

# 第 3 講

---

---



---

## 目次

第3講 .....	1
1. データから見る地域特性 .....	1
(1) 地域の課題解決と統計活用 .....	1
(2) 地域課題の分析事例 .....	2



# 第3講

## 1. データから見る地域特性

本章の Point	地理総合において必要な「地理情報の収集および諸資料の情報化」の具体的方法について学びます。さらに、統計的手法を活用した世界の状況や地域の現状の理解と地域の課題解決に向けた取り組みや探究方法について学びます。
-----------	---

### (1) 地域の課題解決と統計活用

本項における参考情報	◇ 高校からの統計・データサイエンス活用～上級編～ 第6部地域の課題解決と統計活用 ◇ 指導用 高校からの統計・データサイエンス活用 ～上級編～ 第6部地域の課題解決と統計活用
------------	---

地域の課題解決の流れは、以下のとおりです。

STEP1 課題設定	STEP2 必要なデータ と分析計画	STEP3 データ・統計資料 の収集	STEP4 分析	STEP5 結論
課題を設定し、仮説をたてる。	どのようなデータ・統計資料を集めてどのような分析をするか検討する。	必要なデータ・統計資料を集める。 収集したデータに応じ、欠損値の処理等加工する。	グラフや統計量で傾向を捉える。	結論を導く。

出典：指導用 高校からの統計・データサイエンス活用 ～上級編～ 第6部 地域の課題解決と統計活用 より整理

まず、地域の課題を設定し、その課題の周辺情報を調べます。その上で、課題や地域の状況を表すデータや、課題内容に応じ入手可能な情報を検討し、仮説と分析方針を決定します。必要なデータを収集、分析のために欠損値の処理等の加工、取り纏めを行います。統計量を算出、データを可視化し、分析します。次の項目からは、実際の分析事例で説明していきます。

## (2) 地域課題の分析事例

### 事例1：品質の良い米を育成する条件

ここでは、分析の手順および EXCEL を用いた分析の方法について説明します。

課題解決の手順	実施内容例
<b>STEP1:問題</b> 課題の設定	<p>◇「品質」の良い米をつくるには 例) 米の品質とは何か？を調べる</p> <ul style="list-style-type: none"> <li>・品位等検査による格付け 1等米が最高ランク</li> <li>・研究会の分析報告：「2010年の夏、新潟県は記録的な猛暑に襲われたため、コメに栄養分が行き届かず、白濁した粒が増えて品質が低下した」</li> </ul> <p>調査結果の例は、以下を参照 ・指導用 高校からの統計・データサイエンス活用 ～上級編～ (<a href="https://www.jstat.or.jp/mwbhwp/wp-content/uploads/sfs_6-1-1.zip">https://www.jstat.or.jp/mwbhwp/wp-content/uploads/sfs_6-1-1.zip</a>)</p> <p>⇒<b>米の品質と気温に関係ありとの仮説設定</b></p>
<b>STEP2:計画</b> どのようなデータ・統計資料を集めて どのような分析をするか	<p>◇収集するデータ：調べた条件に該当するデータを収集する</p> <ul style="list-style-type: none"> <li>・品質が良い米:1等米 ⇒農林水産省公表の1等米比率</li> <li>・特に品質に影響を与えるとされる登熟期間である8月の気温等 ⇒気象データ 8月の気温（平均、最高、最低）、降水量、日照時間</li> </ul> <p>◇分析内容：仮定した条件（データ）の関係を探る</p> <ol style="list-style-type: none"> <li>a) 1等米比率と各気象データの時系列変化</li> <li>b) 1等米比率と各気象データの相関関係</li> <li>c) 品質が低下する可能性がある気温を回帰分析により予測</li> </ol> <p>⇒<b>1等米比率と8月の気温の関係の検証</b> <b>品質が低下する可能性がある平均気温の予測</b></p>
<b>STEP3:収集</b> 必要なデータ・統計資料を集める	<p><u>農林水産省公表の1等米比率</u></p> <p>農林水産省ホームページ統計情報 <a href="https://www.maff.go.jp/j/tokei/">https://www.maff.go.jp/j/tokei/</a> ホーム &gt; 研究会等 &gt; 水稻の作柄に関する委員会 &gt; 平成30年度 第3回水稻の作柄に関する委員会（平成31年3月14日） 資料 No.2-1 作況指数、10a 当たり収量、平年収量及び一等米比率の推移 <u>新潟の過去の気象データ</u></p> <p>気象庁ホームページ <a href="http://www.jma.go.jp/jma/index.html">http://www.jma.go.jp/jma/index.html</a> ホーム &gt; 各種データ・資料 &gt; 過去の気象データ・ダウンロード 新潟・月平均気温、月最高気温、月最低気温・降水量の月合計・日照時間の月合計、8月、1979年～2015年を選択地点・項目・期間を選択してダウンロードする。（CSV形式）</p> <p>⇒<b>データを収集し、データクリーニングを行う</b></p>

<b>STEP4:分析</b> グラフや統計量で 傾向を捉える	a) 1 等米比率と各気象データの時系列変化 ⇒時系列変化のグラフ（1 等米比率、平均気温、最高気温、最低気温） b) 1 等米比率と各気象データの相関関係 ⇒8 月の平均気温と 1 等米比率の散布図、相関係数の算出 c) 品質が低下する可能性がある気温を回帰分析により予測 ⇒8 月の平均気温と 1 等米比率の散布図、回帰 ⇒EXCEL を使用し、時系列変化の分析、相関関係の分析、回帰分析を行う
<b>STEP5:結論</b> 結論を導く	分析結果を読み解く ・品質は、平均気温と相関関係がある ・平均気温 27℃以上で品質低下が起こる

出典：指導用 高校からの統計・データサイエンス活用 ～上級編～ 第6部 地域の課題解決と統計活用 より整理

上記分析の STEP 3 以降の手順、分析結果を説明する。

### <STEP3：収集 必要なデータ・統計資料を集める>

#### ① データの取得

農林水産省データ：農林水産省ホームページ統計情報 <https://www.maff.go.jp/j/tokei/>  
 ホーム > 研究会等 > 水稻の作柄に関する委員会 > 平成 30 年度 第 3 回水稻の作柄に関する委員会（平成 31 年 3 月 14 日）> 資料 No.2-1 作況指数、10a 当たり収量、平年収量及び一等米比率の推移

気象データ：気象庁ホームページ <http://www.jma.go.jp/jma/>

ホーム > 各種データ・資料 > 過去の気象データ・ダウンロード

選択地点・項目・期間を選択して CSV 形式のファイルをダウンロードする。

ホーム > 各種データ・資料 > 過去の気象データ・ダウンロード

過去の気象データ・ダウンロード

検索条件

データをダウンロード

重要なお知らせ

このページの使い方

よくある質問

CSVファイルの形式

画面に表示 ▶

CSVファイルをダウンロード ▶

選択地点・項目をクリア

選択された地点

観測項目

新潟

削除

選択された項目

月平均気温

月最高気温

月最低気温

降水量の月合計

日照時間の月合計

削除

削除

削除

削除

削除

選択された期間(日本標準時)

1979年1月から

2020年1月までの月別値を表示

選択されたオプション

利用上注意が必要なデータを表示させる

観測環境などの変化以前のデータを表示させる

ダウンロードデータはすべて数値で格納

設定

データの種類

時別値

日別値

2 日別値

半旬別値

旬別値

月別値

3か月別値※

最初に選択してください

過去の平均値との比較オプション

平年値も表示

平年値からの差(比)も表示

(平年値1981年から2010年の30年平均値)

前年までの 1 年平均も表示

前年までの 1 年平均からの差(比)も表示

項目

気温

降水

日照/日射

積雪/降雪

風

湿度/気圧

雲量/天気

月平均気温

日最高気温の月平均

日最低気温の月平均

月最高気温

月最低気温

日最高気温の月最高※

日最低気温の月最低※

日平均気温 25℃以上の日数(月)

日平均気温 0℃未満の日数(月)

日最高気温 25℃以上の日数(月)

日最高気温 0℃未満の日数(月)

日最低気温 25℃以上の日数(月)

日最低気温 0℃未満の日数(月)

※気象庁(気象台等)のみ値があります

最高・最低(最大・最小)値の発生時刻を表示

出典：気象庁ホームページ 過去の気象データ・ダウンロード <http://www.data.jma.go.jp/gmd/risk/obsdl/index.php>

## ② 必要データ項目を整理する。

1等米比率と新潟市8月の気象データ（1979年～2015年）

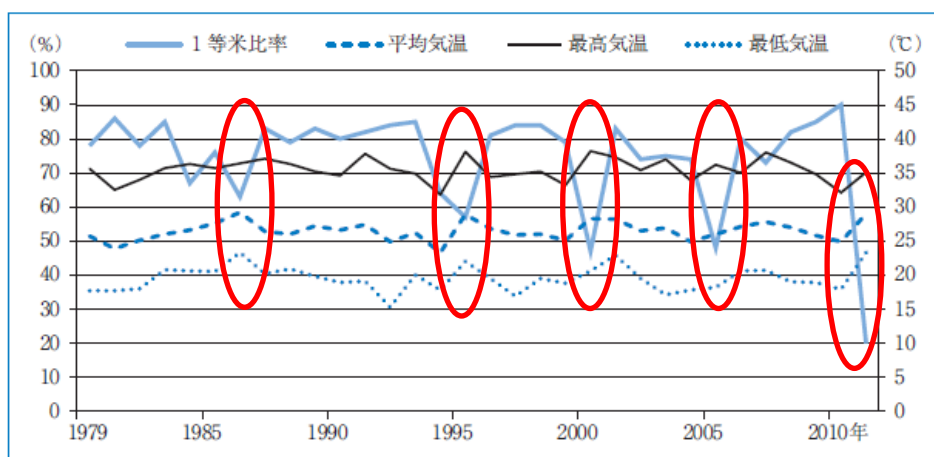
年	平均気温（℃）	最高気温（℃）	最低気温（℃）	降水量（mm）	日照時間（時間）	1等米比率（％）
1979	25.7	35.6	17.7	130.0	217.3	78
80	23.9	32.5	17.7	250.0	175.8	86
81	25.1	34.0	18.0	213.0	214.7	78
82	26.0	35.7	20.8	38.5	202.5	85
83	26.6	36.3	20.6	89.5	216.2	67
84	27.6	35.7	20.5	17.5	301.7	76
85	29.2	36.4	23.3	3.5	342.6	63
86	26.4	37.1	20.2	67.5	236.4	83
87	26.0	36.3	20.9	166.0	168.2	79

出典：指導用 高校からの統計・データサイエンス活用 ～上級編～ 第6部 地域の課題解決と統計活用

## &lt;STEP4、STEP5の詳細&gt;

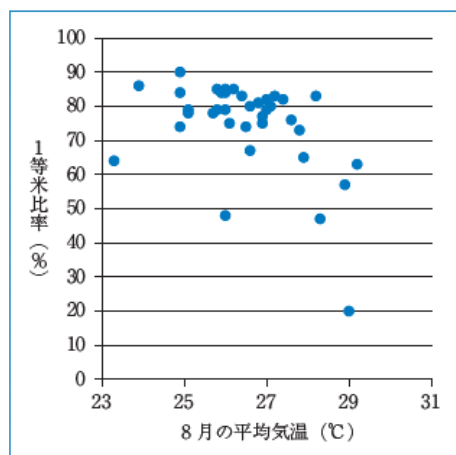
## a) 1等米比率と各気象データの時系列変化

1等米比率と気温の推移



1等米比率と気温の経年変化のグラフから、1等米比率の割合が落ち込む時は、平均気温、最高気温、最低気温が高くなる傾向が読み取れる。気温の上昇と米の品質には関係があると思われる。

平均気温と1等米比率の関係



## b) 1等米比率と各気象データの相関関係

EXCELのCORREL関数により、相関係数は以下のとおりであり、弱い相関があると言える。

相関係数  $r = -0.460$

相関係数の絶対値による相関の判定は、経験的に以下のようになっていることが多い。

強い相関がある：0.8以上

弱い相関がある：0.4以上

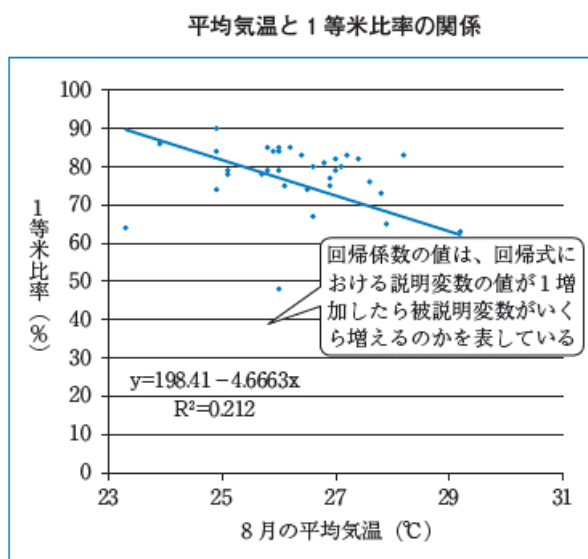
ほぼ相関がない：0.2未満

また、 $r$ は負の値であり、これは右肩下がりの直線で、

1等米比率と平均気温は負の相関関係を示している。



### c) 品質が低下する可能性がある気温を回帰分析により予測



散布図に近似曲線を追加したところ左図のとおり

である。回帰直線を表す式は、

$$y = 198.41 - 4.6663x$$

$$R^2 = 0.212$$

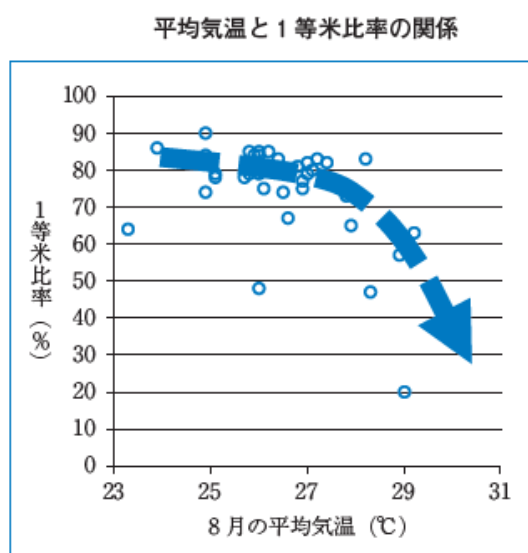
回帰分析とは、ある変数に他の変数がどのように影響するかを量的に求める分析手法です。

この式より、気温から、1等米比率をある程度、予想することができます。 $R^2$ は決定係数と呼ばれ、被説明変数のばらつきが、説明変数のばらつきによってどの程度説明できるかを示します。常に0～1の値になります。1に近いほど、見かけ上当てはまりの良い回帰式と言えます。

## 第3講

出典：指導用 高校からの統計・データサイエンス活用 ～上級編～  
第6部 地域の課題解決と統計活用 より整理

この散布図で、説明変数（8月の平均気温）と被説明変数（1等米比率）が直線関係かを確認すると、平均気温が上がると、1等米比率が下がる傾向にありますが、27.0℃あたりを境に1等米比率が急に低下する傾向があります。平均気温27.0℃以上の年データのみで、相関係数を計算すると相関係数は $r = -0.735$ となり、マイナス値が大きくなり27.0℃以上だとより負の相関が強くなります。



平均気温27.0℃以上の年の1等米比率

年	平均気温(°C)	1等米比率(%)
1985	29.2	63
2010	29.0	20
1994	28.9	57
1999	28.3	47
2000	28.2	83
2012	27.9	65
2006	27.8	73
1984	27.6	76
1990	27.4	82
1988	27.2	83
2005	27.1	80
2007	27.0	82
2011	27.0	79
平均	27.9	68.5

出典：指導用 高校からの統計・データサイエンス活用 ～上級編～ 第6部 地域の課題解決と統計活用 より整理

## 事例2：人口減少に向かう地域の把握：将来人口推計

課題解決の手順	実施内容例
<b>STEP1:問題</b> 課題の設定	◇人口減少に向かう地域の把握 例) 全国で最も早く人口減少が進んだ島根県の将来人口をコーホート変化率法 <sup>1</sup> で推計する
<b>STEP2:計画</b> どのようなデータ・統計資料を集めて どのような分析をするか	◇収集するデータ：5年後の人口の推計 ・島根県の男女別5歳階級別の人口 ⇒国勢調査 出生 2005年、2010年、2015年 ・母親の年齢別出生数 ⇒厚生労働省 人口動態調査 2005年、2010年、2015年  ◇分析内容：コーホート変化率法での推計 a)5歳刻みの年齢階級別男女別人口推計値を算出 b)0～4歳の男女別人口推計値を算出
<b>STEP3:収集</b> 必要なデータ・統計資料を集める	<u>島根県の男女別5歳階級別の人口</u> 国勢調査 e-Stat <a href="https://www.e-stat.go.jp/">https://www.e-stat.go.jp/</a> e-Stat > 統計データ > すべて > 主要な統計から探す > 国勢調査 > 3-1 年齢(各歳), 男女別人口, 年齢別割合, 平均年齢及び年齢中位数(総数及び日本人) - 全国※, 全国市部・郡部, 都道府県, 21 大都市, 特別区, 人口 50 万以上の市  <u>母親の年齢別出生数</u> 厚生労働省 人口動態調査 e-Stat <a href="https://www.e-stat.go.jp/">https://www.e-stat.go.jp/</a> e-Stat > 統計データ > すべて > 主要な統計から探す > 人口動態調査 > 出生 年次 > 2015 年 中巻 5 出生数, 性・母の年齢(5歳階級)・都道府県(21 大都市再掲) 別

<sup>1</sup> 人口推計 <https://www.mhlw.go.jp/topics/bukyoku/seisaku/syousika/030819/2b.html>

「コーホート」とは、同じ年（又は同じ期間）に生まれた人々の集団のことを指す

「コーホート変化率法」とは、各コーホートについて、過去における実績人口の動勢から「変化率」を求め、それに基づき将来人口を推計する方法である。

STEP4:分析 グラフや統計量で 傾向を捉える	<p>コーホート変化率法での推計（2005年～2010年）</p> <p>手順1：5歳刻みの年齢階級別 2005年～2010年の男女別コーホート変化率を算出</p> <p>手順2：算出したコーホート変化率に基づき、5歳刻みの年齢階級別人口の2015年推計値を算出</p> <p>手順3：0～4歳人口の推計のための、人口指標を算出</p> <ul style="list-style-type: none"> <li>－女性年齢5歳階級別出生率の算出</li> <li>－出生性比の算出</li> </ul> <p>手順4：5年間の累積出生数の算出</p> <p>（女性年齢5歳階級別人口×年齢5歳階級別出生率）×5年</p> <p>手順5：男女別の0～4歳人口の推計値の算出（出生数×出生性比）</p>
	<p>STEP5:結論 結論を導く</p> <ul style="list-style-type: none"> <li>・2015年の島根県の男女別5歳階級別の人口推計値を分析する</li> <li>・2015年の国勢調査結果との比較を分析する</li> </ul>

### <STEP4の詳細>

手順1：5歳刻みの年齢階級別 2005年～2010年の男女別コーホート変化率を算出

算出例）2005年～2010年の25～29歳のコーホート変化率

= 2010年25～29歳の年齢階級別人口 ÷ 2005年20～24歳の年齢階級別人口

手順2：算出したコーホート変化率に基づき、5歳刻みの年齢階級別人口の2015年推計値を算出

算出例）2015年25～29歳の年齢階級別推計人口

= 手順1算出25～29歳のコーホート変化率 × 2010年20～24歳の年齢階級別人口

島根県の年齢階級別の男性人口					島根県の年齢階級別の女性人口				
総数(年齢)	国勢調査男性人口		手順1 2005～10年までの コーホート変化率	手順2 ステップ1×2010年コホート人口 =2015年推計人口	総数(年齢)	国勢調査女性人口		手順1 2005～10年までの コーホート変化率	手順2 ステップ1×2010年コホート人口 =2015年推計人口
	2005年	2010年				2005年	2010年		
総数(年齢)	353,703	342,991			総数(年齢)	388,520	374,406		
0～4歳	15,710	14,519			0～4歳	15,006	13,906		
5～9歳	17,062	15,734	1.002	14,548	5～9歳	16,236	15,068	1.004	13,962
10～14歳	18,632	16,892	0.990	15,577	10～14歳	17,896	16,099	0.992	14,947
15～19歳	19,788	17,111	0.918	15,507	15～19歳	18,080	15,788	0.882	14,199
20～24歳	<b>16,239</b>	13,464	0.680	11,635	20～24歳	<b>16,186</b>	13,484	0.746	11,778
25～29歳	19,844	<b>16,779</b>	<b>1.033</b>	<b>13,908</b>	25～29歳	19,288	<b>16,347</b>	<b>1.010</b>	<b>13,619</b>
30～34歳	22,314	20,414	1.029	17,266	30～34歳	21,967	19,231	0.997	16,298
35～39歳	19,318	22,758	1.020	20,822	35～39歳	19,316	21,685	0.987	18,981
40～44歳	19,921	19,583	1.014	23,077	40～44歳	20,508	19,187	0.993	21,533
45～49歳	22,860	20,000	1.004	19,661	45～49歳	22,455	20,234	0.987	18,938
50～54歳	27,270	22,854	1.000	20,000	50～54歳	25,929	22,207	0.989	20,011
55～59歳	31,116	26,896	0.986	22,534	55～59歳	29,970	25,714	0.992	22,029
60～64歳	22,475	30,526	0.981	26,385	60～64歳	24,627	29,891	0.997	25,637
65～69歳	20,961	21,343	0.950	29,000	65～69歳	25,035	24,120	0.979	29,263
70～74歳	22,301	19,121	0.912	19,465	70～74歳	27,942	24,078	0.962	23,203
75～79歳	18,203	19,063	0.855	16,348	75～79歳	26,254	26,162	0.936	22,537
80～84歳	11,153	13,795	0.758	14,450	80～84歳	20,403	23,184	0.883	23,101
85～89歳	5,324	6,861	0.615	8,484	85～89歳	12,615	15,911	0.780	18,084
90～94歳	2,043	2,425	0.455	3,122	90～94歳	6,310	7,764	0.615	9,785
95～99歳	428	600	0.294	713	95～99歳	1,807	2,456	0.389	3,020
100歳以上	62	69	0.141	85	100歳以上	262	446	0.216	530
不詳	679	2,184			不詳	428	1,444		

資料：総務省「国勢調査」

## 手順3：0～4歳人口の推計のための、人口指標を算出

女性の年齢5歳階級別出生率を算出します。

$$\text{出生率} = \text{出生数} \div \text{各年齢階級の女性人口}$$

次に、出生性比を求めます。安定的な数値を得るために、2009年、2010年、2011年の3カ年の平均を取ります。

$$\text{出生性比} = \text{出生男児数} \div \text{出生女児数} \times 100$$

(女子100人に対して生まれる男子の数)

2010年における島根県の出生率と出生性比

女性の年齢5歳階級別出生率

出生率	
～14歳	0.00002
15～19歳	0.00464
20～24歳	0.05023
25～29歳	0.10677
30～34歳	0.10256
35～39歳	0.04753
40～44歳	0.00754
45～49歳	0.00012
50歳～	0

出生性比	
出生性比	104.0

## 手順4：5年間の累積出生数の算出

$$\text{累積出生数} = (\text{女性年齢5歳階級別人口} \times \text{年齢5歳階級別出生率}) \times 5 \text{年}$$

2010年女性人口より算出された年間出生数は、5,645人。5年の累積出生数は、28,225人となります。なお、2005年と2010年の女子人口の平均値を取る方が望ましいです。

年齢	出生率	2010年 女性人口	出生数推計値	
			算出値	少数点以下切捨て
～14歳	0.00002	16,099	0.32198	0
15～19歳	0.00464	15,788	73.25632	73
20～24歳	0.05023	13,484	677.3013	677
25～29歳	0.10677	16,347	1,745.369	1,745
30～34歳	0.10256	19,231	1,972.331	1,972
35～39歳	0.04753	21,685	1,030.688	1,031
40～44歳	0.00754	19,187	144.67	145
45～49歳	0.00012	20,234	2.42808	2
50歳～	0	22,207	0	0
計				5645

## 手順5：2015年男女別の0～4歳人口の推計値の算出

出生性比により男女を振り分け、2015年の男女別の0～4歳人口を推計します。

$$\text{男子 } 28,225 \times 104.0 / 204.0 = 14,389.2156 \Rightarrow 14,389 \text{ 人}$$

$$\text{女子 } 28,225 - 14,389 = 13,836 \text{ 人}$$

## &lt;STEP5の詳細&gt;

2015年の島根県の人口推計値および2015年国勢調査人口は、以下のとおり。実際の人口と比較すると、全体的に推計値よりも人口は少なかった。乖離が大きかったのは、0～4歳、15～19歳、20～24歳であった。「コーホート変化率法」は、推計に必要なデータを入手しやすい点や、他の地域でも数値を入れ替えて計算が容易な点である。ただし、元となる人口構成や変化率が、今後も同様に続くことが前提となっている推計方法であるため、長期の推計を行う場合は誤差が大きくなる可能性が高い。また、小さな地域の推計値を利用する際は、人口移動等の影響が強く表れる可能性がある。

	2015年人口推計値		2015年国勢調査人口		国勢調査一推計値		誤差率	
	男性	女性	男性	女性	男性	女性	男性	女性
0～4歳	14389	13836	13822	12995	▲ 567	▲ 841	▲ 3.9	▲ 6.1
5～9歳	14548	13962	14566	13919	18	▲ 43	0.1	▲ 0.3
10～14歳	15577	14947	15553	14864	▲ 24	▲ 83	▲ 0.2	▲ 0.6
15～19歳	15507	14199	16276	14601	769	402	5.0	2.8
20～24歳	11635	11778	12481	11704	846	▲ 74	7.3	▲ 0.6
25～29歳	13908	13619	14297	13563	389	▲ 56	2.8	▲ 0.4
30～34歳	17266	16298	16907	16146	▲ 359	▲ 152	▲ 2.1	▲ 0.9
35～39歳	20822	18981	20119	18851	▲ 703	▲ 130	▲ 3.4	▲ 0.7
40～44歳	23077	21533	22261	21300	▲ 816	▲ 233	▲ 3.5	▲ 1.1
45～49歳	19661	18938	19074	18705	▲ 587	▲ 233	▲ 3.0	▲ 1.2
50～54歳	20000	20011	19526	19905	▲ 474	▲ 106	▲ 2.4	▲ 0.5
55～59歳	22534	22029	22221	21976	▲ 313	▲ 53	▲ 1.4	▲ 0.2
60～64歳	26385	25637	26057	25597	▲ 328	▲ 40	▲ 1.2	▲ 0.2
65～69歳	29000	29263	28933	29305	▲ 67	42	▲ 0.2	0.1
70～74歳	19465	23203	19588	23182	123	▲ 21	0.6	▲ 0.1
75～79歳	16348	22537	16612	22515	264	▲ 22	1.6	▲ 0.1
80～84歳	14450	23101	14784	23218	334	117	2.3	0.5
85～89歳	8484	18084	8595	18344	111	260	1.3	1.4
90～94歳	3122	9785	3079	9778	▲ 43	▲ 7	▲ 1.4	▲ 0.1
95～99歳	713	3020	613	3086	▲ 100	66	▲ 14.0	2.2
100歳以上	85	530	94	528	9	▲ 2	10.6	▲ 0.4

注 1) 集計には、国勢調査の大項目 koku2A.0001 を使用

注 2) 日本人の値を推計値より実測値が 3 % 以上乖離がある場合、背景色が桃色

「指導用 高校からの統計・データサイエンス活用 ～上級編～」でも同様に島根県の人口推計を行っているが、使用している国勢調査の大項目が異なることから若干結果が異なっている。

### 事例 3 : 世界の平均気温と緯度の関係 : 重回帰分析

課題解決の手順	実施内容例
<b>STEP1:問題</b> 課題の設定	◇世界の平均気温と緯度の関係を明らかにする。
<b>STEP2:計画</b> どのようなデータ・統計資料を集めて どのような分析をするか	◇利用するデータ : ・世界の気温に関する理科年表データ ◇分析内容 : 重回帰分析での変数間の相関関係理解 a) 平均気温と緯度の関係を単回帰分析で明らかにする。 b) 平均気温を説明する要因として標高を追加した分析を行う。
<b>STEP3:収集</b> 必要なデータ・統計資料を集める	緯度の再計算 緯度は 0 度と 90 度の間の値を取り、数字が多くなるほど寒くなることから、北緯 60 度 12 分(60°12'N)といった形式の値を、 「 $\cos(60 + (12 * 1/60)) * 2 * (\text{円周率}) / 360 * (\text{南緯(S)の場合} - 1)$ 」で再計算する。

<b>STEP4:分析</b> グラフや統計量で 傾向を捉える	a)平均気温と緯度の関係を単回帰分析で明らかにする。 単回帰分析でそもその傾向を調べる。 b)平均気温を説明する要因として標高を追加した分析を行う。 単回帰分析で相関が見られた場合は、更に対応する説明変数を追加する。
<b>STEP5:結論</b> 結論を導く	変数を増やし重回帰分析にすることでモデルの精度が上がる。単回帰分析で出た結果だけでなく、変数を追加して結果を比較することでより良いモデルが得られる場合がある。

### <STEP4 の詳細>

#### a)平均気温と緯度の関係を単回帰分析で明らかにする。

世界の平均気温と緯度の関係を探る。被説明変数を平均気温、説明変数を緯度とした。以下の実線囲い部分が分析結果である。平均気温と緯度は正の相関を持ち、t 値（t-Value）も高く、有意な結果が得られている。モデルの説明力を示す自由度調整済み決定係数（Adjusted R-squared）は 0.8 と高く平均気温と緯度の間には一定の相関があることが見受けられる。

```
Call:
lm(formula = temp ~ c.lon, data = dt1)

Residuals:
    Min       1Q   Median       3Q      Max
-37.772  -1.382   0.546   2.424   8.987

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  -28.571      1.560    -18.32  <2e-16 ***
c.lon         55.657      1.868     29.79  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.488 on 212 degrees of freedom
(26 observations deleted due to missingness)
Multiple R-squared:  0.8072,    Adjusted R-squared:  0.8063
F-statistic: 887.5 on 1 and 212 DF,  p-value: < 2.2e-16
```

#### b)平均気温を説明する要因として標高を追加した分析を行う。

平均気温を説明する変数として緯度と標高を追加する。以下実線の囲い部分が分析結果である。平均気温と、緯度は正の相関を持ち、標高（高度）は負の相関を持つ。t 値は高く、有意な結果が得られている。モデルの説明力を示す自由度調整済み決定係数は、先ほどの単回帰モデルよりも結果が改善しており a)の単回帰分析よりも望ましい結果となっている。



```

Call:
lm(formula = temp ~ c.lon + 高度, data = dt1)

Residuals:
    Min       1Q   Median       3Q      Max
-21.1342  -1.5395   0.0741   1.7064   7.6163

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -2.701e+01  1.079e+00  -25.02  <2e-16 ***
c.lon        5.587e+01  1.287e+00   43.40  <2e-16 ***
高度       -5.230e-03  3.407e-04  -15.35  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.092 on 211 degrees of freedom
(26 observations deleted due to missingness)
Multiple R-squared:  0.9089, Adjusted R-squared:  0.908
F-statistic: 1053 on 2 and 211 DF, p-value: < 2.2e-16

```