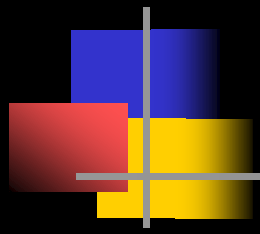
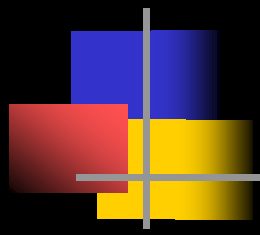


Chapter 2

Data Visualization

Professor Jung Jin Lee
Soongsil University, Korea
New Uzbekistan University, Uzbekistan



Chapter 2 Data visualization

2.1 Visualization of qualitative data

- 2.1.1 Visualization of raw data of a categorical variable
- 2.1.2 Visualization of frequency data of a categorical variable
- 2.1.3 Visualization of text data

2.2 Visualization of quantitative data

- 2.2.1 Visualization of a single quantitative variable
- 2.2.2 Visualization of two or more quantitative variables

2.1 Visualization of qualitative data

❖ **Visualization of categorical raw data**

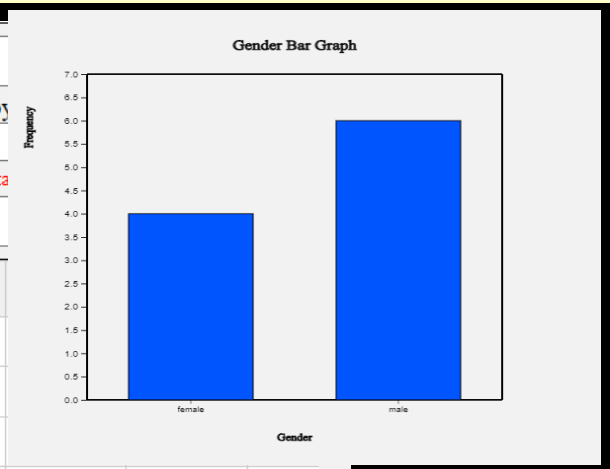
File: Gender.csv

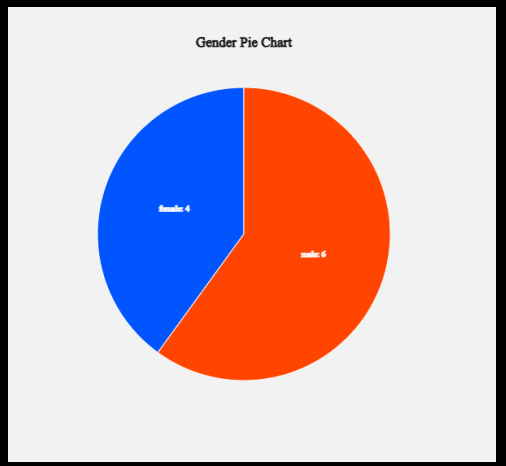
Analysis Var: 1: Gender | by: ---

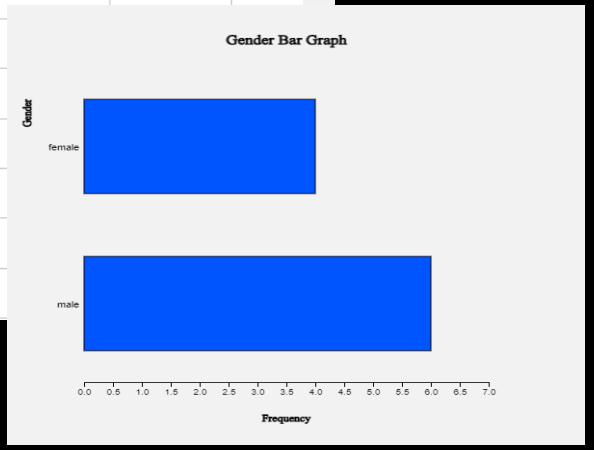
(Selected data: Raw Data) (Summary Data

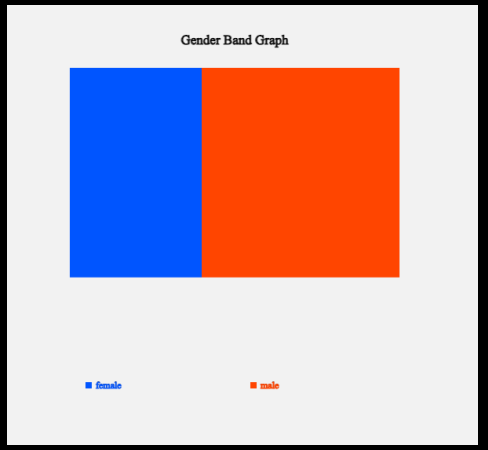
SelectedVar: V1

	Gender	V2	V3
1	male		
2	female		
3	male		
4	female		
5	male		
6	male		
7	male		
8	female		
9	male		
10	female		









2.1 Visualization of qualitative data

❖ Visualization of categorical raw data

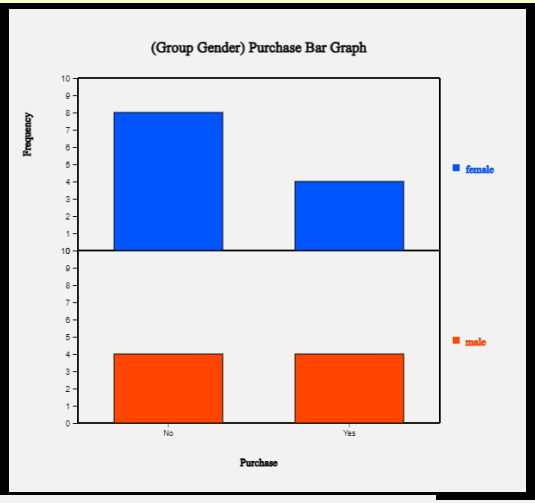
File PurchaseByCredit20.csv EditVar

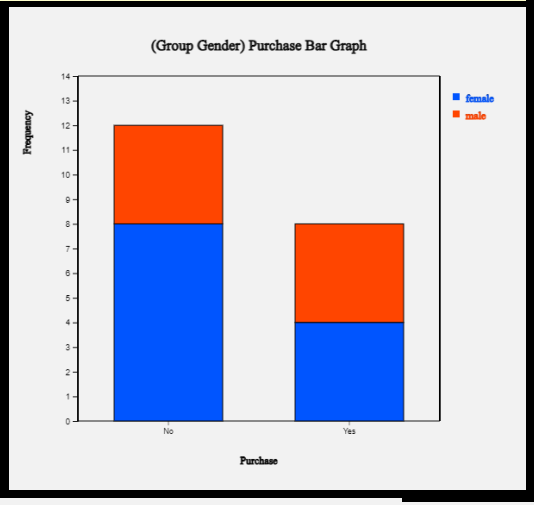
Analysis Var --- Group ---

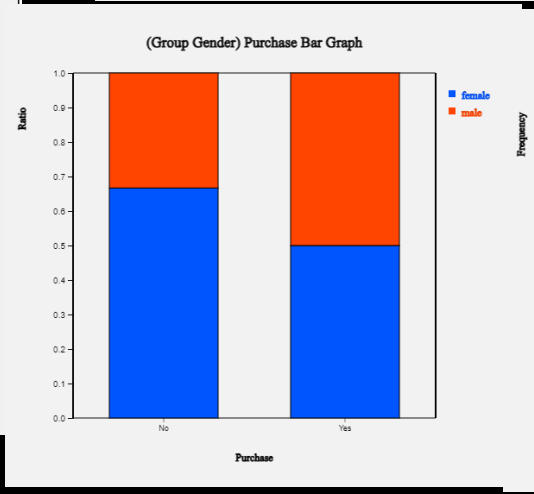
(Select variables by click var name) (Summary Data: Multiple Selection)

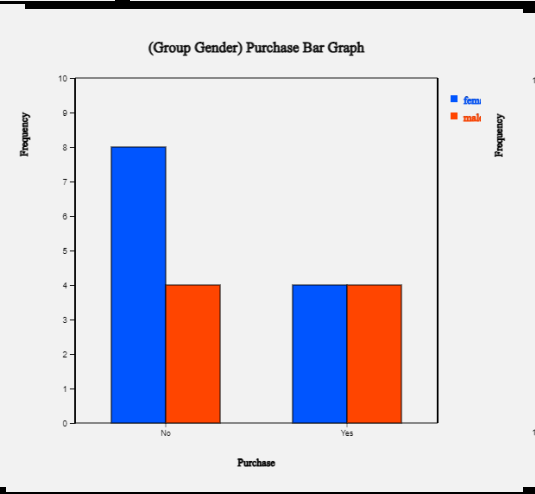
SelectedVar Cancel

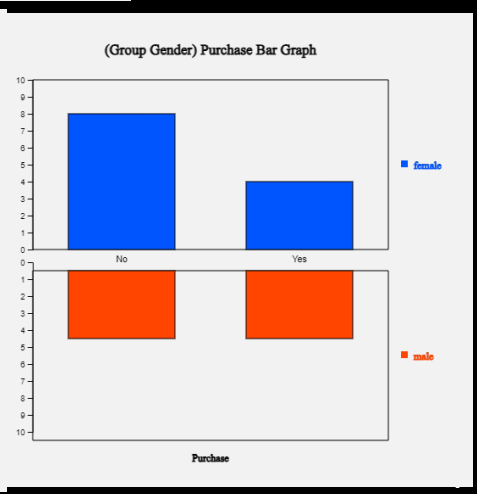
	id	Gender	Age	Income	Credit	Purch
1	1	male	20s	LT2000	Fair	Yes
2	2	female	30s	GE2000	Good	No
3	3	female	20s	GE2000	Fair	No
4	4	female	20s	GE2000	Fair	Yes
5	5	female	20s	LT2000	Bad	No
6	6	female	30s	GE2000	Fair	No
7	7	female	30s	GE2000	Good	Yes
8	8	male	20s	LT2000	Fair	No
9	9	female	20s	GE2000	Good	No
10	10	male	30s	GE2000	Fair	Yes
11	11	female	30s	GE2000	Good	Yes
12	12	female	20s	LT2000	Fair	No
13	13	male	30s	GE2000	Fair	No
14	14	male	30s	LT2000	Fair	Yes
15	15	female	30s	GE2000	Good	Yes
16	16	female	30s	GE2000	Fair	No
17	17	female	20s	GE2000	Bad	No
18	18	male	20s	GE2000	Bad	No
19	19	male	30s	GE2000	Good	Yes
20	20	male	20s	LT2000	Fair	No





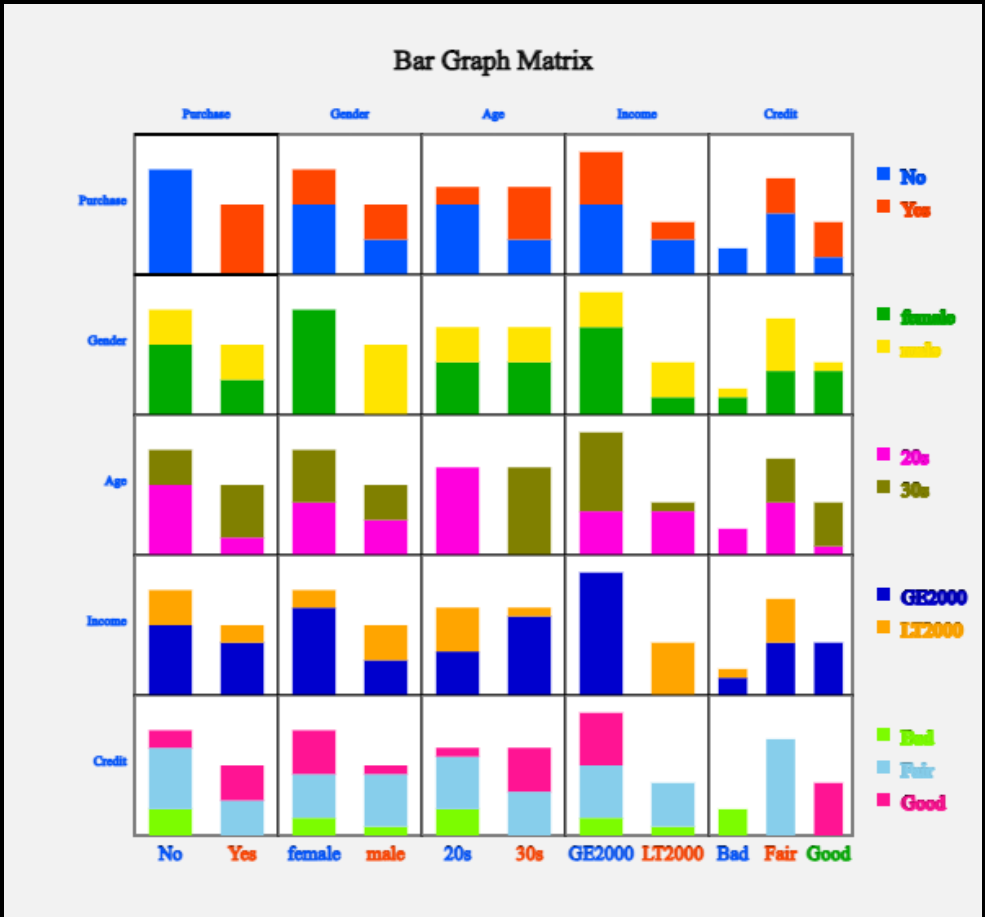






2.1 Visualization of qualitative data

❖ **Visualization of categorical raw data**

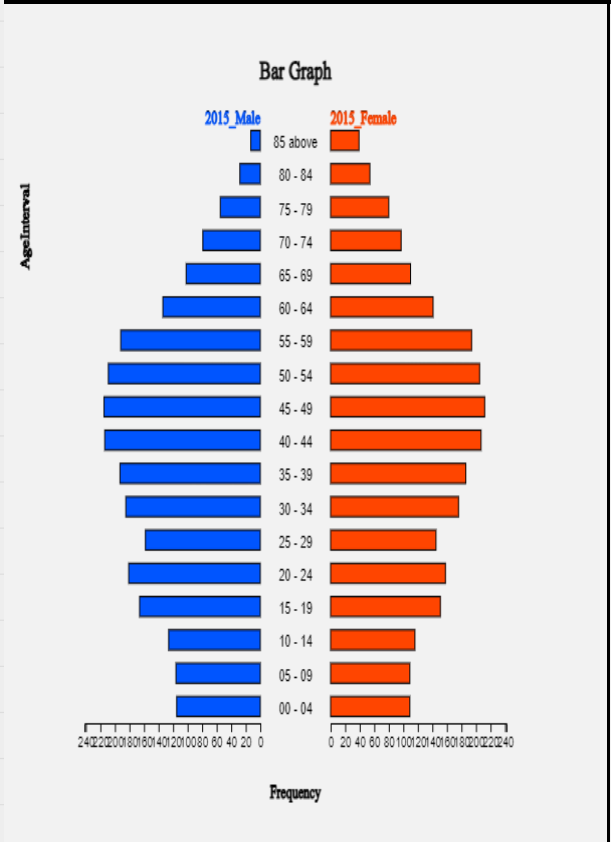


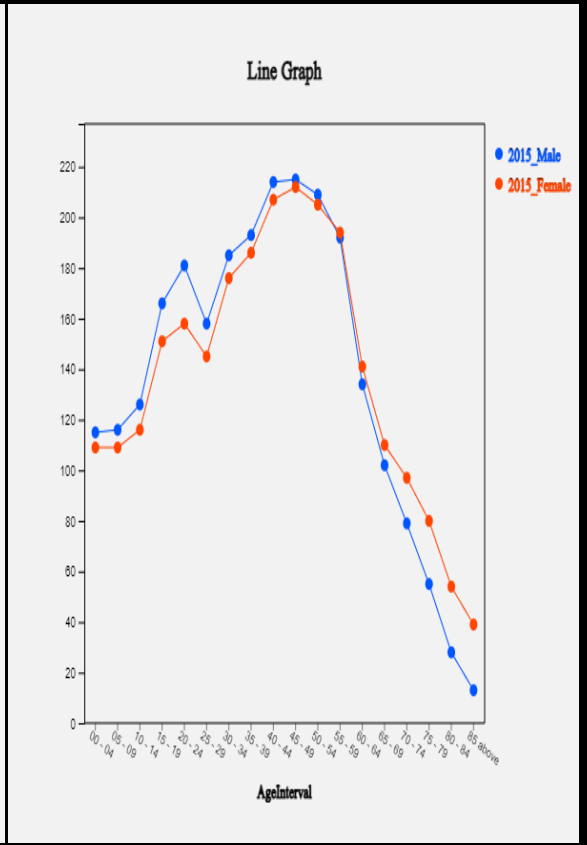
2.1 Visualization of qualitative data

❖ Visualization of frequency data

Table 2.1.5 male and female populations by age group in Korea (KOSTAT Census 2015, unit 10,000 persons)

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	158	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





2.1 Visualization of qualitative data

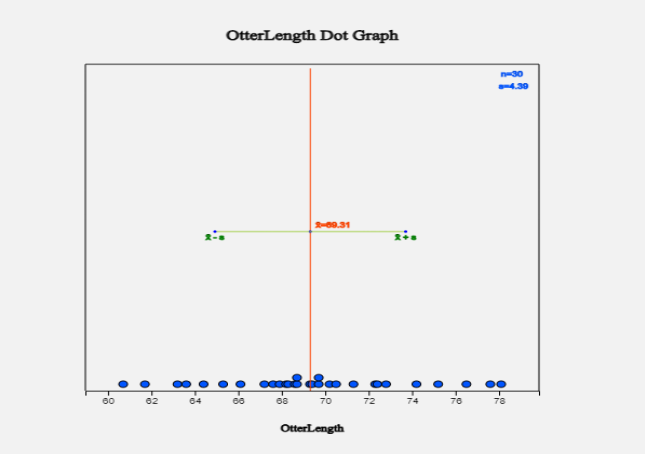
❖ Visualization of text data

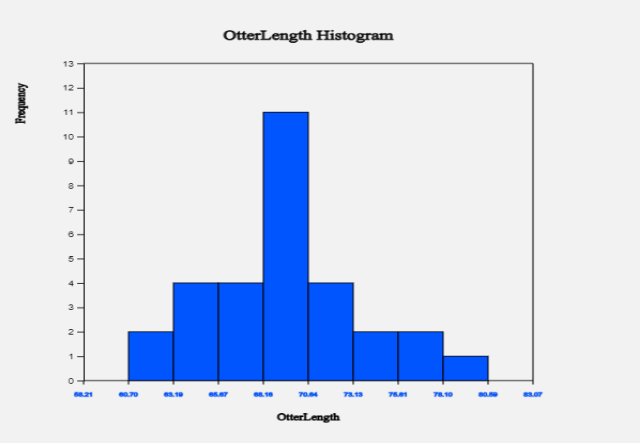
The development of these technologies has created massive data, simply called ā€Bīg Dataā€™, that were unimaginable in the past. Typical examples of the big data include data from Google's search records, social media messages by mobile phones, web logs by internet connections, and telephone records of global telecom companies. The big data are expected to grow and increase exponentially in the future and the hyper-forecasting is also expected to be possible. The success or failure of each individual, group, company and even country would depend on how to utilize the big data efficiently. The analysis of the big data that emerged this century is so enormous and diverse in the amount of data that can not be fully utilized just by traditional statistical approaches. For the analysis and utilization of the big data, theories of statistics, computer science, mathematics, management or related disciplines must also be applied simultaneously. Data Science is a new area of study in which statistics, mathematics, computer science and other disciplines are fused to analyze and utilize the big data

2.2 Visualization of quantitative data

❖ Visualization of a single variable

	OtterLeng	V2
1	63.2	
2	65.3	
3	67.6	
4	68.7	
5	69.7	
6	60.7	
7	72.4	
8	75.2	
9	64.4	
10	76.5	
11	68.3	
12	69.3	
13	70.2	
14	71.3	
15	74.2	
16	63.6	
17	66.1	
18	67.9	
19	68.7	
20	70.5	





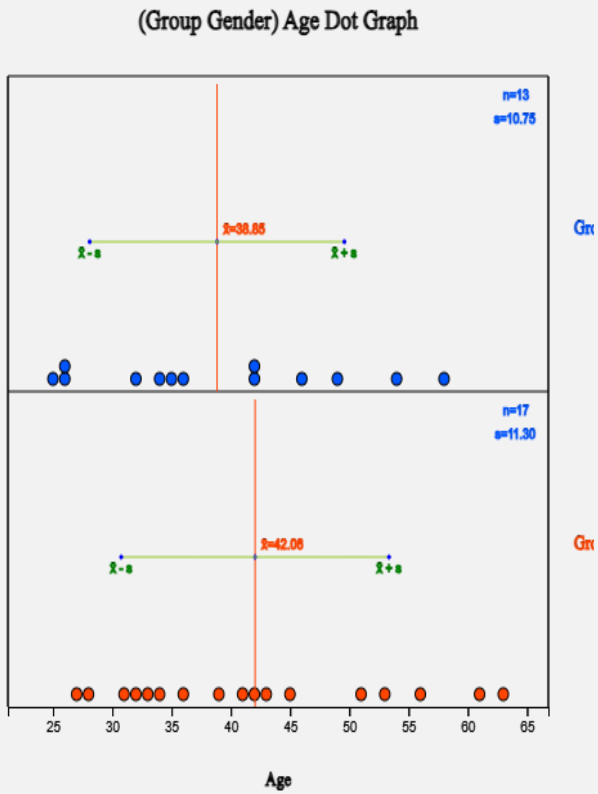
OtterLength Stem and Leaf Plot

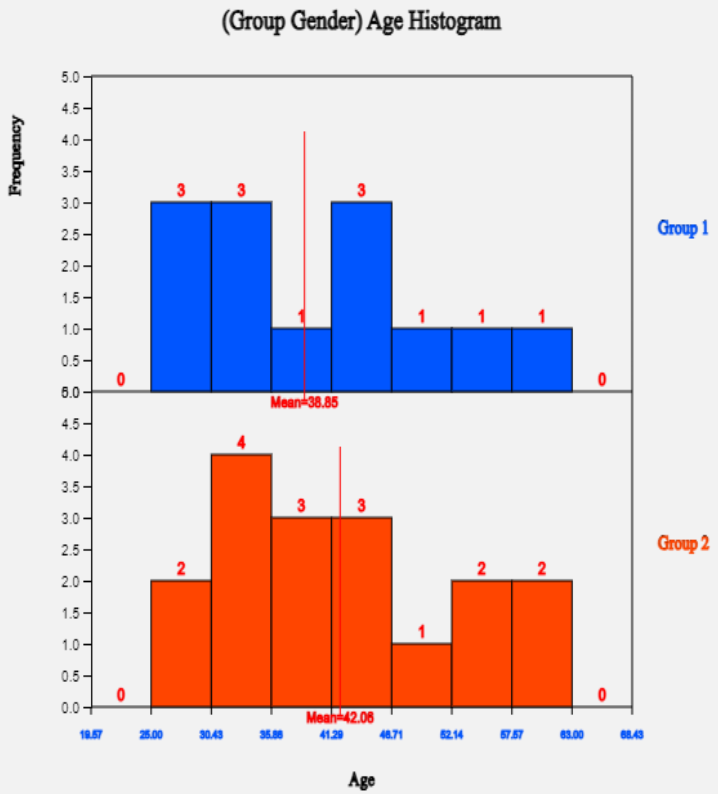
Stem	Leaf
60	7
61	7
62	
63	26
64	4
65	3
66	1
67	269
68	23677
69	3477
70	25
71	3
72	348
73	
74	2
75	2
76	5
77	6
78	1

2.2 Visualization of quantitative data

❖ **Visualization of a single variable with group**

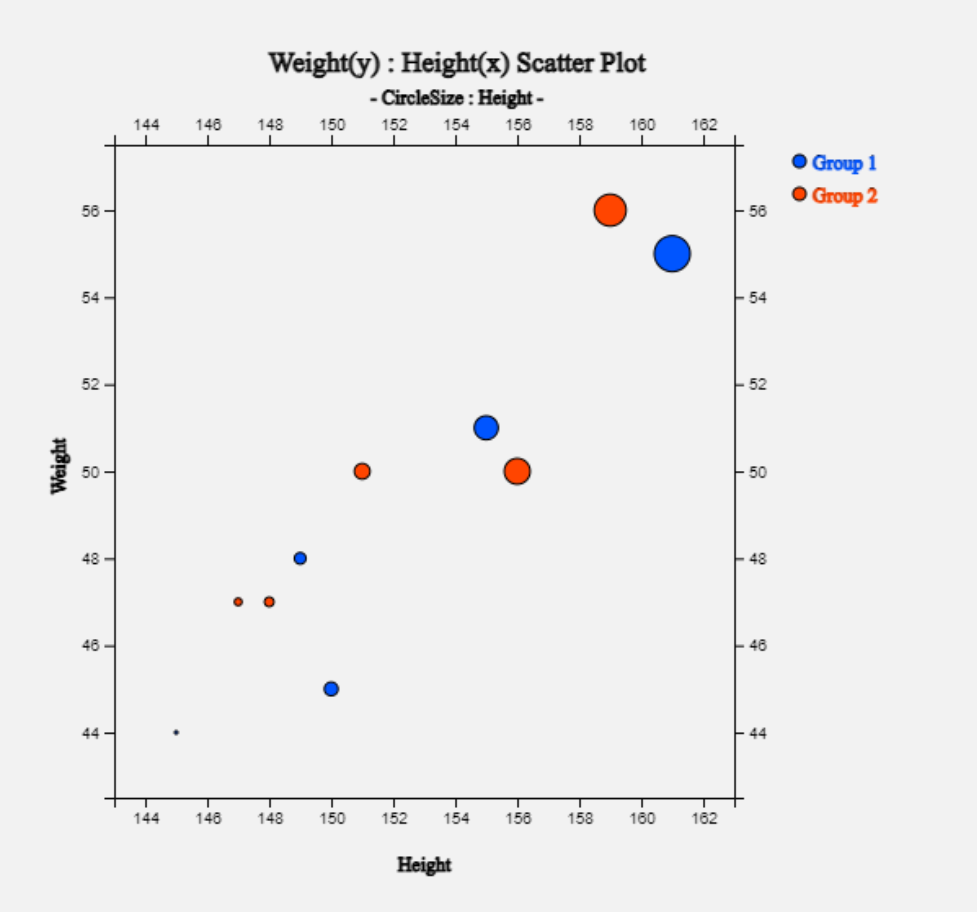
	Gender	Age
1	1	26
2	1	34
3	2	28
4	2	39
5	1	32
6	1	36
7	2	41
8	2	42
9	1	26
10	1	25
11	2	33
12	2	43
13	1	54
14	1	49
15	2	56
16	2	31
17	2	27
18	1	42
19	2	32
20	2	36





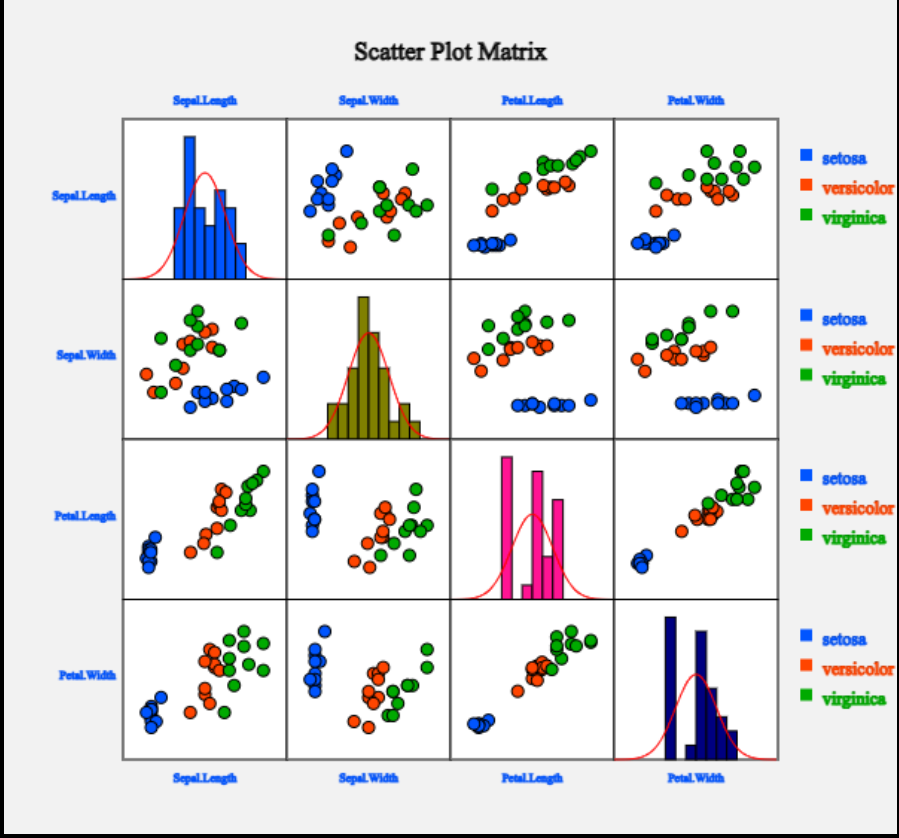
2.2 Visualization of quantitative data

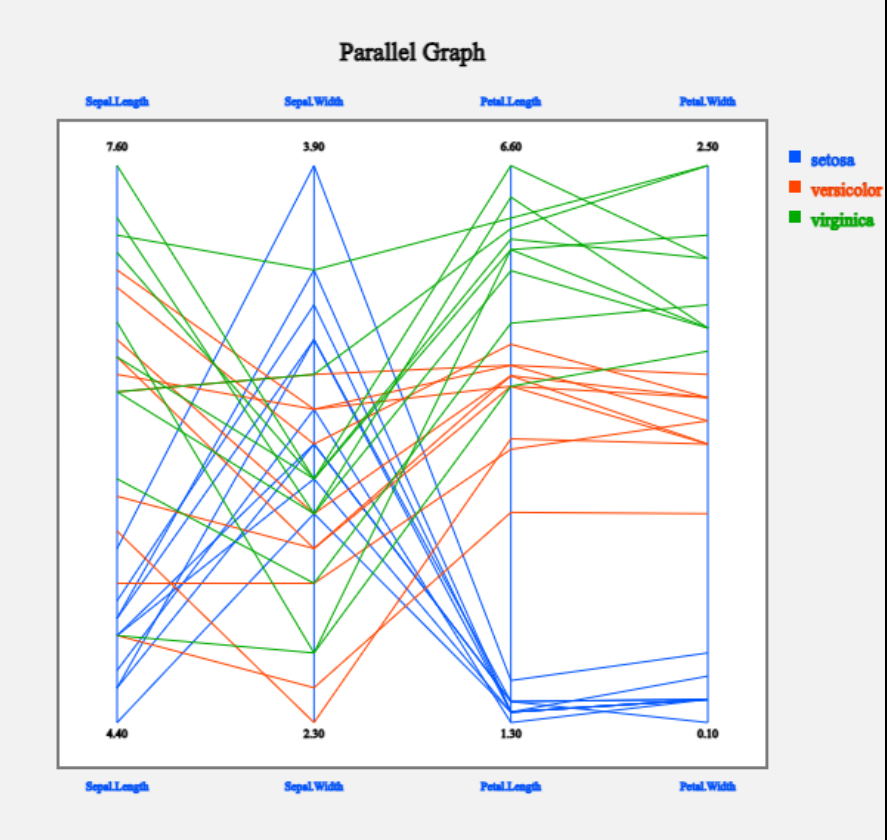
❖ **Visualization of two variables**

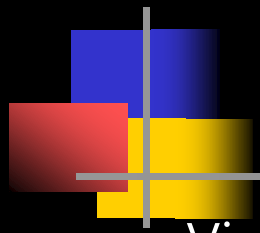


2.2 Visualization of quantitative data

❖ **Visualization of several variables**







Summary

- Visualization of qualitative data:
  - bar graph, pie chart, band graph
  - group variable
  - word cloud for text data
  - bar graph matrix

- Visualization of quantitative data:
  - dot graph, histogram, stem and leaf plot
  - scatterplot for two variables
  - scatterplot matrix, parallel coordinate plot

Thank you !!!