

Lab Lecture 2 Data visualization and summary statistic with R

Data 4: Risk factors & Colorectal cancer

	patient_no	age	fh	dm	smoking	bmi	fit	hematocrit	coloscopy
1	1	58	No	Yes	No	32.6	Negative	43.1	Normal
2	2	59	No	No	No	20	Negative	39.1	Normal
3	3	57	No	No	No	27.7	Negative	40.4	Normal
4	4	55	No	No	No	29.9	Negative	40.8	Normal
5	5	63	No	No	No	24.6	Negative	43.6	Abnormal
6	6	50	No	No	No	16.9	Negative	43.1	Normal
7	7	52	No	No	No	32.7	Negative	43.7	Abnormal
8	8	51	No	No	No	21.8	Negative	37.2	Normal
9	9	57	No	No	No	25.2	Negative	44.5	Normal
10	10	57	No	No	No	20	Negative	38.3	Normal
11	11	60	No	No	No	26.6	Negative	43	Normal
12	12	60	No	No	No	23.3	Positive	34.7	Colorectal cancer
13	13	54	Yes	No	No	39.5	Negative	42.7	Abnormal
14	14	56	No	No	No	17.4	Negative	43.7	Abnormal
15	15	50	No	No	Yes	29	Negative	41.4	Abnormal

1. Data Description

Column	ความหมาย	Data Types	
		Categorical data	Numerical data
patien_no	ลำดับผู้ป่วย		✓ (แบบ Discrete)
age	อายุ		✓
fh	ประวัติครอบครัวเป็นมะเร็ง	✓	
dm	ประวัติโรคเบาหวาน	✓	
smoking	ประวัติการสูบบุหรี่	✓	
bmi	BMI		✓
fit	Fecal immunochemical test (FIT)	✓	
hematocrit	Hematocrit		✓
coloscopy	ผลการส่องกล้อง	✓	

2. Measurement Scales

Column	ความหมาย	Levels of measurement scales			
		Nominal	Ordinal	Interval	Ratio
patien_no	ลำดับผู้ป่วย				✓
age	อายุ				✓
fh	ประวัติครอบครัวเป็นมะเร็ง	✓			
dm	ประวัติโรคเบาหวาน	✓			
smoking	ประวัติการสูบบุหรี่	✓			
bmi	BMI				✓
fit	Fecal immunochemical test (FIT)	✓			
hematocrit	Hematocrit				✓
coloscopy	ผลการส่องกล้อง	✓			

3. Descriptive Statistics

- Categorical data

Column	ความหมาย	Frequency	Relative Frequency	Percentage
fh	ประวัติครอบครัวเป็นมะเร็ง	No = 549 คน Yes = 50 คน	No = 0.91652755 Yes = 0.08347245	No = 92% Yes = 8%
dm	ประวัติโรคเบาหวาน	No = 561 คน Yes = 38 คน	No = 0.93656093 Yes = 0.06343907	No = 94% Yes = 6%
smoking	ประวัติการสูบบุหรี่	No = 576 คน Yes = 23 คน	No = 0.96160267 Yes = 0.03839733	No = 96% Yes = 4%
fit	Fecal immunochemical test (FIT)	Negative = 553 คน Positive = 46 คน	Negative = 0.92320534 Positive = 0.07679466	Negative = 92% Positive = 8%
coloscopy	ผลการส่องกล้อง	Abnormal = 170 คน Colorectal cancer = 11 คน Normal = 418 คน	Abnormal = 0.28380634 Colorectal cancer = 0.01836394 Normal = 0.69782972	Abnormal = 28% Colorectal cancer = 2% Normal = 70%

หมายเหตุ 1: สำหรับ patien_no ไม่มีในตารางนี้ เพราะ patien_no เป็นเพียงลำดับหมายเลขที่สามารถบ่งบอกตัวของตัวผู้ป่วยได้เท่านั้น ไม่สามารถเอามาคำนวณต่อได้

หมายเหตุ 2: เนื่องจากพื้นที่มีจำกัด จึงทำให้การแสดงผลข้อมูลในตารางนี้อาจดูไม่เรียบร้อยมากนัก

- Numerical data

Column	ความหมาย	Measures of Location			Measures of Spread			
		Mean	Median	Mode	Range	SD	Variance	Quantile
age	อายุ	56.25876	56	53	16	4.18927	17.54999	0% = 50 25% = 53 50% = 56 75% = 59 100% = 65
bmi	BMI	24.44508	23.9	22.9	29.56	3.972121	15.77775	0% = 15.20 25% = 21.80 50% = 23.90 75% = 26.70 100% = 44.76
hematocrit	Hematocrit	39.8177	39.8	40.099998	25.1	3.3076	10.94022	0% = 25.7 25% = 37.9 50% = 39.8 75% = 41.8 100% = 50.8

หมายเหตุ 1: เนื่องจากพื้นที่มีจำกัด จึงทำให้การแสดงผลข้อมูลในตารางนี้อาจดูไม่เรียบร้อยมากนัก

4. Data Visualization

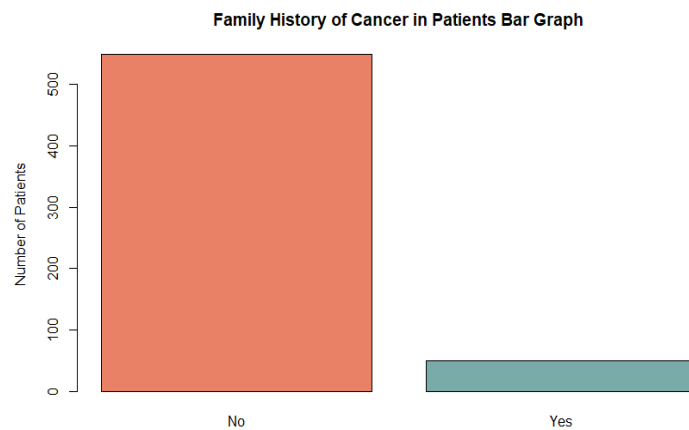
- Categorical data

1. fh: ประวัติครอบครัวเป็นมะเร็ง

- Frequency Table

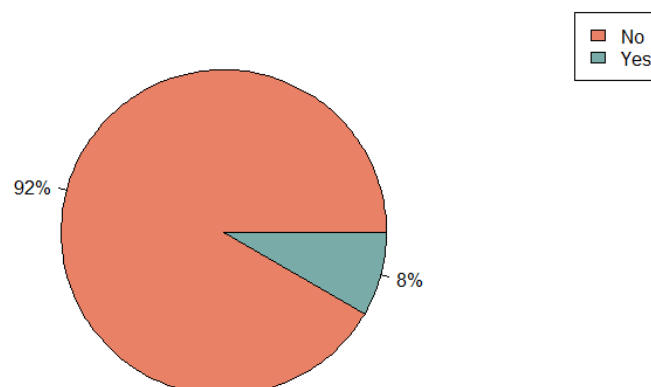
No (Number of Person)	Yes (Number of Person)
549	50

- Bar Graph



- Pie Chart

Family History of Cancer in Patients Pie Chart

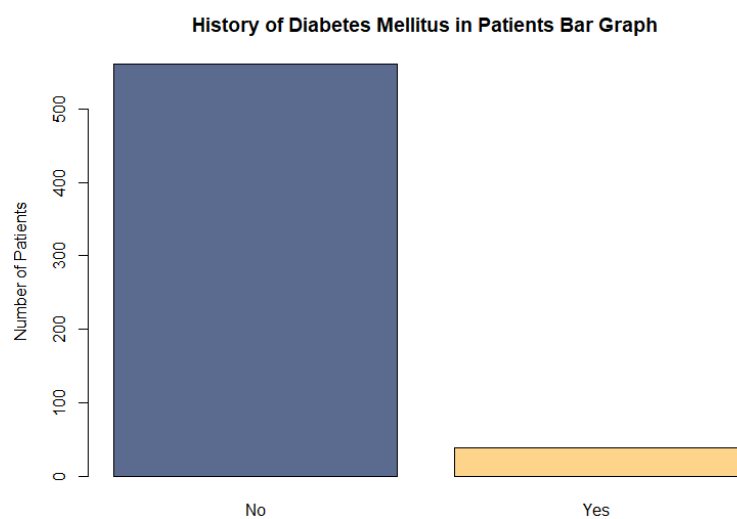


2. dm: ประวัติโรคเบาหวาน

- Frequency Table

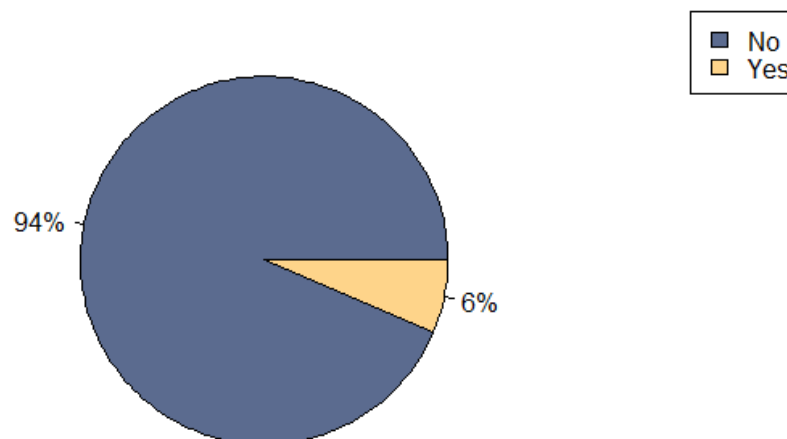
No (Number of Person)	Yes (Number of Person)
561	38

- Bar Graph



- Pie Chart

History of Diabetes Mellitus in Patients Pie Chart

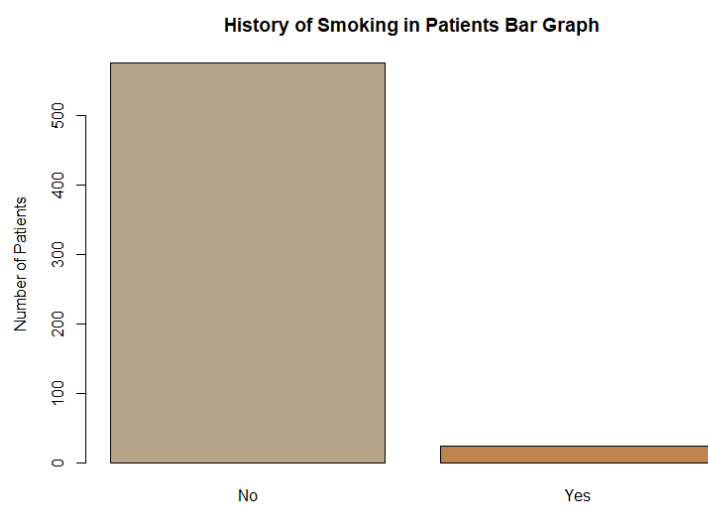


3. smoking: ประวัติการสูบบุหรี่

- Frequency Table

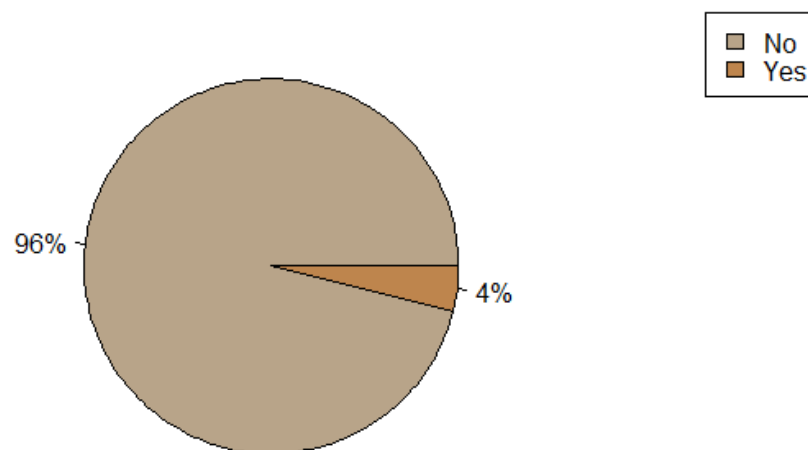
No (Number of Person)	Yes (Number of Person)
576	23

- Bar Graph



- Pie Chart

History of Smoking in Patients Pie Chart

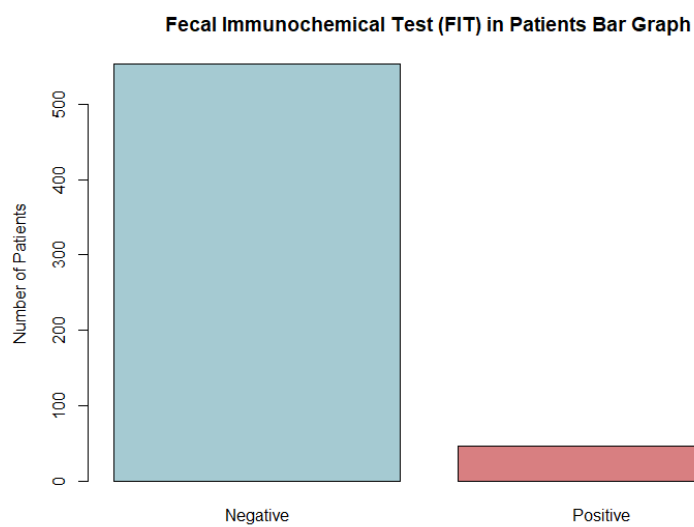


4. fit: Fecal immunochemical test (FIT)

- Frequency Table

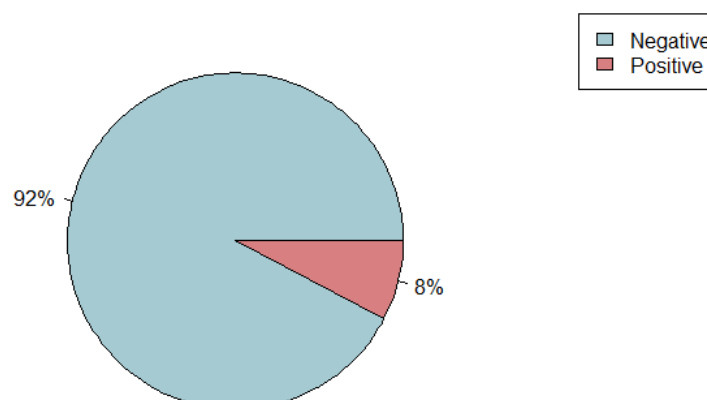
Negative (Number of Person)	Positive (Number of Person)
553	46

- Bar Graph



- Pie Chart

Fecal Immunochemical Test (FIT) in Patients Pie Chart

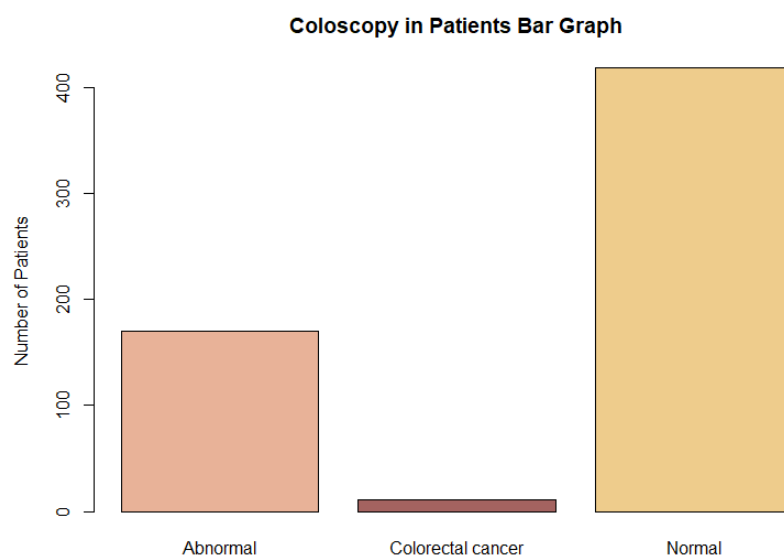


5. colonoscopy: ผลการส่องกล้อง

- Frequency Table

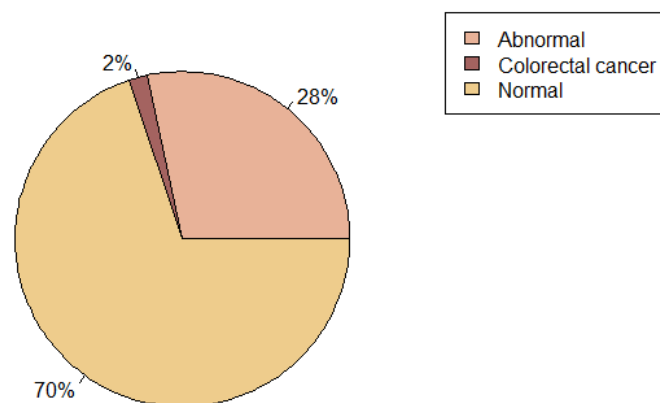
Abnormal (Number of Person)	Colorectal cancer (Number of Person)	Normal (Number of Person)
170	11	418

- Bar Graph



- Pie Chart

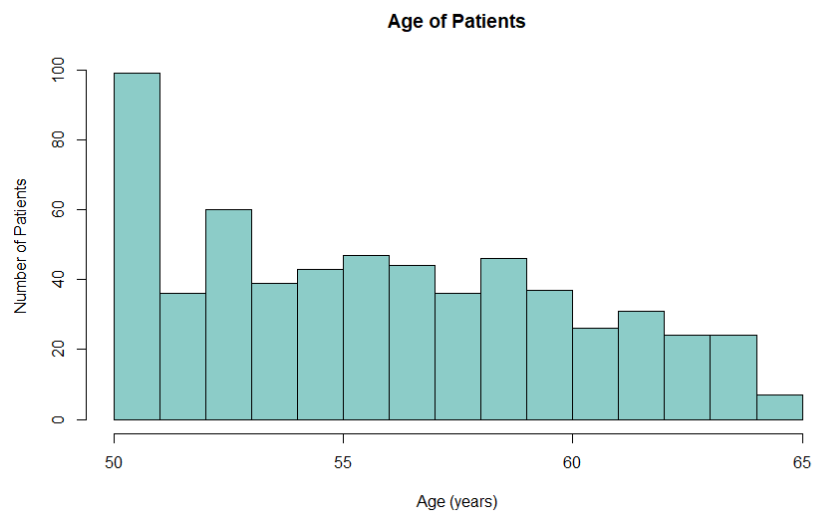
Colonoscopy in Patients Pie Chart



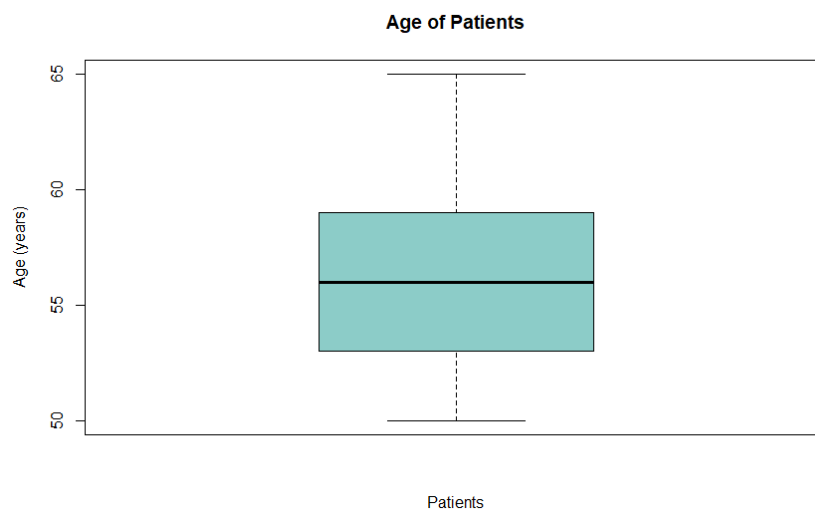
- Numerical data

- age: อายุ

- Histograms

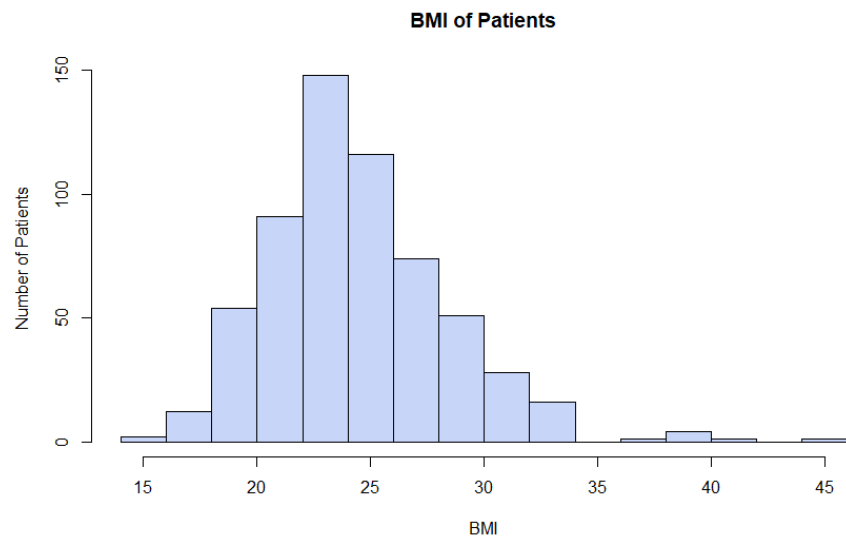


- Boxplot

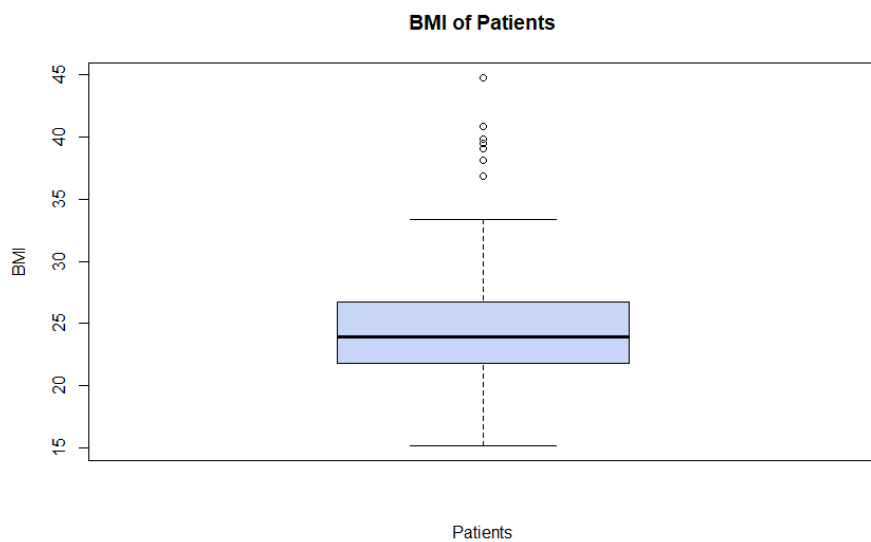


2. bmi

- Histograms

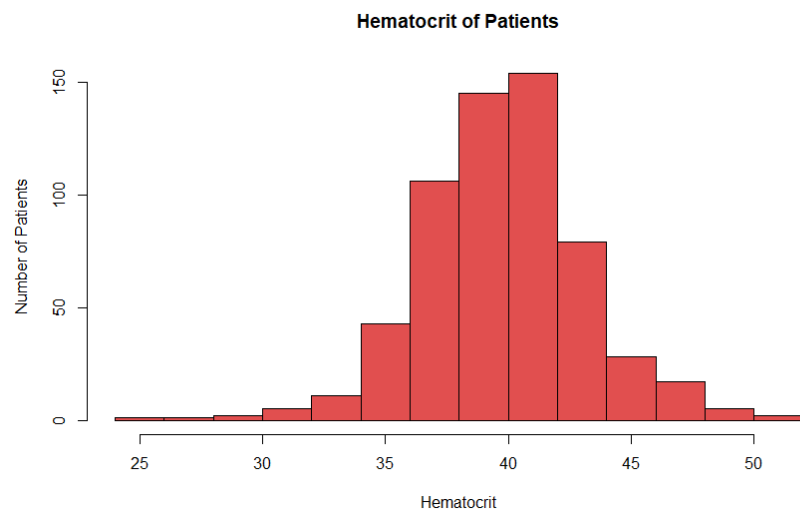


- Boxplot



3. hematocrit

- Histograms



- Boxplot

