Int bigData team6

CORRELATION OF MARRIAGE RATE WITH EDUCATION AND INCOME

TEAM PROJECT PRESENTATION

SPEAKER BYEORI MOON
YEONSOO OH
SUBIN KIM

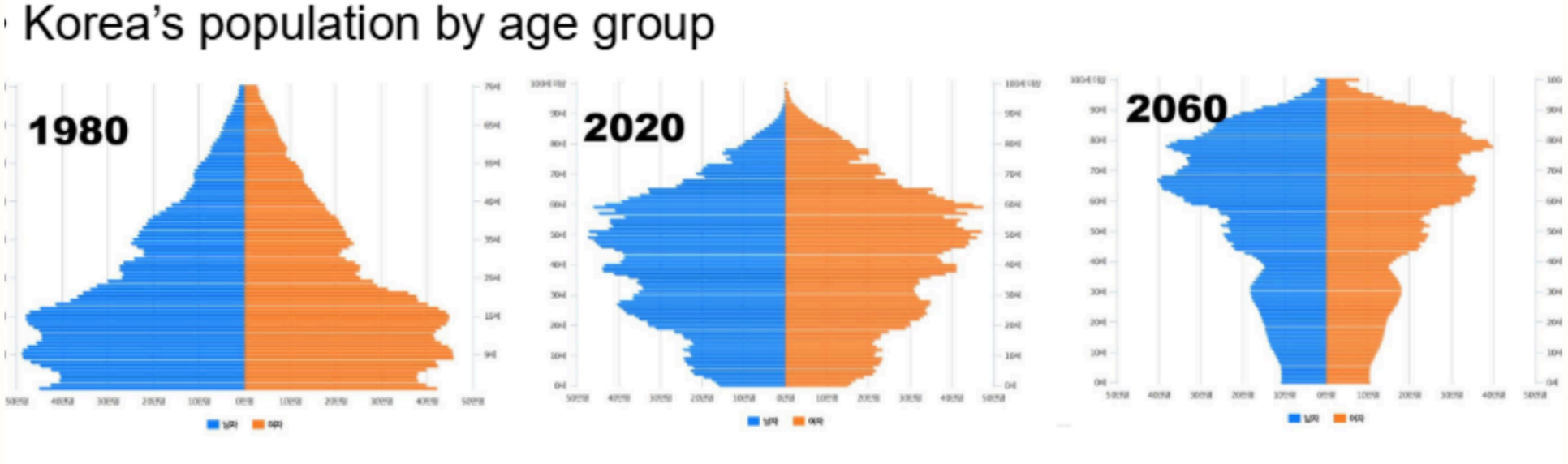
CONTENTS

INTRODUCTION

Visualization Example

KOREAN BIRTH RATE AND POPULATION DECLINE

Korea's population by age group



reference by IntBBigdata lceture5 visualization

DATA PREPARATION

Int bigData team6

KOREA INCOME AND WELFARE

WITH KEGGLE

	id	year	wave	region	income	family_member	gender	year_born	education_level	marriage	religion
0	10101	2005	1	1	614.0	1	2	1936	2	2	2
1	10101	2011	7	1	896.0	1	2	1936	2	2	2
2	10101	2012	8	1	1310.0	1	2	1936	2	2	2
3	10101	2013	9	1	2208.0	1	2	1936	2	2	2
4	10101	2014	10	1	864.0	1	2	1936	2	2	2

INCLUDES THE INCOME, WELFARE, AND LIVING STANDARDS OF KOREAN HOUSEHOLDS.

A TOTAL OF 10046 PEOPLE WERE SURVEYED AND RECORDED THEIR BIRTH YEAR, INCOME LEVEL, EDUCATION LEVEL, PLACE OF BIRTH, GENDER, NUMBER OF FAMILY MEMBERS, RELIGIOUS AFFILIATION, ETC.

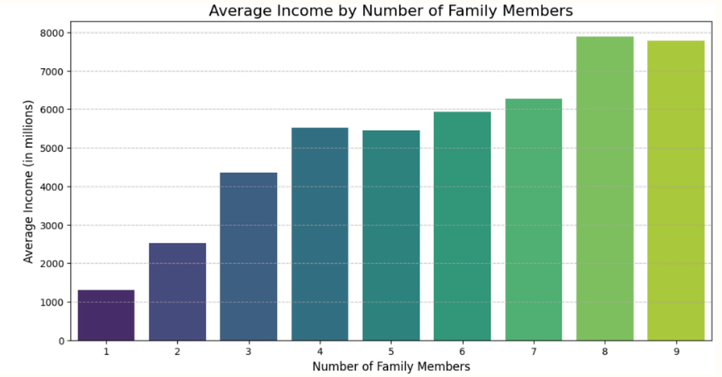
Int bigData team6

DATA PREPROCESSING

REMOVE INCOME DATA OUTLIERS

```
# Remove the top 1%
upper_limit = data['income'].quantile(0.99)
data = data[data['income'] <= upper_limit]
```

NORMALIZE INCOME OUTLIERS BY NUMBER OF FAMILY MEMBERS



GENERATE DATA ABOUT MARITAL STATUS

```
# Remove data to exclude: marriage == 1 (minor), marriage == 6 (other)
data = data[~data['marriage'].isin([1, 6])]

# Remove data to exclude: marriage == 1 (minor), marriage == 6 (other)
data['has_marriage_experience'] = data['marriage'].apply(lambda x: 1 if x in [2, 3, 4] else 0)
```

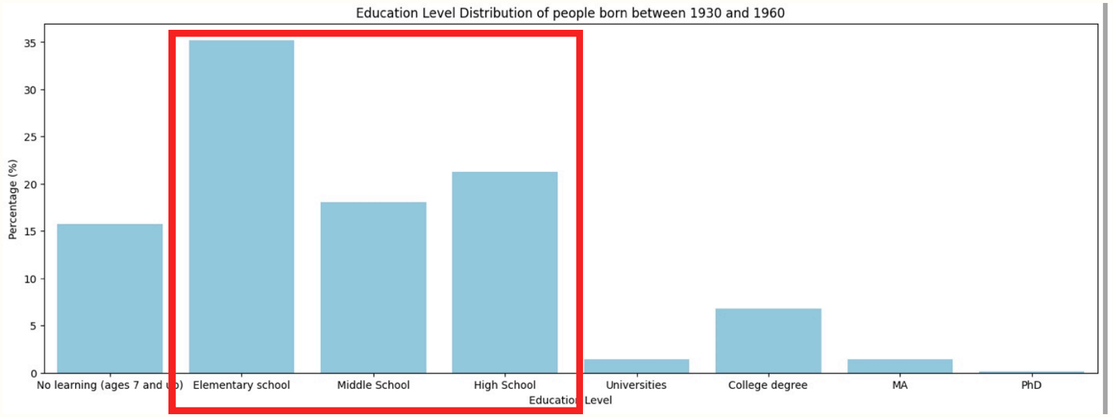
SEPARATE DATA BY GENERATION

```
# Create a generation variable
analyzed_data['generation'] = pd.cut(data['year_born'], bins=[1930, 1960, 1980, 2000, 2020],
                                     labels=['1930s-60s', '1960s-80s', '1980s-2000s', '2000s+'])
```

Int bigData team6

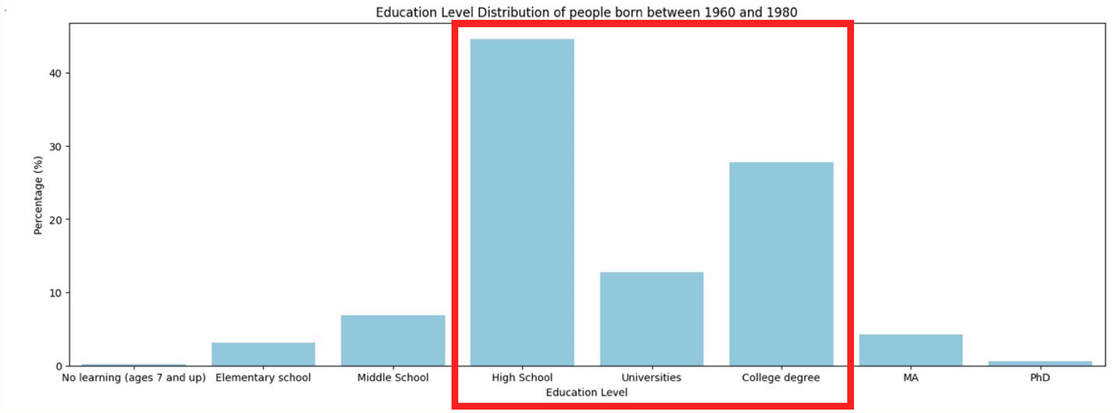
DATA ANALYSIS

FIRST GROUP
(1930~1960)



education_level	percentage
No learning (ages 7 and up)	28.830680
Elementary school	41.927752
Middle School	14.620784
High School	11.123951
Universities	0.749016
College degree	2.465331
MA	0.226845
PhD	0.055641

SECOND GROUP
(1960~1980)



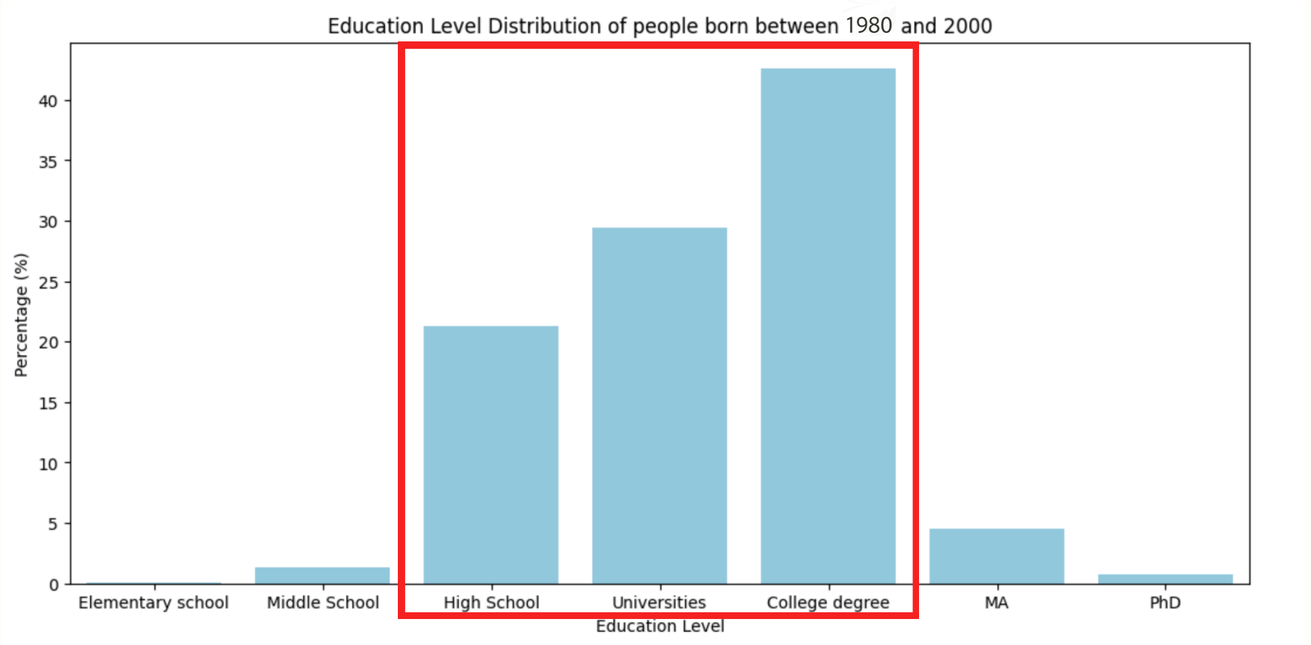
education_level	percentage
No learning (ages 7 and up)	0.429017
Elementary school	6.500260
Middle School	12.142486
High School	51.092044
Universities	11.895476
College degree	15.574623
MA	2.028081
PhD	0.338014

Int bigData team6

DATA ANALYSIS

THIRD GROUP
(1980~2000)

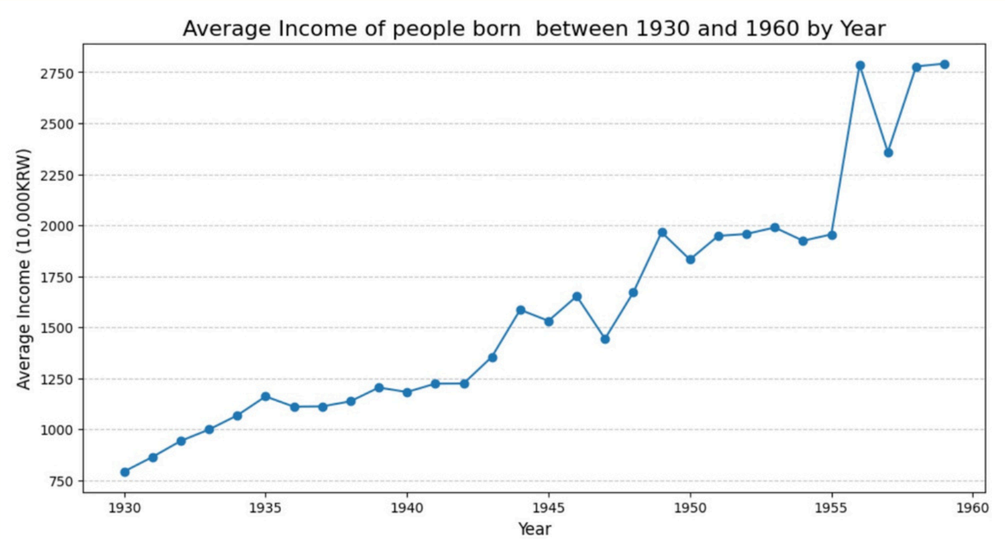
education_level	percentage
Elementary school	0.100100
Middle School	2.402402
High School	22.022022
Universities	27.977978
College degree	42.492492
MA	4.354354
PhD	0.650651



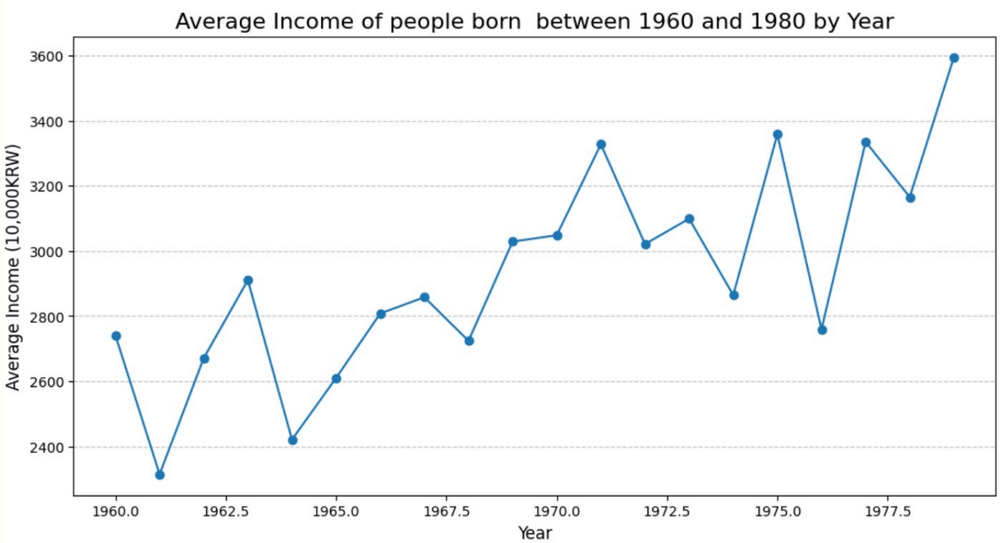
Int bigData team6

DATA ANALYSIS

FIRST GROUP
(1930~1960)



SECOND GROUP
(1960~1980)

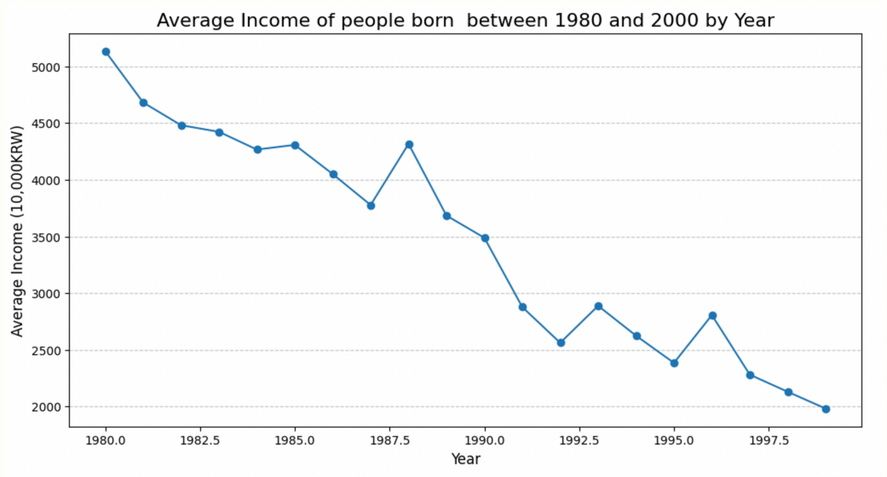


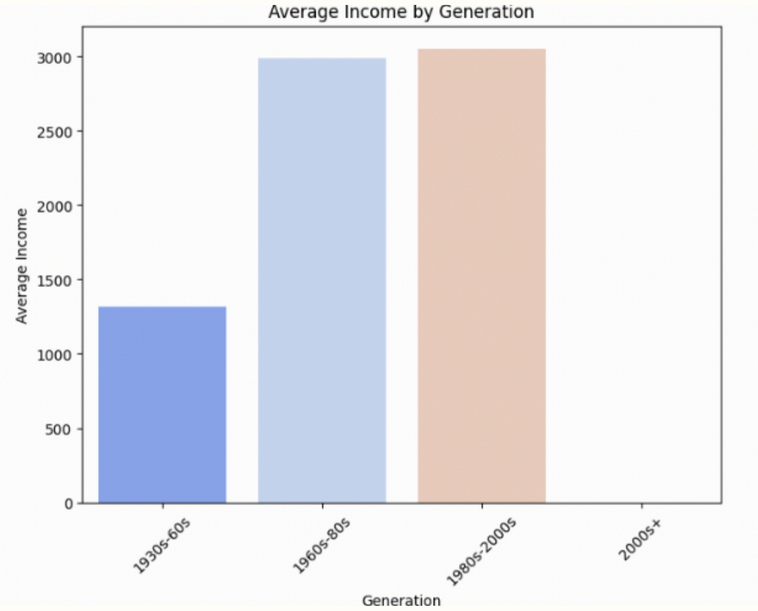
Int bigData team6

DATA ANALYSIS

THIRD GROUP (1980~2000)

AVERAGE INCOME BY GENERATION

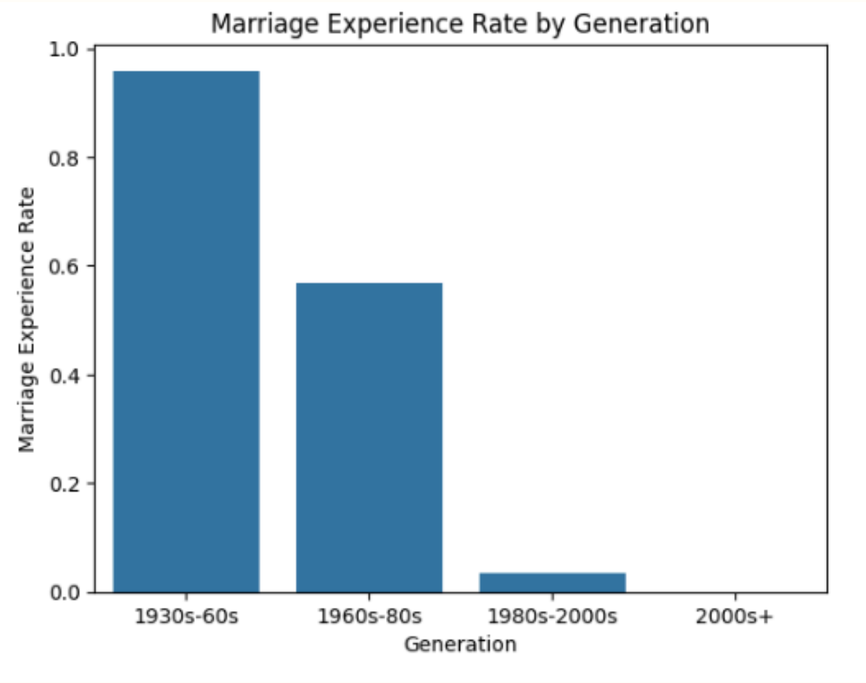




Int bigData team6

DATA ANALYSIS

OVERALL GENERATION



CONCLUSION

The overall leveling up of educational standards has made it increasingly difficult to secure stable employment and satisfactory income with just a college degree.

This impacts individuals' economic conditions and social stability, ultimately tying into declining marriage rates.

THANK YOU!