Chapter 2

Visualization of Categorical Data

Professor Jung Jin Lee

Chapter 2 Visualization of Categorical Data

2.1 Graphs of Categorical Data

2.2 Visualization of Categorical Summary Data

2.3 Visualization of Categorical Raw Data

2.1 Graphs of Categorical Data

- **Categorical Data** : Data of a variable on finite possible value

Ex) Data on gender in a class

=〉 {male, male, female, … } raw data

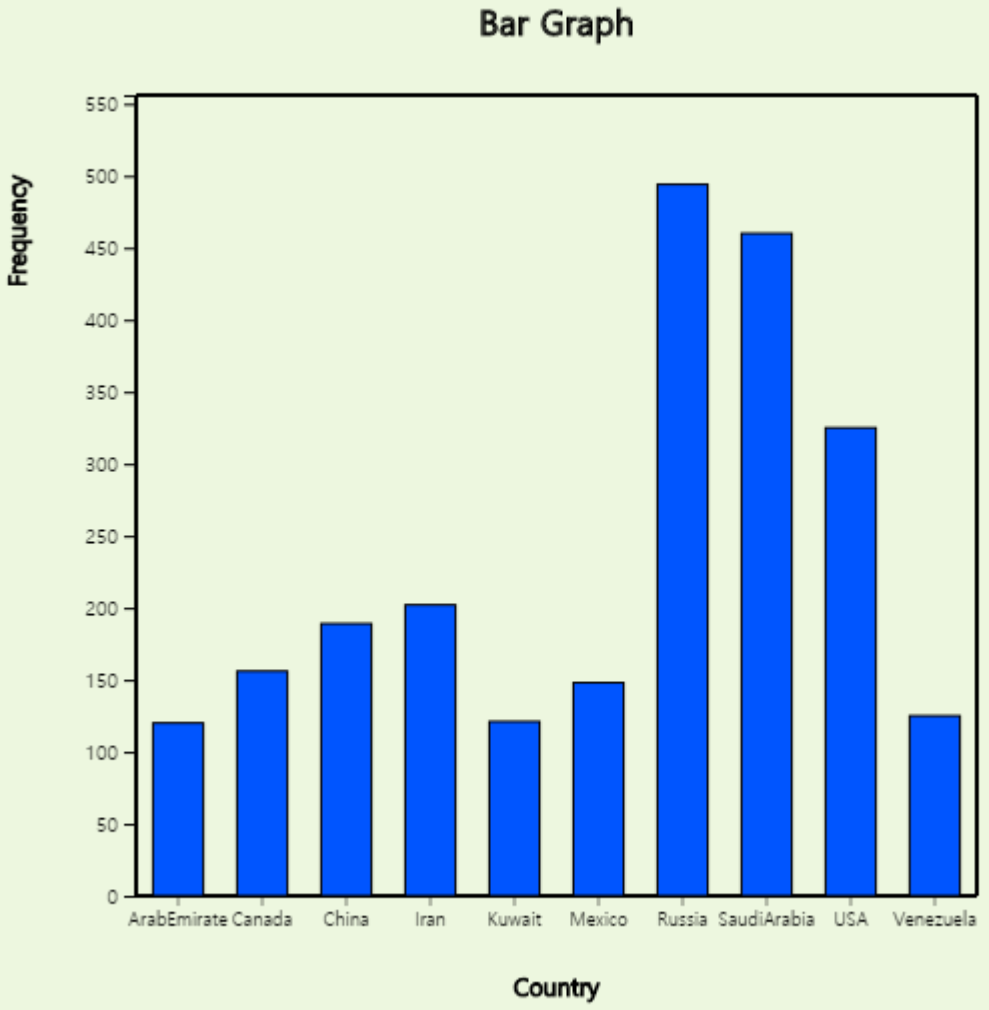
=〉 male 10 summary data

female 15

- **Bar chart, pie chart, band graph, and line graph** are used for visualizing categorical data.

2.1 Graphs of Categorical Data

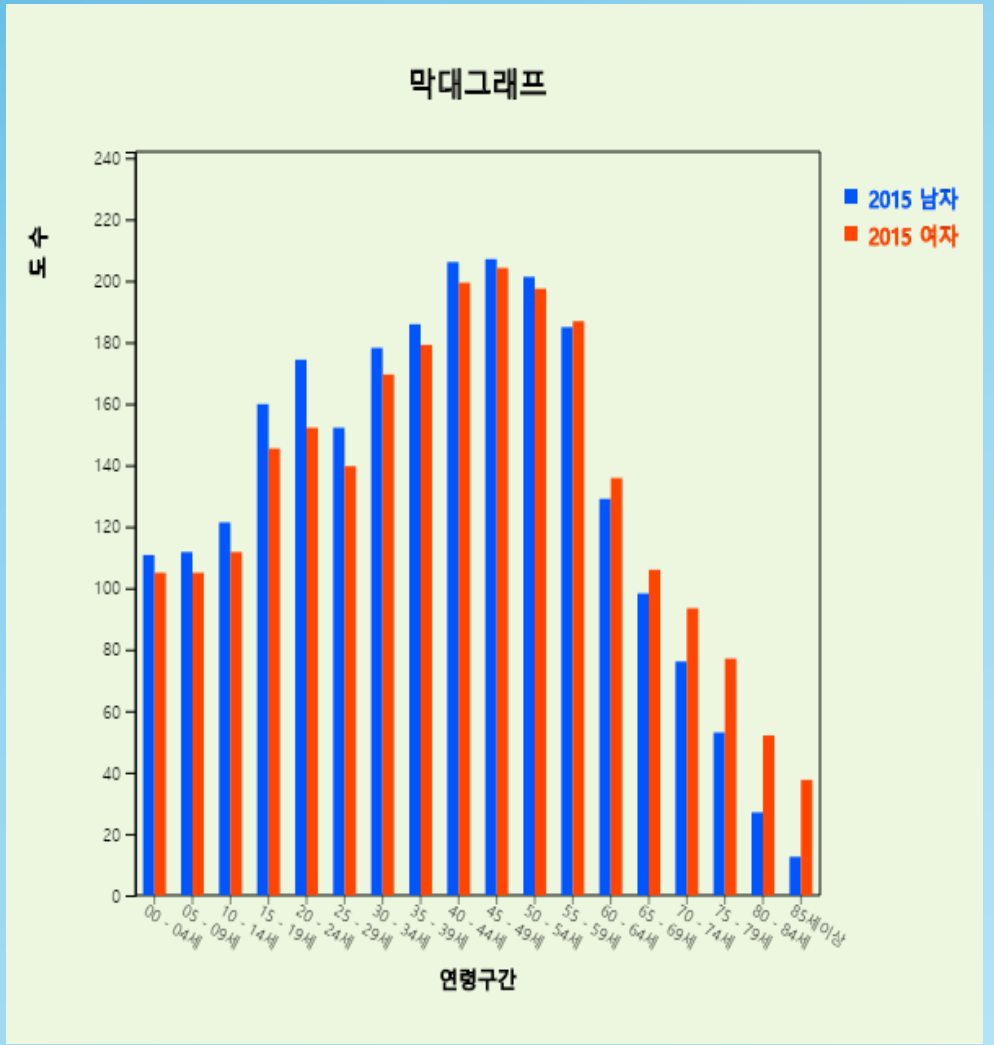
- Bar chart shows the frequency of each category as the height of the bar to compare the frequency of the data for each category.
  - Spacing the bars to emphasize that they are categorical data.



2.1 Graphs of Categorical Data

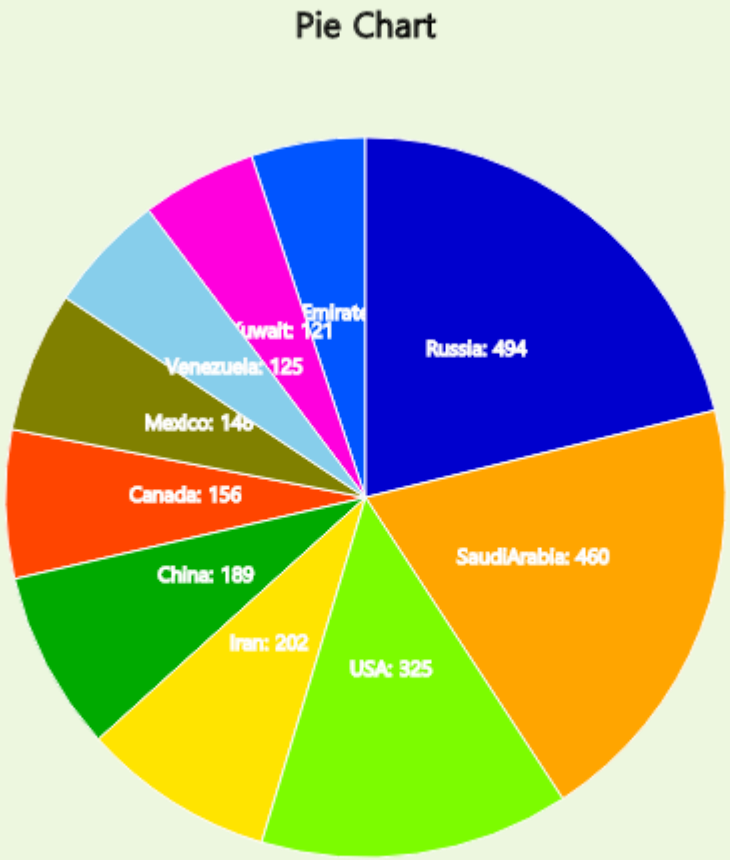
- Bar chart can be drawn by dividing the frequency of one categorical variable by another (called a group variable) such as gender.

- separated bar chart
- stacked bar chart
- ratio bar chart
- side by side bar chart



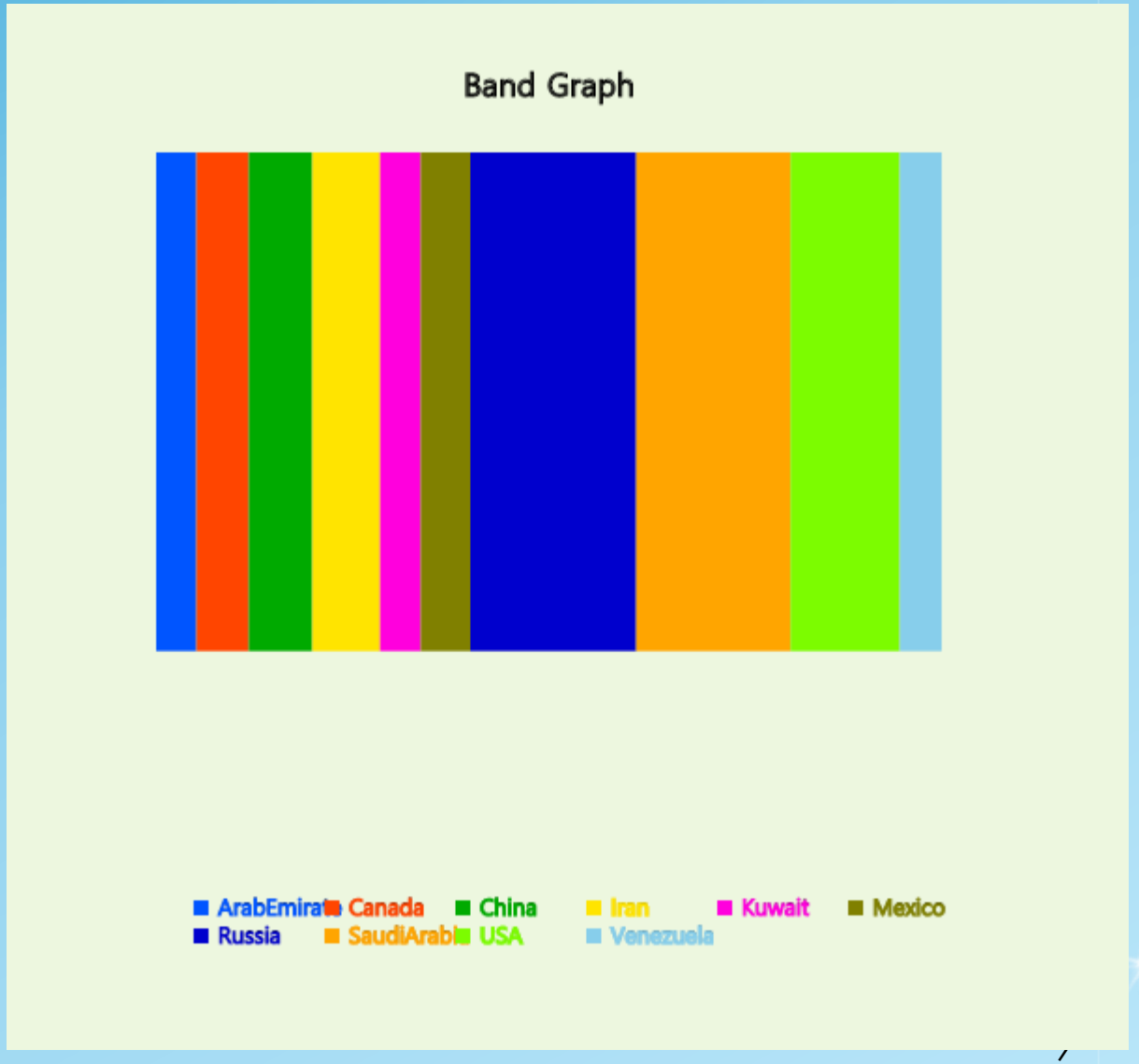
2.1 Graphs of Categorical Data

- Pie Chart shows the frequency of each category divided by the amount of aid angle.
- Draw the largest category of aid angles in a clockwise order so that the ratio can be compared well.
- Doughnut graph with the center of the circle empty is also used.



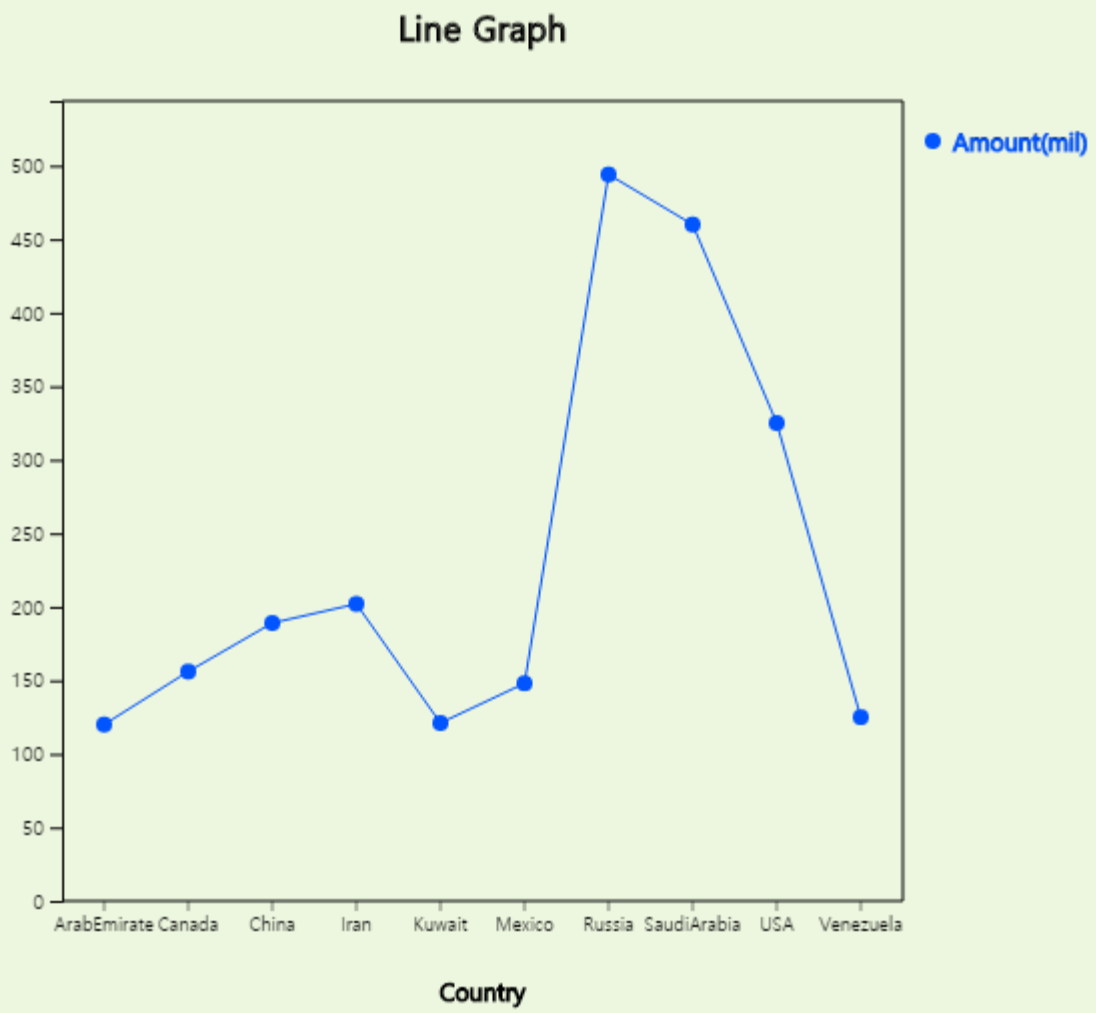
2.1 Graphs of Categorical Data

- **Band graph** is a graph that shows the ratio of frequency in each category to total data divided by square pieces in the form of a pie chart.



2.1 Graphs of Categorical Data

- Line graph shows the X-axis as a category value and the value of the other variables as a Y-value in relation to the values of each category, and then connects them as a line.
- It is similar to the bar chart, however, as the amount of exports per year, the change in data over time can be easily observed.



2.2 Visualization of Summary Data

2.2.1 Summary Data

- Investigate the gender of a class student and examine them for 'male', 'female', 'male' and 'male'...
- The recorded data is referred to as raw data, and the compilation of these in the form of a frequency distribution table is referred to as summary data.

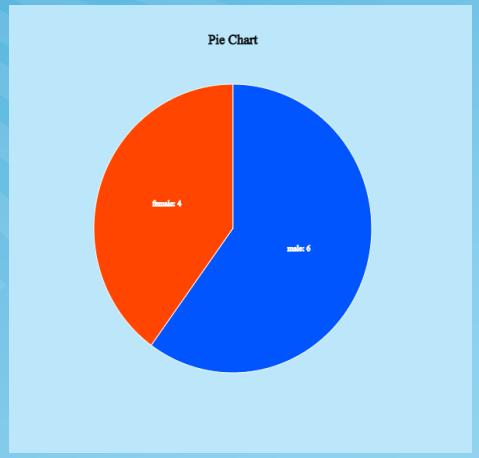
[EX 2.2.1] Summary data on gender of a class

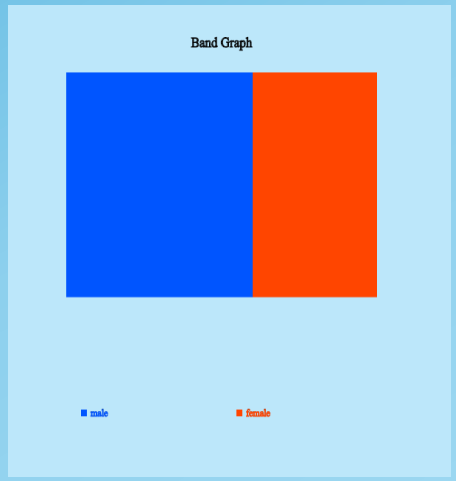
Gender	Student
Male	6
Female	4

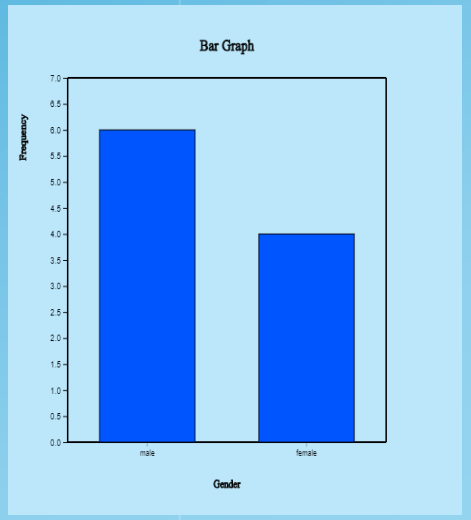
2.2 Visualization of Summary Data

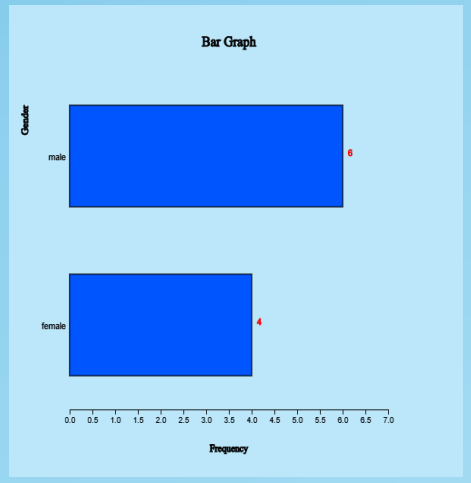
[EX 2.2.1] Summary data on gender of a class

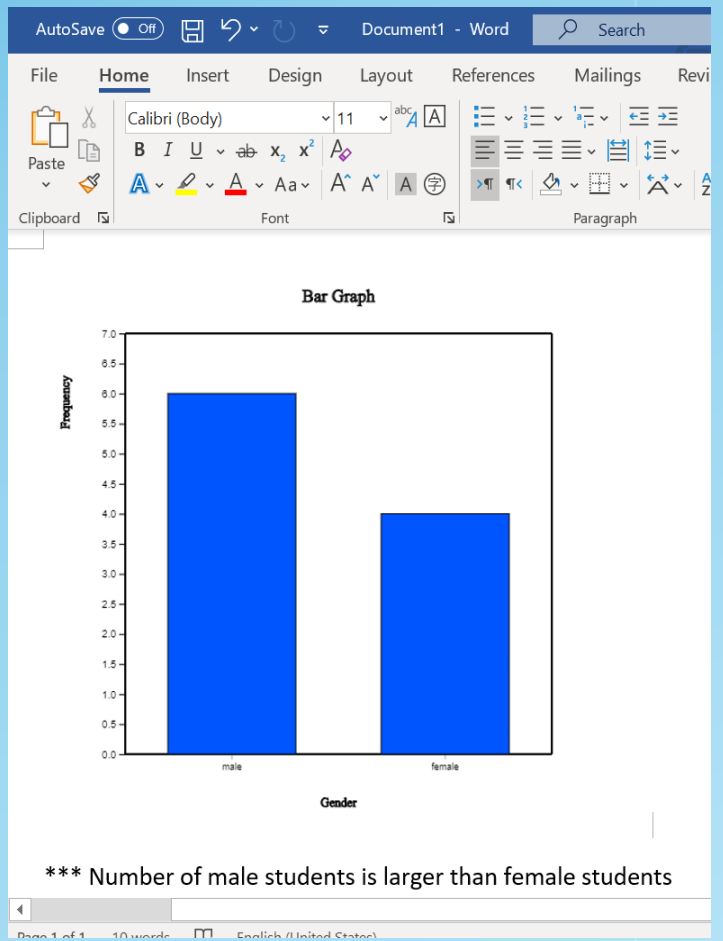
Gender	Student
Male	6
Female	4







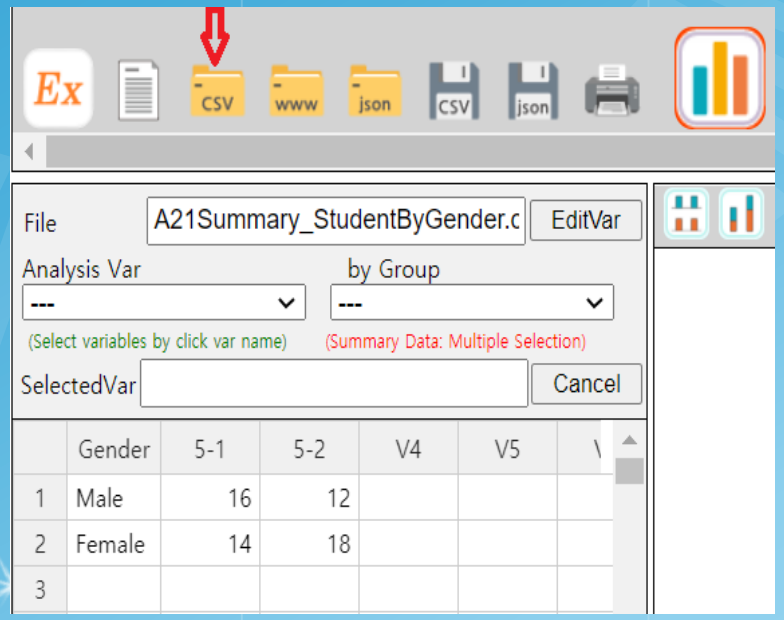


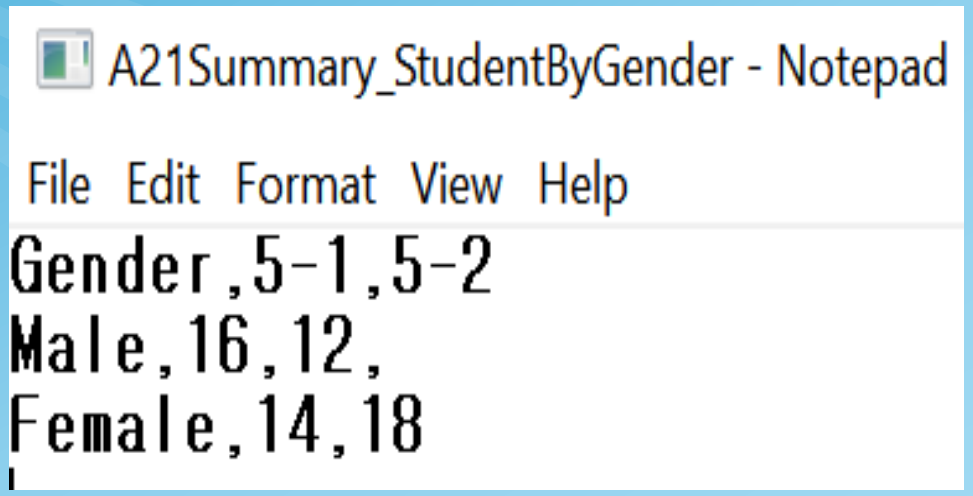


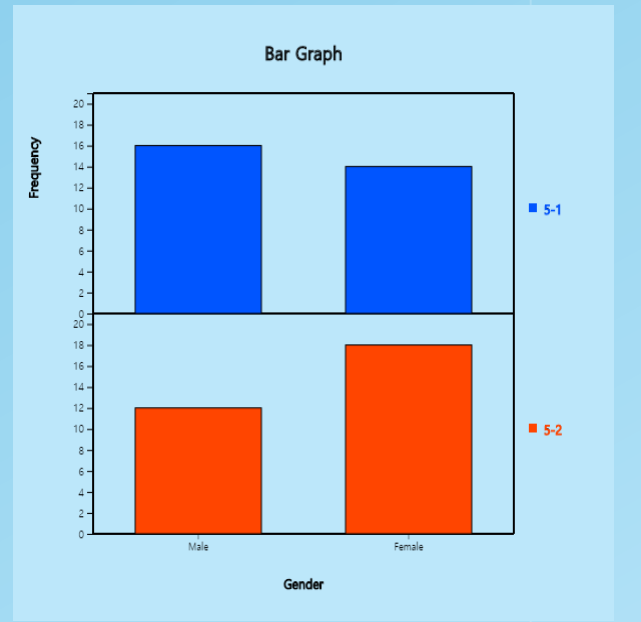
*** Number of male students is larger than female students

2.2 Visualization of Summary Data

[EX 2.2.3] Summary data on gender of two class







2.2 Visualization of Summary Data

2.2.1 Summary Data

- As the number of data increases, organizing raw data into summary data is not an easy task for the general public.
- One of the main functions of the statistical package is to organize raw data into summary data.
- Usually, textbooks or public institutions provide information to the public in the form of summary data.

2.2 Visualization of Summary Data

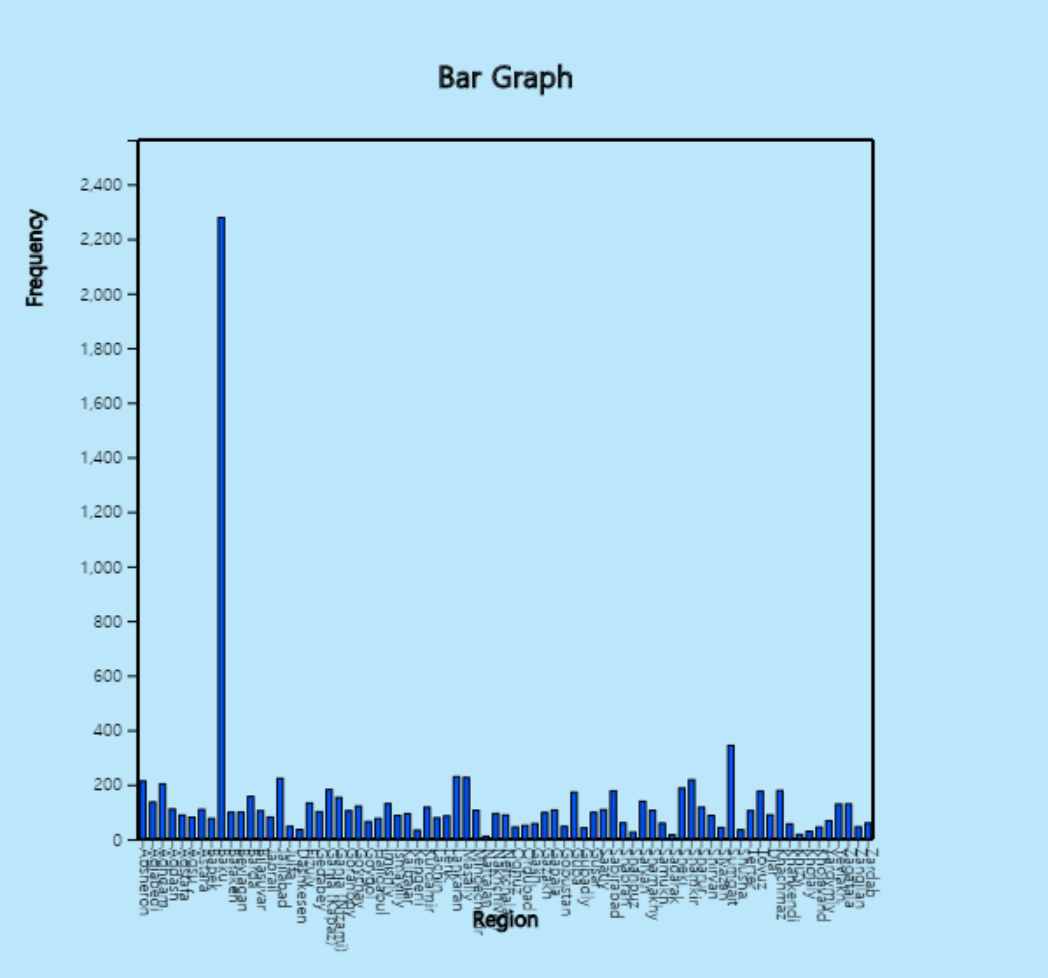
[Ex] (Regional Population of Azerbaijan)

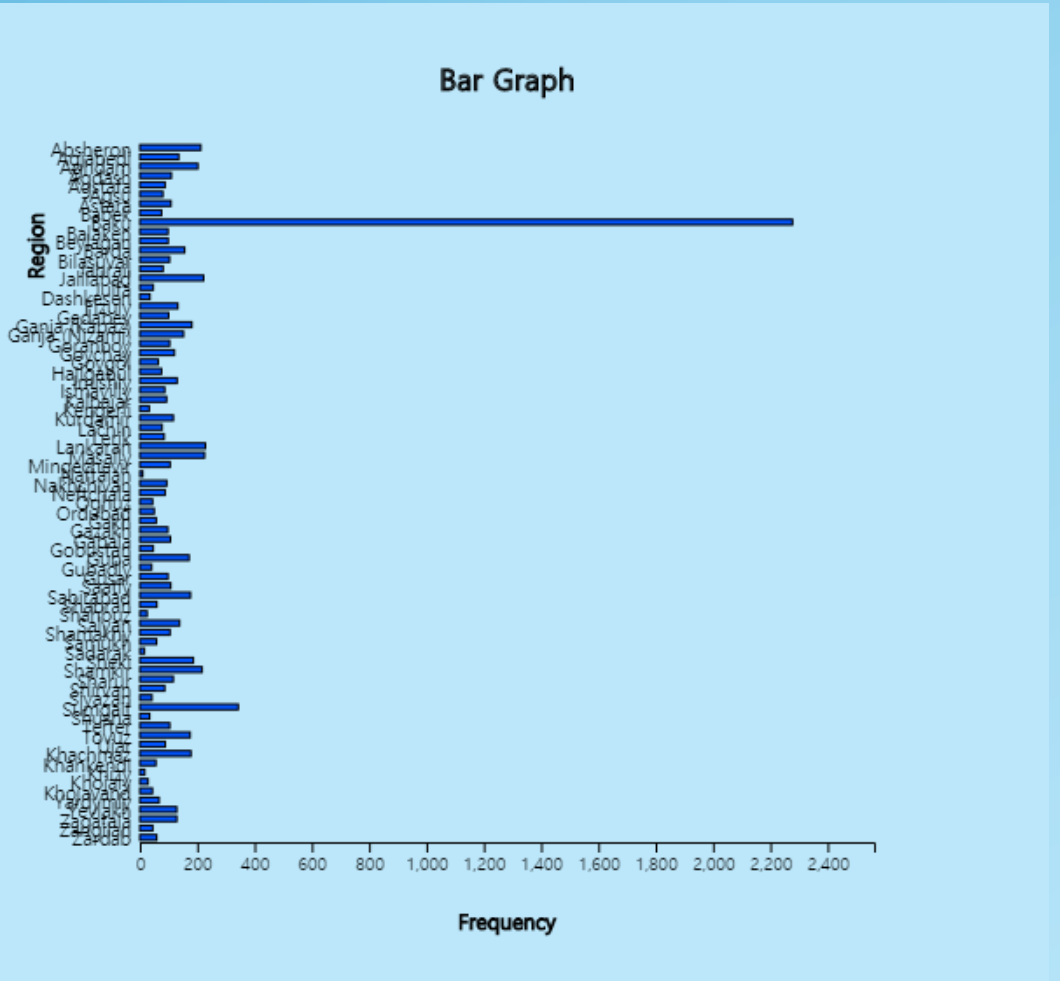
Using『eStat』draw bar graph, pie chart, band graph, line graph of regional population of Azerbaijan.

	Region	Populati
1	Abshero	212.6
2	Agjabed	135.5
3	Aghdam	202.2
4	Agdash	109.9
5	Agstafa	87.9
6	Agsu	80.1
7	Astara	108.6
8	Babek	75.6
9	Baku	2277.5
10	Balaken	98.3
11	Beylagar	98.6
12	Barda	156.3
13	Bilasuva	103.8
14	Jabrail	80.8
15	Jalilabac	222.4
16	Julfa	46.7
17	Dashkes	35.1
18	Fizuly	132.4

2.2 Visualization of Summary Data

[Ex] (Regional Population of Azerbaijan)

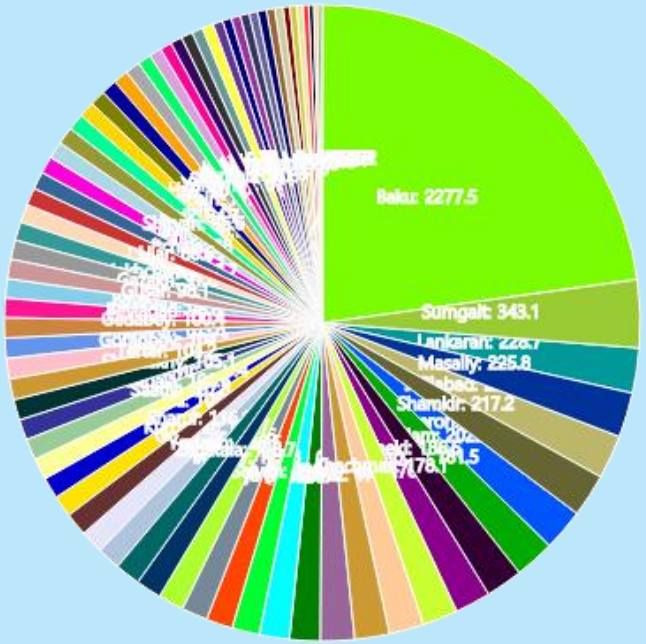




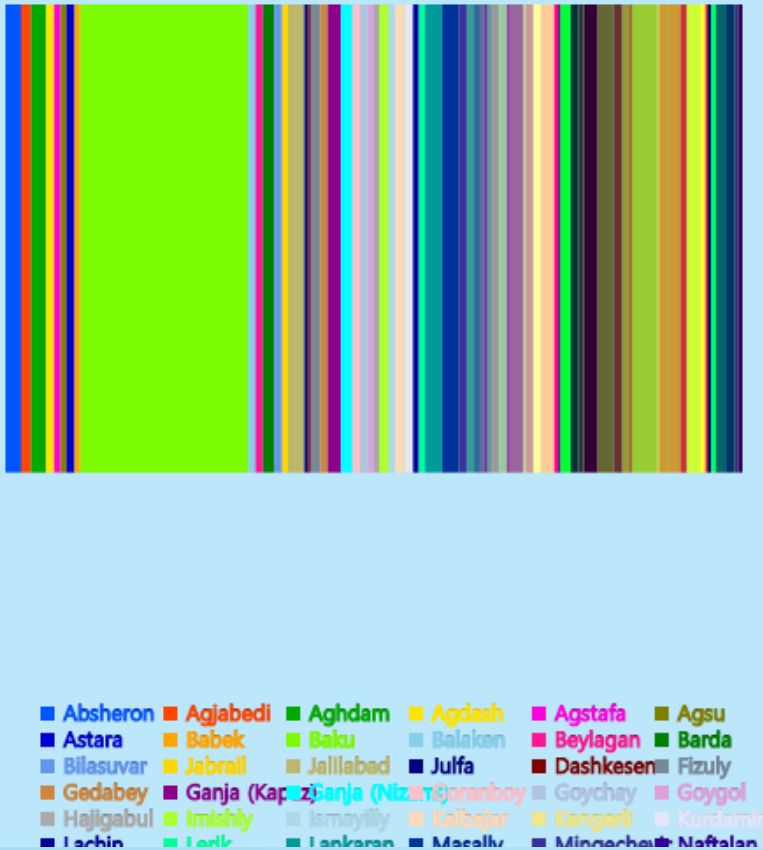
2.2 Visualization of Summary Data

[Ex (Regional Population of Azerbaijan)

Pie Chart

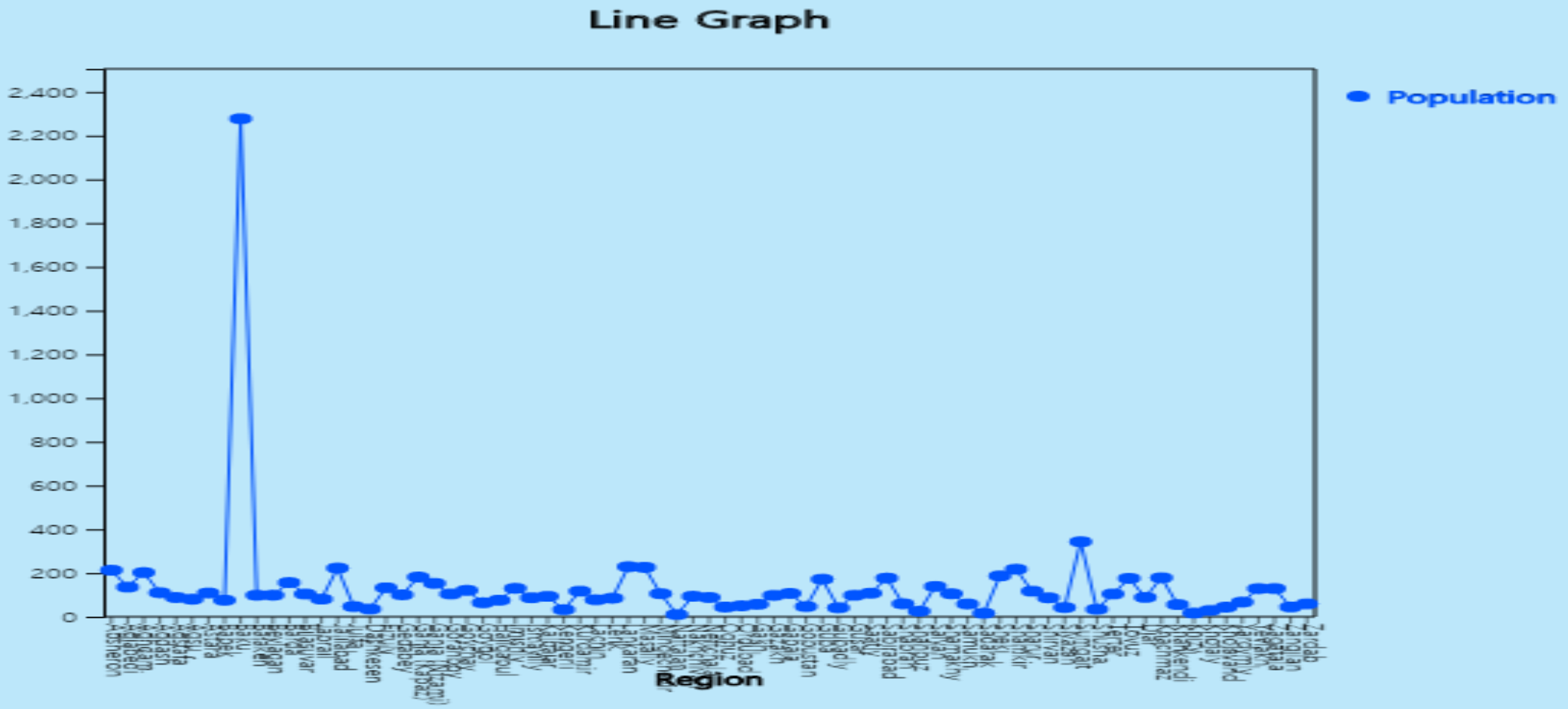


Band Graph



2.2 Visualization of Summary Data

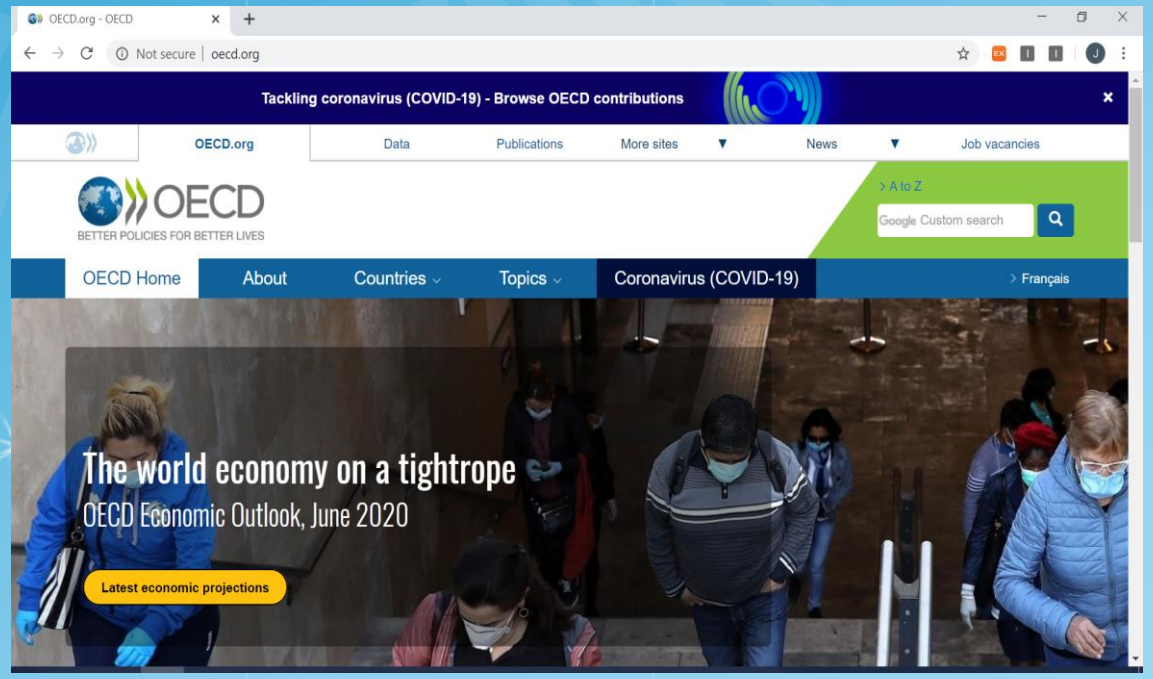
[Ex] (Regional Population of Azerbaijan)

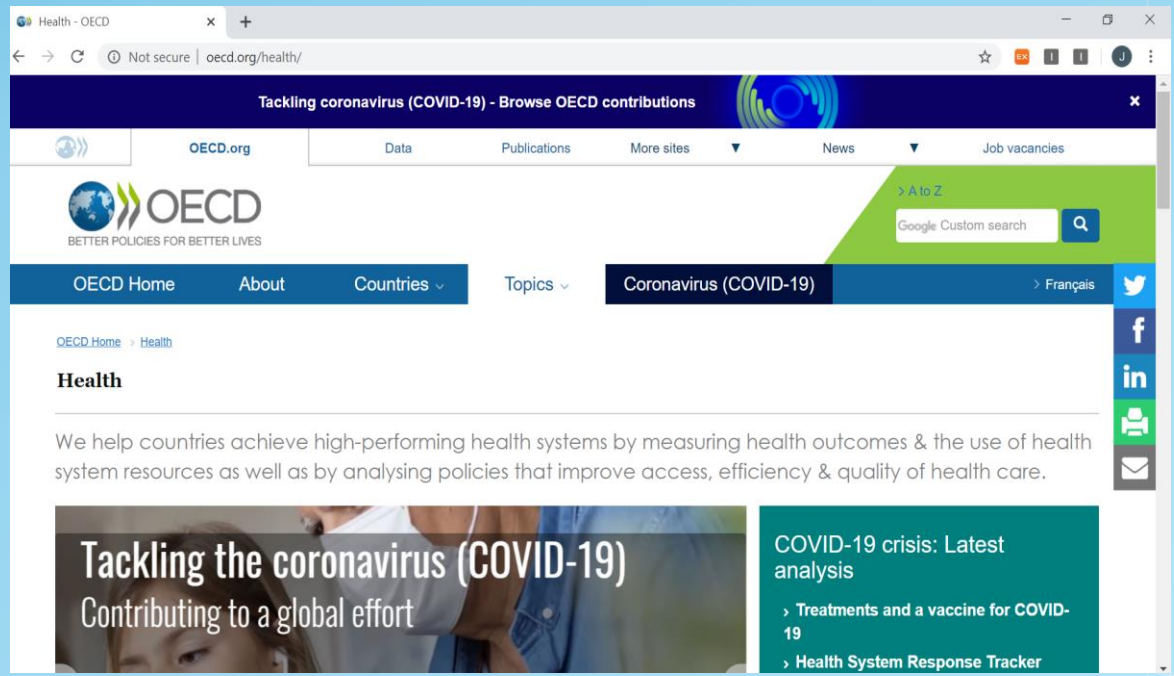


2.2 Visualization of Summary Data

[Ex 2.2.2] (Life Expectancy at Birth : OECD)

https://www.oecd.org

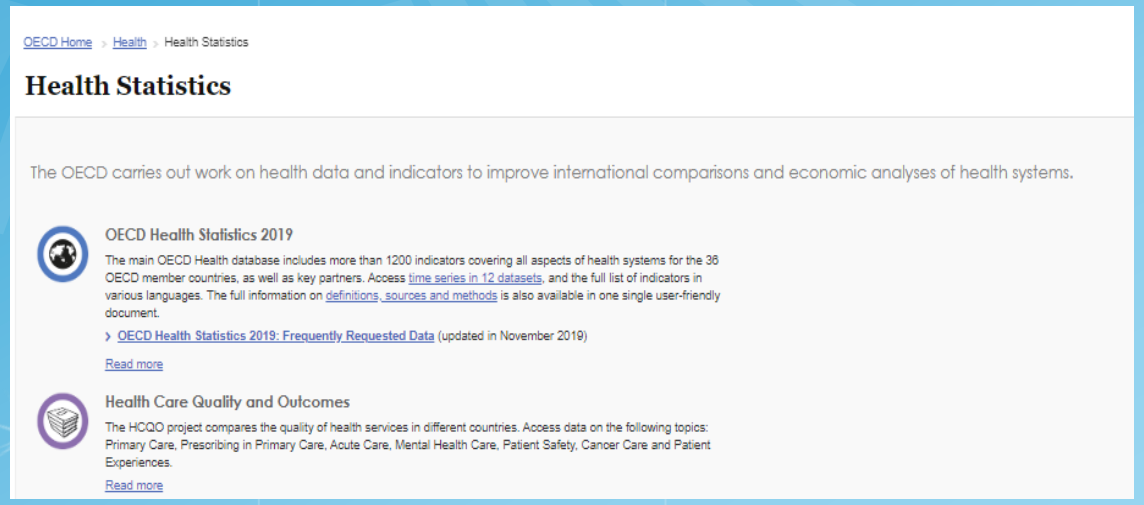


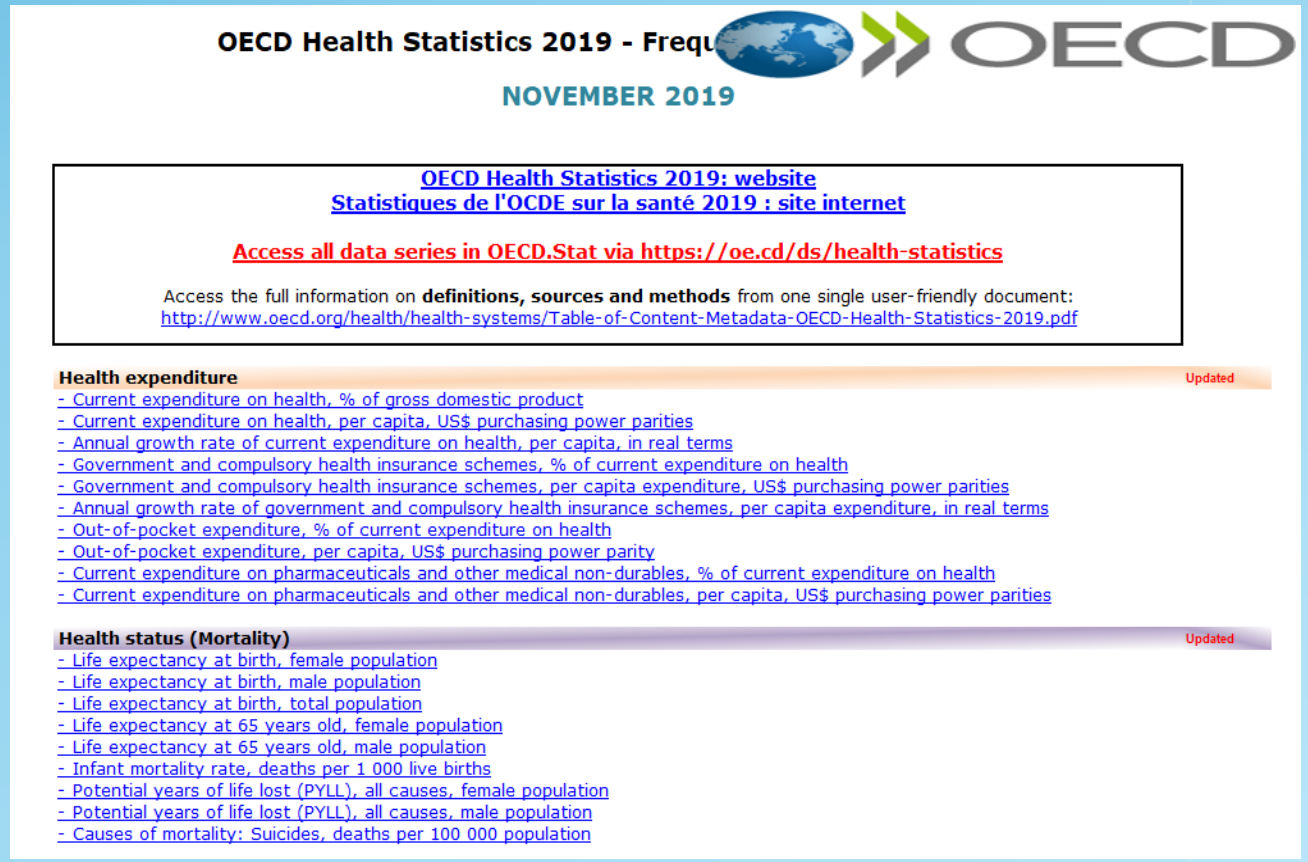


2.2 Visualization of Summary Data

[Ex 2.2.2] (Life Expectancy at Birth : OECD)

https://www.oecd.org

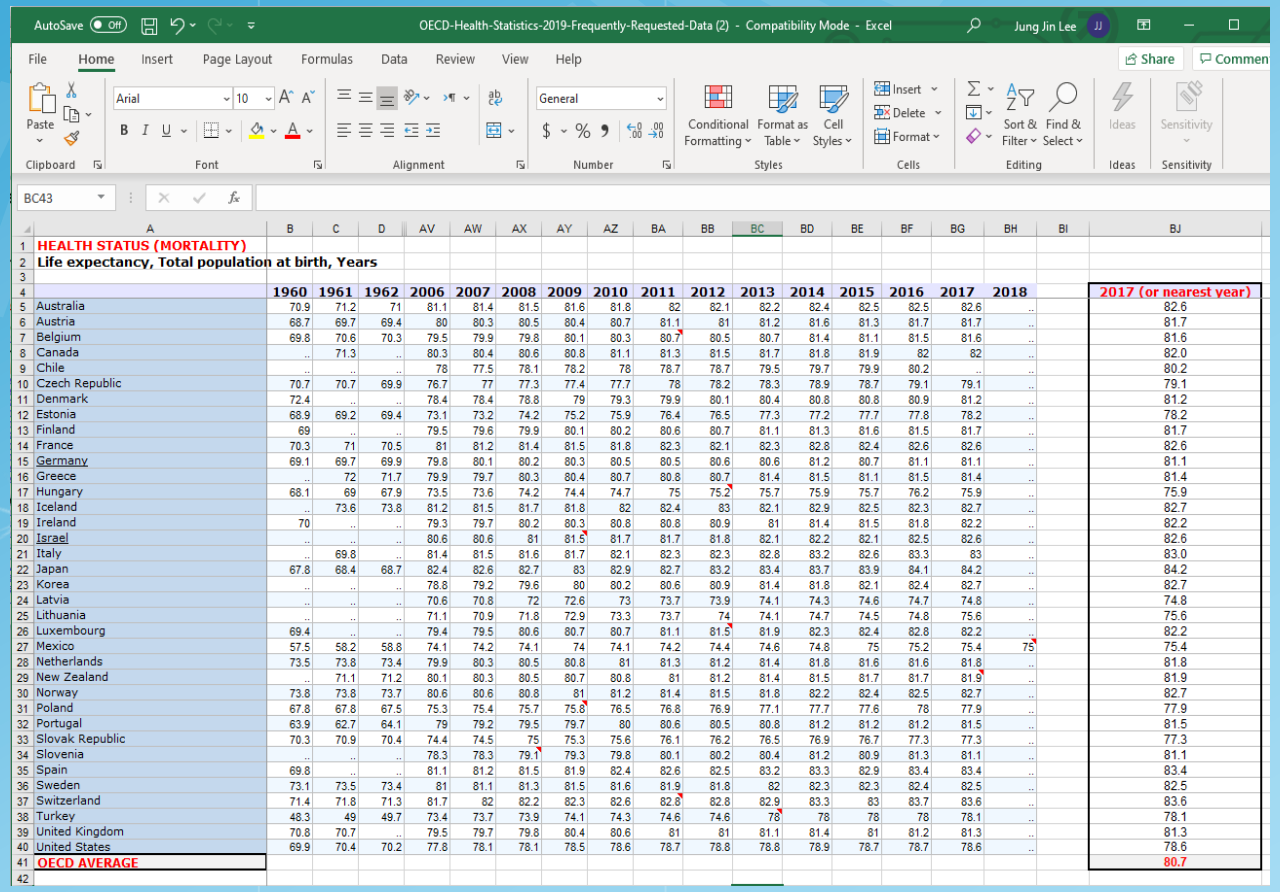




2.2 Visualization of Summary Data

[Ex 2.2.2] (Life Expectancy at Birth : OECD)

https://www.oecd.org



AutoSave Off — OECD-Health-Statistics-2019-Frequently-Requested-Data (2) - Compatibility Mode - Excel — Jung Jin Lee

File Home Insert Page Layout Formulas Data Review View Help — Share — Comment

Arial 10 — General — Paste — Clipboard — Font — Alignment — Number — Conditional Formatting — Format as Table — Cell Styles — Styles — Insert — Delete — Format — Cells — Sort & Filter — Find & Select — Editing — Ideas — Sensitivity

BC43

	A	B	C	D	AV	AW	AX	AY	AZ	BA	BB	BC	BD	BE	BF	BG	BH	BI	BJ
1	HEALTH STATUS (MORTALITY)																		
2	Life expectancy, Total population at birth, Years																		
3																			
4		1960	1961	1962	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018		2017 (or nearest year)
5	Australia	70.9	71.2	71	81.1	81.4	81.5	81.6	81.8	82	82.1	82.2	82.4	82.5	82.5	82.6	..		82.6
6	Austria	68.7	69.7	69.4	80	80.3	80.5	80.4	80.7	81.1	81	81.2	81.6	81.3	81.7	81.7	..		81.7
7	Belgium	69.8	70.6	70.3	79.5	79.9	79.8	80.1	80.3	80.7	80.5	80.7	81.4	81.1	81.5	81.6	..		81.6
8	Canada	..	71.3	..	80.3	80.4	80.6	80.8	81.1	81.3	81.5	81.7	81.8	81.9	82	82	..		82.0
9	Chile	..	..	..	78	77.5	78.1	78.2	78	78.7	78.7	79.5	79.7	79.9	80.2	..	..		80.2
10	Czech Republic	70.7	70.7	69.9	76.7	77	77.3	77.4	77.7	78	78.2	78.3	78.9	78.7	79.1	79.1	..		79.1
11	Denmark	72.4	..	..	78.4	78.4	78.8	79	79.3	79.9	80.1	80.4	80.8	80.8	80.9	81.2	..		81.2
12	Estonia	68.9	69.2	69.4	73.1	73.2	74.2	75.2	75.9	76.4	76.5	77.3	77.2	77.7	77.8	78.2	..		78.2
13	Finland	69	..	..	79.5	79.6	79.9	80.1	80.2	80.6	80.7	81.1	81.3	81.6	81.5	81.7	..		81.7
14	France	70.3	71	70.5	81	81.2	81.4	81.5	81.8	82.3	82.1	82.3	82.8	82.4	82.6	82.6	..		82.6
15	Germany	69.1	69.7	69.9	79.8	80.1	80.2	80.3	80.5	80.5	80.6	80.6	81.2	80.7	81.1	81.1	..		81.1
16	Greece	..	72	71.7	79.9	79.7	80.3	80.4	80.7	80.8	80.7	81.4	81.5	81.1	81.5	81.4	..		81.4
17	Hungary	68.1	69	67.9	73.5	73.6	74.2	74.4	74.7	75	75.2	75.7	75.9	75.7	76.2	75.9	..		75.9
18	Iceland	..	73.6	73.8	81.2	81.5	81.7	81.8	82	82.4	83	82.1	82.9	82.5	82.3	82.7	..		82.7
19	Ireland	70	..	..	79.3	79.7	80.2	80.3	80.8	80.8	80.9	81	81.4	81.5	81.8	82.2	..		82.2
20	Israel	..	..	..	80.6	80.6	81	81.5	81.7	81.7	81.8	82.1	82.2	82.1	82.5	82.6	..		82.6
21	Italy	..	69.8	..	81.4	81.5	81.6	81.7	82.1	82.3	82.3	82.8	83.2	82.6	83.3	83	..		83.0
22	Japan	67.8	68.4	68.7	82.4	82.6	82.7	83	82.9	82.7	83.2	83.4	83.7	83.9	84.1	84.2	..		84.2
23	Korea	..	..	..	78.8	79.2	79.6	80	80.2	80.6	80.9	81.4	81.8	82.1	82.4	82.7	..		82.7
24	Latvia	..	..	..	70.6	70.8	72	72.6	73	73.7	73.9	74.1	74.3	74.6	74.7	74.8	..		74.8
25	Lithuania	..	..	..	71.1	70.9	71.8	72.9	73.3	73.7	74	74.1	74.7	74.5	74.8	75.6	..		75.6
26	Luxembourg	69.4	..	..	79.4	79.5	80.6	80.7	80.7	81.1	81.5	81.9	82.3	82.4	82.8	82.2	..		82.2
27	Mexico	57.5	58.2	58.8	74.1	74.2	74.1	74	74.1	74.2	74.4	74.6	74.8	75	75.2	75.4	75		75.4
28	Netherlands	73.5	73.8	73.4	79.9	80.3	80.5	80.8	81	81.3	81.2	81.4	81.8	81.6	81.6	81.8	..		81.8
29	New Zealand	..	71.1	71.2	80.1	80.3	80.5	80.7	80.8	81	81.2	81.4	81.5	81.7	81.7	81.9	..		81.9
30	Norway	73.8	73.8	73.7	80.6	80.6	80.8	81	81.2	81.4	81.5	81.8	82.2	82.4	82.5	82.7	..		82.7
31	Poland	67.8	67.8	67.5	75.3	75.4	75.7	75.8	76.5	76.8	76.9	77.1	77.7	77.6	78	77.9	..		77.9
32	Portugal	63.9	62.7	64.1	79	79.2	79.5	79.7	80	80.6	80.5	80.8	81.2	81.2	81.2	81.5	..		81.5
33	Slovak Republic	70.3	70.9	70.4	74.4	74.5	75	75.3	75.6	76.1	76.2	76.5	76.9	76.7	77.3	77.3	..		77.3
34	Slovenia	..	..	..	78.3	78.3	79.1	79.3	79.8	80.1	80.2	80.4	81.2	80.9	81.3	81.1	..		81.1
35	Spain	69.8	..	..	81.1	81.2	81.5	81.9	82.4	82.6	82.5	83.2	83.3	82.9	83.4	83.4	..		83.4
36	Sweden	73.1	73.5	73.4	81	81.1	81.3	81.5	81.6	81.9	81.8	82	82.3	82.3	82.4	82.5	..		82.5
37	Switzerland	71.4	71.8	71.3	81.7	82	82.2	82.3	82.6	82.8	82.8	82.9	83.3	83	83.7	83.6	..		83.6
38	Turkey	48.3	49	49.7	73.4	73.7	73.9	74.1	74.3	74.6	74.6	78	78	78	78	78.1	..		78.1
39	United Kingdom	70.8	70.7	..	79.5	79.7	79.8	80.4	80.6	81	81	81.1	81.4	81	81.2	81.3	..		81.3
40	United States	69.9	70.4	70.2	77.8	78.1	78.1	78.5	78.6	78.7	78.8	78.8	78.9	78.7	78.7	78.6	..		78.6
41	OECD AVERAGE																		80.7
42																			

File: EX020202_OECD_LifeExpectancy — EditVar

Analysis Var: 1: Country — by Group: 2: LifeExp

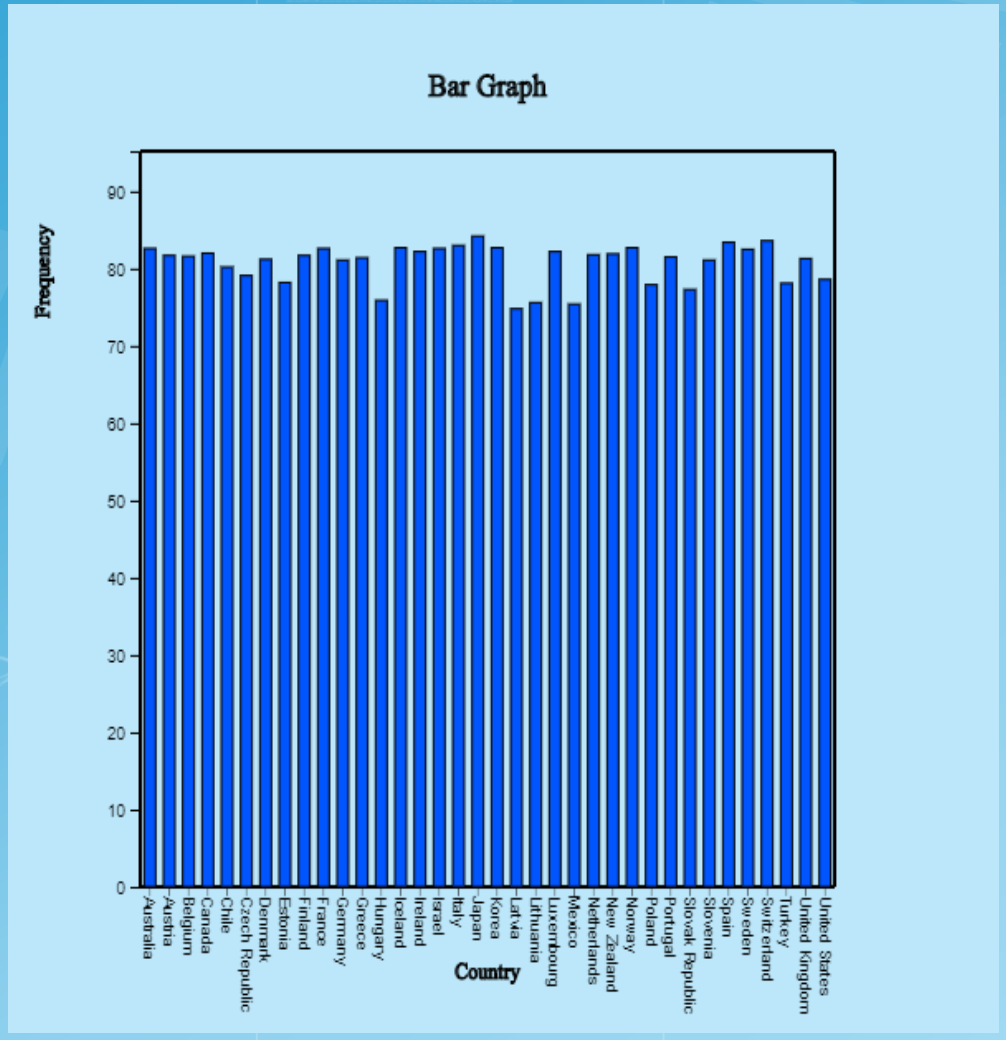
(Selected data: Summary Data) (Summary Data: Multiple Selection)

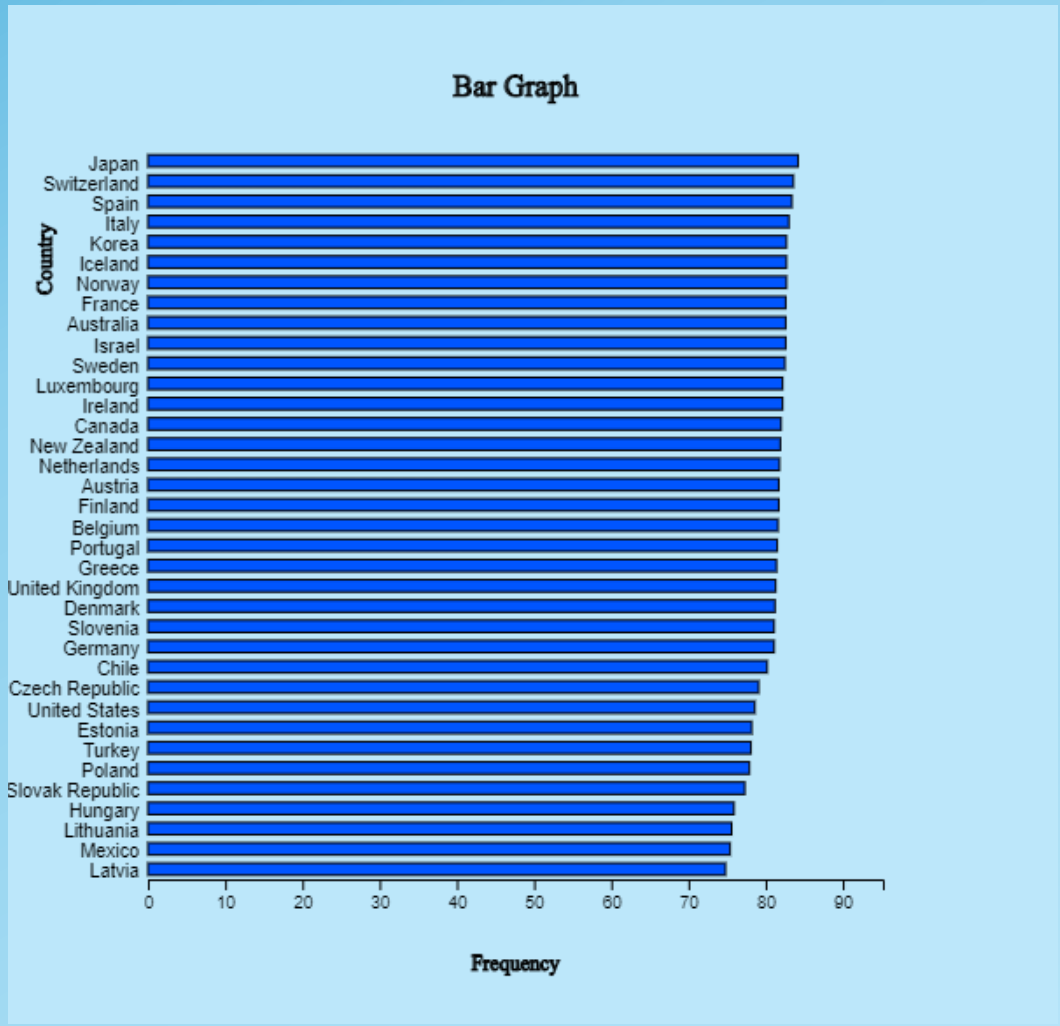
SelectedVar: V1 by V2, — Cancel

	Country	LifeExp	V3	V4
1	Australia	82.6		
2	Austria	81.7		
3	Belgium	81.6		
4	Canada	82.0		
5	Chile	80.2		
6	Czech Republic	79.1		
7	Denmark	81.2		
8	Estonia	78.2		
9	Finland	81.7		
10	France	82.6		
11	Germany	81.1		
12	Greece	81.4		
13	Hungary	75.9		
14	Iceland	82.7		
15	Ireland	82.2		
16	Israel	82.6		
17	Italy	83.0		
18	Japan	84.2		
19	Korea	82.7		
20	Latvia	74.8		
21	Lithuania	75.6		
22	Luxembourg	82.2		
23	Mexico	75.4		
24	Netherlands	81.8		
25	New Zealand	81.9		

2.2 Visualization of Summary Data

[Ex 2.2.2] (Life Expectancy at Birth : OECD)





2.2 Visualization of Summary Data

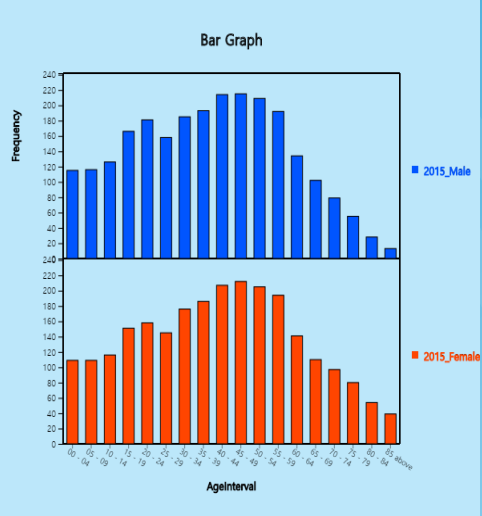
[Ex 2.2.4] (Male and Female Population by Age Group – Two Group Summary Data)
In 2015, the male and female populations by age group in Korea are shown in Table 2.2.3.

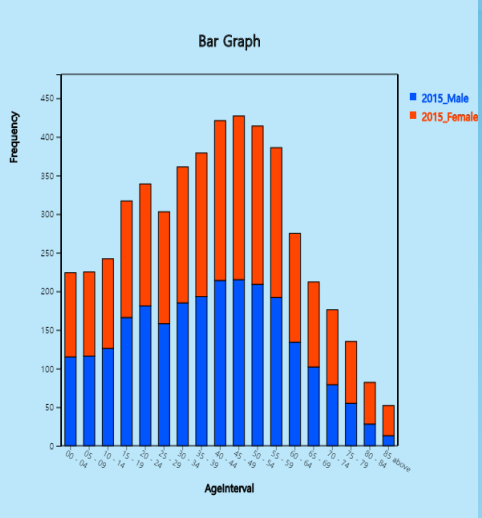
Using this data, draw bar chart, pie chart, band graph, and line graph of the population by age group and examine their characteristics.

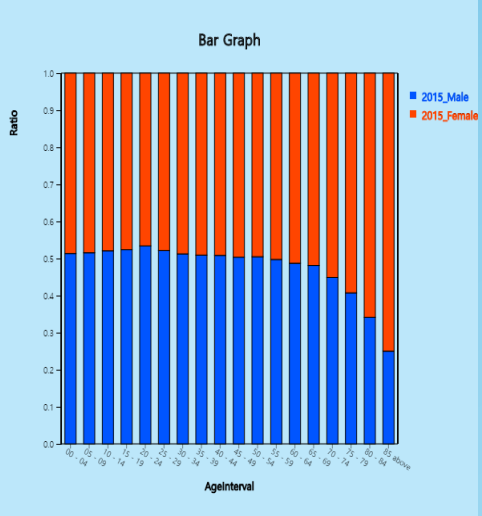
Age Interval	2015 Male	2015 Female
00 - 04	115	109
05 - 09	116	109
10 - 14	126	116
15 - 19	166	151
20 - 24	181	158
25 - 29	158	145
30 - 34	185	176
35 - 39	193	186
40 - 44	214	207
45 - 49	215	212
50 - 54	209	205
55 - 59	192	194
60 - 64	134	141
65 - 69	102	110
70 - 74	79	97
75 - 79	55	80
80 - 84	28	54
over 85	13	39

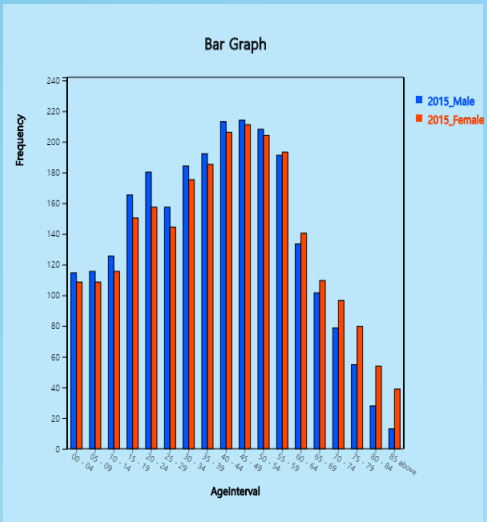
2.2 Visualization of Summary Data

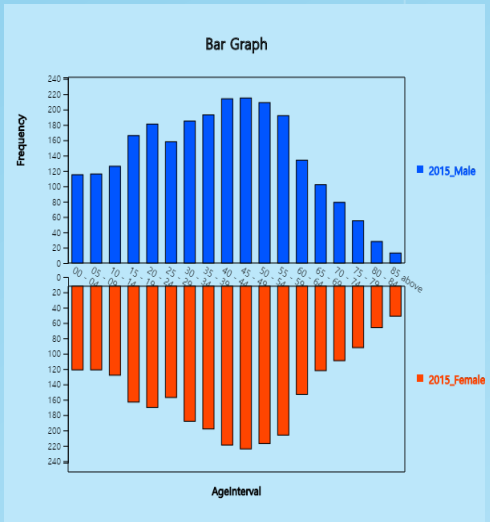
[Ex 2.2.4] (Male and Female Population by Age Group) - Two Group





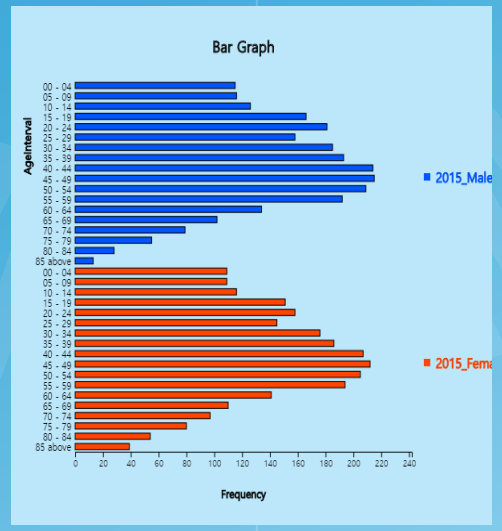


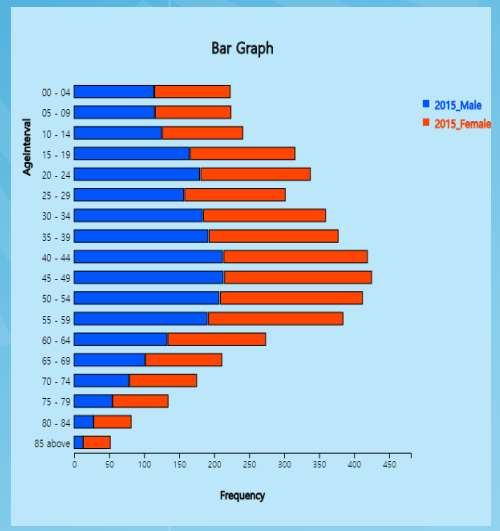


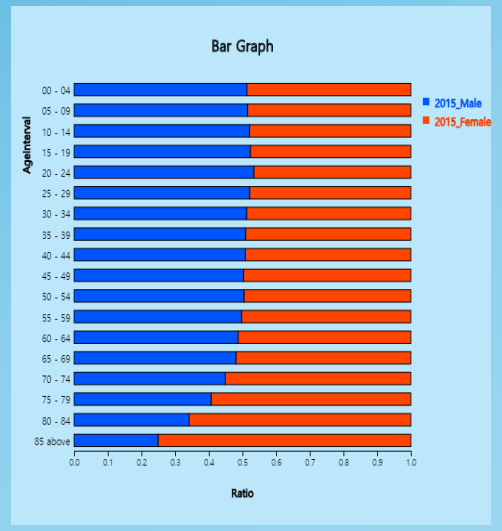


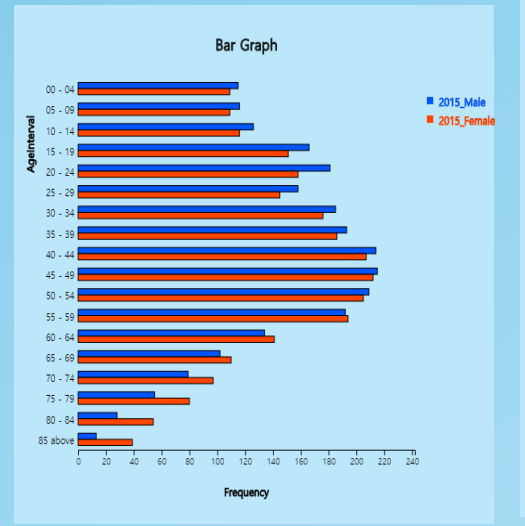
2.2 Visualization of Categorical Summary Data

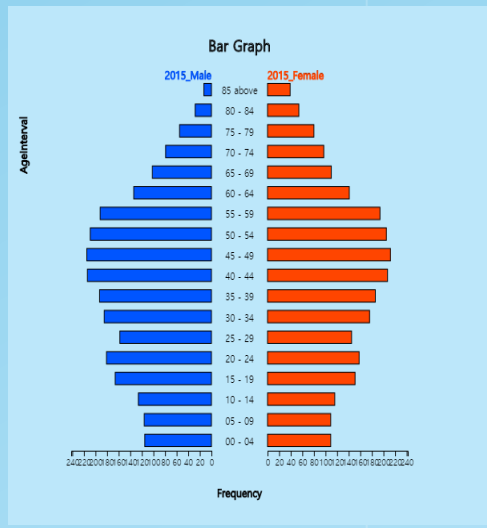
[Ex 2.2.4] (Male and Female Population by Age Group) - Two Group





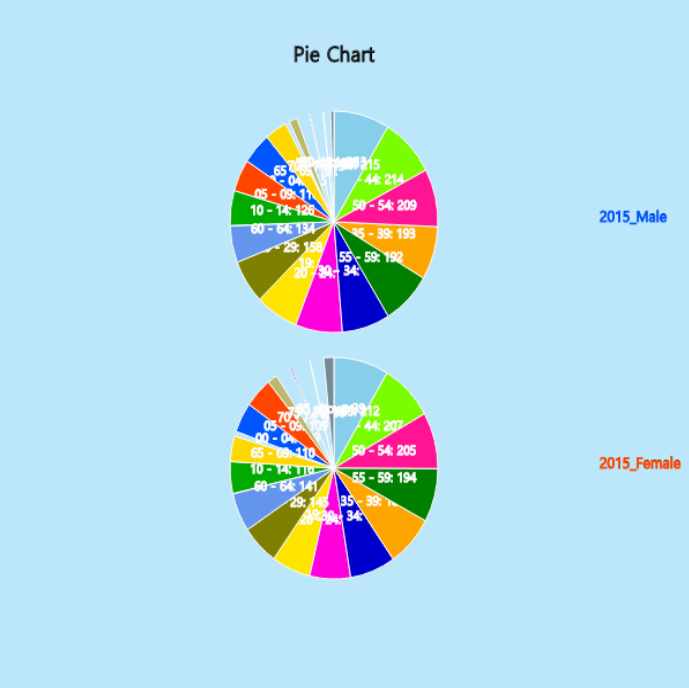


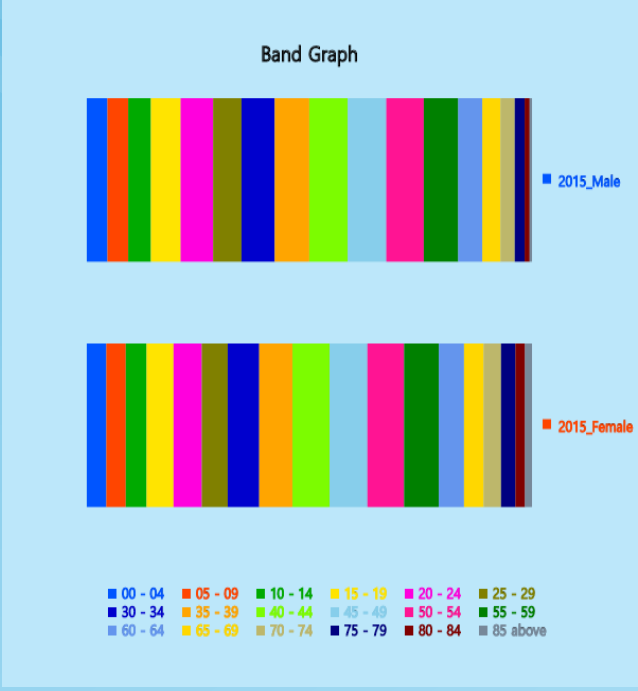


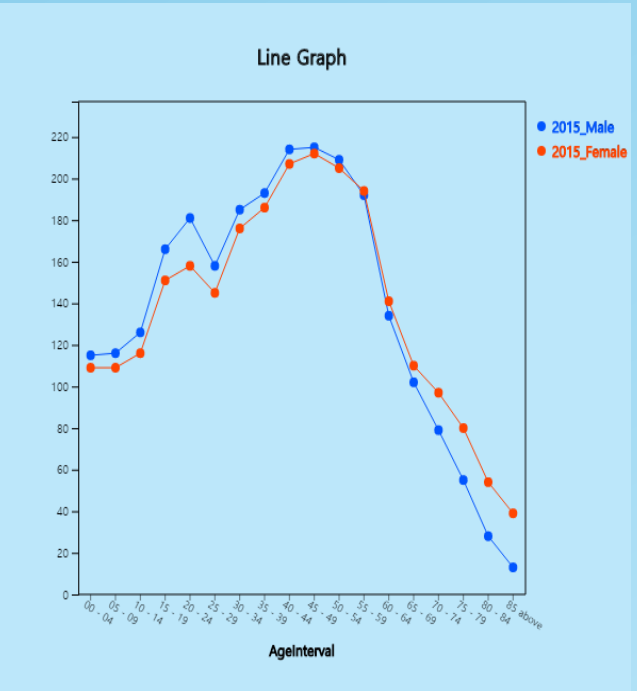


2.2 Visualization of Categorical Summary Data

[Ex 2.2.4] (Male and Female Population by Age Group) - Two Group







2.2 Visualization of Categorical Summary Data

[Example 2.2.5] (OECD Export - Import by Country in 2017) - two group summary data

Ex ⇨ eBook ⇨ EX02020_OECD_ExportImport_2017.csv.

Use a line graph to find out the characteristics of export and import of each countriytemperature change.

File 011OECD_ExportImport__2017.cs EditVar

X Var: 1: Country ▾ (Selected data: Summary Data)

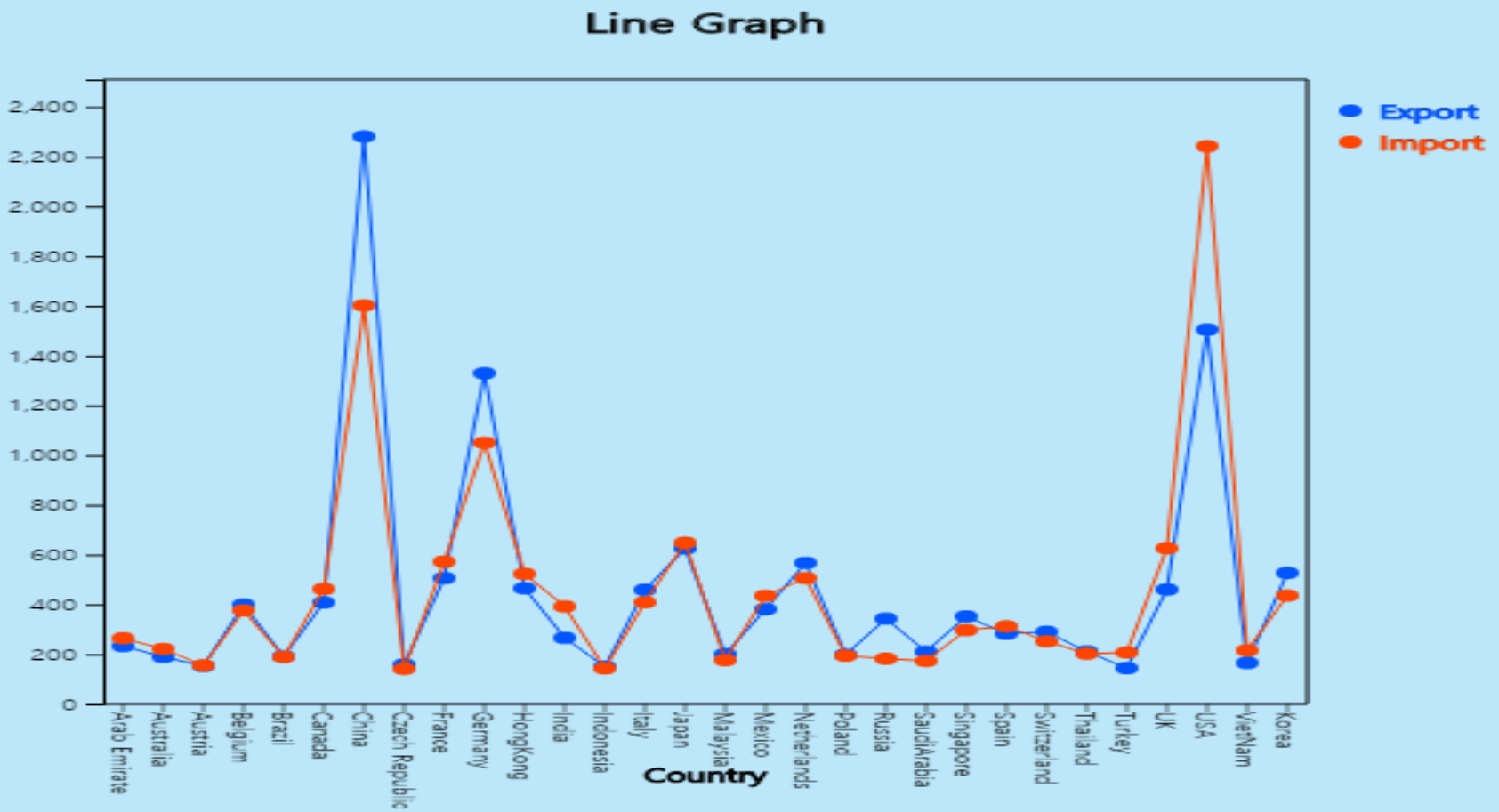
by Group: 3: Import ▾ (Summary Data: Multiple Selection)

SelectedVar: V1 by V2,V3, Cancel

	Country	Export	Import	V4	V
1	Arab Emirate	233	264		
2	Australia	189	221		
3	Austria	153	155		
4	Belgium	399	376		
5	Brazil	191	189		
6	Canada	408	462		
7	China	2281	1602		
8	Czech Republic	158	141		
9	France	506	572		
10	Germany	1329	1050		
11	HongKong	465	523		
12	India	266	392		
13	Indonesia	150	143		
14	Italy	459	409		
15	Japan	625	648		
16	Malaysia	200	176		
17	Mexico	381	435		
18	Netherlands	567	506		
19	Poland	198	194		

2.2 Visualization of Summary Data

[Example 2.2.5] (OECD Export - Import by Country)



2.3 Visualization of Raw Data

[Example 2.3.1] (Gender Survey – Raw data without group)
For the census, gender data is compiled by coding (1: male, 2: female) as shown in Table 2.3.1
(Ex ⇨ eBook ⇨ Ex020301_Raw_Gender.csv).

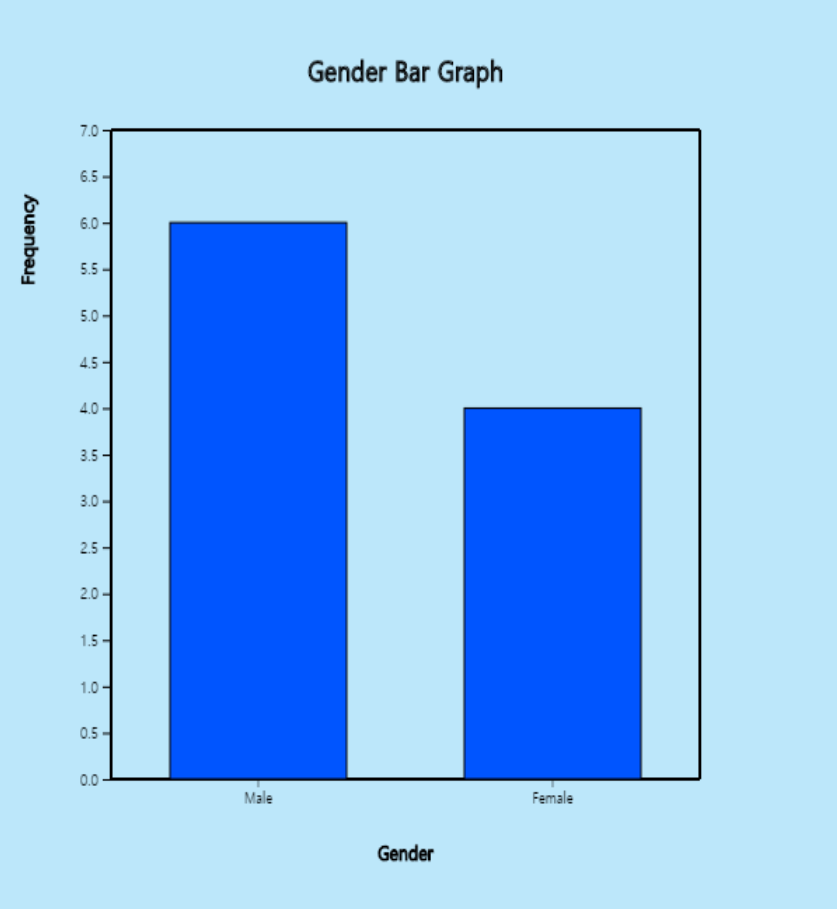
These data are called raw data.

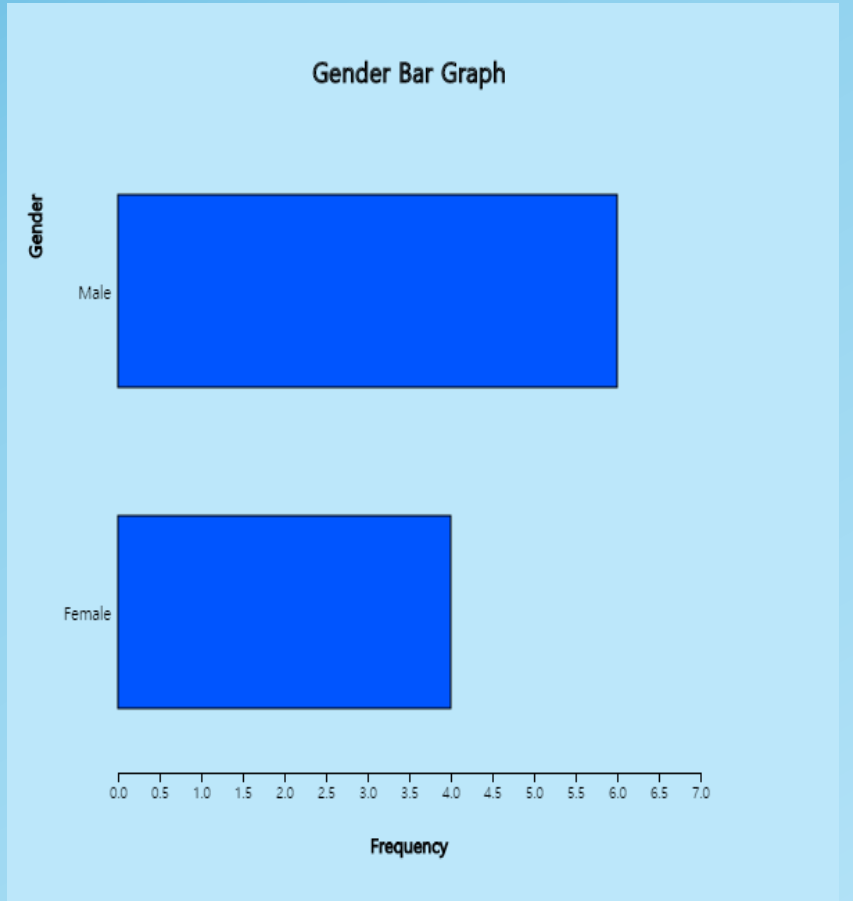
Use 『eStat』 to draw a bar chart, pie chart, and a band graph to find out the characteristics.

Gender
1
2
1
2
1
1
1
2
1
2

2.3 Visualization of Raw Data

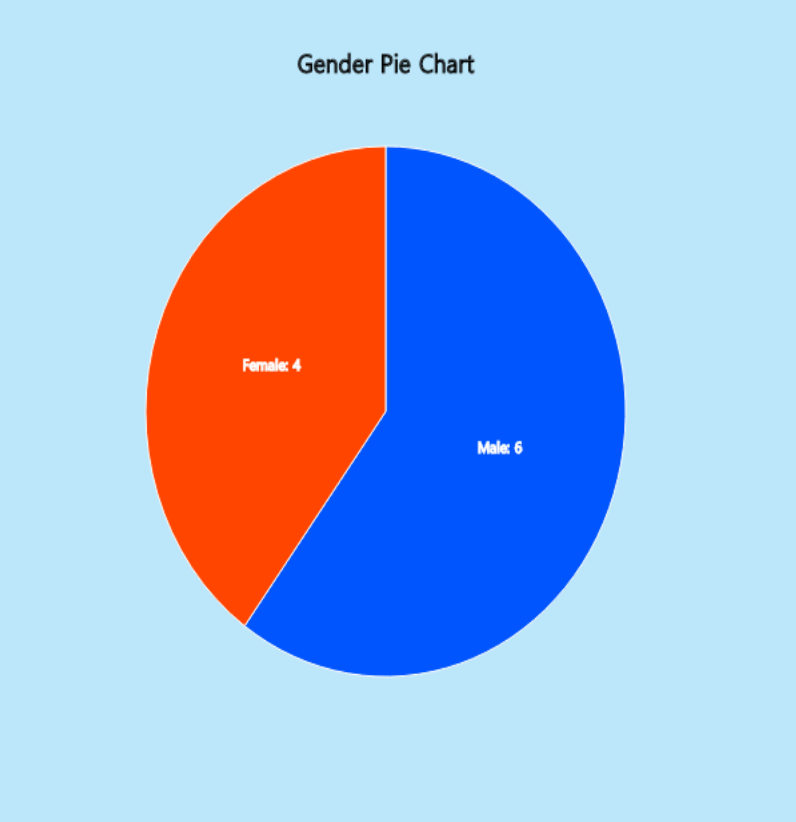
[Example 2.3.1] (Gender Survey – Raw data without group)

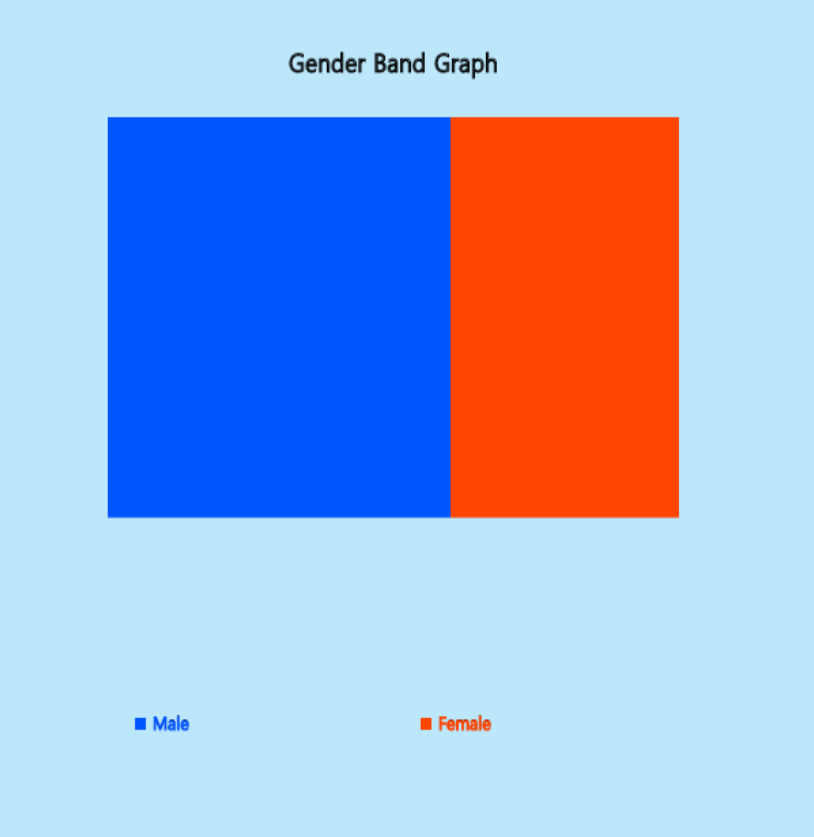




2.3 Visualization of Raw Data

[Example 2.3.1] (Gender Survey – Raw data without group)





2.3 Visualization of Raw Data

[Example 2.3.2] (Population – raw data with group)
In addition to the gender data of [Example 2.3.1], marital status is also surveyed as Table 2.3.2 shows the data of marriage status along with gender (1: Single 2: Married, 3: Other).

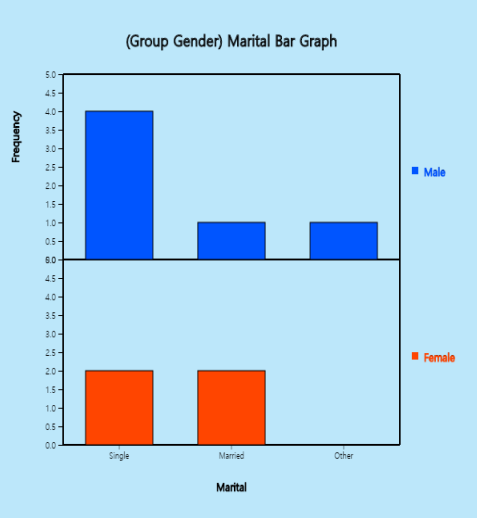
Ex ⇨ eBook ⇨ EX020302_Raw_MaritalByGender.csv)

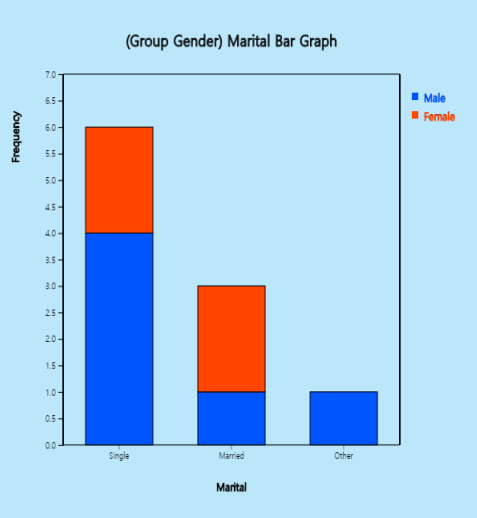
In this data, compare the characteristics of marital status by gender with bar charts, pie chart, band graph, and line graph for each male and female.

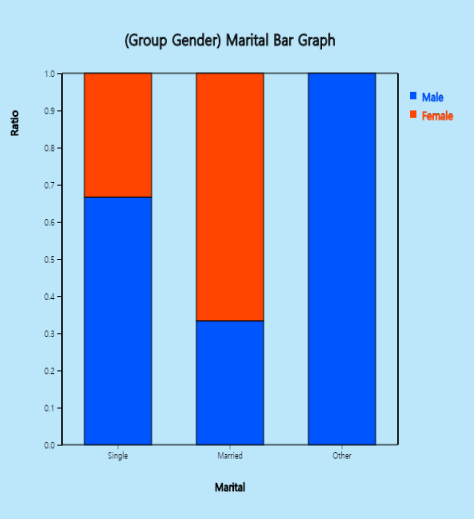
Gender (1:Male, 2:Female)	Marital Status (1:Single, 2:Married, 3:Other)
1	1
2	2
1	1
2	1
1	2
1	1
1	1
2	2
1	3
2	1

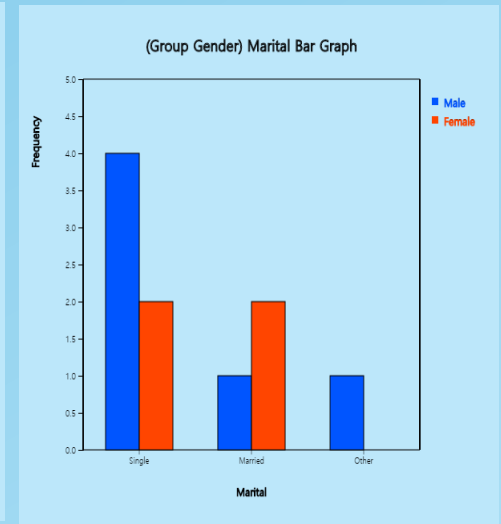
2.3 Visualization of Raw Data

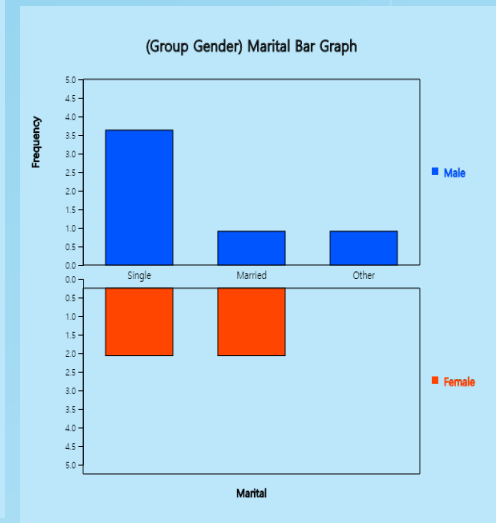
[Example 2.3.2] (Population – raw data with group)





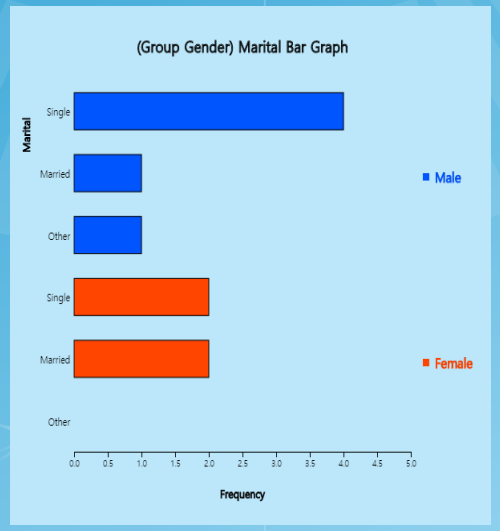


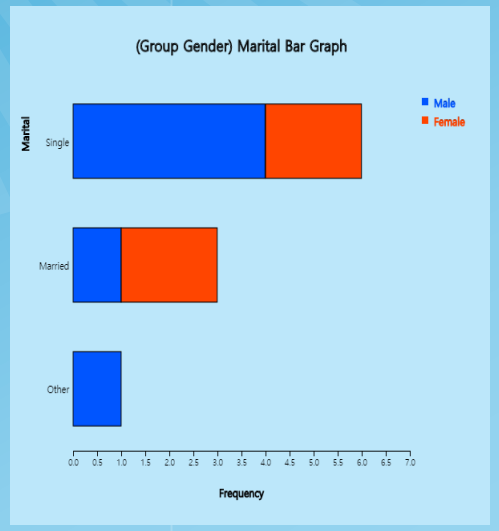


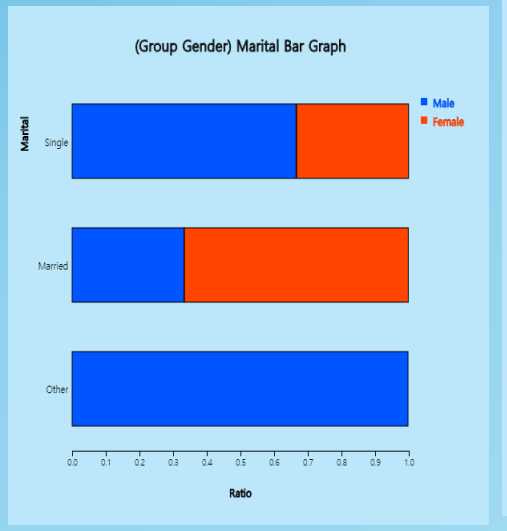


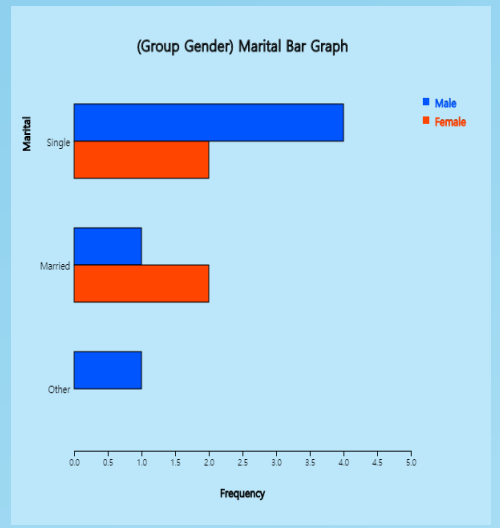
2.3 Visualization of Raw Data

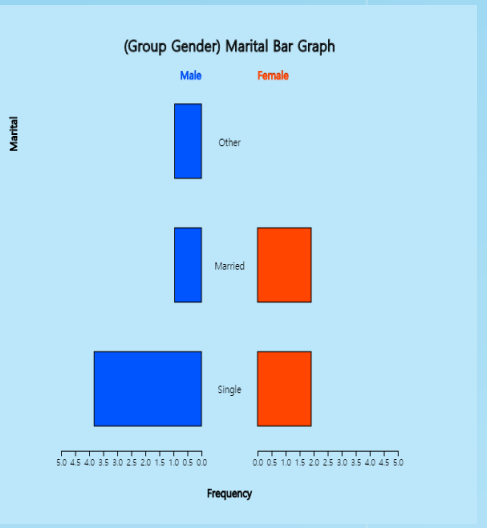
[Example 2.3.2] (Population – raw data with group)





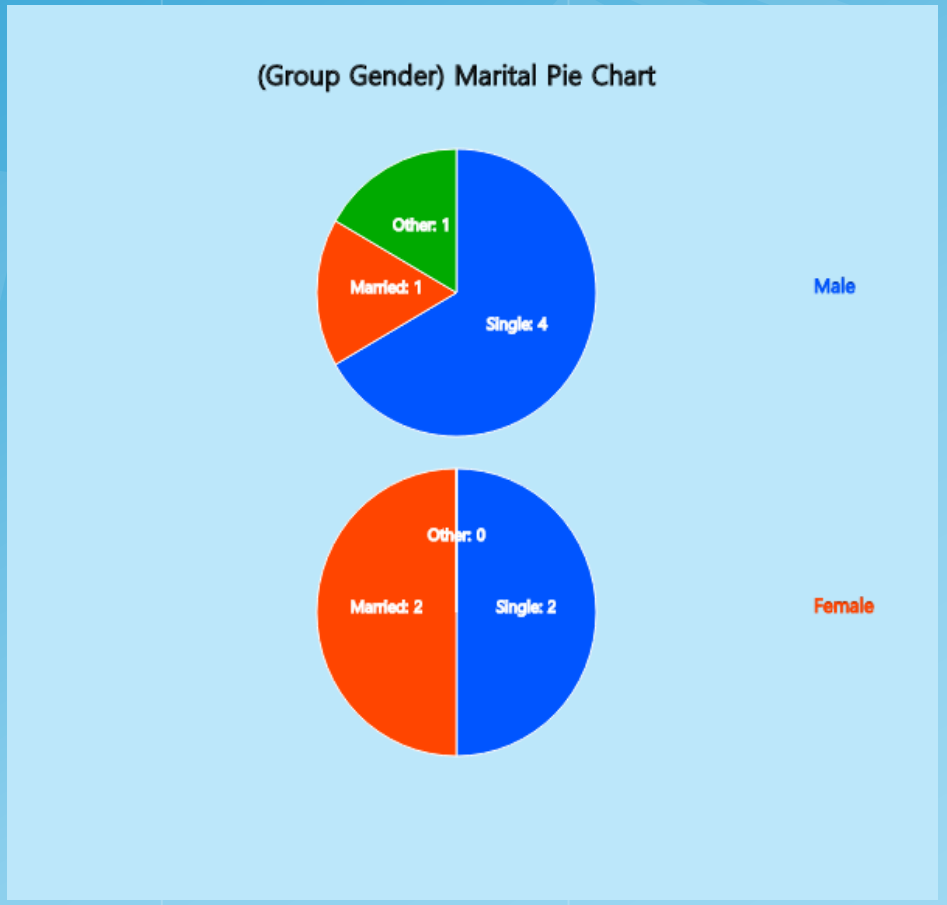


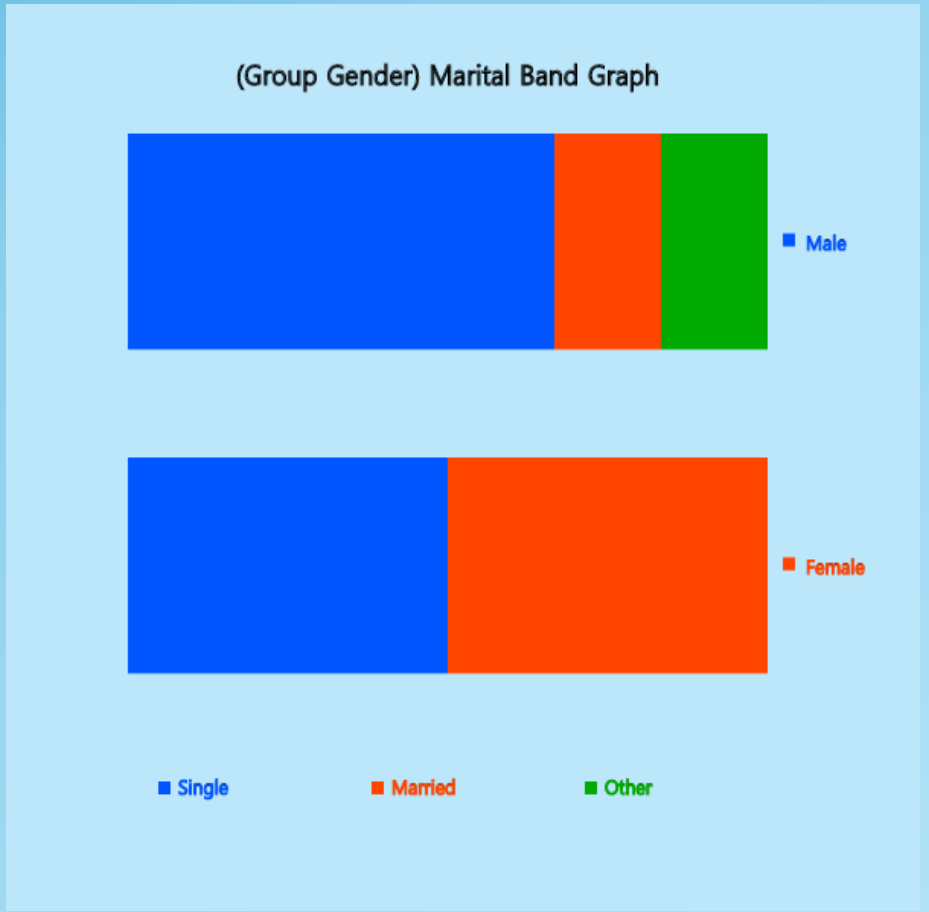




2.3 Visualization of Raw Data

[Example 2.3.2] (Population – raw data with group)





2.4 Summary

- Graphs for Visualizing Categorical Data
  - Bar Graph, Pie Chart, Band Graph, Line Graph
- Visualization of Summary Data
  - Visualization of single group summary data
  - Visualization of multi-group summary data
- Visualization of Raw Data
  - Visualization of single group raw data
  - Visualization of multi-group raw data

Thank you