Exploratory Data Analysis

- StudentID: 21900338

- Name: Yunjin Bae

- 1st Major: Economics

- 2nd Major: AI convergence

Here's a concise summary of my project: I used General Social Survey (GSS) data to analyze two main aspects. First, I investigated how the survey year relates to the average age at which people marry, separately for men and women. I calculated yearly averages and standard deviations and visualized the relationship with scatter plots.

Second, I explored the link between individuals' highest education level and their age at marriage. I calculated the average age at marriage based on education level and visualized it using scatter plots.

1. Data overview

- sample size: 59600 11232

- number of variables: 11232

- data type: object

2. Univariate analysis

2.1 Survey Yearwise first Marriage Age Statistics

This is a statistical analysis of the average first marriage age by survey year. The analysis is conducted separately for women and men, and it includes the calculation of mean, median, and standard deviation for each group

	Year	Mean	Max	Min	Median	StdDev
1	1974	20.73810	99	13	20.0	4.906992
2	1975	21.16963	99	12	20.0	7.183978
3	1976	20.93976	99	13	20.0	5.869480
4	1977	21.26486	99	13	20.0	6.566830
5	1978	21.29566	99	13	20.0	8.159631
6	1980	21.09944	99	12	20.0	7.797214
7	1982	20.96337	99	13	20.0	6.125338
8	1983	21.55155	99	12	20.0	8.027689
9	1984	22.17520	99	12	20.0	11.057647
10	1985	20.84732	99	14	20.0	5.525554
11	1986	21.81692	99	13	20.0	8.276794
12	1987	22.03584	99	13	20.0	9.447950
13	1988	21.78267	99	13	20.0	8.856189
14	1989	22.08607	99	13	20.0	9.597479
15	1990	22.45611	99	14	20.5	10.253327
16	1991	22.41484	99	13	20.0	10.603438
17	1993	22.29313	99	13	21.0	7.868378
18	1994	22.45730	99	13	21.0	8.462019
19	2006	22.31661	99	13	21.0	7.350396

Figure 1. Statistics of women's first marriage age by survey year

	Year	Mean	Max	Min	Median	StdDev
1	1974	23.91980	99	15	22.5	6.013550
2	1975	24.38504	99	16	22.0	8.898550
3	1976	24.13211	98	16	23.0	5.812399
4	1977	24.24425	99	16	23.0	7.514560
5	1978	23.54869	99	16	22.0	5.875863
6	1980	24.28516	99	13	23.0	9.496561
7	1982	24.55921	99	16	22.0	9.551031
8	1983	23.82258	99	16	22.0	7.774589
9	1984	24.06236	99	16	23.0	6.945299
10	1985	24.10575	99	15	23.0	7.987324
11	1986	23.28571	99	15	22.0	5.834966
12	1987	25.04355	99	14	23.0	11.791966
13	1988	23.91858	99	16	22.0	8.916609
14	1989	24.12500	99	14	23.0	8.595519
15	1990	25.46055	99	15	23.0	11.518706
16	1991	24.44516	99	16	23.0	7.788814
17	1993	25.09346	99	16	23.0	9.760362
18	1994	25.03151	99	16	23.0	10.157137
19	2006	25.54717	99	16	24.0	9.303872

Figure 2. Statistics of men's first marriage age by survey year

Both men and women show a gradual increase in the average first marriage age over the survey years

2.2 Average age at first marriage statistics categorized by highest degree

	Year	Mean	Max	Min	Median	StdDev
1	0	22.05358	99	12	20	9.036639
2	1	22.41050	99	13	21	7.801599
3	2	23.01837	99	14	22	7.291373
4	3	24.43591	99	14	23	6.814871
5	4	25.42426	99	14	24	7.728244

Figure 3. Statistics of average marriage age by Highest Degree

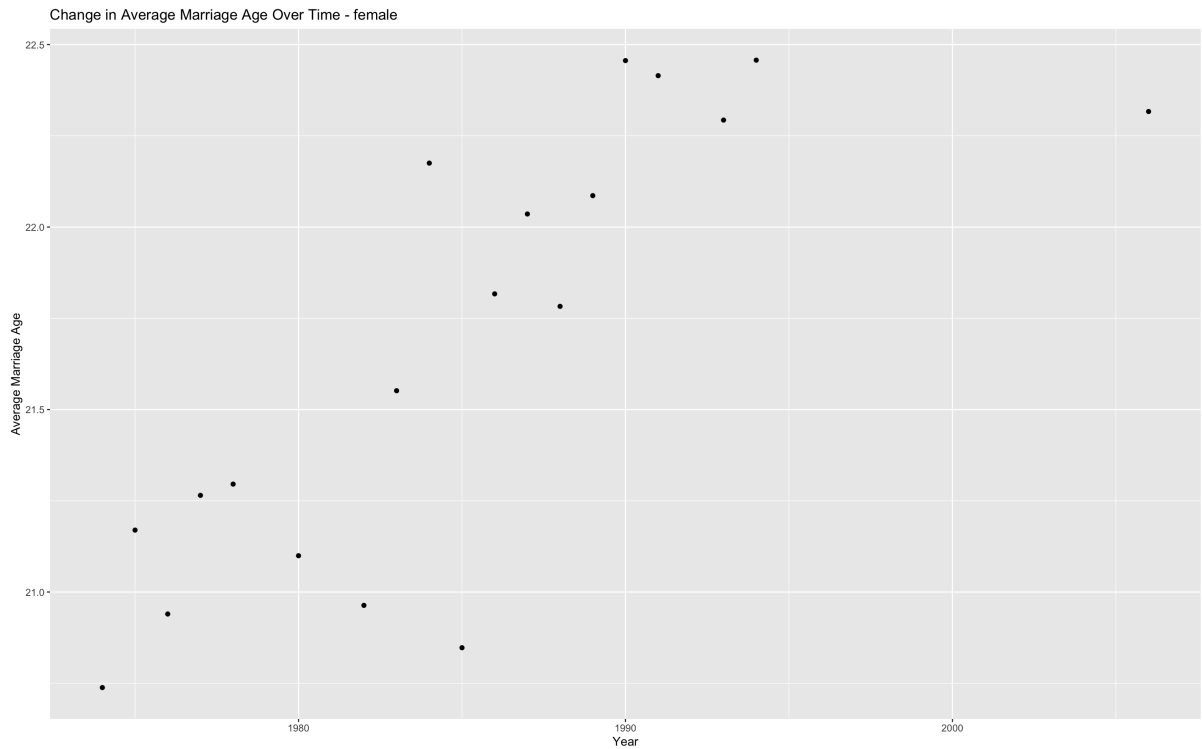## 3.1 Correlation between survey year and age at first marriage

Change in Average Marriage Age Over Time - female

Figure 4. Change in Average Women's Marriage age over time

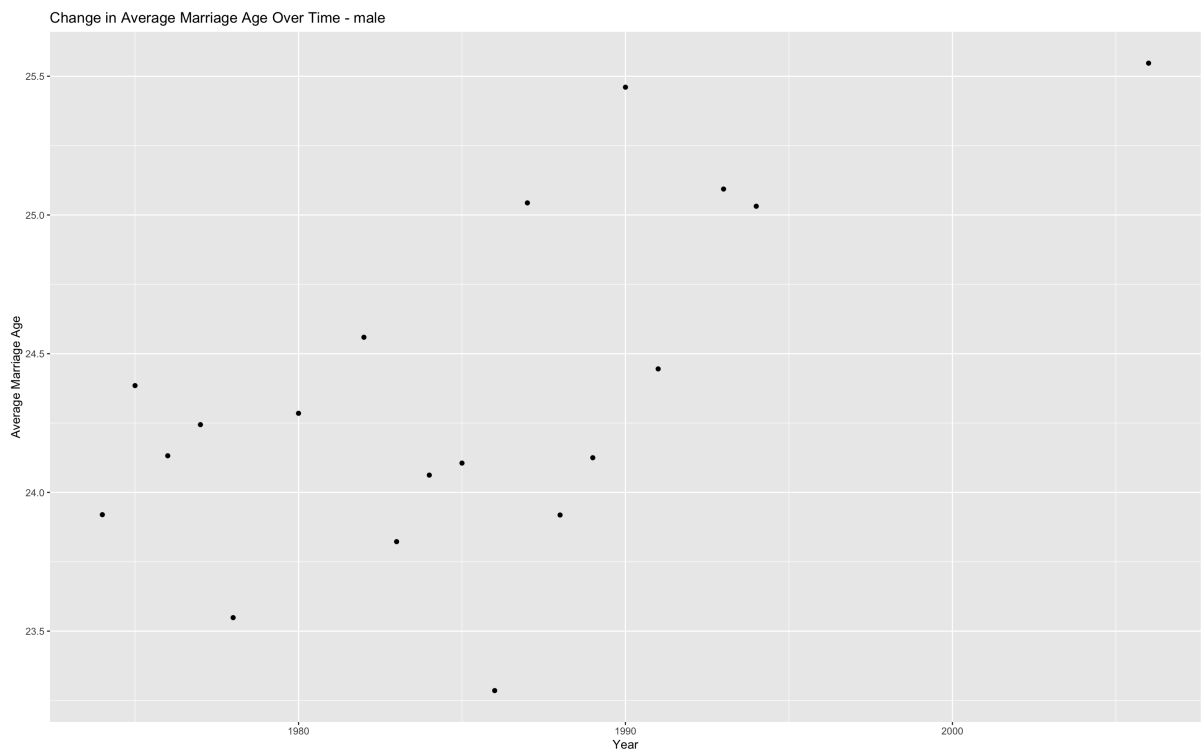Change in Average Marriage Age Over Time - male

Figure 4. Change in Average Men's Marriage age over time

Upon analyzing the data, it was observed that the average marriage age increases over time when separated by gender, but when examining the entire dataset, the correlation coefficient was 0.05. This indicates that there is a weak relationship between the survey year and the average marriage age, suggesting that survey year and marriage age are not strongly related.

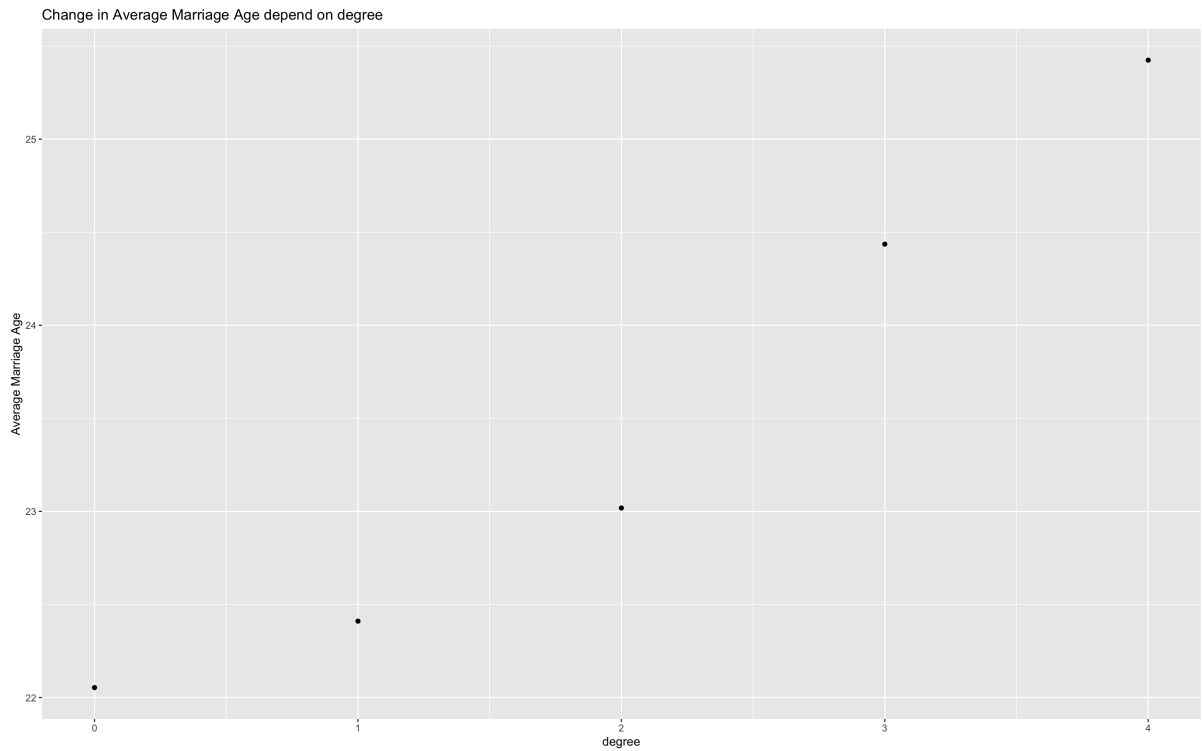## 3.2 Correlation between highest education and age at first marriage

Figure 4. Change in Average Marriage age depend on degree

This is a plot depicting the relationship between marriage age and the highest level of education. It was observed that as the level of education increases, the average marriage age also increases. When calculating the correlation coefficient for the entire dataset, it was found to be 0.13, indicating a very weak positive correlation.

4. Suggestion

In this analysis, it was found that survey year and educational level had limited impact on marriage age. Therefore, in a new project, investigating other variables that may influence marriage age is a consideration.

For instance, factors such as economic indicators, regional influences, family composition, cultural elements, changes in marriage rates, and factors like home purchases after marriage can be examined. Through this additional investigation, the aim is to uncover complex patterns and causal relationships affecting marriage age and identifying potential determinants that may impact future marriage trends.

Exploring additional variables provides insights into the factors influencing marriage age, potentially shedding light on future marriage trends.