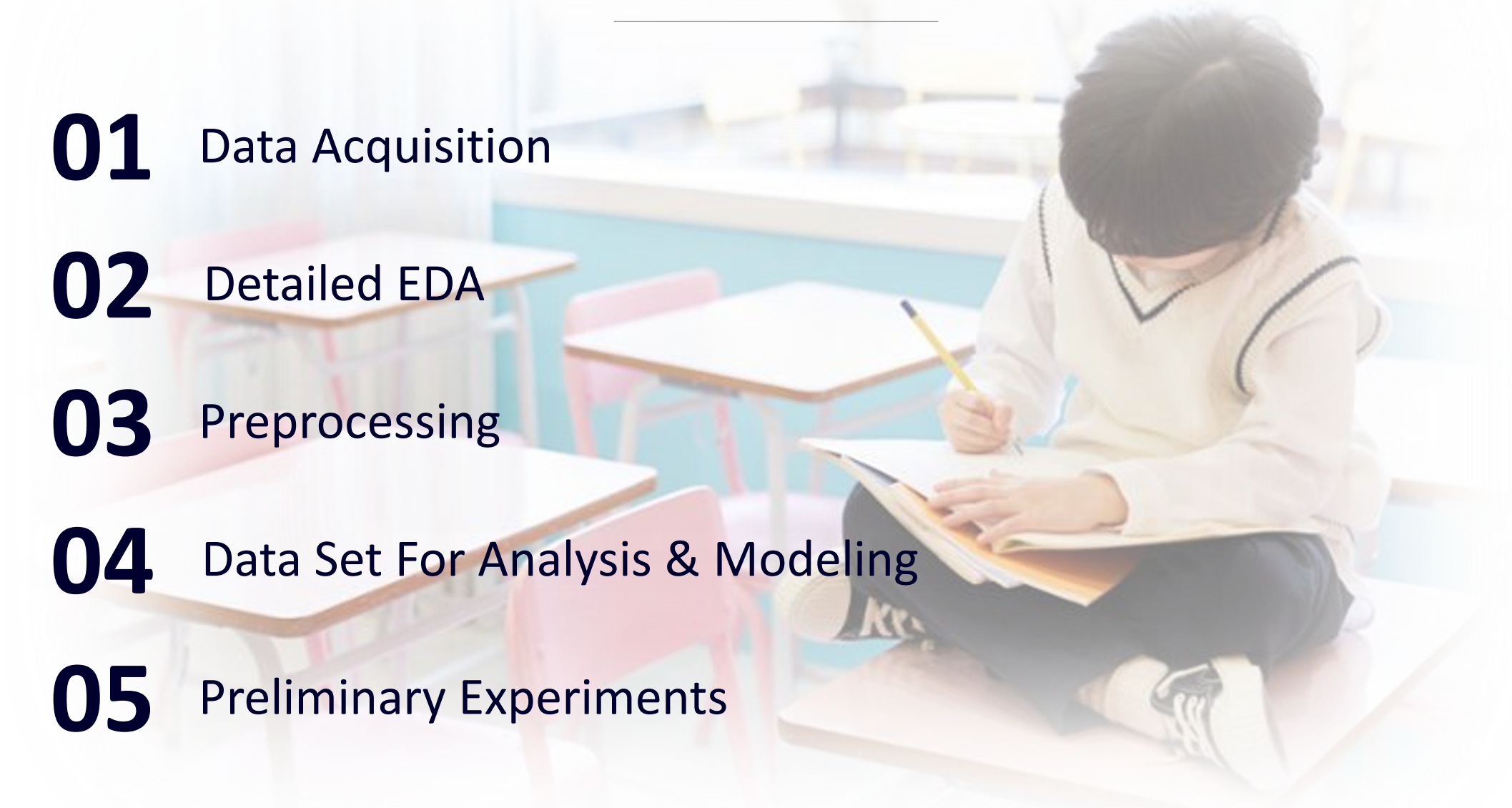
Korean Birthrate Problem

Team 8 - Progress

19102085 Park Sehong
19102088 Park Jinsoo
21102044 Oh Seyeon

CONTENTS

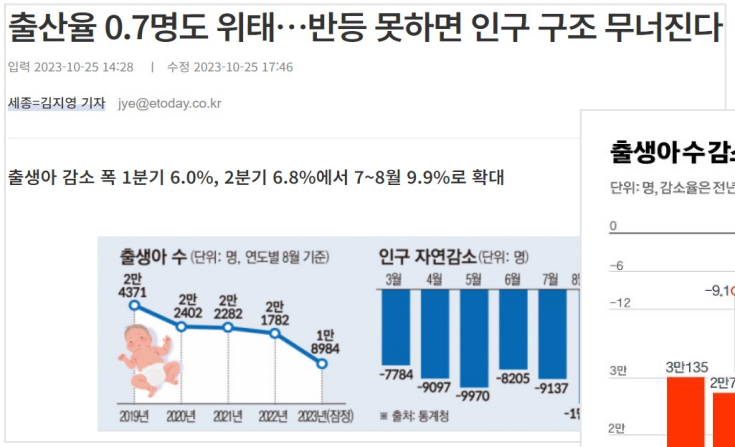


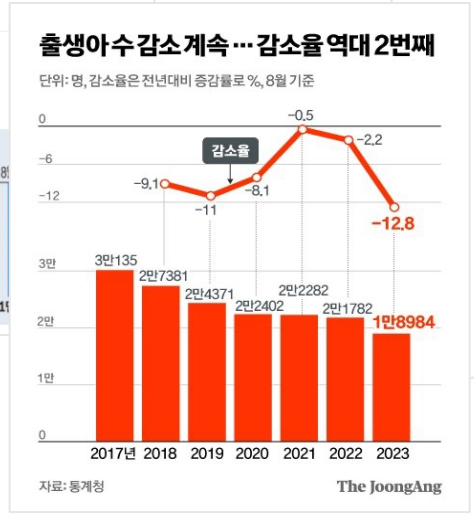
00. Introduction

Purpose of Analysis

1) The purpose of creating a birth rate prediction model is to analyze and identify the factors that most significantly influence birth rates.

2) Proposing Solutions from a Socio-Economic Standpoint Based on the Identified Causes.





01. Data Acquisition

- Period: 1990 ~ 2021
- 15 countries

Country Code	Country Name
AUT	AUSTRIA
CAN	Canada
CHE	Switzerland
DEU	Germany
ESP	Spain

Country Code	Country Name
FIN	Finland
FRA	France
GRC	Greece
HUN	Hungary
ITA	Italy

Country Code	Country Name
JPN	Japan
KOR	Korea
LUX	Luxembourg
POL	Poland
PRT	Portugal

01. Data Acquisition

Feature Name	Data (Description)
BirthRate (Target Y)	Birth rate, crude (per 1,000 people)
FemaleLaborParticipationRate	Female Labor force participation rate (% of population)
AvgHourWorked	Average annual hours worked (= the total number of hours actually worked per year divided by the average number of people in employment per year.)
BothWorking	Proportion of all adults are in employment(working)
FirstBirthAge	Mean age of women at first childbirth
MarriageAge	Mean age at first marriage
MarriageRate	Crude marriage rate (per 1,000 inhabitants)
EmploymentRate	Employment rate
UnEmploymentRate	Unemployment rate

01. Data Acquisition

Feature Name	Data (Description)
HousingPrice	Housing prices (including housing rent prices indices, real and nominal house prices indices, and ratios of price to rent and price to income.)
InterestRate	Short-term interest rates
PartTimeRate	Part-time employment rate
FamilyExpenditure	Public spending on family benefits
HealthExpenditure	Current health expenditure (% of GDP)
LaborMarketExepnditure	Public expenditure for labor market
UnemploymentExpenditure	Public expenditure to compensate for unemployment
GDI	Gross Domestic Income
GDP	Gross Domestic Product

01. Data Acquisition

Feature Name	Data (Description)
GNI	Gross National Income
PovertyGap	Poverty gap at $6.85 a day
EduExpenditureOfGDP	Government expenditure on education (% of GDP)
EduExpenditureOfGov	Proportion of government expenditure on education
WeeksPaidLeaveForMothers	Length of paid maternity, parental and home care leave available to mothers in weeks
TotalLaborParticipationRate	Total Labor force participation rate (% of population)
DivorceRate	Crude divorce rate (divorces per 1000 people)
InflationRate	Inflation rate
Population	The number of people

02. Detailed EDA

- Data Merge
  - Merge each feature based on country ID.

	ID	Year	Country Name	BirthRate	LaborParticipationRate	AvgHoursWorked	BothWorking
0	AUT1990	1990	Austria	11.8	55.172	NaN	NaN
1	AUT1991	1991	Austria	12.2	55.979	NaN	NaN
2	AUT1992	1992	Austria	12.2	57.604	NaN	NaN
3	AUT1993	1993	Austria	12.0	58.208	NaN	NaN
4	AUT1994	1994	Austria	11.6	61.895	NaN	NaN
...	...	...	...	...	...	...	...
475	PRT2017	2017	Portugal	8.4	71.674	1727.140	69.3
476	PRT2018	2018	Portugal	8.5	72.376	1737.704	71.0
477	PRT2019	2019	Portugal	8.4	72.901	1743.930	NaN
478	PRT2020	2020	Portugal	8.2	71.790	1611.193	NaN
479	PRT2021	2021	Portugal	7.7	72.851	1648.943	NaN

	ID	Country	Year	InflationRate
0	AUT1990	AUT	1990	3.261872
1	AUT1991	AUT	1991	3.337426
2	AUT1992	AUT	1992	4.020847
3	AUT1993	AUT	1993	3.631786
4	AUT1994	AUT	1994	2.953407
...	...	...	...	...
490	CAN2018	CAN	2018	2.268226
491	CAN2019	CAN	2019	1.949269
492	CAN2020	CAN	2020	0.717000
493	CAN2021	CAN	2021	3.395193
494	CAN2022	CAN	2022	6.802801

02. Detailed EDA

- Data Merge

  - Result

There are missing values!

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 480 entries, 0 to 479
Data columns (total 30 columns):
 #   Column                        Non-Null Count  Dtype
---  ------                        --------------  -----
 0   ID                            480 non-null    object
 1   Year                          480 non-null    int64
 2   Country Name                  480 non-null    object
 3   BirthRate                     480 non-null    float64
 4   FemaleLaborParticipationRate  480 non-null    float64
 5   AvgHoursWorked                453 non-null    float64
 6   BothWorking                   191 non-null    float64
 7   FirstBirthAge                 455 non-null    float64
 8   MarriageAge                   174 non-null    float64
 9   MarriageRate                  406 non-null    float64
 10  EmploymentRate                465 non-null    float64
 11  UnemploymentRate              434 non-null    float64
 12  HousingPrice                  414 non-null    float64
 13  InterestRate                  432 non-null    float64
 14  PartTimeRate                  462 non-null    float64
 15  FamilyExpenditure             445 non-null    float64
 16  HealthExpenditure             318 non-null    float64
 17  LaborMarketExpenditure        415 non-null    float64
 18  UnemploymentExpenditure       446 non-null    float64
 19  GDI                           470 non-null    float64
 20  GDP                           479 non-null    float64
 21  GNI                           455 non-null    float64
 22  PovertyGap                    308 non-null    float64
 23  EduExpenditureOfGDP           415 non-null    float64
 24  EduExpenditureOfGov           370 non-null    float64
 25  WeeksPaidLeaveForMothers      480 non-null    float64
 26  TotalLaborParticipationRate   480 non-null    float64
 27  DivorceRate                   441 non-null    float64
 28  InflationRate                 480 non-null    float64
 29  Population                    480 non-null    float64
dtypes: float64(27), int64(1), object(2)
memory usage: 112.6+ KB
```

03. Preprocessing

- Handling Missing Values
  - List up the missing ratio in descending order

	missing_ratio
MarriageAge	63.750000
BothWorking	60.208333
PovertyGap	35.833333
HealthExpenditure	33.750000
EduExpenditureOfGov	22.916667
MarriageRate	15.416667
HousingPrice	13.750000
LaborMarketExpenditure	13.541667
EduExpenditureOfGDP	13.541667
InterestRate	10.000000
UnemploymentRate	9.583333
DivorceRate	8.125000
FamilyExpenditure	7.291667
UnemploymentExpenditure	7.083333
AvgHoursWorked	5.625000
FirstBirthAge	5.208333
GNI	5.208333
PartTimeRate	3.750000
EmploymentRate	3.125000
GDI	2.083333
GDP	0.208333
WeeksPaidLeaveForMothers	0.000000
TotalLaborParticipationRate	0.000000
InflationRate	0.000000
ID	0.000000
Year	0.000000
FemaleLaborParticipationRate	0.000000
BirthRate	0.000000
Country Name	0.000000
Population	0.000000

03. Preprocessing

- Handling Missing Values
  - List up the missing ratio in descending order
  - Drop features with a missing value over 20%

결측치 20% 이상 제거

```
columns_to_drop = data.columns[data.isnull().mean() > 0.2]
data = data.drop(columns_to_drop, axis=1)
```

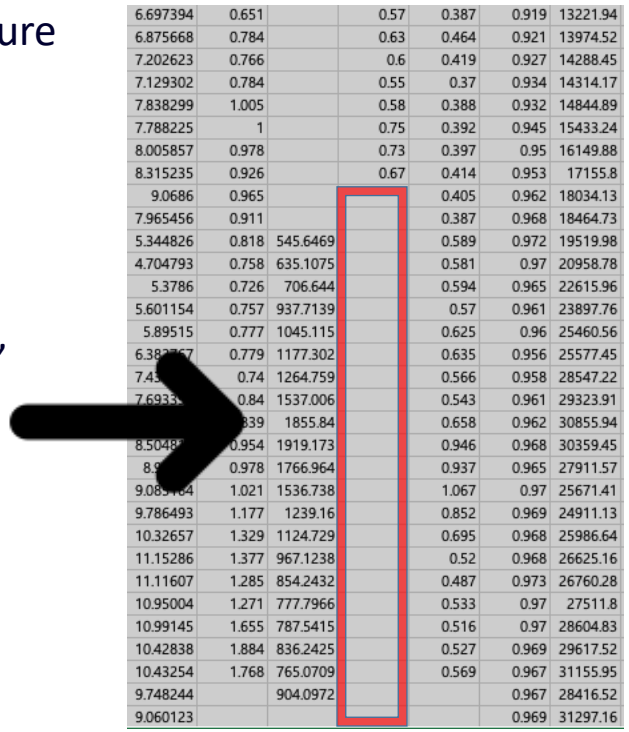
```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 480 entries, 0 to 479
Data columns (total 25 columns):
 #   Column                          Non-Null Count  Dtype
---  ------                          --------------  -----
 0   ID                              480 non-null    object
 1   Year                            480 non-null    int64
 2   Country Name                    480 non-null    object
 3   BirthRate                       480 non-null    float64
 4   FemaleLaborParticipationRate    480 non-null    float64
 5   AvgHoursWorked                  453 non-null    float64
 6   FirstBirthAge                   455 non-null    float64
 7   MarriageRate                    406 non-null    float64
 8   EmploymentRate                  465 non-null    float64
 9   UnemploymentRate                434 non-null    float64
 10  HousingPrice                    414 non-null    float64
 11  InterestRate                    432 non-null    float64
 12  PartTimeRate                    462 non-null    float64
 13  FamilyExpenditure               445 non-null    float64
 14  LaborMarketExpenditure          415 non-null    float64
 15  UnemploymentExpenditure         446 non-null    float64
 16  GDI                             470 non-null    float64
 17  GDP                             479 non-null    float64
 18  GNI                             455 non-null    float64
 19  EduExpenditureOfGDP             415 non-null    float64
 20  WeeksPaidLeaveForMothers        480 non-null    float64
 21  TotalLaborParticipationRate     480 non-null    float64
 22  DivorceRate                     441 non-null    float64
 23  InflationRate                   480 non-null    float64
 24  Population                      480 non-null    float64
dtypes: float64(22), int64(1), object(2)
memory usage: 93.9+ KB
```

03. Preprocessing

- Handling Missing Values
  - List up the missing ratio in descending order
  - Drop features with a missing value over 20%
  - If missing values are concentrated in a country, remove the feature

“LaborMarketExpenditure” data in Greece



6.697394	0.651		0.57	0.387	0.919	13221.94
6.875668	0.784		0.63	0.464	0.921	13974.52
7.202623	0.766		0.6	0.419	0.927	14288.45
7.129302	0.784		0.55	0.37	0.934	14314.17
7.838299	1.005		0.58	0.388	0.932	14844.89
7.788225	1		0.75	0.392	0.945	15433.24
8.005857	0.978		0.73	0.397	0.95	16149.88
8.315235	0.926		0.67	0.414	0.953	17155.8
9.0686	0.965			0.405	0.962	18034.13
7.965456	0.911			0.387	0.968	18464.73
5.344826	0.818	545.6469		0.589	0.972	19519.98
4.704793	0.758	635.1075		0.581	0.97	20958.78
5.3786	0.726	706.644		0.594	0.965	22615.96
5.601154	0.757	937.7139		0.57	0.961	23897.76
5.89515	0.777	1045.115		0.625	0.96	25460.56
6.38[illegible]	0.779	1177.302		0.635	0.956	25577.45
7.43	0.74	1264.759		0.566	0.958	28547.22
7.6933	0.84	1537.006		0.543	0.961	29323.91
[illegible]	[illegible]39	1855.84		0.658	0.962	30855.94
8.5048[illegible]	0.954	1919.173		0.946	0.968	30359.45
8.9	0.978	1766.964		0.937	0.965	27911.57
9.08[illegible]	1.021	1536.738		1.067	0.97	25671.41
9.786493	1.177	1239.16		0.852	0.969	24911.13
10.32657	1.329	1124.729		0.695	0.968	25986.64
11.15286	1.377	967.1238		0.52	0.968	26625.16
11.11607	1.285	854.2432		0.487	0.973	26760.28
10.95004	1.271	777.7966		0.533	0.97	27511.8
10.99145	1.655	787.5415		0.516	0.97	28604.83
10.42838	1.884	836.2425		0.527	0.969	29617.52
10.43254	1.768	765.0709		0.569	0.967	31155.95
9.748244		904.0972			0.967	28416.52
9.060123					0.969	31297.16

한 나라에 결측치가 몰려있는 경우 해당 Feature 제거

```
data = data.drop('LaborMarketExpenditure', axis=1)
```

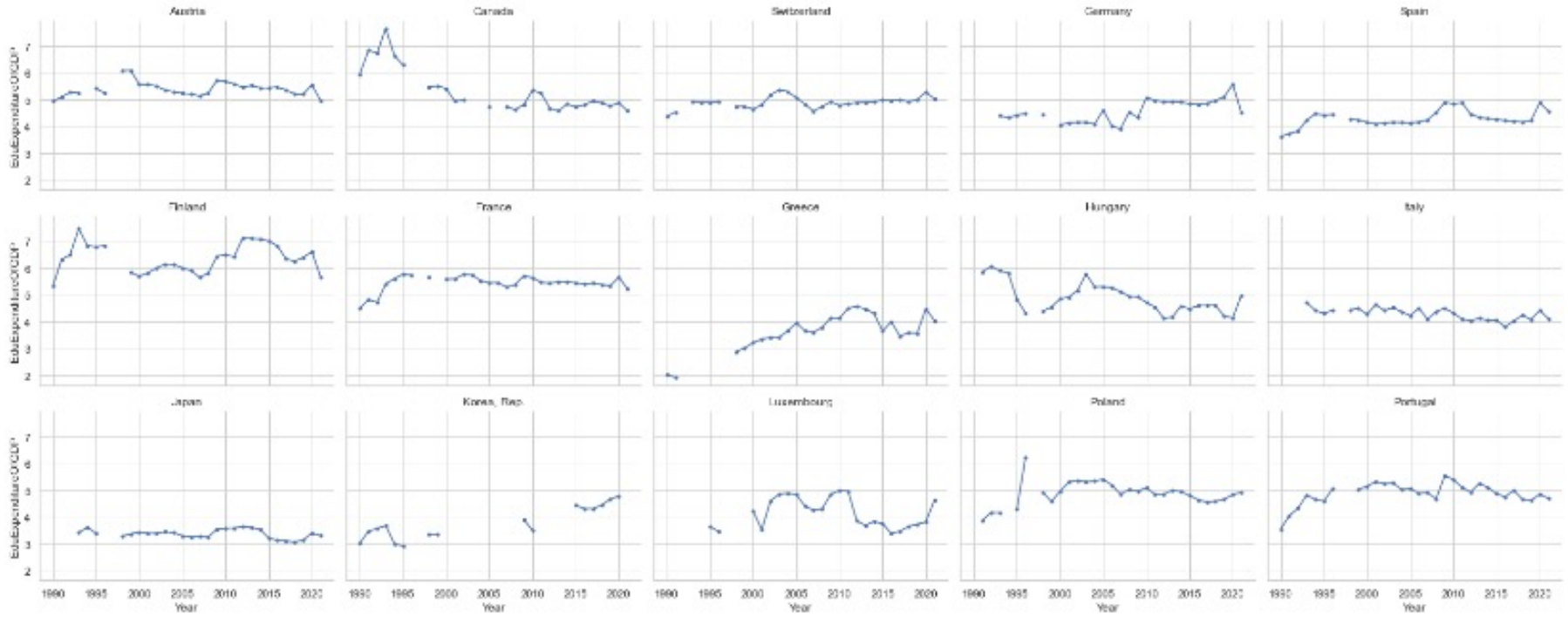
```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 480 entries, 0 to 479
Data columns (total 24 columns):
 #   Column                          Non-Null Count  Dtype
---  ------                          --------------  -----
 0   ID                              480 non-null    object
 1   Year                            480 non-null    int64
 2   Country Name                    480 non-null    object
 3   BirthRate                       480 non-null    float64
 4   FemaleLaborParticipationRate    480 non-null    float64
 5   AvgHoursWorked                  453 non-null    float64
 6   FirstBirthAge                   455 non-null    float64
 7   MarriageRate                    406 non-null    float64
 8   EmploymentRate                  465 non-null    float64
 9   UnemploymentRate                434 non-null    float64
 10  HousingPrice                    414 non-null    float64
 11  InterestRate                    432 non-null    float64
 12  PartTimeRate                    462 non-null    float64
 13  FamilyExpenditure               445 non-null    float64
 14  UnemploymentExpenditure         446 non-null    float64
 15  GDI                             470 non-null    float64
 16  GDP                             479 non-null    float64
 17  GNI                             455 non-null    float64
 18  EduExpenditureOfGDP             415 non-null    float64
 19  WeeksPaidLeaveForMothers        480 non-null    float64
 20  TotalLaborParticipationRate     480 non-null    float64
 21  DivorceRate                     441 non-null    float64
 22  InflationRate                   480 non-null    float64
 23  Population                      480 non-null    float64
dtypes: float64(21), int64(1), object(2)
memory usage: 90.1+ KB
```

03. Preprocessing

- Graphs after removing features of missing values

  - Feature : EduExpenditureOfGDP



03. Preprocessing

- Time Series Data Preprocessing Methods
  - Time-based Interpolation
  - Moving Average
  - Time Series Modeling
  - Linear Interpolation
  - Fill with Time – based Mean or Median

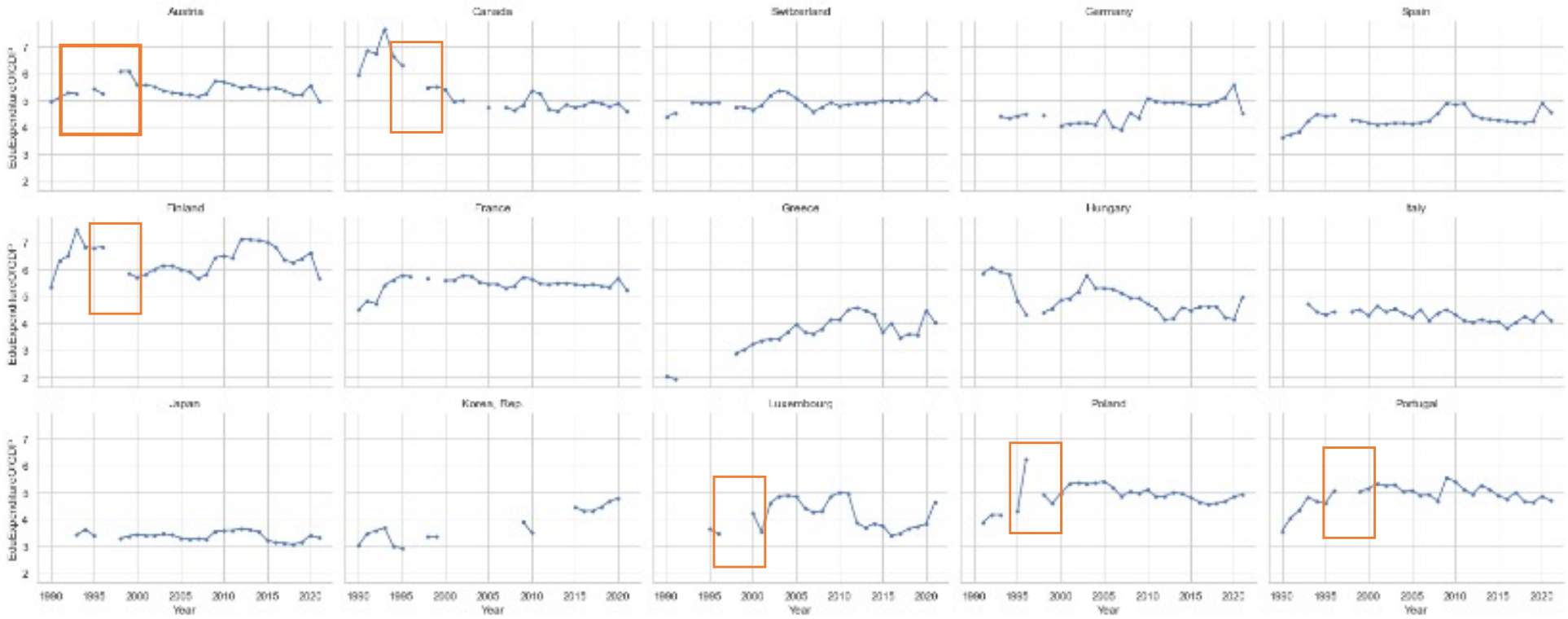
03. Preprocessing

- Method : Time – based interpolation
  - First + ends = NaN
  - First = NaN
  - Ends = NaN
  - Middle = NaN

```python
def interpolate_with_direction(column):
    if column.isna().iloc[0] and column.isna().iloc[-1]:
        # 양 끝이 모두 NaN인 경우: 양방향 보간
        return column.interpolate(method='time', limit_direction='both')
    elif column.isna().iloc[0]:
        # 시작 부분이 NaN인 경우: 앞에서 뒤로 보간
        return column.interpolate(method='time', limit_direction='backward')
    elif column.isna().iloc[-1]:
        # 끝 부분이 NaN인 경우: 뒤에서 앞으로 보간
        return column.interpolate(method='time', limit_direction='forward')
    else:
        # 중간에 NaN이 있는 경우: 기본 보간
        return column.interpolate(method='time')
```

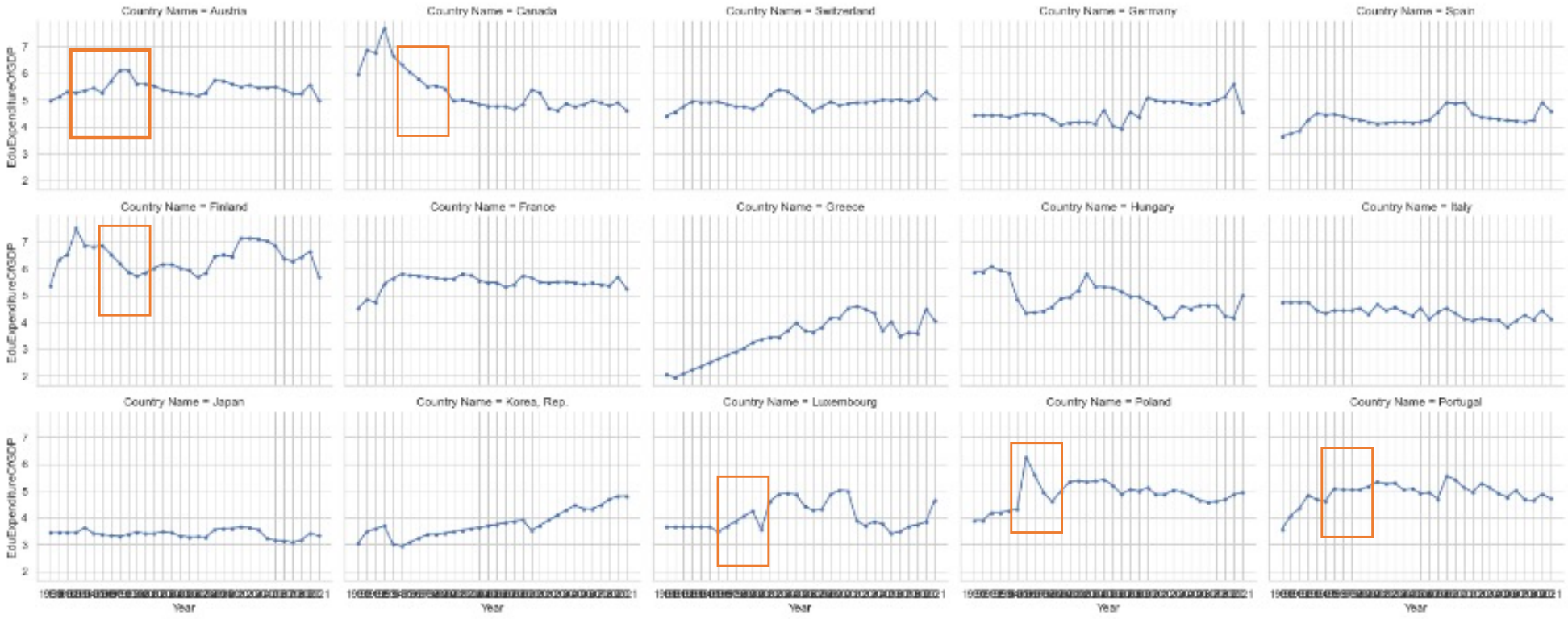
03. Preprocessing

- Check that the missing values are filled.



03. Preprocessing

- Check that the missing values are filled.



03. Preprocessing

- Feature Extraction : extracting new features as a combination of all the original features

1) **Work-Leisure Balance Index** : AvgHoursWorked / LaborParticipationRate

```
data['WorkLeisureBalanceIndex'] = data['AvgHoursWorked'] / data['TotalLaborParticipationRate']
```

- It provides a detailed look at the **characteristics of the labor market**. This ratio indicates the average working hours relative to the labor participation rate, **offering insights into the intensity of work and leisure time** available, which can impact birth rates.

- **High Value**: Long working hours combined with a low participation rate, suggesting **a high work intensity** with less leisure time, potentially leading to lower birth rates.
- **Low Value**: Shorter working hours with a high participation rate, suggesting **more leisure time is available**, which might be conducive to higher birth rates.

03. Preprocessing

2) LaborMarketStability : the ratio of the Employment Rate to the Unemployment Rate

```
data['LaborMarketStability'] = data['EmploymentRate'] / data['UnemploymentRate']
```

- This feature helps **in assessing the overall health and stability of the labor market**, where a higher ratio indicates stability (high employment, low unemployment), and a lower ratio indicates instability (low employment, high unemployment).

3) MarriageStabilityIndex : The ratio of the Marriage Rate to the Divorce Rate

```
data['MarriageStabilityIndex'] = data['MarriageRate'] / data['DivorceRate']
```

- This index serves as an indicator of marital stability.
- **High value** : suggests that marriages are generally **more stable or lasting**, which may be connected to greater domestic stability and could encourage childbirth.
- **Low value** : indicate **less stable or shorter marriages**, potentially affecting birth rates negatively.

03. Preprocessing

4) HousingAffordabilityIndex : the ratio of the HousingPriceIndex to the PerGDP

```
data['PerCapitaGDP'] = data['GDP'] / data['Population']
data['HousingAffordabilityIndex'] = data['HousingPrice'] / data['PerCapitaGDP']
```

- This index quantifies the **economic burden of purchasing a home** by comparing the cost of buying a house to the income level of individuals or households.
- **Low value:** indicates that **buying a house is relatively affordable**, which may lead to less financial strain on starting and expanding families, potentially increasing birth rates.
- **High value:** implies that **buying a house is expensive** relative to personal income, which could increase financial pressure and act as a deterrent to having children.

03. Preprocessing

- Correlation Analysis (over 0.7)

Exclusion features:

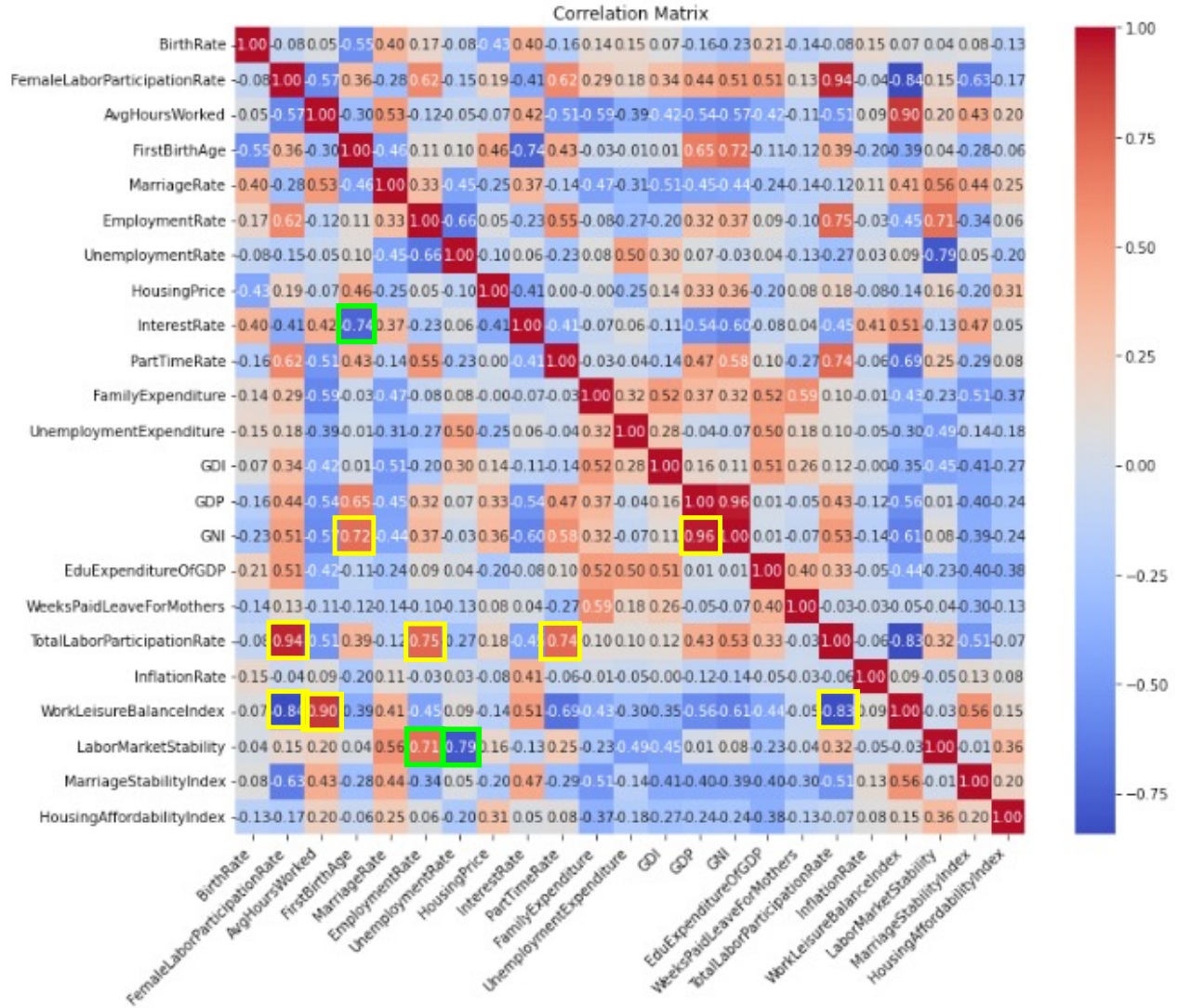
- 'TotalLaborParticipationRate'
- 'WorkLeisureBalanceIndex'
- 'GNI'

"The three features displaying high correlation with the **various other features** were excluded."

Features for make various dataset:

- LaborMarkeyStability / UnemploymentRate, EmploymentRate
- FirstBirthAge / InterestRate

"These features only show high correlation within **their respective pairs**.
Therefore, it is believed that the choice of features could impact future performance. Hence, **four new datasets** have been formed using combinations of the above four features."



03. Preprocessing

Creating different datasets for comparing the performance of models made from different datasets

- **data01**: Exclusion = 'EmploymentRate', 'UnemploymentRate', 'InterestRate'

- **data02**: Exclusion = 'EmploymentRate', 'UnemploymentRate', 'FirstBirthAge'

- **data03**: Exclusion = 'LaborMarketStability', 'InterestRate'

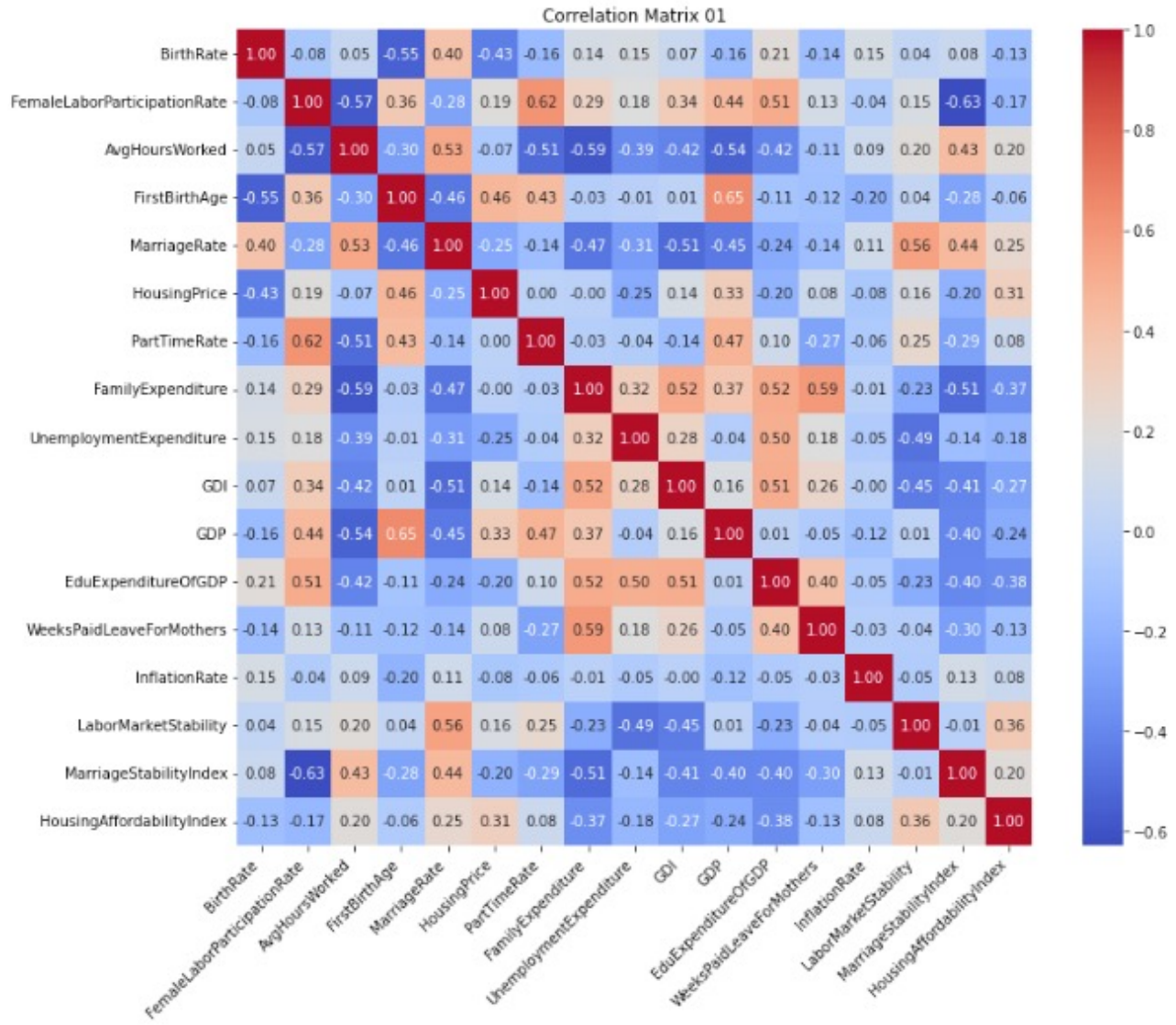
- **data04**: Exclusion = 'LaborMarketStability', 'FirstBirthAge'

Data Set For Analysis & Modeling

- Correlation Matrix – Data 01

Exclusion features:

- 'EmploymentRate'
- 'UnemploymentRate'
- 'InterestRate'

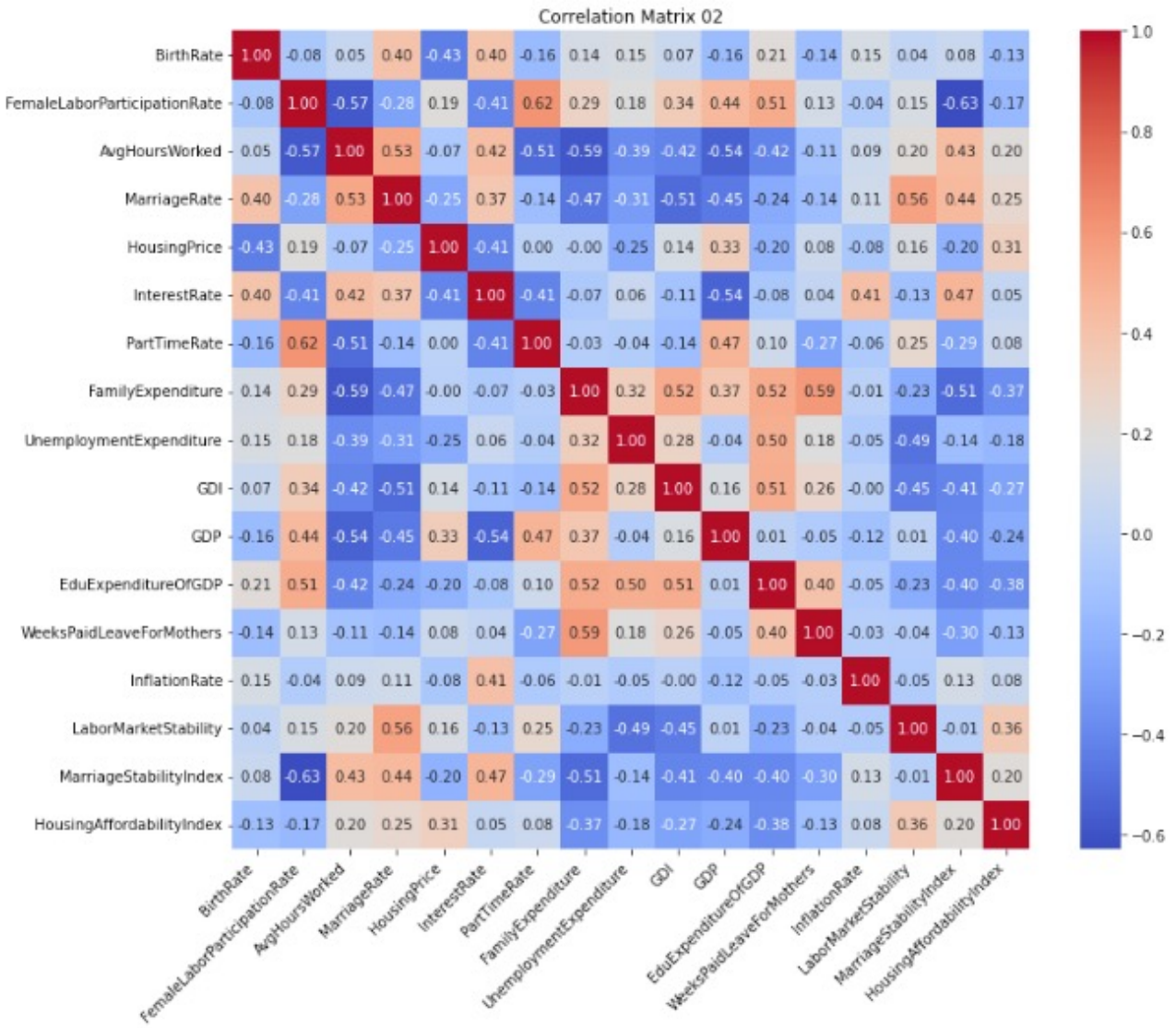


Data Set For Analysis & Modeling

- Correlation Matrix – Data 02

Exclusion features:

- 'EmploymentRate'
- 'UnemploymentRate'
- 'FirstBirthAge'

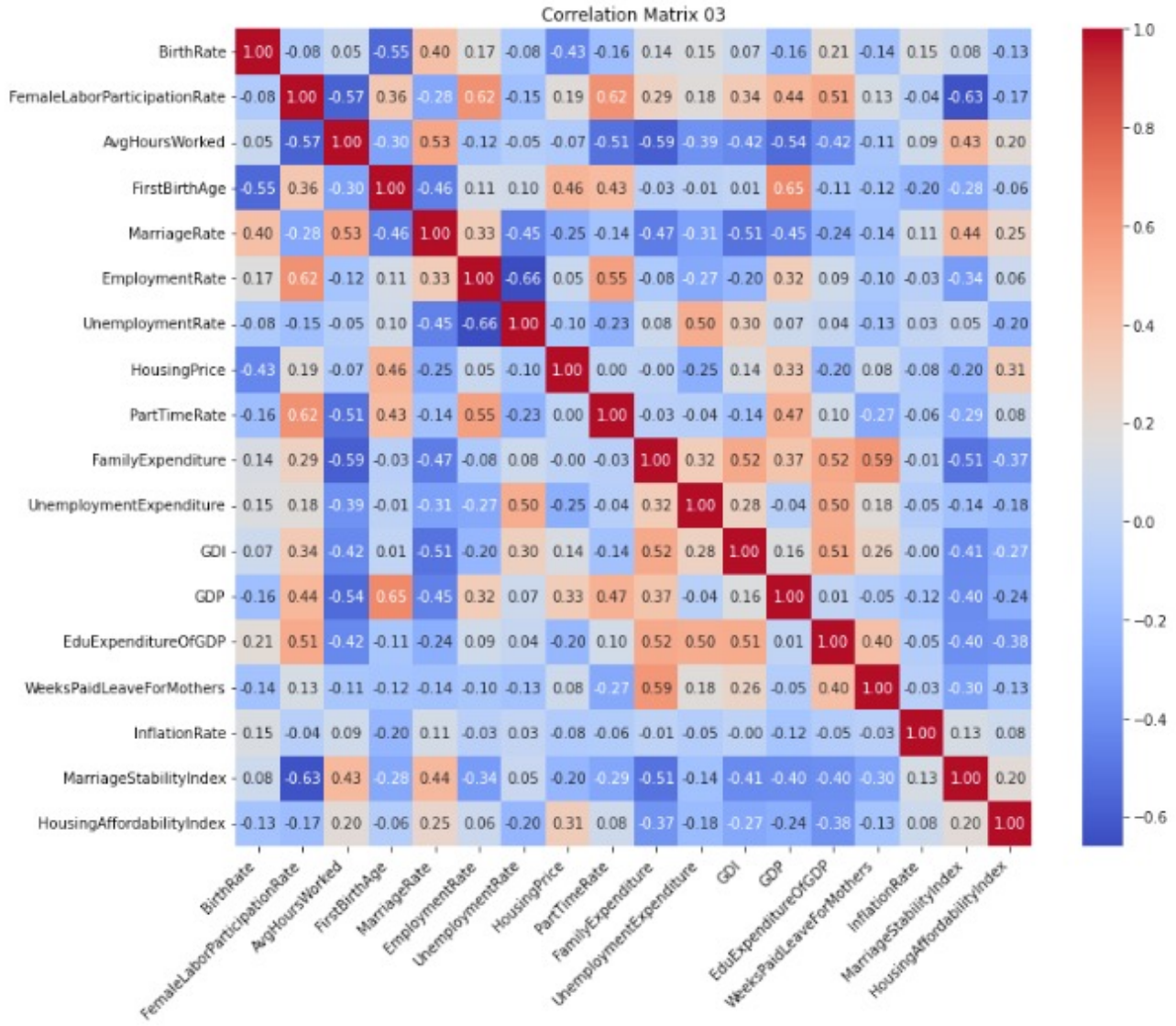


Data Set For Analysis & Modeling

- Correlation Matrix – Data 03

Exclusion features:

- 'LaborMarketStability'
- 'InterestRate'

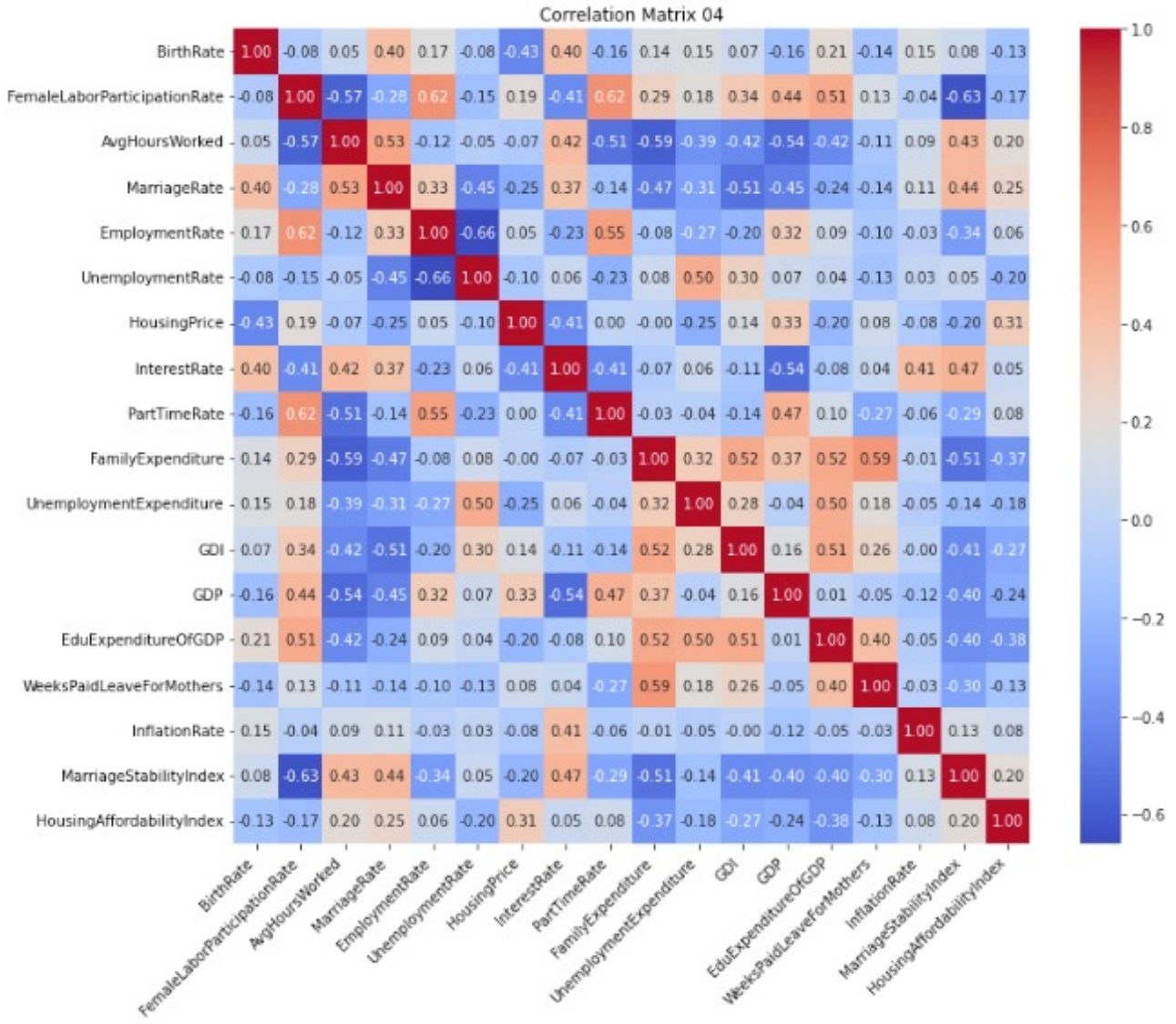


Data Set For Analysis & Modeling

- Correlation Matrix – Data 04

Exclusion features:

- 'LaborMarketStability'
- 'FirstBirthAge'



05. Preliminary Experiments

1) Check if there is a performance difference between datasets created through Correlation Analysis.

- Used model: XGBoost (the model proposed in Proposal Presentation)

- XGBoost: Advanced implementation of gradient boosting algorithms. XGBoost can handle various types of data

[data01]

```
XGBoost Model MSE: 1.0353283823561519
XGBoost Model R² : 0.5221714462379863
```

[data04]

```
XGBoost Model MSE: 0.9298940139183386
XGBoost Model R² : 0.5708319028100357
```

“There is a performance difference between the two datasets, confirming that feature selection impacts model performance.”

Therefore, in the Final Presentation, We plan to decide which dataset to choose based on a comparison of their performances.

05. Preliminary Experiments

2) Check if the proposed model performs well

```
from sklearn.ensemble import RandomForestRegressor
from sklearn.metrics import mean_squared_error
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
import xgboost as xgb
from sklearn.metrics import mean_squared_error, r2_score

target = data01['BirthRate']
features = data01.drop(['BirthRate'], axis=1)

X_train, X_test, y_train, y_test = train_test_split(features, target, test_size=0.3, random_state=42

scaler = StandardScaler()
X_train_scaled = scaler.fit_transform(X_train)
X_test_scaled = scaler.transform(X_test)

xgboost_model = xgb.XGBRegressor(objective ='reg:squarederror',
                                 colsample_bytree = 0.3,
                                 learning_rate = 0.1,
                                 max_depth = 5,
                                 alpha = 10,
                                 n_estimators = 10)

xgboost_model.fit(X_train_scaled, y_train)

y_pred_xgb = xgboost_model.predict(X_test_scaled)
mse_xgb = mean_squared_error(y_test, y_pred_xgb)
r2_xgb = r2_score(y_test, y_pred_xgb)

print(f"XGBoost Model MSE: {mse_xgb}")
print(f"XGBoost Model R² : {r2_xgb}")
```

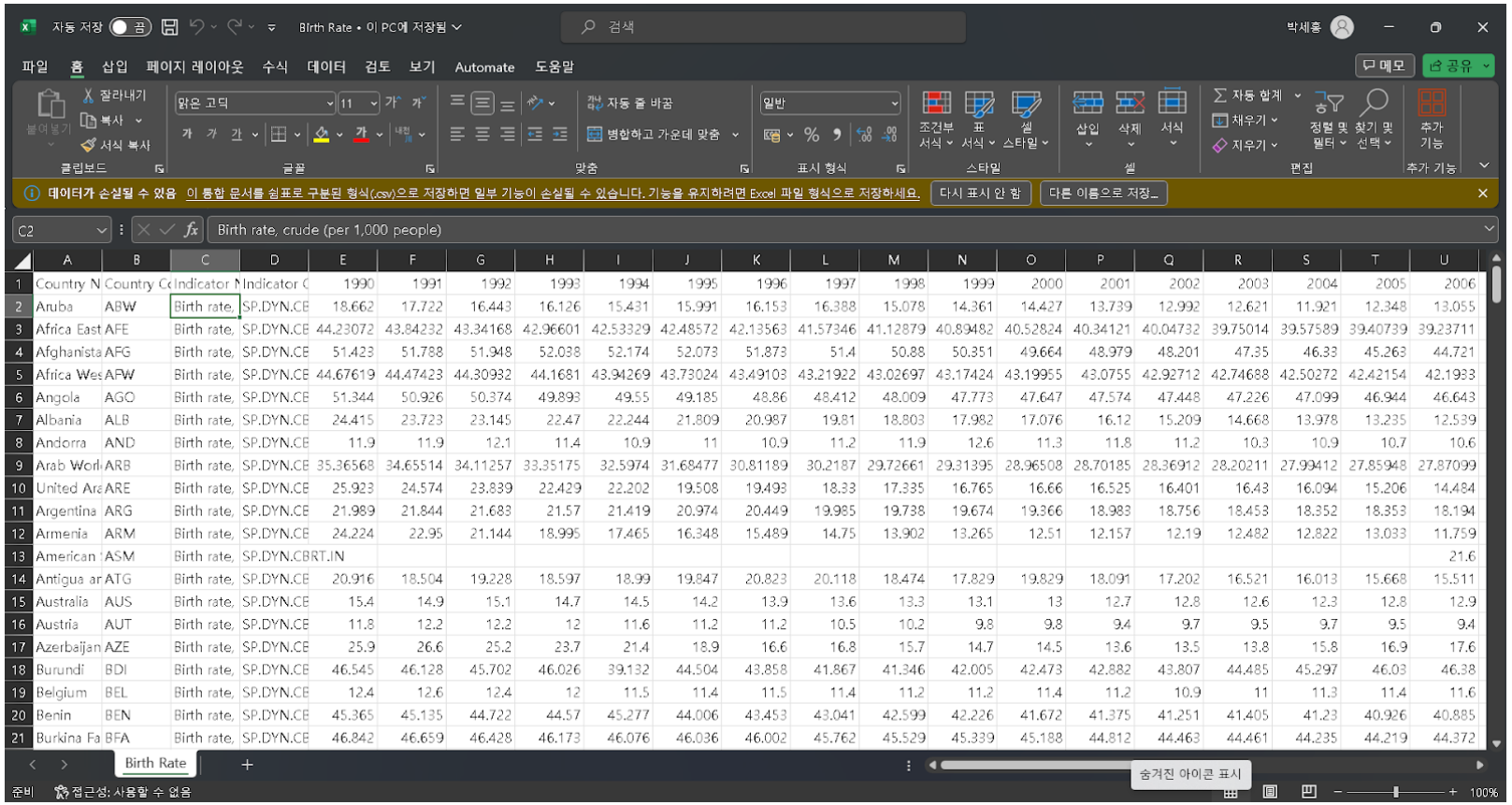
```
XGBoost Model MSE: 1.0353283823561519
XGBoost Model R² : 0.5221714462379863
```

"It has been determined that the XGBoost model, initially proposed during the Proposal phase, is not performing adequately. (High MSE, Low R^2)"

↓↓

Consequently, for the Final Presentation, the plan is to experiment with various models, including RandomForest, LinearRegression and etc, to identify the best performing model.

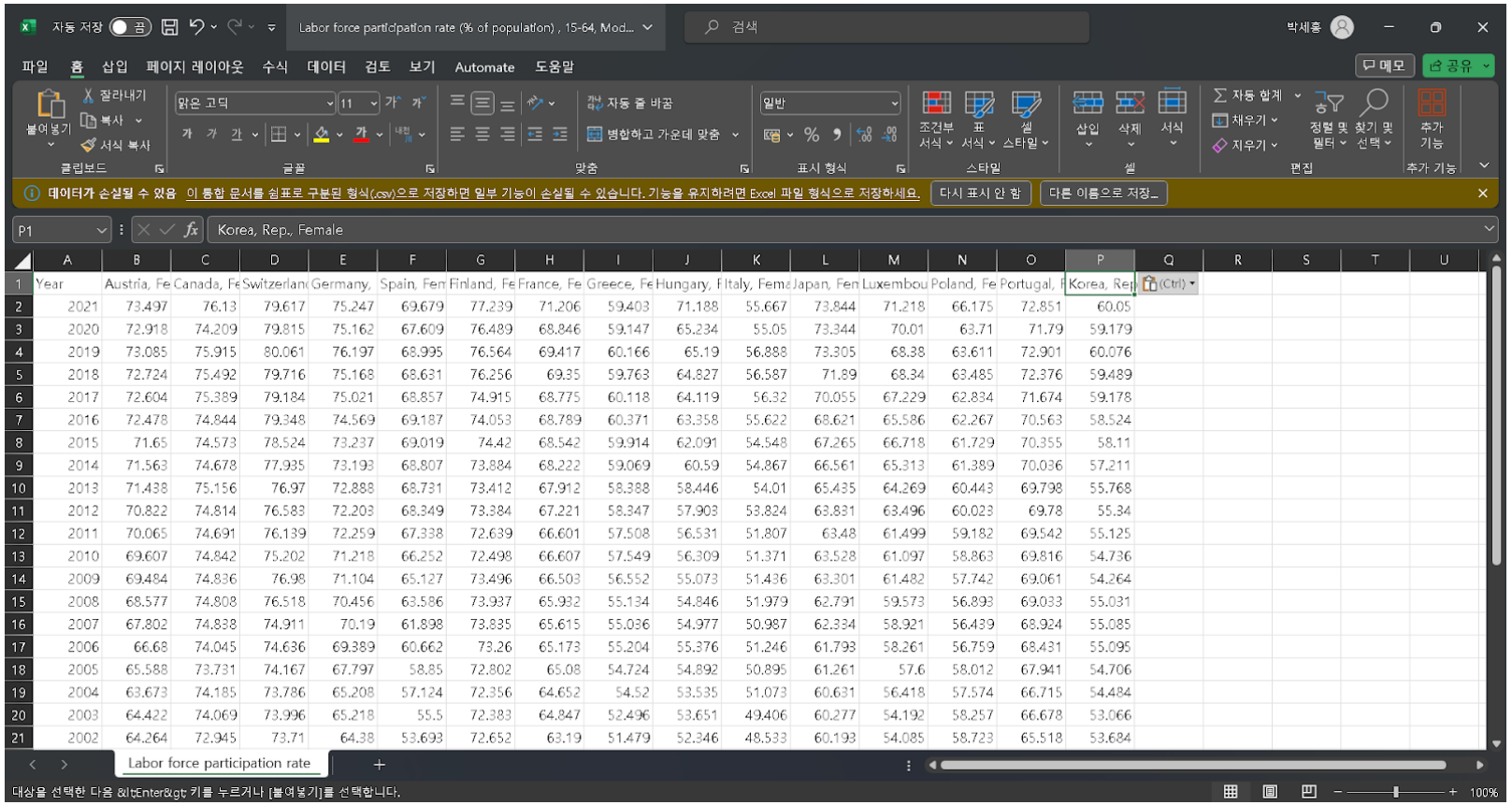
Appendix



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	Country N	Country Co	Indicator N	Indicator C	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
2	Aruba	ABW	Birth rate,	SP.DYN.CE	18.662	17.722	16.443	16.126	15.431	15.991	16.153	16.388	15.078	14.361	14.427	13.739	12.992	12.621	11.921	12.348	13.055
3	Africa East	AFE	Birth rate,	SP.DYN.CE	44.23072	43.84232	43.34168	42.96601	42.53329	42.48572	42.13563	41.57346	41.12879	40.89482	40.52824	40.34121	40.04732	39.75014	39.57589	39.40739	39.23711
4	Afghanista	AFG	Birth rate,	SP.DYN.CE	51.423	51.788	51.948	52.038	52.174	52.073	51.873	51.4	50.88	50.351	49.664	48.979	48.201	47.35	46.33	45.263	44.721
5	Africa Wes	AFW	Birth rate,	SP.DYN.CE	44.67619	44.47423	44.30932	44.1681	43.94269	43.73024	43.49103	43.21922	43.02697	43.17424	43.19955	43.0755	42.92712	42.74688	42.50272	42.42154	42.1933
6	Angola	AGO	Birth rate,	SP.DYN.CE	51.344	50.926	50.374	49.893	49.55	49.185	48.86	48.412	48.009	47.773	47.647	47.574	47.448	47.226	47.099	46.944	46.643
7	Albania	ALB	Birth rate,	SP.DYN.CE	24.415	23.723	23.145	22.47	22.244	21.809	20.987	19.81	18.803	17.982	17.076	16.12	15.209	14.668	13.978	13.235	12.539
8	Andorra	AND	Birth rate,	SP.DYN.CE	11.9	11.9	12.1	11.4	10.9	11	10.9	11.2	11.9	12.6	11.3	11.8	11.2	10.3	10.9	10.7	10.6
9	Arab Worl	ARB	Birth rate,	SP.DYN.CE	35.36568	34.65514	34.11257	33.35175	32.5974	31.68477	30.81189	30.2187	29.72661	29.31395	28.96508	28.70185	28.36912	28.20211	27.99412	27.85948	27.87099
10	United Ara	ARE	Birth rate,	SP.DYN.CE	25.923	24.574	23.839	22.429	22.202	19.508	19.493	18.33	17.335	16.765	16.66	16.525	16.401	16.43	16.094	15.206	14.484
11	Argentina	ARG	Birth rate,	SP.DYN.CE	21.989	21.844	21.683	21.57	21.419	20.974	20.449	19.985	19.738	19.674	19.366	18.983	18.756	18.453	18.352	18.353	18.194
12	Armenia	ARM	Birth rate,	SP.DYN.CE	24.224	22.95	21.144	18.995	17.465	16.348	15.489	14.75	13.902	13.265	12.51	12.157	12.19	12.482	12.822	13.033	11.759
13	American	ASM	Birth rate,	SP.DYN.CBRT.IN																	21.6
14	Antigua ar	ATG	Birth rate,	SP.DYN.CE	20.916	18.504	19.228	18.597	18.99	19.847	20.823	20.118	18.474	17.829	19.829	18.091	17.202	16.521	16.013	15.668	15.511
15	Australia	AUS	Birth rate,	SP.DYN.CE	15.4	14.9	15.1	14.7	14.5	14.2	13.9	13.6	13.3	13.1	13	12.7	12.8	12.6	12.3	12.8	12.9
16	Austria	AUT	Birth rate,	SP.DYN.CE	11.8	12.2	12.2	12	11.6	11.2	11.2	10.5	10.2	9.8	9.8	9.4	9.7	9.5	9.7	9.5	9.4
17	Azerbaijan	AZE	Birth rate,	SP.DYN.CE	25.9	26.6	25.2	23.7	21.4	18.9	16.6	16.8	15.7	14.7	14.5	13.6	13.5	13.8	15.8	16.9	17.6
18	Burundi	BDI	Birth rate,	SP.DYN.CE	46.545	46.128	45.702	46.026	39.132	44.504	43.858	41.867	41.346	42.005	42.473	42.882	43.807	44.485	45.297	46.03	46.38
19	Belgium	BEL	Birth rate,	SP.DYN.CE	12.4	12.6	12.4	12	11.5	11.4	11.5	11.4	11.2	11.2	11.4	11.2	10.9	11	11.3	11.4	11.6
20	Benin	BEN	Birth rate,	SP.DYN.CE	45.365	45.135	44.722	44.57	45.277	44.006	43.453	43.041	42.599	42.226	41.672	41.375	41.251	41.405	41.23	40.926	40.885
21	Burkina Fa	BFA	Birth rate,	SP.DYN.CE	46.842	46.659	46.428	46.173	46.076	46.036	46.002	45.762	45.529	45.339	45.188	44.812	44.463	44.461	44.235	44.219	44.372

Ex) 'BirthRate': Birth rate crude (per 1,000 people)
https://data.worldbank.org/indicator/SP.DYN.CBRT.IN?end=2021&most_recent_year_desc=false&start=1960&view=chart

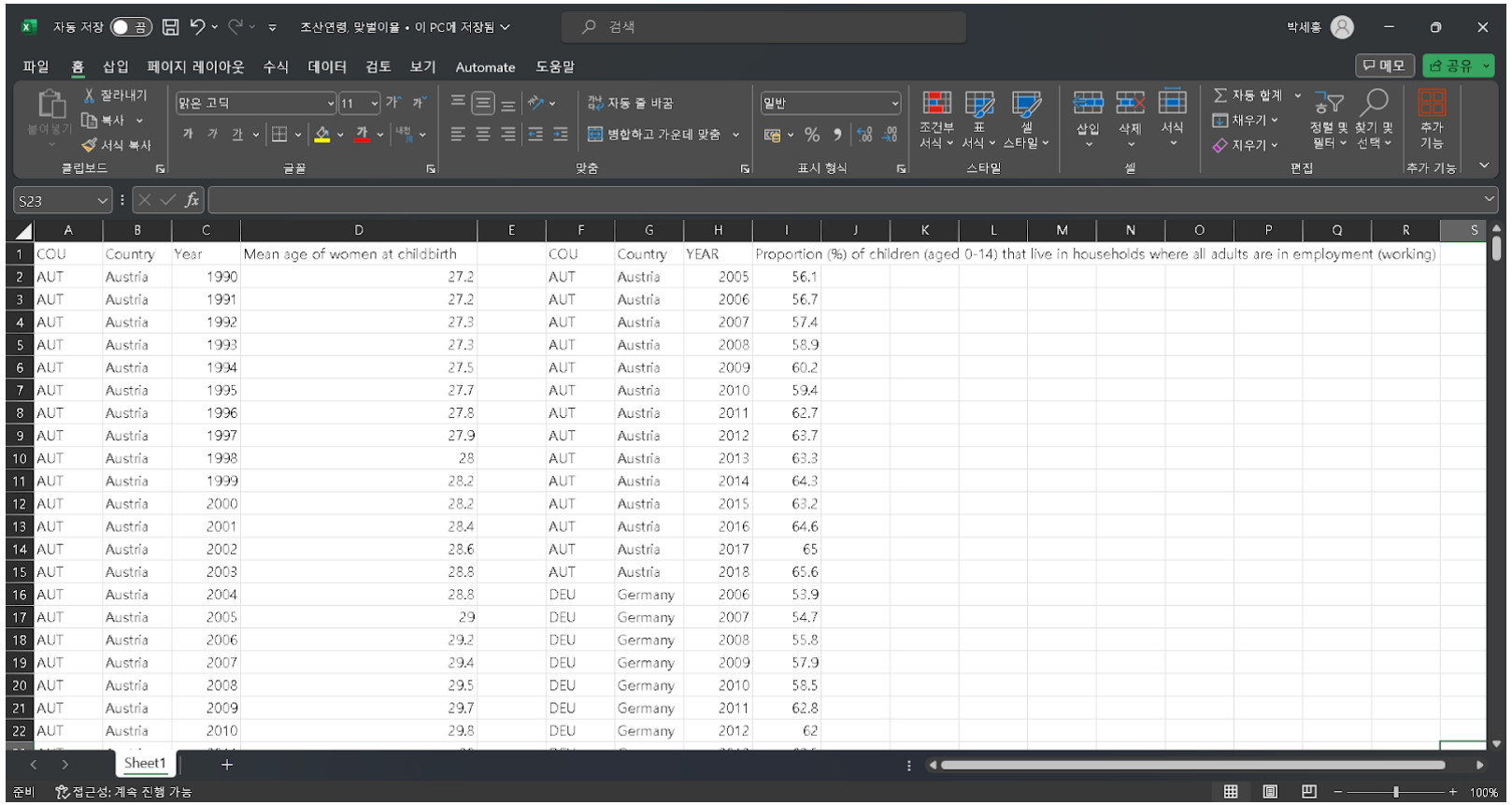
Appendix



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	Year	Austria, Fe	Canada, Fe	Switzerlan	Germany,	Spain, Fen	Finland, Fe	France, Fe	Greece, Fe	Hungary, F	Italy, Fema	Japan, Fen	Luxembou	Poland, Fe	Portugal, F	Korea, Rep	(Ctrl)
2	2021	73.497	76.13	79.617	75.247	69.679	77.239	71.206	59.403	71.188	55.667	73.844	71.218	66.175	72.851	60.05	
3	2020	72.918	74.209	79.815	75.162	67.609	76.489	68.846	59.147	65.234	55.05	73.344	70.01	63.71	71.79	59.179	
4	2019	73.085	75.915	80.061	76.197	68.995	76.564	69.417	60.166	65.19	56.888	73.305	68.38	63.611	72.901	60.076	
5	2018	72.724	75.492	79.716	75.168	68.631	76.256	69.35	59.763	64.827	56.587	71.89	68.34	63.485	72.376	59.489	
6	2017	72.604	75.389	79.184	75.021	68.857	74.915	68.775	60.118	64.119	56.32	70.055	67.229	62.834	71.674	59.178	
7	2016	72.478	74.844	79.348	74.569	69.187	74.053	68.789	60.371	63.358	55.622	68.621	65.586	62.267	70.563	58.524	
8	2015	71.65	74.573	78.524	73.237	69.019	74.42	68.542	59.914	62.091	54.548	67.265	66.718	61.729	70.355	58.11	
9	2014	71.563	74.678	77.935	73.193	68.807	73.884	68.222	59.069	60.59	54.867	66.561	65.313	61.389	70.036	57.211	
10	2013	71.438	75.156	76.97	72.888	68.731	73.412	67.912	58.388	58.446	54.01	65.435	64.269	60.443	69.798	55.768	
11	2012	70.822	74.814	76.583	72.203	68.349	73.384	67.221	58.347	57.903	53.824	63.831	63.496	60.023	69.78	55.34	
12	2011	70.065	74.691	76.139	72.259	67.338	72.639	66.601	57.508	56.531	51.807	63.48	61.499	59.182	69.542	55.125	
13	2010	69.607	74.842	75.202	71.218	66.252	72.498	66.607	57.549	56.309	51.371	63.528	61.097	58.863	69.816	54.736	
14	2009	69.484	74.836	76.98	71.104	65.127	73.496	66.503	56.552	55.073	51.436	63.301	61.482	57.742	69.061	54.264	
15	2008	68.577	74.808	76.518	70.456	63.586	73.937	65.932	55.134	54.846	51.979	62.791	59.573	56.893	69.033	55.031	
16	2007	67.802	74.838	74.911	70.19	61.898	73.835	65.615	55.036	54.977	50.987	62.334	58.921	56.439	68.924	55.085	
17	2006	66.68	74.045	74.636	69.389	60.662	73.26	65.173	55.204	55.376	51.246	61.793	58.261	56.759	68.431	55.095	
18	2005	65.588	73.731	74.167	67.797	58.85	72.802	65.08	54.724	54.892	50.895	61.261	57.6	58.012	67.941	54.706	
19	2004	63.673	74.185	73.786	65.208	57.124	72.356	64.652	54.52	53.535	51.073	60.631	56.418	57.574	66.715	54.484	
20	2003	64.422	74.069	73.996	65.218	55.5	72.383	64.847	52.496	53.651	49.406	60.277	54.192	58.257	66.678	53.066	
21	2002	64.264	72.945	73.71	64.38	53.693	72.652	63.19	51.479	52.346	48.533	60.193	54.085	58.723	65.518	53.684	

Ex) 'FemaleLaborParticipationRate' : Labor force participation rate (% of population) , 15-64, Modeled
https://genderdata.worldbank.org/indicators/sl-tlf-acti-zs/?age=15-64&gender=female&geos=JPN_LUX_POL_PRT_KOR&view=trend

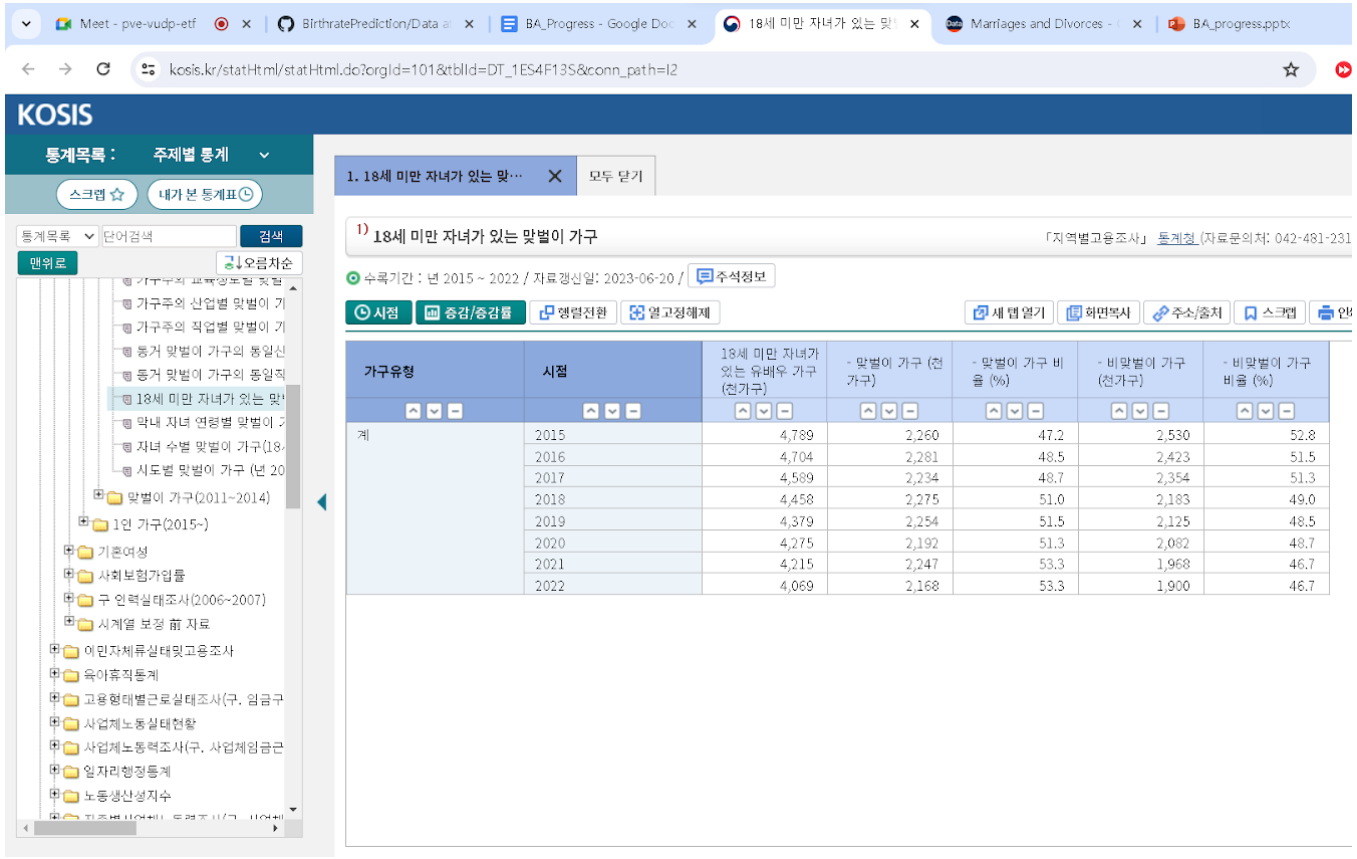
Appendix

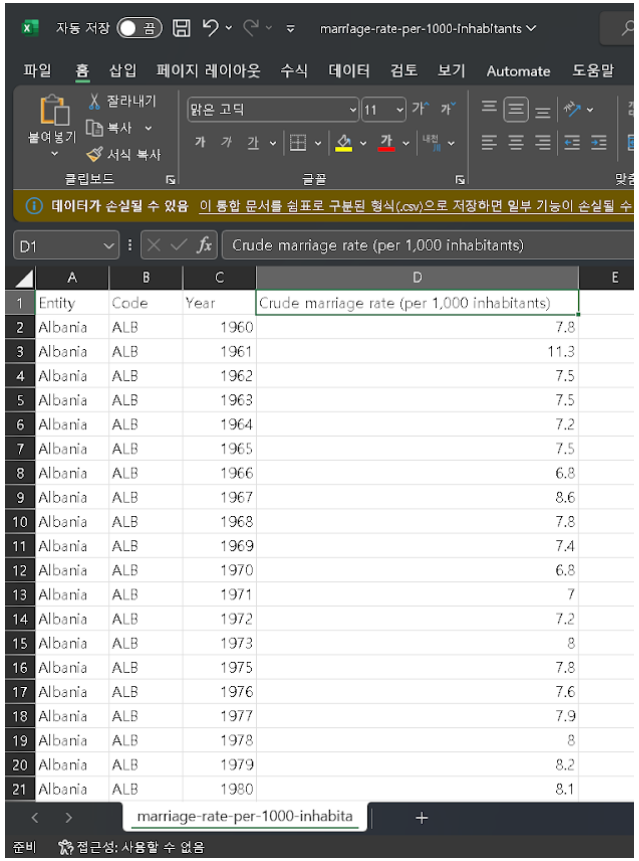


	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	COU	Country	Year	Mean age of women at childbirth		COU	Country	YEAR	Proportion (%) of children (aged 0-14) that live in households where all adults are in employment (working)										
2	AUT	Austria	1990	27.2		AUT	Austria	2005	56.1										
3	AUT	Austria	1991	27.2		AUT	Austria	2006	56.7										
4	AUT	Austria	1992	27.3		AUT	Austria	2007	57.4										
5	AUT	Austria	1993	27.3		AUT	Austria	2008	58.9										
6	AUT	Austria	1994	27.5		AUT	Austria	2009	60.2										
7	AUT	Austria	1995	27.7		AUT	Austria	2010	59.4										
8	AUT	Austria	1996	27.8		AUT	Austria	2011	62.7										
9	AUT	Austria	1997	27.9		AUT	Austria	2012	63.7										
10	AUT	Austria	1998	28		AUT	Austria	2013	63.3										
11	AUT	Austria	1999	28.2		AUT	Austria	2014	64.3										
12	AUT	Austria	2000	28.2		AUT	Austria	2015	63.2										
13	AUT	Austria	2001	28.4		AUT	Austria	2016	64.6										
14	AUT	Austria	2002	28.6		AUT	Austria	2017	65										
15	AUT	Austria	2003	28.8		AUT	Austria	2018	65.6										
16	AUT	Austria	2004	28.8		DEU	Germany	2006	53.9										
17	AUT	Austria	2005	29		DEU	Germany	2007	54.7										
18	AUT	Austria	2006	29.2		DEU	Germany	2008	55.8										
19	AUT	Austria	2007	29.4		DEU	Germany	2009	57.9										
20	AUT	Austria	2008	29.5		DEU	Germany	2010	58.5										
21	AUT	Austria	2009	29.7		DEU	Germany	2011	62.8										
22	AUT	Austria	2010	29.8		DEU	Germany	2012	62										

Ex) 'FirstBirthAge', 'BothWorking': Mean age of women at childbirth, Proportion (%) of children (aged 0-14) that live in households where all adults are in employment (working)
https://stats.oecd.org/Index.aspx?DataSetCode=FAMILY

Appendix





Ex) 'BothWorking' for Korea missing value: KOSIS 맞벌이 가구
https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1ES4F13S&conn_path=I2

Ex) 'MarriageRate' : Crude marriage rate (per 1,000 inhabitants)
https://ourworldindata.org/marriages-and-divorces

Appendix

- Sources

- https://data.worldbank.org/indicator/SP.DYN.CBRT.IN?end=2021&most_recent_year_desc=false&start=1960&view=chart
- https://genderdata.worldbank.org/indicators/sl-tlf-acti-zs/?age=15-64&gender=female&geos=JPN_LUX_POL_PRT_KOR&view=trend
- https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1ES4F13S&conn_path=I2
- https://stats.oecd.org/Index.aspx?DataSetCode=FAMILY
- https://genderdata.worldbank.org/indicators/sp-dyn-smam/?geos=WLD_JPN_AUS_CHN_KOR&view=trend
- https://ourworldindata.org/marriages-and-divorces
- https://data.oecd.org/unemp/unemployment-rate.htm#indicator-chart
- https://data.oecd.org/price/housing-prices.htm
- https://data.oecd.org/interest/short-term-interest-rates.htm#indicator-chart
- https://data.oecd.org/emp/part-time-employment-rate.htm#indicator-chart

Appendix

- Sources
  - https://data.oecd.org/socialexp/family-benefits-public-spending.htm
  - https://data.worldbank.org/indicator/SH.XPD.CHEX.GD.ZS?end=2021&start=2000&view=chart
  - https://ourworldindata.org/grapher/gender-development-index?tab=table
  - https://data.oecd.org/gdp/gross-domestic-product-gdp.htm
  - https://data.oecd.org/natincome/gross-national-income.htm
  - https://data.worldbank.org/indicator/SI.POV.UMIC.GP
  - https://data.worldbank.org/indicator/SE.XPD.TOTL.GD.ZS
  - https://data.worldbank.org/indicator/SE.XPD.TOTL.GB.ZS
  - https://stats.oecd.org/Index.aspx?DataSetCode=FAMILY#
  - https://genderdata.worldbank.org/indicators/sl-tlf-acti-zs/?age=15-64&gender=female&geos=JPN_LUX_POL_PRT_KOR&view=trend

Q & A