

2020年11月28日

食と生活についての相関

コンピュータグラフィックス中間レポート

1. テーマ

都道府県別に食料の品目別消費量と、平均寿命や美肌ランキングなどをグラフに可視化することで面白い相関が得られるのではないかと考え、今回のテーマとして「食と生活についての相関」を選んだ。

「食と生活」の相関がメインテーマではあるが、グラフを作成して相関関係を見ていったところ「食と食」、「生活と生活」の相関も興味深かったため、そちらについても言及していくこととする。

注： 文章中のグラフは上下反転している。

2. 統計データ一覧

2.1.1. 食のデータ

今回は生活に大きく関わりそうな食料という観点から、主食・健康に良さそうな食料・健康に悪そうな食料の3種類の食材を選出した。

以下がその一覧である。

- ・米
- ・食パン
- ・豚肉
- ・カツオ
- ・カップラーメン
- ・チョコレート
- ・牛乳
- ・ビール
- ・発泡酒

(データ引用：総務省統計局「家計調査」)

2.1.2. 生活のデータ

生活では、食生活との関わりが深そうなものと都道府県別で差が大きそうなものを選出した。

以下がその一覧である。

- ・ 60歳以上有業率

(データ引用：総務省統計局「都道府県別60歳以上の有業率」)

- ・ 100歳以上人口

(データ引用：厚生労働省「百歳以上高齢者数について」)

- ・ 平均寿命

(データ引用：厚生労働省「都道府県別生命表の概況」)

- ・ 未婚率

(データ引用：総務省統計局「国勢調査人口等基本集計」)

- ・ 美肌度合い

(データ引用：株式会社ポーラ「美肌県グランプリ総集編」)

- ・ 通院率

(データ引用：厚生労働省「国民生活基礎調査の概況」)

- ・ 学力テストの結果

(データ引用：文部科学省「全国的な学力調査(全国学力・学習状況調査等)」)

- ・ 高齢化率

(データ引用：総務省「人口推計」)

3. 結果に関する見当

「食と生活」に関しては、基本的に「健康に良い食材」だと考えて選んだ米、豚肉、カツオ、牛乳は平均寿命や100歳以上の人口が高い地域で多く消費されているのではないかという見当をつけた。逆に、そのような地域では「健康に悪い食材」だと考えて選んだカップラーメン、チョコレート、ビール、発泡酒は消費量が少ないと考えられる。

ただし、上記のような相関が得られたとしてもその食材が本当に「健康に良い」から消費量が高くなったとは言い切れない。逆も然りである。なぜなら、高齢化が進んでいる地域では「健康かどうか」は関係なく単にカップラーメンやチョコレートを好まない高齢者が多いのではないだろうか。さらに、若い世代が多い地域ではカップラーメンやチョコレートの消費量が高くなるのは必然だと言える。このように年代による嗜好の偏りから消費量が変わってくることを考えると、都道府県人口のどのくらいの割合を高齢者が占めているのかが重要になってくると感じた。そのため、統計データに高齢者率の列を追加し、分析していくこととした。

「食と食」の相関については、個人的な意見ではあるがパンと牛乳は正の相関があると考えている。また、米とカツオも正の相関がありそうであるし、カップラーメンの消費量が多い地域ではチョコレートの消費量も多いのではないかと見当をつけている。

「生活と生活」の相関関係に関しては、平均寿命と100歳以上人口に関しては相関関係がありそうであるが、その2つと60歳以上有業率は負の相関があると考えている。これは60歳を超えても元気に働くお年寄りがいることで平均寿命が上がりそうな気もするが、実際には老人アパートなどでゆったりと生きるお年寄りの方が長生きをしているのではないか、という予測である。

他の美肌偏差値や学力テストの相関に関してはあまり想像がつかないため、結果を見てから考察をしたいと思う。

4. Csv データの作成

今回は Google SpreadSheet を使用し、集めてきたデータを1つの csv ファイルにまとめた。データは一番上から北海道、青森県・・・、と北から順番に並んでいる。

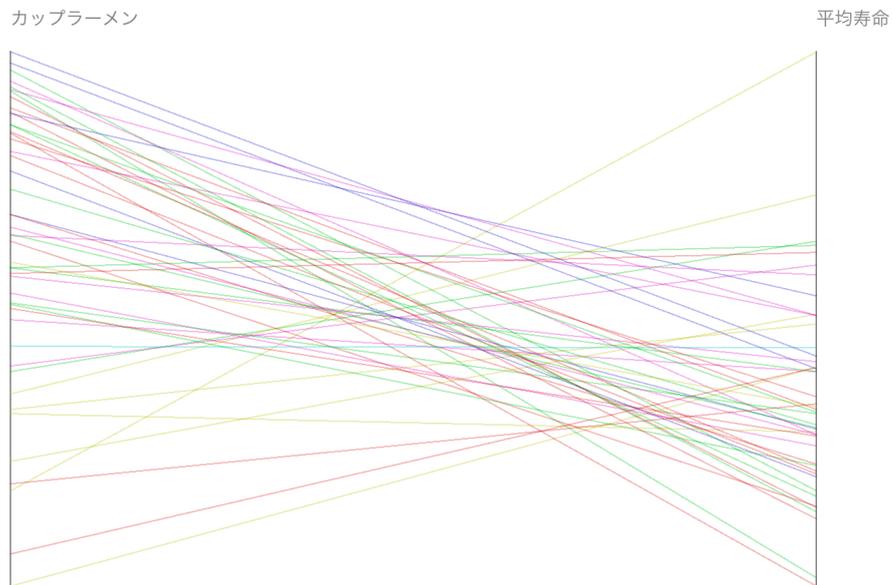
平均寿命は男女別のデータだったため、平均を取った値を採用している。同様に学力は国語A、国語B、数学A、数学B、理科の5科目だったため、こちらも5科目の平均を取って値を算出している。

データを数値で見たときにも相関があるかをなんとなく見たかったため、数値は0~100 程度になるように調整してある。そのため、単位は品目によってバラバラである。

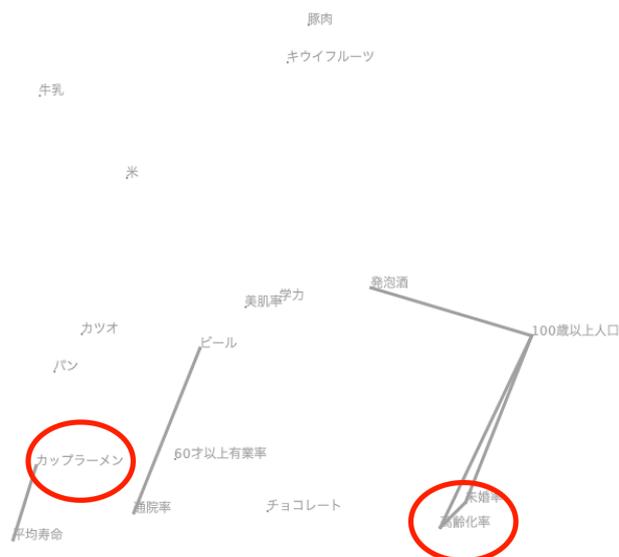
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	
1	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric										
2	米	パン	カップラーメン	キウイフルーツ	豚肉	カツオ	チョコレート	牛乳	ビール	発泡酒	60才以上有業率	100歳以上人口	平均寿命	未婚率	美乳率	通関率	学力	高齢化率	
3	87.72	79.825	80	63.3	92.76	26.6	69.55	53.425	27.3	39.5	27.9	64.2	83.525	26.2	47.04	39.76	73.25	78.25	
4	77.87	71.125	87.16	70.8	91.88	32.2	46.25	54.875	25.1	48.5	30.8	46.4	82.3	24.5	51.49	37.27	78	81.5	
5	74.84	73.875	76.6	101.5	86.70	82.2	55.6	73.05	25.8	35	30.9	62.9	83.15	24	43.27	42.14	76.5	81.25	
6	61.70	71.275	69.76	74.4	89.46	97.7	47.725	83.2	27.3	36.5	28.9	46.4	84.075	26.9	42.56	38.69	73	69.5	
7	73.59	62.325	78.68	80.2	97.95	45.5	56.525	60.4875	26	39	27.6	64.5	82.945	21.7	68.33	40.52	82	91	
8	85.75	72.9	72.12	49.7	78.84	64.7	66.2	54.7375	25.1	27	31.1	59.7	83.74	22.4	53.28	40.04	75.25	82.25	
9	77.65	58.275	76.16	67.5	93.44	99.3	57.175	72.825	25.5	33	29.2	58.9	83.26	24.2	44.28	39.4	75.25	77.25	
10	60.20	72.8	64.6	51.1	90.42	89	63.85	81.4125	20	24.5	30.8	49.2	83.305	26.4	38.98	37	76.75	72.25	
11	70.47	79.525	67.48	65.7	86.49	53.4	55.25	65.0375	21	24	32.3	46.9	83.17	26.1	43.53	38.68	75.25	70	
12	75.21	67.175	95.6	80.8	84.17	38	55.35	58.8375	19.7	28	32.5	61.1	83.725	25.7	19.57	35.61	74.5	73.5	
13	71.64	75.875	67.48	90	99.36	44.7	66.125	67.35	18.2	26	31.6	32.9	83.74	29.1	41.38	36.65	74	66	
14	65.47	78.45	70.64	76.7	80.37	44.1	83.475	67.2	19.4	26	31.6	39.3	83.935	28	48.95	37.15	74.25	68.75	
15	59.32	93.025	70.76	83.8	94.63	54.9	67.725	67.9875	43	32.5	34.1	43.5	84.165	34.8	56.56	37.92	78.75	57.75	
16	69.78	90.875	50	99.3	103.52	38	60.9	73.1625	22.2	29	30.8	42.3	84.28	30.3	49.81	37.23	75	62.75	
17	85.30	81.175	62.76	78.3	98.21	50.7	66.2	75.69	28.7	39.5	30.9	76.5	84.005	24.5	50.66	38.66	76.5	79.75	
18	79.18	82.2	80.4	98.1	90.57	20.6	64.2	68.325	27.7	40.5	32.8	71.9	84.015	22.8	63.97	37.94	79.25	80	
19	74.15	86.525	62.8	61.9	88.20	32.1	68.125	61.425	26.6	36	32.1	69.6	84.16	24.9	65	36.93	82.5	73	
20	69.37	78.075	53.88	67.4	81.02	31.6	64.325	60.2375	26.2	49	34.1	66	84.405	22.8	46.77	34.45	79.75	75.5	
21	76.28	76.55	71.04	53.7	96.38	22.3	58.475	66.6	21.5	25	36.0	81.4	84.035	25.9	58.88	36.02	74.25	75.75	
22	65.72	69.25	55.56	57.1	88.04	31.2	59.45	71.9	24.3	27.5	36.3	83.2	84.71	23.9	44.77	39.52	75	78.75	
23	71.27	80.725	53.32	61.5	77.67	39.9	60.7	55.6125	20.3	30	33.3	55.7	83.91	24	42.78	38.32	75.75	74.5	
24	83.43	83.175	51.04	51.6	96.18	70.1	63.9	62.35	22.2	28	34.4	54.2	84.025	25.2	44.87	36.76	74.75	73.75	
25	74.69	83.2	60.56	76.2	82.30	27.7	78.7	66.775	25.2	33.5	32.5	36.6	83.98	27.9	47.33	37.64	72.25	62.25	
26	75.72	77.425	54.84	118.7	82.15	67.1	62.2	83.9	21.3	31	32.4	57.1	83.925	24.1	52.5	37.53	74	73.5	
27	75.58	93	51.8	71.6	91.87	56.2	83.45	64.0375	17.9	28.5	31.5	52	84.675	26.3	44.73	34.96	72	64.25	
28	74.78	94.1	51.44	88.9	87.28	61.2	61.9	82	28.5	41.5	30.7	65.4	84.375	29.7	62.89	38.93	77.25	72.25	
29	75.21	89.25	84.64	60.9	81.99	43.7	66.3	66.875	30.4	46.5	28.8	40.1	83.48	29.5	53.74	39.31	72.75	68.75	
30	64.65	106.25	64.48	103.3	90.55	41.5	67.1	83.575	23.8	41.5	27.4	51.6	83.995	26.6	45.55	37.41	74.5	72	
31	80.82	100.175	54.84	96.1	92.74	50.9	77	73.375	17.6	31	26.6	56.6	84.305	25.7	59.86	38.27	73.5	77.25	
32	74.61	105.35	67.96	54.6	97.19	73.7	49.2	79.5	22.4	51	33.4	76.8	83.205	23.4	59.43	40.13	75.25	81.75	
33	70.34	97.5	92.72	62.5	78.15	44.7	54.775	85.5	23.9	42	30.9	97.9	83.72	23.9	42.09	39.26	74.5	79	
34	72.29	100.5	58.88	94.9	87.98	27.1	55.025	97.0625	23.4	39	33.0	101	84.215	22.2	78.64	38.77	72.75	85	
35	61.17	92.325	65.12	61.3	79.03	49.4	62.275	73.2	18.8	34	30.3	73.3	84.35	25.2	58.29	37.79	74.25	75.25	
36	70.45	96.25	57.52	41.4	89.22	42.4	70.45	75.325	24.8	48	30.5	72.9	84.205	25.2	55.63	39.02	78.75	72.5	
37	70.27	79.55	68.2	35.3	66.76	150	405.4	56.0875	21.8	46	30.1	83.7	83.695	23.2	45.36	39.02	76.75	84.75	
38	70.58	76.4	57.16	52.6	75.46	86.9	61.825	73.7125	21.9	40.5	29.7	74.3	83.49	23.4	57.04	36.69	73.5	82.75	
39	69.33	97.15	63.92	55.7	80.91	76.5	64.175	78.3375	23.3	33	31.8	84.6	84.03	23	43.81	39.75	76	78.75	
40	71.05	80.375	51.76	48	86.09	75	54.625	60.875	23.9	39	29.0	83.6	83.49	23.9	59.94	38.54	76.75	81.5	
41	67.48	98.225	74.4	54.1	77.45	330.1	60.575	45.2	29.4	84.5	30.8	96.5	83.635	25.2	54.11	39.05	75.75	87	
42	70.82	72.6	67	69.5	84.68	18.7	61.75	57.4875	24.3	45	28.1	60.1	83.9	28.6	50.54	36.51	75.25	69	
43	84.05	82.05	86.52	39.1	84.39	22.7	63.55	62.8875	21.4	45	32.6	78.9	83.885	25	48.5	36.58	75	74.25	
44	82.12	87.4	48.44	96.6	83.66	38.7	52.2	67.8375	22.4	45	28.1	79.6	83.675	24.7	47.59	36.99	73.75	80	

5.1.1. 食と生活の相関

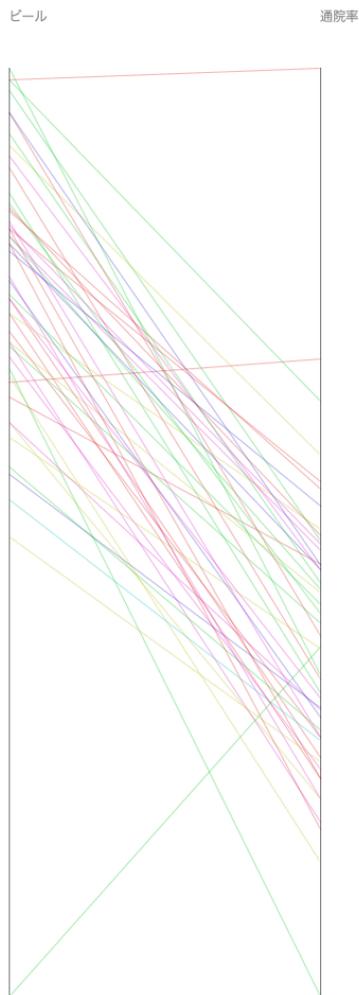
まず最初に、平均寿命とカップラーメンの相関について注目した。平均寿命が高い都道府県ほどカップラーメンの消費量が少ないと言える。



散布図を見てみると、高齢化率とカップラーメンの消費量との相関関係はあまり見られないため、この情報量ではカップラーメンは「健康に悪い食材」かもしれないと言える。



次に、ビールの消費量と通院率の相関関係について調べてみる。



ビールの消費量と通院率はあまり関係がないと思っていたが、実際は正の相関が見られた。つまり、ビールを多く消費する地域では病院に通院している人が多いということである。

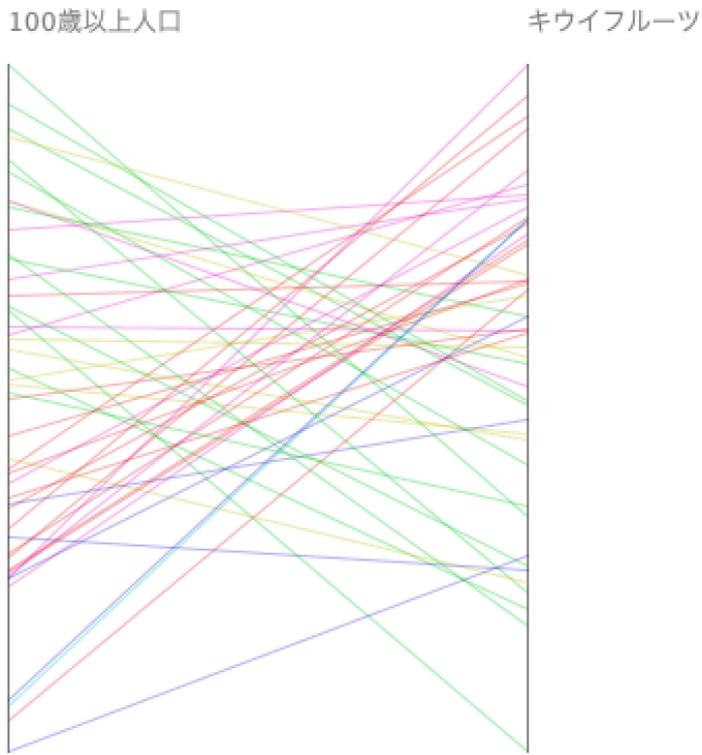
この理由に関しては何も見当がつかず、ネットで調べたところ

”傷病別にみると、男では「高血圧症」での通院者率が最も高く、次いで「糖尿病」、「歯の病気」、女では「高血圧症」が最も高く、次いで「眼の病気」、「歯の病気」となっている”

(引用：厚生労働省「世帯員の健康状況」より)
との記述を見つけた。

過度な飲酒は高血圧を引き起こす上に、ビールは糖質を含んでいる。そのため、ビールの飲み過ぎで高血圧症もしくは糖尿病を引き起こしてしまうことが通院率を上げる原因になっているのかもしれない。

3つ目は、100歳以上人口とキウイフルーツの相関関係について見てみる。



緑とピンクの線に注目して見れば、少し負の相関が見られると言える。つまり、100歳以上人口が少ない地域ではキウイフルーツの消費量は多く、100歳以上人口が多い地域ではキウイフルーツの消費量は少ないのである。

初め、キウイフルーツは”健康に良い食べ物”であると仮定して選んだが、お年寄りにはあまり食べないのかもしれない。ただ、後述するように100歳以上人口が高い

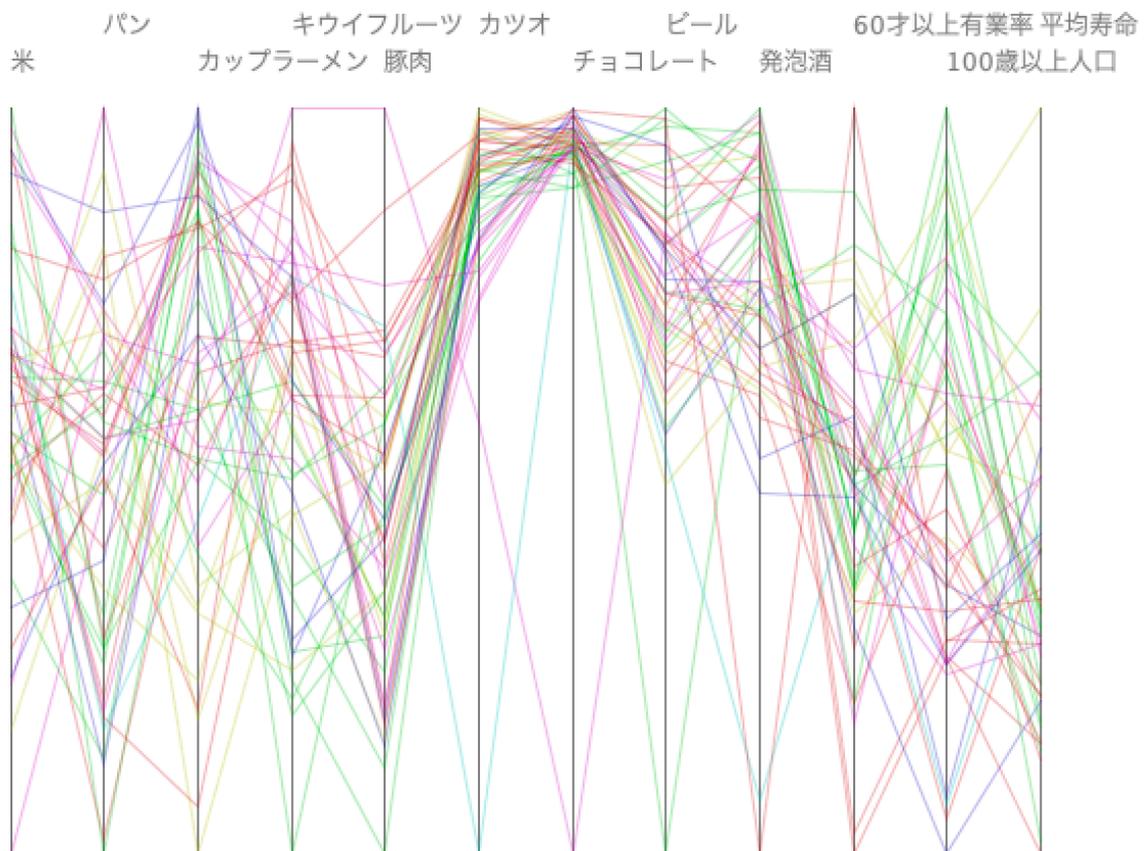
から高齢化している地域とは言えないため、このグラフからはあまり情報を得られなさそうであるという結論に至った。

5.1.2. 食と食の相関

食と食の相関について、何かしらの相関が見られるのではないかと期待していたが、実際はめぼしい相関が得られなかった。結果は以下のようにになっている。

残念ながらカツオとチョコレートの消費量については一部の地域が大量に消費していたため、グラフが見づらくなってしまった。

ちなみに、チョコレートの消費量が異常に多いのは山口県で、2位の5倍ほど差をつけてのトップであった。



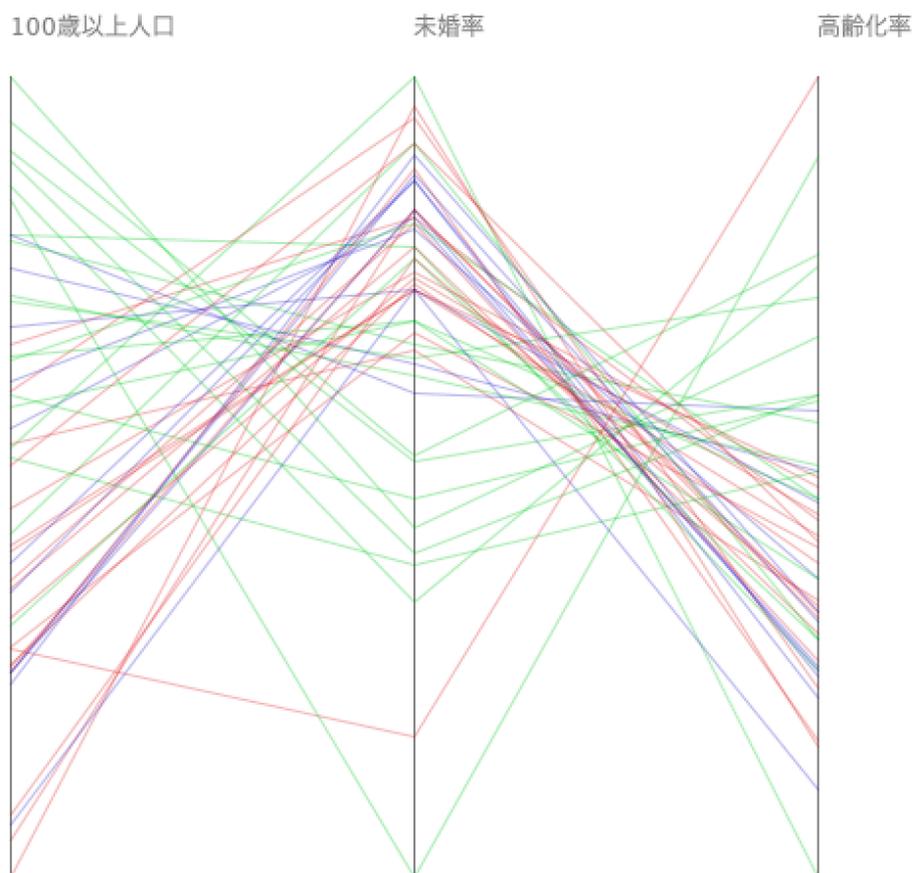
5.1.3. 生活と生活の相関

ここからは生活と生活の相関について見ていく。

下記の図は100歳以上人口と未婚率、高齢化率に関する折れ線グラフである。

このグラフから、100歳以上人口と未婚率、未婚率と高齢化率はそれぞれ負の相関を持っていると言える。つまり、100歳以上人口が高いもしくは高齢化率が高い地域は未婚率が低いのである。

これは、地方などの地域密着型社会を思い浮かべれば未婚率が低そうなのは容易に想像ができる。逆に、都心では働いている独身の人が少なくないと思うのが妥当だろう。



また、このグラフからはもう一つの興味深い事実が読み取れる。それは、100歳以上人口と高齢化率が全く同じではないことである。

確かに100歳以上人口と高齢化率の2軸には正の相関が見られ、多くの線はほぼ同じ高さの位置につながっている。しかし、100歳以上人口と未婚率が共に高い1本の赤い線に注目してみると、他の都道府県とは違い高齢化率が意外にも低いのがわかる。この地域はおそらく、都心に近く発展しているが、老人ホームなどに住む100歳以上の人口が多い地域なのではないだろうか。

また、少し見づらくはあるが100歳以上人口と未婚率が共に低い緑の1本の線は高齢化率が高くなっている。この地域は高齢化が進んでいて未婚率が低い、100歳以上の人口が少ない土地のようだ。

ここで、高齢化と平均寿命の折れ線グラフを見てみる。



先述した地域は高齢化率が高かったため、このグラフで言うと平均寿命が一番低いところから出ている黄色い線の地域ではないかと思われる。

つまり、この地域は高齢化が進んでいて未婚率は低い、平均寿命が短い地域のようなのである。そのため、100歳以上の人口が少なかったのではないだろうか。

6. 感想

今回、「食と生活についての相関関係」をテーマに csv データから折れ線グラフ・散布図に可視化をして見てきた。見当通りのものもあれば、意外な結果も得ることができて非常に興味深かった。

データの情報量が少なかったため判断しきれなかったり、一部外れ値があるせいで求めたかった情報が得られなかったりと、情報を扱うことの難しさも学ぶことができたと思う。少ない情報で結論を出そうとすると事実とは違う結果を導いてしまうかもしれないと恐ろしくなった。このことから世間でよく聞く「〇〇は健康に良い!」「〇〇を食べると脳を衰退化させる!」などの情報が事実であるかどうかは自分で見極めなければならないな、と痛感した。

7. 引用

食のデータ

「都道府県別統計とランキングで見る県民性」

<https://todo-ran.com/t/kiji/11798>

総務省統計局「家計調査」<http://www.stat.go.jp/data/kakei/index.html>

生活のデータ

総務省統計局「都道府県別60歳以上の有業率」

<http://www.stat.go.jp/data/shugyou/topics/topi740.html>

厚生労働省「百歳以上高齢者数について」

https://www.mhlw.go.jp/stf/houdou/0000177628_00001.html

厚生労働省「都道府県別生命表の概況」

<https://www.mhlw.go.jp/toukei/saikin/hw/life/tdfk15/index.html>

総務省統計局「国勢調査人口等基本集計」

<https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00200521&tstat=000001039448&cycle=0&tclass1=000001045009&tclass2=000001046265>

株式会社ポーラ「美肌県グランプリ総集編」

<https://www.pola.co.jp/special/bihadaken/>

厚生労働省「国民生活基礎調査の概況」

<https://www.mhlw.go.jp/toukei/saikin/hw/k-tyosa/k-tyosa13/index.html>

文部科学省「全国的な学力調査（全国学力・学習状況調査等）」

https://www.mext.go.jp/a_menu/shotou/gakuryoku-chousa/index.html

内閣府「高齢化の状況」

https://www8.cao.go.jp/kourei/whitepaper/w-2019/html/zenbun/sl1_14.html