ノ下7丁目	126	150	145	28	28	郷	699	885	948	1,831	郷	224	192	246	433
丸ノ下8丁目	134	169	160	32	32	郷	240	323	323	646	郷	284	283	297	589
丸ノ下9丁目	276	317	371	69	69	郷	1,559	1,840	1,978	3,924	郷	232	259	341	631
丸ノ下10丁目	139	149	159	32	32	郷	1,102	1,349	1,395	2,712	郷	877	839	891	1,721
丸ノ下11丁目	132	137	173	31	31	郷	342	431	437	605	郷	630	675	617	1,791
丸ノ下12丁目	192	198	244	44	44	郷	844	905	903	1,805	郷	1,259	1,618	1,742	3,353
丸ノ下13丁目	499	505	547	1,035	1,035	郷	830	1,189	1,226	2,293	郷	863	1,201	1,293	2,481
丸ノ下14丁目	445	405	531	93	93	郷	1,298	1,791	1,840	3,935	郷	295	440	432	881
丸ノ下15丁目	434	479	515	99	99	郷	2,180	2,981	3,039	6,090	郷	475	670	691	1,351
丸ノ下16丁目	373	403	458	88	88	郷	1,762	2,332	2,341	4,673	郷	1,124	1,703	1,702	3,401
丸ノ下17丁目	439	432	520	97	97	郷	1,387	1,839	1,777	3,613	郷	1,196	1,637	1,729	3,381
丸ノ下18丁目	293	299	291	57	57	郷	184	247	257	304	郷	1,331	1,764	1,804	3,641
丸ノ下19丁目	214	235	292	49	49	郷	657	627	699	1,720	郷	129	191	201	391
丸ノ下20丁目	218	251	294	58	58	郷	47	67	95	123	郷	6,647	11,720	12,326	24,041
丸ノ下21丁目	293	304	393	67	67	郷	638	695	695	1,698	郷	430	632	633	1,301
丸ノ下22丁目	392	442	452	92	92	郷	972	1,039	1,015	2,054	郷	1,090	1,335	1,336	3,091
丸ノ下23丁目	293	349	349	69	69	郷	667	822	838	1,690	郷	677	1,037	1,381	2,731
丸ノ下24丁目	135	144	159	30	30	郷	384	729	735	1,454	郷	1,448	2,168	2,198	4,399
丸ノ下25丁目	117	143	185	31	31	郷	1,020	1,372	1,414	2,785	郷	3,633	5,710	5,819	11,324
丸ノ下26丁目	164	211	221	43	43	郷	531	1,139	1,212	2,351	郷	132	191	206	381
丸ノ下27丁目	210	243	263	50	50	郷	283	379	393	772	郷	174	253	297	581
丸ノ下28丁目	292	353	377	74	74	郷	354	787	756	1,513	郷	449	628	670	1,291
丸ノ下29丁目	339	449	452	90	90	郷	235	314	329	643	郷	217	324	330	671
丸ノ下30丁目	641	656	1,054	1,944	1,944	郷	392	644	557	1,071	郷	79	105	116	221
丸ノ下31丁目	315	353	403	78	78	郷	132	172	184	355	郷	204	264	288	591
丸ノ下32丁目	313	366	626	1,224	1,224	郷	135	205	211	419	郷	99	130	131	281
上田	2,383	3,040	3,140	6,169	6,169	郷	415	564	603	1,189	郷	83	97	89	181
郷	145	160	197	35	35	郷	284	418	397	813	郷	1,416	2,615	2,739	4,441
郷	232	319	330	61	61	郷	91	154	169	343	郷	61,129	79,112	82,492	161,601

▼表単位変換



▼通し変換



# イタレート (ブロック漸移) 型

## 年別統計:104 法規別労働組合数及び組合員数

年別統計:104 法規別労働組合数及び組合員数

ファイル:年次 2015~ 表:月次

年度ブロックが年毎に右側に漸移していくパターン

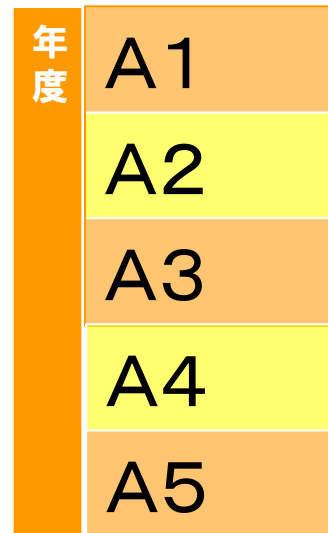
データは年度をブロック抽出し通し変換では差分を追加する

105 産業別職業紹介状況

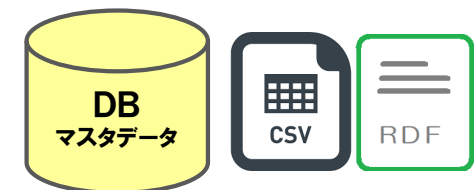
産業別	58年度		59年度		60年度		61年度		62年度	
	求人	充足	求人	充足	求人	充足	求人	充足	求人	充足
総数	A	3,571	A	3,773	A	3,610	A	3,347	A	4,650
農林漁業	22	13	13	10	10	11	11	17	17	17
鉱業	15	23	19	19	19	5	2	24	24	24
建設業	194	194	194	194	194	194	194	194	194	194
製造業	1,789	8,794	2	6,855	3	4	4	4	5	5
電気・ガス・熱供給・水道業	-	-	-	-	-	-	-	-	-	-
運輸・通信業	652	660	-	547	632	721	721	721	721	721
卸売・小売業、飲食店	3,422	3,434	-	2,835	2,887	2,990	2,990	2,990	2,990	2,990
金融・保険業	151	172	-	174	255	265	265	265	265	265
サービス業	1,681	1,804	-	1,976	1,990	2,036	2,036	2,036	2,036	2,036
公務	22	15	-	13	21	10	10	10	10	10

資料 上田公共職業安定所

### ▼表単位変換



### ▼通し変換



# 複数表かつイタレート型 上田市自治会別人口・世帯統計

上田市自治会別人口・世帯統計

ファイル:年次 2015～ 表:月次

表1が上田地域、表2が丸子・真田・武石地域  
それぞれ1枚に収めるため、横3列に整列させている

上田市自治会別人口・世帯統計【上田地域】

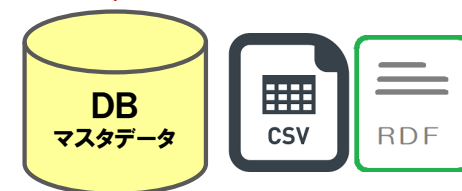
上田市自治会別人口・世帯統計【丸子・真田・武石地域】

▼表単位変換

A1  
A2  
A3  
B1  
B2  
B3

▼通し変換

年月  
2015.1  
2015.2  
2018.1  
2018.2



# 標準型 (行漸移型)

## 年別統計:XX06 56 簡易保険の取り扱い状況

年別統計:XX06 56 簡易保険の取り扱い状況

ファイル:年次 1972～ 表:年次

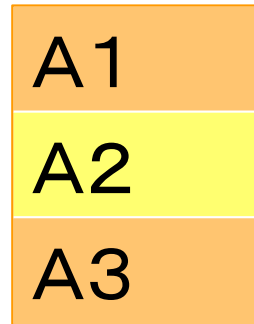
直近の5年度分を合計値で、続けて最新年度の郵便局別データを掲載

56 簡易保険の取り扱い状況

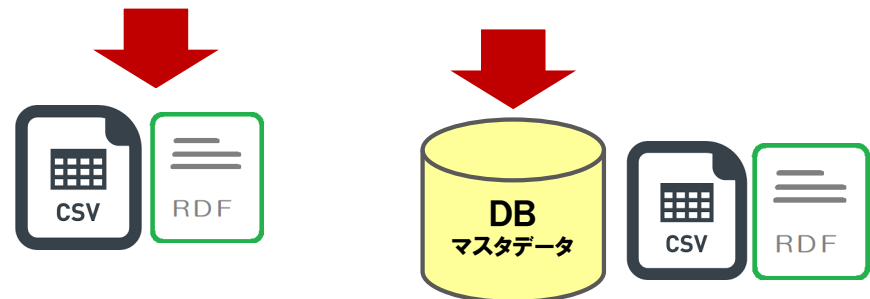
年度	新規契約			消滅			年末現在高		
	件数	保険料 万円	保険金 万円	件数	保険料 万円	保険金 万円	件数	保険料 万円	保険金 万円
(上田郵便局)									
平成元年度	2,155	2,776	399,086	2,188	2,059	346,715	26,135	24,790	4,406,076
平成2年度	3,134	44,797	624,315	3,543	44,478	640,398	26,863	265,762	4,594,542
平成3年度	4,159	6,280	827,122	4,047	4,280	691,678	31,789	34,239	5,657,663
平成4年度	4,669	7,194	1,054,809	4,130	4,614	758,646	34,097	39,288	6,376,411
平成5年度	4,293	6,865	957,495	4,029	3,102	767,379	35,038	41,944	6,706,151
5年度									
川辺郵便局	985	2,043	387,260	669	515	90,640	6,907	8,439	1,430,346
川島郵便局	593	1,018	168,559	475	605	92,327	4,855	5,985	1,040,653
塩田郵便局	487	698	100,175	4	4	420	3,513	3,859	586,241
東塩田郵便局	339	531	72,470	290	328	48,083	2,497	2,726	402,093
別所郵便局	331	466	6,059	258	280	34,152	3,098	3,539	500,917
浦里郵便局	500	615	86,315	285	586	76,996	3,327	3,120	439,613

資料 各郵便局

### ▼表単位変換



### ▼通し変換



# 近々の公開方法

## 統計オープンデータサイト

