

## 実習 5 属性の結合・属性値による出力先の振り分け

これまでの実習と同じ"mmm20151001.shp"（Esri Shapefile 形式の市区町村区域ポリゴンデータ）の市区町村ポリゴンに、平成 27 年国勢調査結果に基づく市区町村別人口等を結合した後、都道府県別の Esri Shapefile 形式のファイルに振り分けて出力するためのワークスペースを作成します。

### 5-1. ソースデータ（平成 27 年国勢調査市区町村別人口等）の準備

はじめに、ソースデータとして平成 27 年国勢調査市区町村別人口等のテーブルデータ（CSV 形式）を準備します。次の URL から zip 圧縮したファイルをダウンロード、解凍してください。

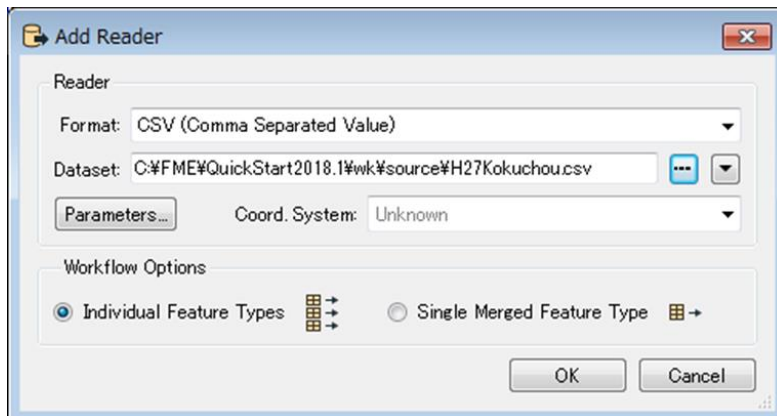
<http://www.pragmatica.jp/fme/support/quickstart201801/source/H27Kokuchou.zip>

このファイルは、e-Stat（政府統計の総合窓口）よりダウンロードした「平成 27 年国勢調査（総務省統計局）都道府県・市区町村別主要統計表」（Excel 形式）に基づいて作成したものです。e-Stat からダウンロードしたデータに手を加えずに利用することもできますが、そのためのワークスペースの内容はこのガイドで説明できる範囲を超えるため、簡略化したデータをあらかじめ用意しました。

### 5-2. CSV リーダーの追加とデータ内容の確認

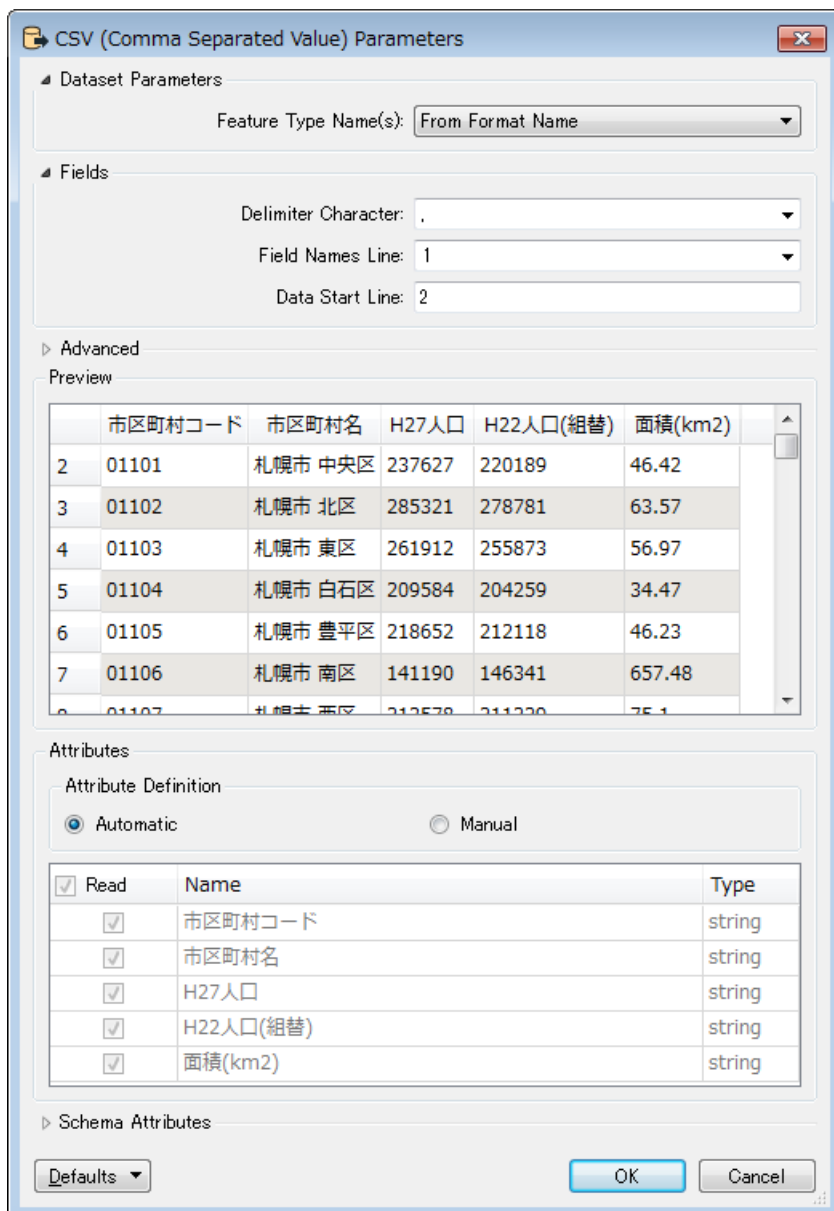
FME Workbench を起動し、ワークスペースに国勢調査データを読み込むための CSV リーダーとフィーチャータ입を追加します。

Workbench メニュー Readers > Add Reader で Add Reader 画面を開き、次のように Format と Dataset を設定してください（Dataset は、ダウンロード・解凍した csv ファイルの実際の保存先パスとします）。

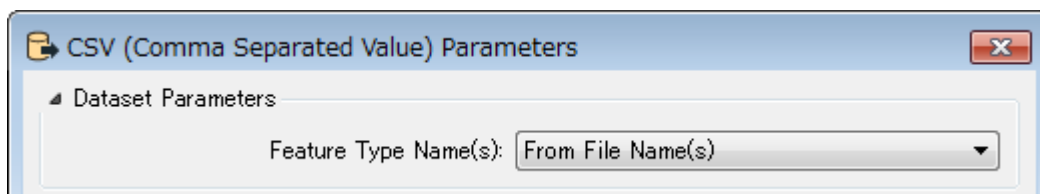


次に、Parameters ボタンをクリックしてください。次ページの図のようなパラメーター設定画面が開きます。

CSV 形式のテーブルには、区切り文字の種類、ヘッダ行（列名が記述されている行）の有無などによるバリエーションがあるため、リーダーをワークスペースに追加する際に、このパラメーター設定画面でプレビューを見ながら、実際のフォーマットに適合するような設定をします。



今回使用するデータの場合、これらのパラメーターは全てデフォルトのままでも良いですが、キャンバス上のフィーチャータイプ名を分かり易くするために、**Feature Type Name(s)** パラメーターを **"From File Name(s)"**に変更しておきます。



このパラメーターが **"From Format Name"** (デフォルト) の場合、フィーチャータイプ名は常に **"CSV"**、**"From File Name(s)"**とした場合はファイル名となります。

この設定でパラメーター設定画面、**Add Reader** 画面を **OK** で閉じると、ワークスペースに **CSV** リーダーとリーダーフィーチャータイプが追加されます。



Inspector を接続してワークスペースを実行し、FME Data Inspector でデータの内容を確認しておきましょう。

View 1 x

inspector [FFS] - H27Kokuchou Columns...

	市区町村コード	市区町村名	H27人口	H22人口(組替)	面積(km2)
1	01101	札幌市 中央区	237627	220189	46.42
2	01102	札幌市 北区	285321	278781	63.57
3	01103	札幌市 東区	261912	255873	56.97
4	01104	札幌市 白石区	209584	204259	34.47
5	01105	札幌市 豊平区	218652	212118	46.23
6	01106	札幌市 南区	141190	146341	657.48
7	01107	札幌市 西区	213578	211229	75.1
8	01108	札幌市 厚別区	127767	128492	24.38
9	01109	札幌市 手稲区	140999	139644	56.77
10	01110	札幌市 清田区	115726	116619	59.87
11	01202	函館市	265979	279127	677.86
12	01203	小樽市	121924	131928	243.83
13	01204	旭川市	339605	347095	747.66
14	01205	室蘭市	88564	94535	80.88
15	01206	釧路市	174742	181169	1362.92
16	01207	帯広市	169327	168057	619.34
17	01208	北見市	121226	125689	1427.41
18	01209	夕張市	8843	10922	763.07
19	01210	岩見沢市	84400	90145	481.02

Q in any column 1896 row(s)

FME Data Inspector では、通常、View ウィンドウにジオメトリ、Table View ウィンドウに属性テーブルが表示されますが、ツールバーの次のボタンの操作によって、View ウィンドウに属性テーブルを表示することもできます。

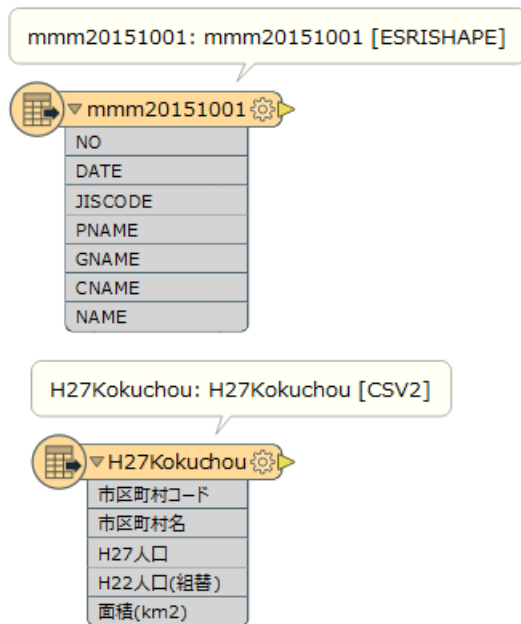


左から: 2D 表示、3D 表示、テーブル表示

上の図は「テーブル表示」の状態です。

### 5-3. ESRISHAPE リーダーの追加

これまでの実習と同じ要領で、**Esri Shapefile** 形式の市区町村区域ポリゴンデータを読み込むためのリーダーとフィーチャータ입を追加します。キャンバス上の配置は次のようにしてください。

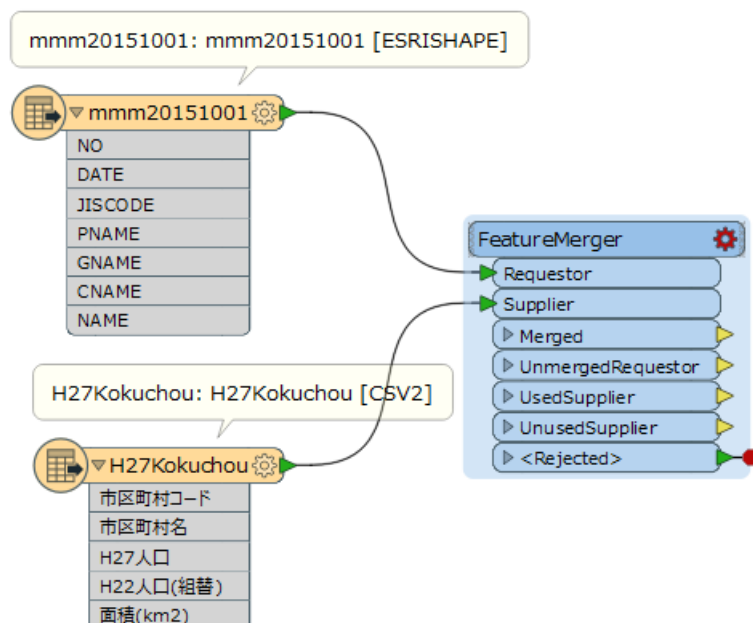


### 5-4. 属性の結合 FeatureMerger

これまでに見てきたように、市区町村区域ポリゴンには、市区町村コードを格納した「**JISCODE**」属性があり、国勢調査結果の**CSV** テーブルにも「市区町村コード」属性があります。したがって、それらが一致することを条件として、国勢調査の市区町村別の人口等を、対応する市区町村ポリゴンに結合することができます。

これを行うためにはいくつかのトランスフォーマーが使えますが、ここでは、最も基本的で、かつ応用範囲も広い **FeatureMerger** トランスフォーマーを使用します。

**FeatureMerger** トランスフォーマーをワークスペースに追加して、次のように接続してください。

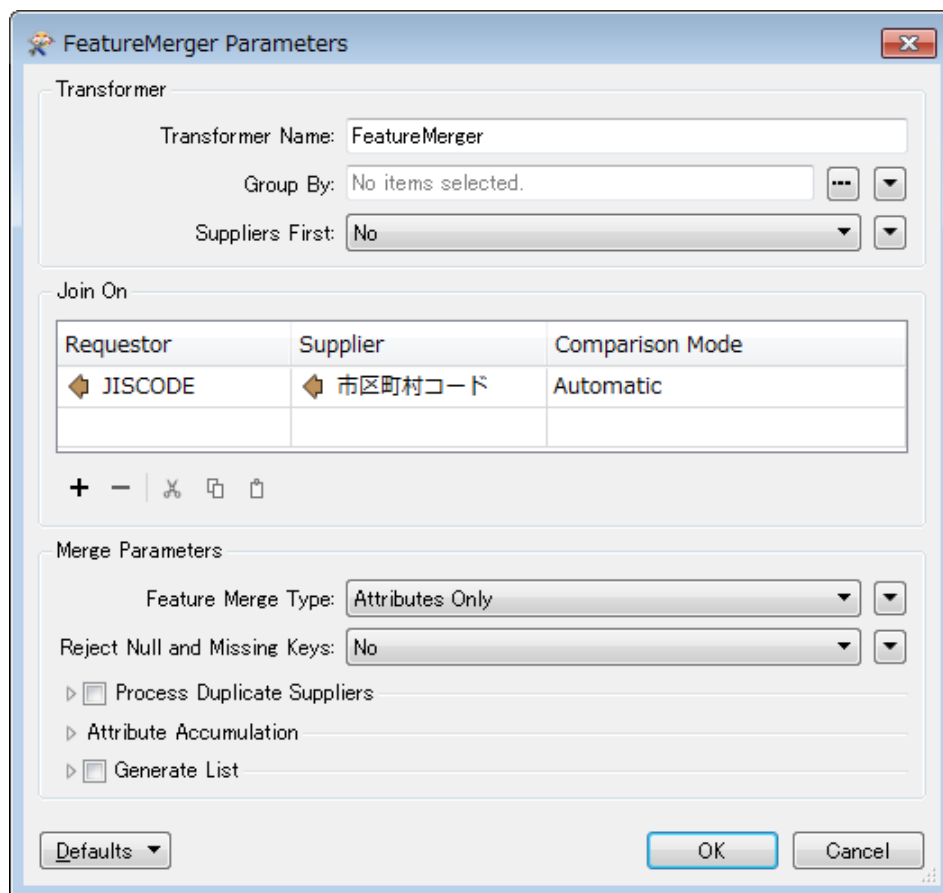


FeatureMerger には 2 つの入力ポートがあり、通常は、主たるフィーチャーが送られてくるデータフローを Requestor ポート、それに結合する属性を供給するフィーチャーが送られてくるデータフローを Supplier ポートに接続します。

属性を結合するための条件は、Join On パラメーターで指定します。パラメーター設定画面を開き、次のように、市区町村コードの一致を結合条件とするための設定をします。

Requestor 側: JISCODE

Supplier 側: 市区町村コード



ここで、市区町村ポリゴン (Shapefile) の JISCODE は数値型であり、北海道～栃木県については先頭の 0 が省かれた 4 桁の数字だったことを思い出してください。それに対して、Data Inspector で確認した国勢調査データ (CSV) の方では、先頭に 0 が付加された 5 桁の数字でした。これらは一致するでしょうか？

トランスフォーマーで属性値の比較をするときに、数値として評価できる文字列 (数字だけ、あるいは、小数点数の形式など) を数値として扱うか、文字列として扱うかを選ぶことができます。

Feature Merger の場合は、Join On パラメーターの Comparison Mode でコントロールします。

Automatic	Requestor 側、Supplier 側の属性値 (文字列) がどちらも数値として評価できるときは数値として比較し、そうでないときは文字列として比較する。
String	常に文字列として比較する。
Numeric	Automatic と同じ。

市区町村コードは数字のみで構成されるので、Comparison Mode を Automatic または Numeric とすれば数値として比較され、その場合は先頭の 0 が無視されるので、北海道～栃木県の市区町村コードも一致します。

FeatureMerger には 5 つの出力ポートがありますが、ここで使うのは Merged ポートのみです。Merged ポートからは、Join On で設定した条件を満たした Requestor フィーチャーが出力され、それらには、一致した Supplier フィーチャーの全ての属性が付加されます。

Merged ポートに Inspector を接続してワークスペースを実行すると、全ての市区町村ポリゴン（ただし、平成 27 年国勢調査結果がない北方四島 3 村を除く）が出力されることが確認できます。

### 5-5. 出力先データのスキーマ（フィーチャータイプ名と属性の構成）設計

ここまでで市区町村ポリゴンに平成 27 年国勢調査結果の人口等を結合することができました。ワークスペースの目的は、その結果を「都道府県別」の Shapefile ファイルに出力することでした。

Shapefile 形式の場合、ファイルがフィーチャータイプに対応するので、都道府県数分 = 47 のフィーチャータイプが必要です。都道府県を識別するための属性値に応じてデータフローを 47 に分岐し、それぞれにライターフィーチャータイプを接続することはできますが、キャンバス上でそのためのデータフローを作成するのはとても面倒なうえに間違いも起こりやすく、現実的とは言えません。

そこで、都道府県ごとのフィーチャータイプ名（Esri Shapefile 形式では出力先ファイル名に対応する）を各フィーチャーに属性として与え、それをライターフィーチャータイプにおけるフィーチャータイプ名に設定することとします。これにより、各フィーチャーの出力先は、その属性の値に応じて振り分けられるようになります。

ここでは、都道府県コード（市区町村コードの先頭 2 桁の数字）と都道府県名をアンダースコアで連結した文字列を出力先のフィーチャータイプ名とし、それを格納する新たな属性を追加することにします。たとえば、北海道は"01\_北海道"、沖縄県は"47\_沖縄県"です。

出力先データには、ソースデータの市区町村ポリゴンの全ての属性と、国勢調査の H27 人口、H22 人口(組替)、面積(km2)を出力することにします。

ここで、出力先のデータフォーマットが Esri Shapefile 形式である場合は、フォーマットの仕様上、属性名は 10 バイト以内でなければならないという制約があることに注意してください。日本語版 Windows の標準文字コード（Shift JIS に準じる）では全角文字は 1 文字 2 バイトであり、属性名「H22 人口(組替)」が 10 バイトを超えてしまうので、それは「H22 人口」に変更することとします。

このことも考慮して、出力先データの属性の構成を次のようにします。

Name	Type	Width	Precision	備考
NO	number	11	0	ソースデータ(市区町村ポリゴン)と同じ
DATE	date			date 型に変更
JISCODE	number	11	0	ソースデータ(市区町村ポリゴン)と同じ
PNAME	char	12		同上
GNAME	char	30		同上
CNAME	char	21		同上
NAME	char	42		同上
H27 人口	number	8	0	データ型設定
H22 人口	number	8	0	属性名変更、データ型設定
面積(km2)	number	8	2	データ型設定

## 5-6. 属性の変換

都道府県別の出力先フィーチャータイプ名として都道府県コード (2桁文字列) と都道府県名をアンダースコアで連結した文字列属性を作成するために、次の 2 つのトランスフォーマーをワークスペースに追加、接続します。

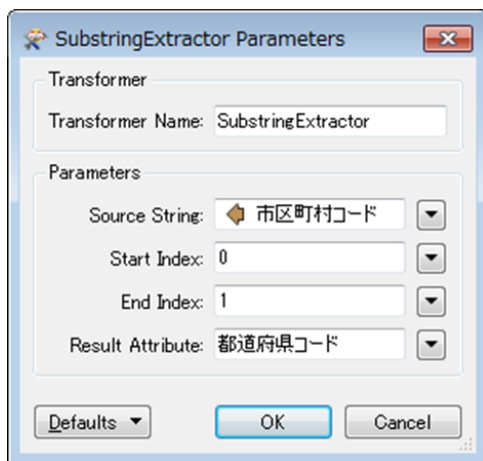
1) **SubstringExtractor**: 次のパラメーター設定によって市区町村コードの先頭 2 文字 (都道府県コード) を抽出し、「都道府県コード」属性に格納します。

**Source Attribute**: 部分文字列を抽出する元の属性

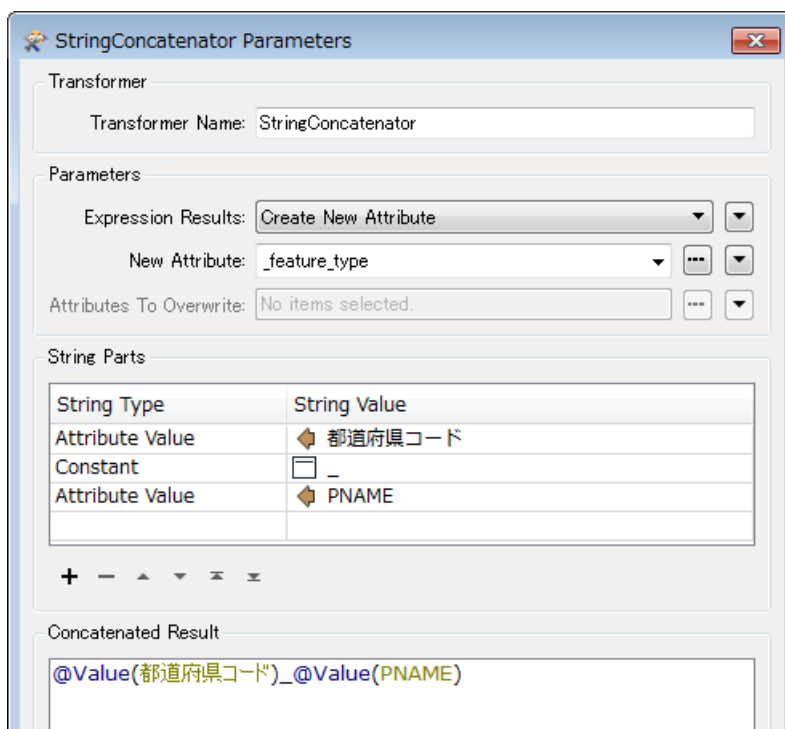
**Start Index**: 抽出する部分文字列の先頭の位置 (元の文字列の先頭を 0 とするインデックス)

**End Index**: 抽出する部分文字列の末尾の位置 (同上)

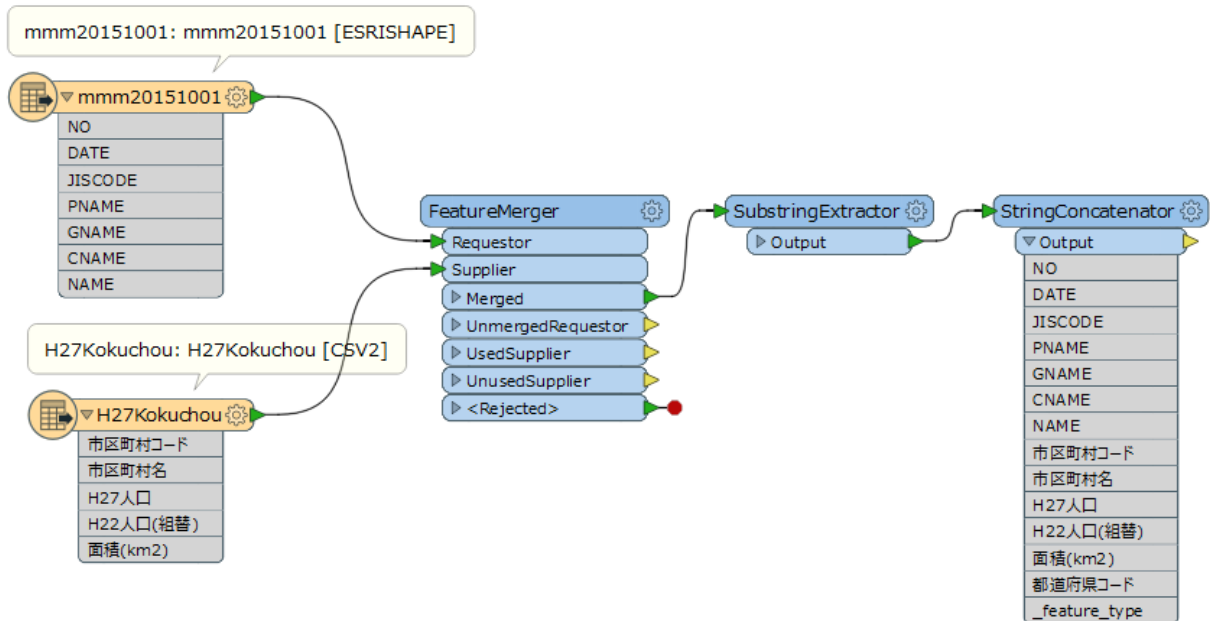
**Result Attribute**: 抽出した部分文字列を格納する属性名



2) **StringConcatenator**: 次のパラメーター設定によって出力先フィーチャータイプ名 (Shapefile ファイル名) とする文字列を作成し、"\_feature\_type"属性に格納します。StringConcatenator については実習 3 も参照してください。



これらを追加、接続した後のキャンバス上のデータフローは次のようになります。

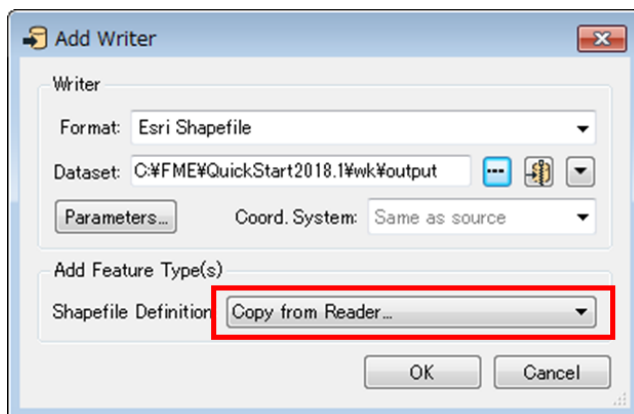


## 5-7. ライターの追加（Copy from Reader モード）と出力先スキーマの編集

出力先の ESRISHAPE ライターとフィーチャータ입を追加します。

### (1) ESRISHAPE ライターの追加

Workbench メニュー Writers > Add Writer を選択して Add Writer 画面を開き、Format と Dataset を設定してください。

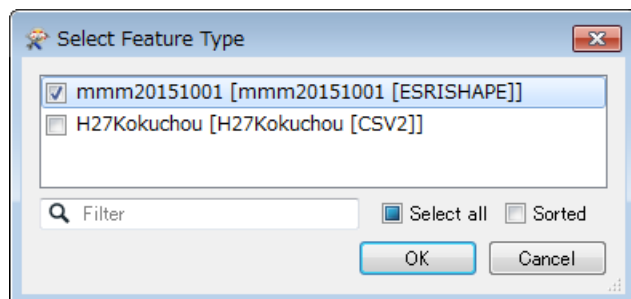


今回は、Add Feature Type(s)として Copy from Reader（ライターフィーチャータ입を追加し、リーダーフィーチャータ입のスキーマをコピーして設定する）を試します。

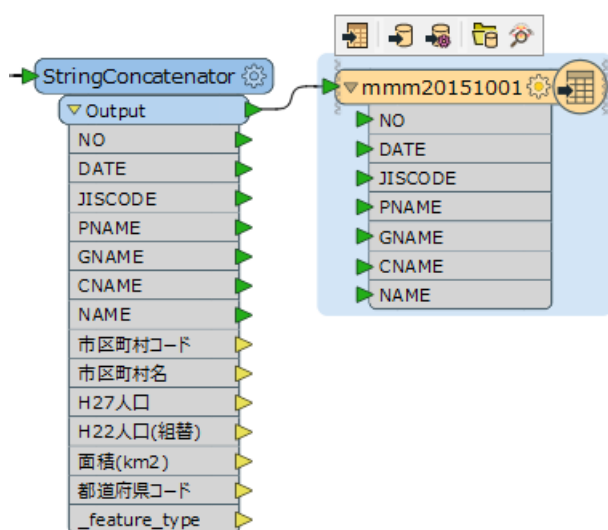
上の図のように Shape File Definition で Copy from Reader を選択してから OK で閉じてください。



ワークスペースには 2 つのリーダーフィーチャータ입 (Shapefile データと CSV データ) があるので、次のように **Select Feature Type** 画面が開き、どちらのスキーマをコピーするかが選択できます。



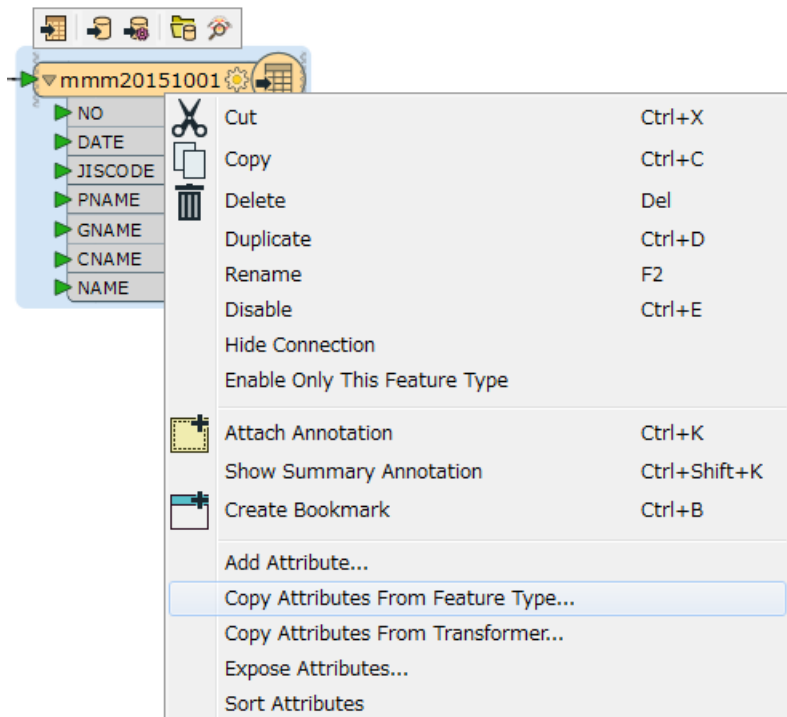
Shapefile データの属性は全て出力することにしたので、不足分は後で追加、編集することとし、ここでは図のように Shapefile データの方のフィーチャータ입を選択して **OK** で閉じ、キャンバスに現れたライターフィーチャータ입を **StringConcatenator** の **Output** ポートに接続してください。



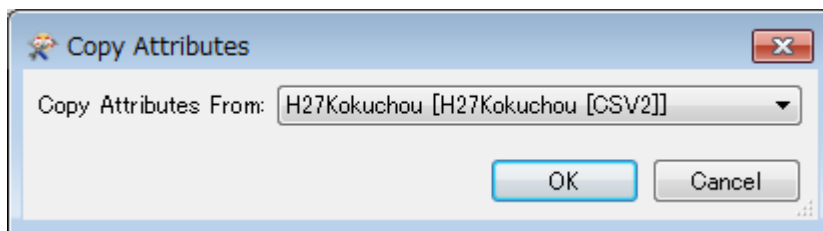
## (2) 出力先スキーマの編集

まず、**CSV** データから読み込んだ人口等を出力するための属性をライターフィーチャータ입に追加します。手動で追加しても良いのですが、ここでは少しでも手間を省くため、**CSV** リーダーフィーチャータ입の属性をコピーして追加し、それを編集することになります。

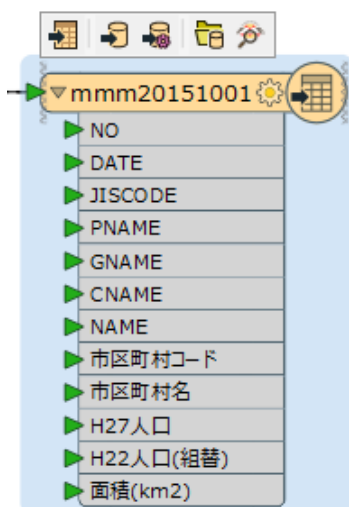
ライターフィーチャータ입を右クリックして表示されるメニューで、**Copy Attributes From Feature Type** を選択してください。



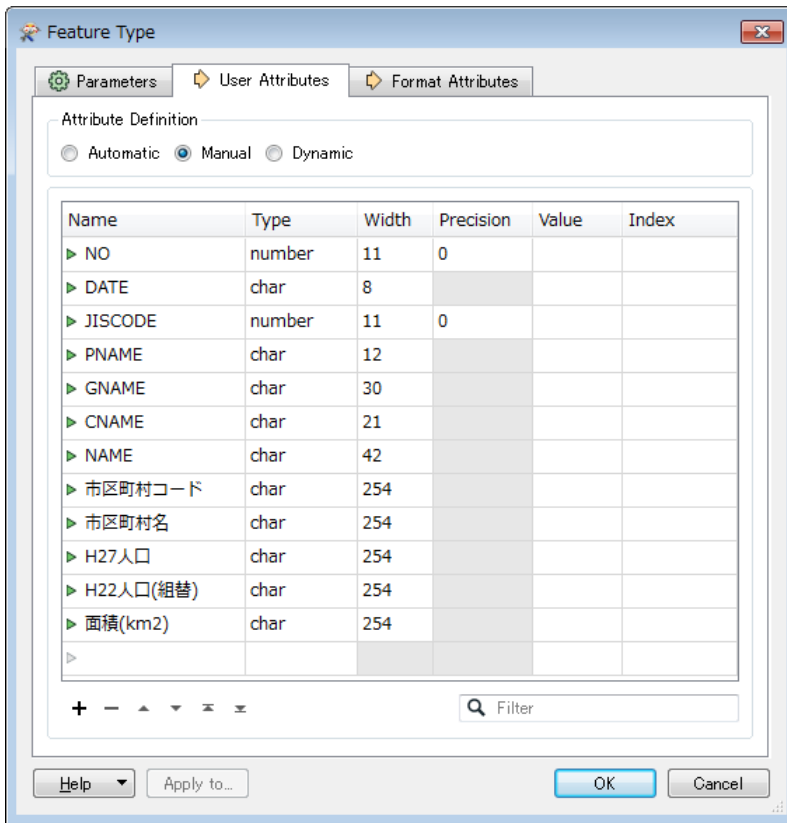
Copy Attributes 画面が表示されるので、CSV データのリーダーフィーチャータイプを選択して OK で閉じます。



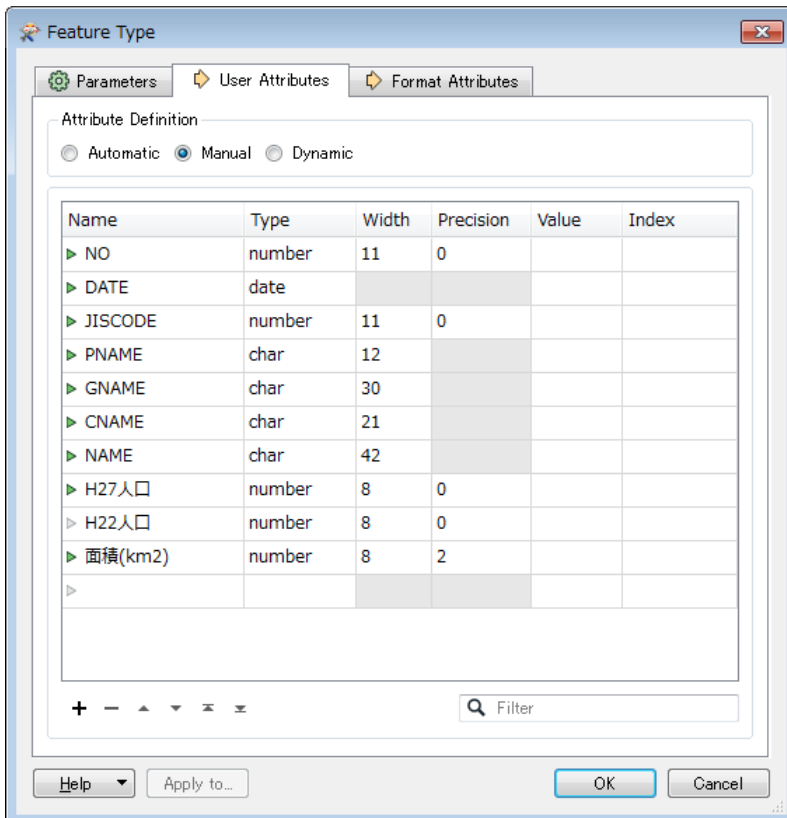
これによって、CSV リーダーフィーチャータイプの属性 "市区町村コード" ~ "面積(km2)" が追加されました。



Feature Type 画面を開き、User Attributes タブを表示させてください。

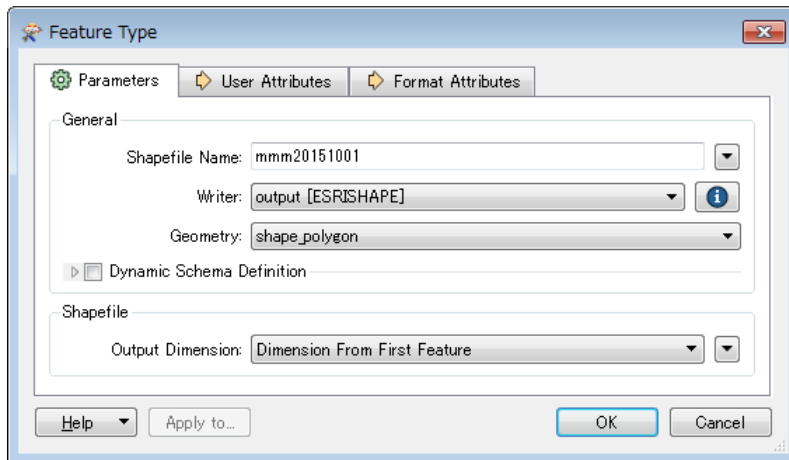


これを編集（データ型の変更、属性名の変更、不要な属性の削除）し、次の図のように、あらかじめ設計しておいたスキーマに変更します。



## 5-8. 出力先フィーチャータイプ名の設定

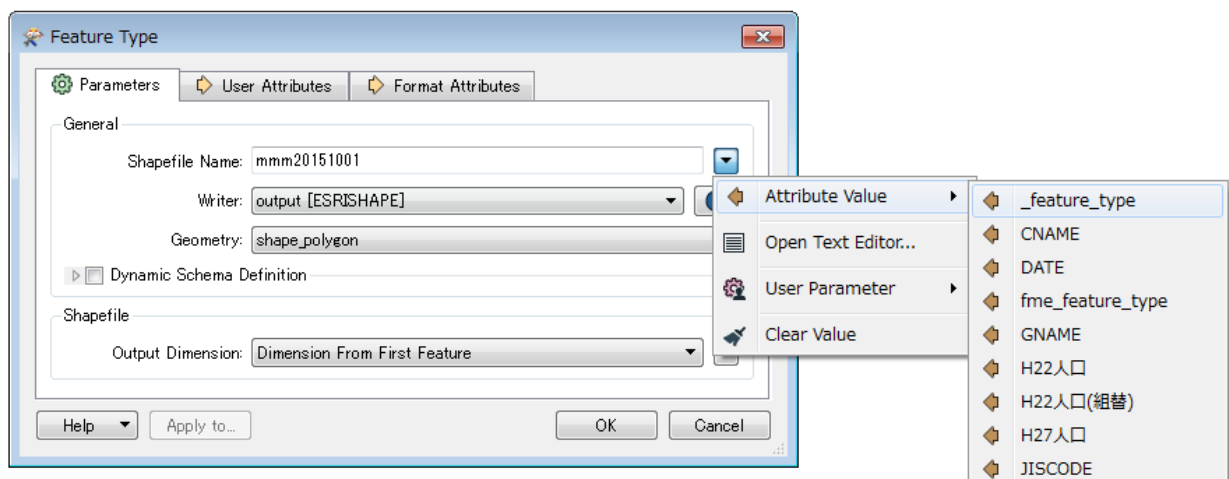
Feature Type 画面の Parameters タブを表示させます。



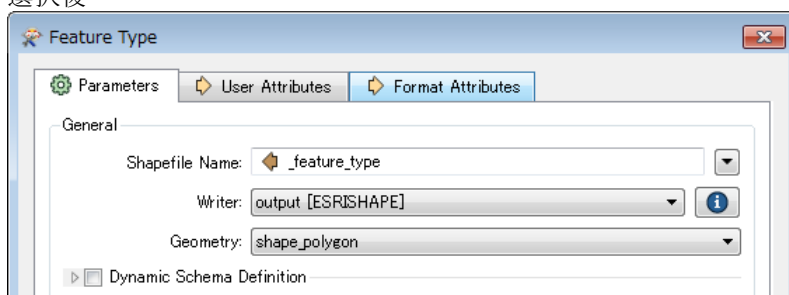
ライターを追加したときに Shapefile リーダーフィーチャータイプのスキーマをコピーしたので、フィーチャータイプ名 (Shapefile Name = mmm20151001) もコピーされています。このまま実行すると、この名前の Shapefile ファイルに「全ての」フィーチャーが出力されます。

このワークスペースは、都道府県別ファイルにフィーチャーを振り分けて出力するのが目的です。出力先のフィーチャータイプ名 (ファイル名) を格納する属性 (\_feature\_type) は作成済みなので、それを "Shapefile Name" フィールドで選択します。

フィーチャータイプ名入力フィールド右側のメニューボタン (▼) をクリックし、"Attribute Value" (属性値) コマンドにマウスカーソルをあわせると、その右側に選択可能な属性のリストが表示されるので、"\_feature\_type" を選択してください。



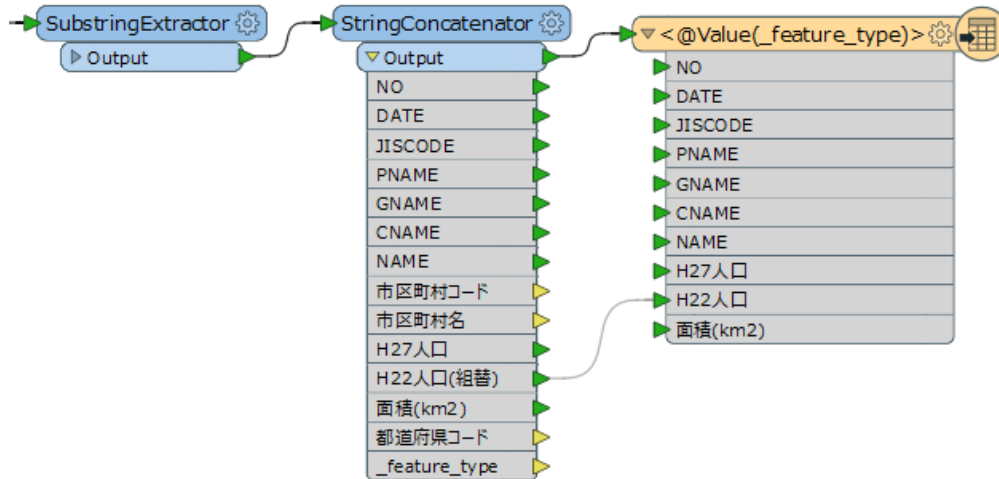
選択後



以上の設定により、各フィーチャーの出力先が"**\_feature\_type**"の値、すなわち都道府県別のファイルに振り分けられるようになります。

**Feature Type** 画面を **OK** で閉じ、名前を変更した属性 "**H22 人口**" だけ明示的にマッピングすればワークスペースは完成です。

H22 人口(組替) → H22 人口



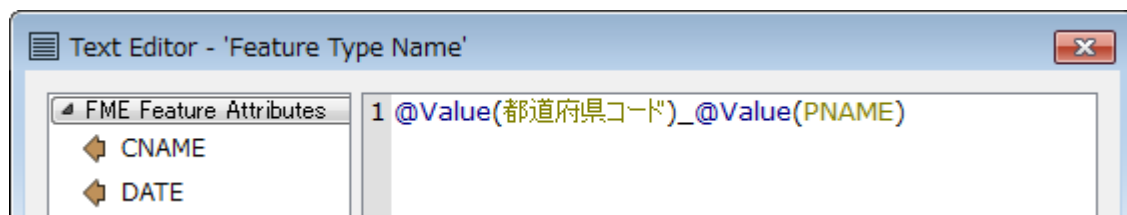
**StringConcatenator** とライターフィーチャータイプの間に **AttributeManager** または **AttributeRenamer** を挿入して属性名を変更しても構いません。

ワークスペースを実行し、出力先データセットとして指定したフォルダーに都道府県別の **Shapefile** ファイルが出力されたことを確認してください。

実習 5 はここまでです。主に次の事項を学びました。

- CSV データの読込
- **FeatureMerger** による属性の結合
- ライターの追加 (**Copy from Reader** モード)
- 属性値による出力先フィーチャータイプの振り分け

この実習のワークスペースでは、フィーチャータイプ名フィールドのメニューボタンで"**Open Text Editor**"を選択してテキストエディタを開き、次の図のように文字列式を入力すると、各フィーチャーの属性値に基づき、ここで出力先フィーチャータイプ名 "<都道府県コード>\_<都道府県名>" が作成されます。この式を設定した場合は、**StringConcatenator** を省くことができます。



属性値や式をフィーチャータイプ名に設定することによって、出力先フィーチャータイプを振り分ける機能は、「**Feature Type Fanout**」と呼ばれます。