

未婚率について

2024-07-21

目次

未婚率について	1
大まかな未婚率のデータについて	1

未婚率について

未婚率は出生率の要因の一つと考えられており、出生率を上げるためには未婚率を下げる事が不可欠である。

未婚率は年々上昇している。未婚率の上昇要因は複雑でさまざまな要因が考えられているが、その1つに女性の社会的地位の向上が挙げられる。

男性と女性は1対1で結婚するため、男性と女性の地域的な未婚率には大きな相関関係が見られそうである。しかし、男性と女性の相関係数はそこまで高くなく、また地域的に一方は高く一方は低い地域も見られた。

男女の未婚率の要因は違うことが考えられ、男女差と地域差を表すことがモチベーションである。またその地域差男女差が何に由来するものかまで知るところまでは行きたい。

大まかな未婚率のデータについて

```
import numpy as np
import pandas as pd

data = pd.read_csv("./data/data1.csv")

data.head()
```

調査年	地域	項目	A1601031_未婚人	A1601032_未婚人	A1601041_未婚人	A1601042_未婚人	A1601051_未婚人	A1601052_未婚人	注釈.5						
			口(25～29歳)(男) 【人】	口(25～29歳)(女) 【人】	口(30～34歳)(男) 【人】	口(30～34歳)(女) 【人】	口(35～39歳)(男) 【人】	口(35～39歳)(女) 【人】							
0	2020年度	北海道	NaN	71,897	NaN	63,280	NaN	54,055	NaN	43,141	NaN	46,785	NaN	36,050	NaN
1	2020年度	青森県	NaN	16,969	NaN	13,000	NaN	13,881	NaN	9,520	NaN	12,898	NaN	8,441	NaN
2	2020年度	岩手県	NaN	16,711	NaN	13,104	NaN	13,613	NaN	8,973	NaN	12,322	NaN	7,654	NaN
3	2020年度	宮城県	NaN	35,506	NaN	30,909	NaN	27,706	NaN	20,818	NaN	23,671	NaN	16,862	NaN
4	2020年度	秋田県	NaN	11,919	NaN	9,050	NaN	10,186	NaN	6,870	NaN	9,711	NaN	5,982	NaN

```
data = data.drop('調査年', axis=1)
data = data.drop('/項目', axis=1)
data = data.drop('注釈', axis=1)
data = data.drop('注釈.1', axis=1)
data = data.drop('注釈.2', axis=1)
data = data.drop('注釈.3', axis=1)
data = data.drop('注釈.4', axis=1)
data = data.drop('注釈.5', axis=1)
data.head()
```

	地域	A1601031_ 未婚人口 (25～29 歳) (男) 【人】	A1601032_ 未婚人口 (25～29 歳) (女) 【人】	A1601041_ 未婚人口 (30～34 歳) (男) 【人】	A1601042_ 未婚人口 (30～34 歳) (女) 【人】	A1601051_ 未婚人口 (35～39 歳) (男) 【人】	A1601052_ 未婚人口 (35～39 歳) (女) 【人】
0	北海道	71,897	63,280	54,055	43,141	46,785	36,050
1	青森県	16,969	13,000	13,881	9,520	12,898	8,441
2	岩手県	16,711	13,104	13,613	8,973	12,322	7,654
3	宮城県	35,506	30,909	27,706	20,818	23,671	16,862
4	秋田県	11,919	9,050	10,186	6,870	9,711	5,982

```
data = data.rename(columns={
    'A1601031_未婚人口 (25～29 歳) (男) 【人】': '25_29_man',
    'A1601032_未婚人口 (25～29 歳) (女) 【人】': '25_29_woman',
    'A1601041_未婚人口 (30～34 歳) (男) 【人】': '30_34_man',
    'A1601042_未婚人口 (30～34 歳) (女) 【人】': '30_34_woman',
    'A1601051_未婚人口 (35～39 歳) (男) 【人】': '35_39_man',
    'A1601052_未婚人口 (35～39 歳) (女) 【人】': '35_39_woman'
})

data.head()
```

	地域	25_29_mar	25_29_woman	30_34_mar	30_34_woman	35_39_mar	35_39_woman
0	北海道	71,897	63,280	54,055	43,141	46,785	36,050
1	青森県	16,969	13,000	13,881	9,520	12,898	8,441
2	岩手県	16,711	13,104	13,613	8,973	12,322	7,654
3	宮城県	35,506	30,909	27,706	20,818	23,671	16,862
4	秋田県	11,919	9,050	10,186	6,870	9,711	5,982

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 47 entries, 0 to 46
Data columns (total 7 columns):
#   Column      Non-Null Count  Dtype
---  -
0   地域        47 non-null    object
1   25_29_man   47 non-null    object
```

```

2  25_29_woman  47 non-null    object
3  30_34_man    47 non-null    object
4  30_34_woman  47 non-null    object
5  35_39_man    47 non-null    object
6  35_39_woman  47 non-null    object
dtypes: object(7)
memory usage: 2.7+ KB

```

```

data['25_29_man'] = data['25_29_man'].str.replace(',', '').astype(int)
data['25_29_woman'] = data['25_29_woman'].str.replace(',', '').astype(int)
data['30_34_man'] = data['30_34_man'].str.replace(',', '').astype(int)
data['30_34_woman'] = data['30_34_woman'].str.replace(',', '').astype(int)
data['35_39_man'] = data['35_39_man'].str.replace(',', '').astype(int)
data['35_39_woman'] = data['35_39_woman'].str.replace(',', '').astype(int)

```

```

data["25_34_39_man"] = data["25_29_man"] + data["30_34_man"] + data["35_39_man"]
data["25_34_39_woman"] = data["25_29_woman"] + data["30_34_woman"] + data["35_39_woman"]
data.head()

```

	地域	25_29_man	25_29_woman	30_34_man	30_34_woman	35_39_man	35_39_woman	25_34_39_man	25_34_39_woman
0	北海道	71897	63280	54055	43141	46785	36050	172737	142471
1	青森県	16969	13000	13881	9520	12898	8441	43748	30961
2	岩手県	16711	13104	13613	8973	12322	7654	42646	29731
3	宮城県	35506	30909	27706	20818	23671	16862	86883	68589
4	秋田県	11919	9050	10186	6870	9711	5982	31816	21902

```
data.info()
```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 47 entries, 0 to 46
Data columns (total 9 columns):
#   Column                Non-Null Count  Dtype
---  -
0   地域                   47 non-null    object
1   25_29_man              47 non-null    int64
2   25_29_woman            47 non-null    int64
3   30_34_man              47 non-null    int64
4   30_34_woman            47 non-null    int64
5   35_39_man              47 non-null    int64
6   35_39_woman            47 non-null    int64
7   25_34_39_man           47 non-null    int64

```

```
8    25_34_39_woman    47 non-null    int64
dtypes: int64(8), object(1)
memory usage: 3.4+ KB
```