

統計データ分析コンペティション 2018

## 優秀賞（高校生の部）

### SSDSE データを活用した全国学習状況調査結果 との相関分析

宮本 雨月、金山 瑠依、門脇 俊樹

（和歌山県立田辺工業高等学校）

#### 審査委員長講評

小学生の学習への関心や取組について、家族構成や自治体の教育費等との関係を分析した結果、大家族世帯が学習状況に良い影響を与えると指摘しました。構成もしっかりとしています。

文部科学省が実施した「平成 21 年度全国学力・学習状況調査」の結果と SSDSE から得られた家族構成や自治体の教育費等のデータをうまく結合しており、相関行列をヒートマップの手法でビジュアル化して表現したところに工夫がみられました。

結論が、文部科学省などが報告していることに関わるので、論文の中に、先行研究、参考資料を参照し、自分たちの分析結果を比較することができれば、一層優れた論文となるでしょう。

# SSDSE データを活用した全国学習状況調査結果との相関分析

宮本雨月\*1・金山瑠依\*1・門脇俊樹\*1

\*1: 和歌山県立田辺工業高等学校

## 1. はじめに

毎年、子供達の学力状況を把握するため文部科学省により全国的な学力調査「全国学力・学習状況調査（全国学力テスト）」が実施されている。小学6年生と中学3年生を対象に実施され、結果として国語A・B、算数（数学）A・Bの正答率および学習状況調査の結果が都道府県別に公表されている。結果については、メディア等でも大きく取り上げられるため世間の関心が高くなっている。

学習に関する意欲や関心等に関しては本人の要因だけでなく、家庭環境や自治体の政策による影響も考えられるため SSDSE のデータの中で、家庭の状況や教育費等の支出に関するデータ項目と学習状況調査結果との関連性を検討、考察することとした。

本分析では、学力である平均正答率には着目せず、学習状況における学習への関心等のデータに着目して分析を行うこととし、特に、中学生よりは小学生の方が家庭の状況による影響が大きいと考え、小学生の都道府県別の学習状況調査結果の数値を基準として、SSDSE のデータ項目を基に加工した関連データとの相関関係を考察した。

## 2. 研究の方法

### 2.1 全国学力・学習状況調査データの活用

平成29年度全国学力・学習状況調査結果資料【小学校・都道府県別】の児童質問紙調査項目<sup>(1)</sup>(表1)における学習に対する関心・意欲・態度（国語への関心等・数学への関心等）、規範意識・自尊感情、学習の基盤となる活動習慣（言語活動・読解力・生活習慣・学習習慣）の質問紙結果を数値化したデータを使用した。

表1 学習状況調査児童質問紙調査使用項目

学習に対する関心・意欲・態度		規範意識・自尊感情		学習の基盤となる活動・習慣		
国語への関心等	数学への関心等	規範意識	自尊感情	言語活動・読解力	生活習慣	学習習慣

### 2.2 SSDSE データの活用

SSDSE のデータの中で、家庭の状況や教育費等の支出に関するデータ項目として以下のデータ項目（表2）を使用し、学習状況調査項目との相関関係を分析した。

表2 SSDSE データ使用項目

一般世帯数	核家族世帯数	単独世帯数	離婚件数	歳出決算総額	民生費	教育費
-------	--------	-------	------	--------	-----	-----

### 2.3 外れ値の検定

データを統計的に分析するために、統計解析ソフトであるエクセル統計（株式会社社会情報サービス）を使用し、関連データの分析を行った。相関関係を考えるために、まずは、外れ値の検定を行い、各項目の外れ値を除外した。

外れ値の検定方法としては、スミルノフ＝グラブズ検定を用い、1回の検定で1つの外れ値を除き、外れ値がなくなるまで繰り返した。また、有意水準の値は、0.05とした。

### 2.4 相関行列の作成

エクセル統計を使用し、外れ値を除外した各項目のデータから相関係数（ピアソンの積率相関係数）を算出し、各変数間の相関係数の行列を作成した。まずは、無相関の検定を行い、相関の有意性を検討し、相関行列を作成した。また、視覚的に判断できるよう相関係数値に応じてセルに色を付けた行列として相関係数のヒートマップも出力した。

相関係数の評価基準としては以下の表3の値を判断基準とした。

表3 相関係数の評価基準値

相関係数の値	相関係数の強弱
1.0 ~ 0.7	強い正の相関
0.7 ~ 0.4	正の相関
0.4 ~ 0.2	弱い正の相関
-0.2 ~ -0.4	弱い負の相関
-0.4 ~ -0.7	負の相関
-0.7 ~ -1.0	強い負の相関

### 3. データの加工

#### 3.1 SSDSE のデータの加工

2.2表2で示したデータ項目について、SSDSE より数値を抽出し、各市町村のデータを基に都道府県別に集計を行った。さらに、各データを使用して以下の(1)~(4)の項目にデータを加工し、再度都道府県別にデータを集計した。数値を整理し、各データ項目を都道府県別にグラフ化した結果を以下の図1~図4に示す。

- (1) 大家族世帯割合：一般世帯数に対する核家族世帯および単独世帯以外の世帯割合（図1）  
 (=一般世帯数-核家族世帯-単独世帯)
- (2) 教育費割合：各都道府県自治体の歳出決算総額に対する教育費の割合（図2）  
 (=教育費/歳出決算総額\*100)
- (3) 離婚件数割合：一般世帯数に対する離婚件数の割合（図3）  
 (=離婚件数/一般世帯数\*100)
- (4) 民生費割合：各都道府県自治体の歳出決算総額に対する民生費の割合（図4）  
 (=民生費/歳出決算総額\*100)

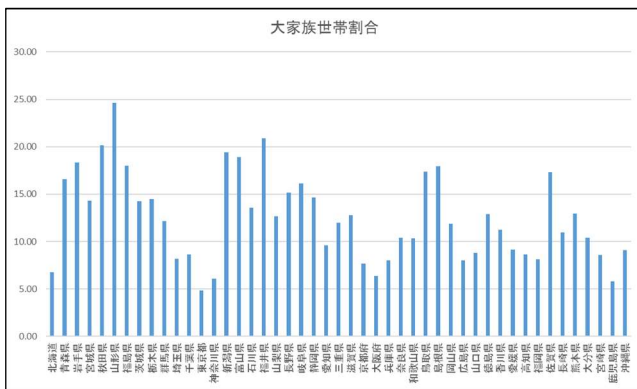


図1 大家族世帯割合

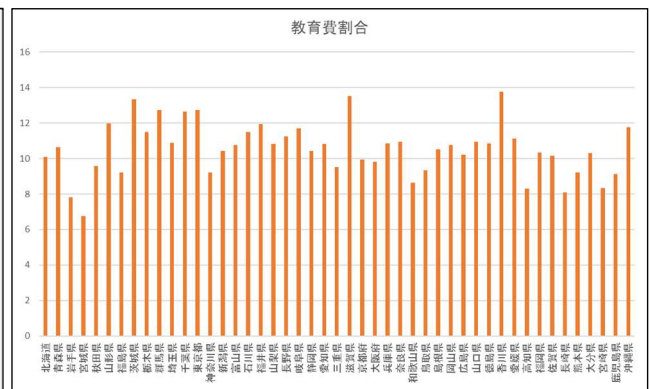


図2 教育費割合

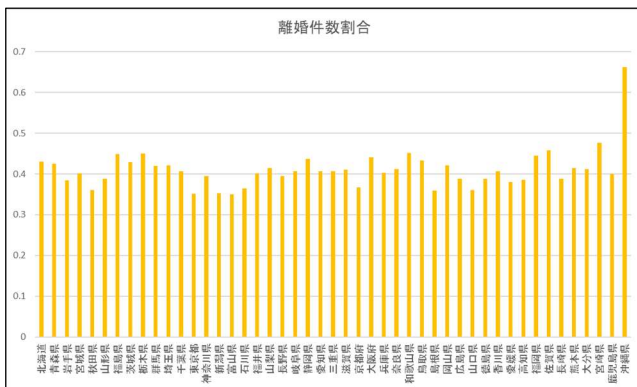


図3 離婚件数割合

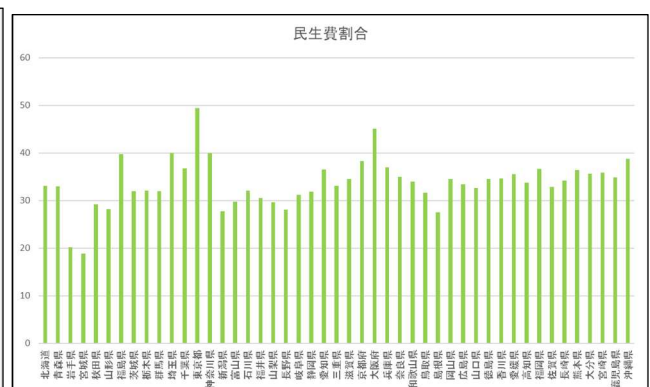


図4 民生費割合

#### 3.2 学習状況調査およびSSDSE 加工データの整理

相関係数を求めるために、平成29年度全国学力・学習状況調査結果資料【都道府県別】小学校<sup>1)</sup>より、児童質問紙調査項目における2.1表1で示したデータ項目についての質問結果が数値化されたデータを抽出した。さらに、3.1で得られた加工データとともに整理し、表4にまとめた。

表4 各項目の集計数値データ一覧【都道府県別】

データ項目	国語への関心等	数学への関心等	規範意識	自尊感情	言語活動・読解力	生活習慣	学習習慣	大家族世帯割合	教育費割合	離婚件数割合	民生費割合
1 北海道	5.0	4.9	5.0	4.8	5.1	4.9	5.3	6.78	10.08	0.43	33.06
2 青森県	5.3	5.3	5.4	5.4	5.4	5.5	5.7	16.60	10.64	0.43	33
3 岩手県	5.4	5.3	5.4	5.1	5.5	5.5	5.6	18.35	7.82	0.38	20.18
4 宮城県	5.0	5.0	4.9	5.0	5.1	5.1	5.5	14.33	6.76	0.40	18.87
5 秋田県	6.2	6.1	5.8	6.0	6.1	5.6	6.6	20.13	9.58	0.36	29.15
6 山形県	5.4	5.0	5.4	5.3	5.3	5.3	5.7	24.66	11.97	0.39	28.17
7 福島県	5.3	5.1	5.0	4.8	5.1	5.2	5.5	17.98	9.20	0.45	39.71
8 茨城県	5.3	5.2	5.3	5.2	5.2	5.1	5.4	14.23	13.33	0.43	32.01
9 栃木県	5.5	5.4	5.4	5.3	5.4	5.5	5.5	14.5	11.49	0.45	32.05
10 群馬県	5.0	5.1	5.2	5.2	5.0	5.3	5.2	12.15	12.73	0.42	31.93
11 埼玉県	5.1	5.1	5.2	5.3	5.3	5.5	5.5	8.20	10.89	0.42	39.94
12 千葉県	4.8	4.7	4.9	4.9	4.7	4.9	4.7	8.66	12.64	0.41	36.76
13 東京都	4.9	5.0	4.6	4.9	5.0	4.9	4.6	4.86	12.74	0.35	49.37
14 神奈川県	5.0	5.0	4.7	4.8	4.8	4.8	4.6	6.11	9.22	0.40	39.92
15 新潟県	5.5	5.3	5.3	5.2	5.6	5.2	5.2	19.44	10.44	0.35	27.73
16 富山県	4.9	5.0	5.1	5.1	5.0	5.0	5.3	18.91	10.75	0.35	29.7
17 石川県	5.5	5.5	5.0	5.1	5.3	5.3	5.5	13.6	11.49	0.37	32.05
18 福井県	5.3	5.4	5.3	5.7	5.2	5.5	5.5	20.91	11.95	0.40	30.52
19 山梨県	5.4	5.4	5.4	5.3	5.4	5.4	5.3	12.69	10.81	0.41	29.67
20 長野県	5.1	4.9	4.9	5.0	5.0	5.4	4.6	15.17	11.26	0.40	28.11
21 岐阜県	5.2	5.0	5.4	5.1	5.2	5.2	5.3	16.11	11.71	0.41	31.21
22 静岡県	4.6	4.9	4.8	5.0	4.8	4.9	4.8	14.64	10.44	0.44	31.85
23 愛知県	4.7	4.8	4.9	4.8	4.7	4.8	4.6	9.59	10.83	0.41	36.54
24 三重県	4.9	5.0	5.1	4.9	4.9	4.9	4.9	11.96	9.53	0.41	33.05
25 滋賀県	4.6	4.5	4.7	4.9	4.6	4.9	4.9	12.78	13.51	0.41	34.53
26 京都府	4.8	4.8	5.0	4.9	4.9	4.8	4.9	7.67	9.93	0.367	38.31
27 大阪府	4.6	4.8	4.6	4.7	4.5	4.6	4.3	6.38	9.81	0.441	45.02
28 兵庫県	4.8	4.8	5.0	5.0	4.8	4.9	4.7	8.04	10.86	0.40	36.92
29 奈良県	4.7	4.7	5.0	4.8	4.7	4.8	4.5	10.39	10.94	0.41	34.91
30 和歌山県	5.0	5.1	5.2	5.0	4.9	5.1	4.8	10.37	8.64	0.45	33.91
31 鳥取県	5.0	4.8	5.0	4.8	4.9	5.1	5.1	17.4	9.35	0.43	31.62
32 島根県	4.9	4.8	5.2	5.0	5.0	5.1	5.0	17.97	10.53	0.36	27.49
33 岡山県	5.1	5.1	5.1	5.1	5.0	5.0	5.2	11.89	10.75	0.42	34.48
34 広島県	5.2	5.2	5.3	5.4	5.2	5.2	5.3	8.00	10.23	0.39	33.43
35 山口県	5.2	5.1	5.3	5.1	5.3	5.1	5.2	8.81	10.95	0.36	32.64
36 徳島県	4.8	5.1	5.2	5.0	4.7	4.9	5.1	12.92	10.86	0.39	34.55
37 香川県	4.7	4.8	4.8	4.7	4.6	4.8	5.1	11.24	13.77	0.41	34.58
38 愛媛県	5.1	5.1	5.3	5.2	5.1	4.9	5.3	9.18	11.12	0.38	35.55
39 高知県	5.6	5.4	5.3	5.2	5.4	5.2	5.2	8.65	8.29	0.39	33.68
40 福岡県	4.8	4.8	4.8	4.9	4.7	4.6	4.6	8.13	10.33	0.45	36.61
41 佐賀県	4.9	5.0	5.1	5.0	4.8	5.0	4.7	17.31	10.15	0.46	32.8
42 長崎県	4.9	5.1	5.0	4.8	4.9	5.1	4.8	10.99	8.10	0.39	34.23
43 熊本県	5.0	4.9	4.9	4.8	5.0	4.9	4.8	12.96	9.22	0.42	36.41
44 大分県	5.0	4.9	4.8	4.9	5.0	4.7	4.7	10.41	10.32	0.41	35.67
45 宮崎県	5.3	5.4	5.3	5.3	5.3	5.2	5.5	8.58	8.33	0.48	35.88
46 鹿児島県	5.0	5.0	5.1	4.9	5.0	4.8	5.1	5.79	9.11	0.40	34.85
47 沖縄県	5.0	5.2	4.5	4.9	4.9	4.6	5.1	9.07	11.76	0.66	38.71

### 3.3 外れ値の除外

3.1 図1~4のグラフからもわかるように、大きく外れている都道府県の値が存在するため、エクセル統計を使用して外れ値の検定を行った。有意水準の値は、0.05とし、スミルノフ＝グラブズ検定を外れ値がなくなるまで繰り返した。それぞれの項目において以下の都道府県の値を外れ値として除外した。

- (1) 国語への関心等 秋田県(6.2)
- (2) 数学への関心等 秋田県(6.1)
- (3) 規範意識 外れ値なし
- (4) 自尊感情 秋田県(6.0) 福井県(5.7)
- (5) 言語活動・読解力 秋田県(6.1)
- (6) 生活習慣 外れ値なし
- (7) 学習習慣 秋田県(6.6)
- (8) 大家族世帯 外れ値なし
- (9) 教育費割合 外れ値なし
- (10) 離婚件数割合 沖縄県(0.66)
- (11) 民生費割合 外れ値なし

### 3.4 無相関の検定

エクセル統計を使用し、相関係数の数字に意味があるかについて無相関の検定を行った。有意水準の値は、0.05とした。

### 3.5 相関行列の作成

エクセル統計を使用し、各変数間の相関係数を算出し、相関行列を出力した。係数の計算式としてはピアソンの積率相関係数の計算式を使用した。また、視覚的にも判断できるように相関行列のヒートマップも同時に出力した。

## 4. 結果

### 4.1 無相関の検定結果

無相関の検定結果表を以下の表5に示す。5%で有意な場合 (P<0.05) は、\*マークが記載され、1%で有意な場合 (P<0.01) は、\*\*マークが記載される。有意水準の判断基準値としては、5%(P<0.05)として判断している。

表5 各項目の母相関係数の無相関の検定 [上三角:P値/下三角:\*, P<0.05 \*\*, P<0.01]

項目	国語への関心等	数学への関心等	規範意識	自尊感情	言語活動・読解力	生活習慣	学習習慣	大家族世帯割合	教育費割合	離婚件数割合	民生費割合
国語への関心等	-	P<0.001	P<0.001	P<0.001	P<0.001	P<0.001	P<0.001	0.0117	0.3574	0.5523	0.0117
数学への関心等	**	-	P<0.001	P<0.001	P<0.001	P<0.001	P<0.001	0.1673	0.2189	0.8698	0.1090
規範意識	**	**	-	P<0.001	P<0.001	P<0.001	P<0.001	P<0.001	0.4095	0.3875	P<0.001
自尊感情	**	**	**	-	P<0.001	P<0.001	P<0.001	0.0560	0.5903	0.7376	0.0170
言語活動・読解力	**	**	**	**	-	P<0.001	P<0.001	0.0144	0.2496	0.2048	0.0023
生活習慣	**	**	**	**	**	-	P<0.001	P<0.001	0.6930	0.5771	P<0.001
学習習慣	**	**	**	**	**	**	-	P<0.001	0.8974	0.7793	P<0.001
大家族世帯割合	*	**	**	*	*	**	**	-	0.7788	0.4414	P<0.001
教育費割合									-	0.4591	0.1016
離婚件数割合										-	0.2492
民生費割合	*	**	**	*	**	**	**	**	**		-

### 4.2 相関行列結果

出力した相関行列の結果を以下の表6に示す。また、視覚的に判断できるように相関係数の値に応じて正の相関は青色、負の相関は赤色の濃淡で出力されている相関行列のヒートマップ(表7)および相関計数値と濃淡の凡例(表8)を以下に示す。

表6 相関行列

項目	国語への関心等	数学への関心等	規範意識	自尊感情	言語活動・読解力	生活習慣	学習習慣	大家族世帯割合	教育費割合	離婚件数割合	民生費割合
国語への関心等	1.0000	0.8439	0.6813	0.6862	0.9241	0.7345	0.7276	0.3687	-0.1389	-0.0910	-0.3688
数学への関心等	0.8439	1.0000	0.5863	0.6792	0.7894	0.6402	0.6581	0.2071	-0.1848	-0.0251	-0.2394
規範意識	0.6813	0.5863	1.0000	0.7696	0.7271	0.7587	0.6804	0.4856	-0.1232	-0.1305	-0.5193
自尊感情	0.6862	0.6792	0.7696	1.0000	0.7567	0.7242	0.6815	0.2869	0.0824	-0.0520	-0.3541
言語活動・読解力	0.9241	0.7894	0.7271	0.7567	1.0000	0.7672	0.7664	0.3587	-0.1732	-0.1927	-0.4394
生活習慣	0.7345	0.6402	0.7587	0.7242	0.7672	1.0000	0.6978	0.5752	-0.0591	-0.0844	-0.5135
学習習慣	0.7276	0.6581	0.6804	0.6815	0.7664	0.6978	1.0000	0.4758	-0.0196	-0.0430	-0.4891
大家族世帯割合	0.3687	0.2071	0.4856	0.2869	0.3587	0.5752	0.4758	1.0000	0.0421	-0.1163	-0.6551
教育費割合	-0.1389	-0.1848	-0.1232	0.0824	-0.1732	-0.0591	-0.0196	0.0421	1.0000	-0.1119	0.2417
離婚件数割合	-0.0910	-0.0251	-0.1305	-0.0520	-0.1927	-0.0844	-0.0430	-0.1163	-0.1119	1.0000	0.1734
民生費割合	-0.3688	-0.2394	-0.5193	-0.3541	-0.4394	-0.5135	-0.4891	-0.6551	0.2417	0.1734	1.0000

表7 相関行列ヒートマップ

項目	国語への関心等	数学への関心等	規範意識	自尊感情	言語活動・読解力	生活習慣	学習習慣	大家族世帯割合	教育費割合	離婚件数割合	民生費割合
国語への関心等	1.0000	0.8439	0.6813	0.6862	0.9241	0.7345	0.7276	0.3687	-0.1389	-0.0910	-0.3688
数学への関心等	0.8439	1.0000	0.5863	0.6792	0.7894	0.6402	0.6581	0.2071	-0.1848	-0.0251	-0.2394
規範意識	0.6813	0.5863	1.0000	0.7696	0.7271	0.7587	0.6804	0.4856	-0.1232	-0.1305	-0.5193
自尊感情	0.6862	0.6792	0.7696	1.0000	0.7567	0.7242	0.6815	0.2869	0.0824	-0.0520	-0.3541
言語活動・読解力	0.9241	0.7894	0.7271	0.7567	1.0000	0.7672	0.7664	0.3587	-0.1732	-0.1927	-0.4394
生活習慣	0.7345	0.6402	0.7587	0.7242	0.7672	1.0000	0.6978	0.5752	-0.0591	-0.0844	-0.5135
学習習慣	0.7276	0.6581	0.6804	0.6815	0.7664	0.6978	1.0000	0.4758	-0.0196	-0.0430	-0.4891
大家族世帯割合	0.3687	0.2071	0.4856	0.2869	0.3587	0.5752	0.4758	1.0000	0.0421	-0.1163	-0.6551
教育費割合	-0.1389	-0.1848	-0.1232	0.0824	-0.1732	-0.0591	-0.0196	0.0421	1.0000	-0.1119	0.2417
離婚件数割合	-0.0910	-0.0251	-0.1305	-0.0520	-0.1927	-0.0844	-0.0430	-0.1163	-0.1119	1.0000	0.1734
民生費割合	-0.3688	-0.2394	-0.5193	-0.3541	-0.4394	-0.5135	-0.4891	-0.6551	0.2417	0.1734	1.0000

表8 ヒートマップ凡例 (青色: 正の相関 赤色: 負の相関)

1.00	0.75~1.00	0.50~0.75	0.25~0.50	0.00~0.25	-1.00~-0.75	-0.75~-0.50	-0.50~-0.25	-0.25~0.00
------	-----------	-----------	-----------	-----------	-------------	-------------	-------------	------------

## 5. 考察

表6の相関行列の値より、大家族世帯割合と学習状況調査の児童質問紙調査項目における規範意識、言語活動・読解力、生活習慣、学習習慣の4項目について正の相関が出力されている。また、この4項目と国語および数学への関心等が正の相関となっている。このことより、大家族世帯として、日常生活を通じて両親だけでなく祖父母等との関わりが、規範意識、言語活動・読解力、生活習慣、学習習慣の高まりを通じて学習への関心を育むのではないかと考察した。学習への関心が高いと、意欲的に学習に取り組み、今後学力が向上する可能性が高いのではないかと考える。

また、各都道府県自治体の歳出費総額に対する教育費の割合は、他のすべてのデータ項目とも相関は有意とはならず、相関がないと判断した。教育費の割合も学習への関心等が高い県で割合が高いというわけではなく、教育費への支出を高めても学力等に及ぼす影響は限定的なのではないかと考察した。

さらに、各都道府県自治体の歳出費総額に対する民生費の割合と、規範意識、自尊感情、言語活動・読解力、生活習慣、学習習慣とは負の相関となっている。民生費の割合が高いということは、福祉や児童手当等の社会福祉に関する支出が多いと考えられるので、公的な支援が必要な家庭が多い自治体では、規範意識、自尊感情、言語活動・読解力、生活習慣、学習習慣が確立されにくく、学習への関心も薄くなってしまっているのではないかと考察した。

以上の考察より、各都道府県において、大家族世帯で生活がしやすい状況を作っていくことが、子供の学習状況に良い影響を及ぼし、学習への関心の高まりを通じて学力向上にもつながって行くのではないかと示唆が得られた。

## 6. 参考文献等

### 6.1 参考文献

- 加藤久和、高校生からの統計入門、p. 118-208、筑摩書房、2016年
- 中室牧子、「学力」の経済学、p. 85-98 p. 114-126、ディスカヴァー・トゥエンティワン、2015年
- 寺内正典(編集)、英語教育学の実証的研究法 Excel で学ぶ統計処理、p. 130-p. 142、研究社、2012年

### 6.2 統計分析使用ソフト

- エクセル統計2015、株式会社社会情報サービス

### 6.3 引用資料

- (1) 文部科学省国立教育政策研究所、平成29年度全国学力・学習状況調査結果資料 【都道府県別】 小学校