

Information about project

- Hi, welcome to my project. I want to inform you about my project. I work with two datasets. One of them began at end of the November (Heart Attack). After you said the dataset is more important for you I create data(Home) by myself then I work with this data.





Bahram Muzaffarli

Project

Two Data

1. Heart Attack (from kaggle.com)
2. Home (create by myself)



Every single code in this presentation was written by myself.



	age	sex	cp	trtbps	chol	fbs	restecg	thalachh	exng	oldpeak	slp	caa	thall	output
1	63	1	3	145	233	1	0	150	0	2.3	0	0	1	1
2	37	1	2	130	250	0	1	187	0	3.5	0	0	2	1
3	41	0	1	130	204	0	0	172	0	1.4	2	0	2	1
4	56	1	1	120	236	0	1	178	0	0.8	2	0	2	1
5	57	0	0	120	354	0	1	163	1	0.6	2	0	2	1
6	57	1	0	140	192	0	1	148	0	0.4	1	0	1	1
7	56	0	1	140	294	0	0	153	0	1.3	1	0	2	1
8	44	1	1	120	263	0	1	173	0	0.0	2	0	3	1
9	52	1	2	172	199	1	1	162	0	0.5	2	0	3	1
10	57	1	2	150	168	0	1	174	0	1.6	2	0	2	1
11	54	1	0	140	239	0	1	160	0	1.2	2	0	2	1
12	48	0	2	130	275	0	1	139	0	0.2	2	0	2	1
13	49	1	1	130	266	0	1	171	0	0.6	2	0	2	1
14	64	1	3	110	211	0	0	144	1	1.8	1	0	2	1
15	58	0	3	150	283	1	0	162	0	1.0	2	0	2	1
16	50	0	2	120	219	0	1	158	0	1.6	1	0	2	1
17	58	0	2	120	340	0	1	172	0	0.0	2	0	2	1
18	66	0	3	150	226	0	1	114	0	2.6	0	0	2	1
19	43	1	0	150	247	0	1	171	0	1.5	2	0	2	1
20	69	0	3	140	239	0	1	151	0	1.8	2	2	2	1

Showing 1 to 21 of 303 entries, 14 total columns

Main points of data

Variable Types

```
HeartAttack 303 obs. of 14 variables
 $ age      : int  63 37 41 56 57 57 56 44 52 57 ...
 $ sex      : int  1 1 0 1 0 1 0 1 1 1 ...
 $ cp       : int  3 2 1 1 0 0 1 1 2 2 ...
 $ trtbps   : int  145 130 130 120 120 140 140 120 172 150 ...
 $ chol     : int  233 250 204 236 354 192 294 263 199 168 ...
 $ fbs      : int  1 0 0 0 0 0 0 0 1 0 ...
 $ restecg  : int  0 1 0 1 1 1 0 1 1 1 ...
 $ thalachh : int  150 187 172 178 163 148 153 173 162 174 ...
 $ exng     : int  0 0 0 0 1 0 0 0 0 0 ...
 $ oldpeak  : num  2.3 3.5 1.4 0.8 0.6 0.4 1.3 0 0.5 1.6 ...
 $ slp      : int  0 0 2 2 2 1 1 2 2 2 ...
 $ caa      : int  0 0 0 0 0 0 0 0 0 0 ...
 $ thall    : int  1 2 2 2 2 1 2 3 3 2 ...
 $ output   : int  1 1 1 1 1 1 1 1 1 1 ...
```

```
← →  Source on Save
1 HeartAttack = read.csv("C:/Users/FX505DT/Desktop/heart.csv", header=TRUE, stringsAsFactors=FALSE)
2 #https://www.kaggle.com/datasets/rashikrahmanpritom/heart-attack-analysis-prediction-datasetlibrary(qqplot2)
3 install.packages('tidyverse')
4 # age - Age of the person
5 # sex - Gender of the person
6 # cp - Chest Pain type chest pain type
7   # 0 = Typical Angina, 1 = Atypical Angina, 2 = Non-anginal Pain, 3 = Asymptomatic
8 # trtbps - resting blood pressure (in mm Hg)
9 # chol - cholesterol in mg/dl fetched via BMI sensor
10 # fbs - (fasting blood sugar > 120 mg/dl) (1 = true; 0 = false)
11 # restecg - resting electrocardiographic results
12 # thalachh - maximum heart rate achieved
13 # exng - exercise induced angina (1 = yes; 0 = no)
14 # oldpeak - Previous peak
15 # slp - Slope
16 # caa - Number of major vessels
17 # thall - Thallium Stress Test result ~ (0,3)
18 # output - Target variable
19
20
```

Heart Attack Data

6.956

25.001

23.795

4.852

1.803

35.891

1.047

18.65

5.004

25.778

41.884

44.421

42.268

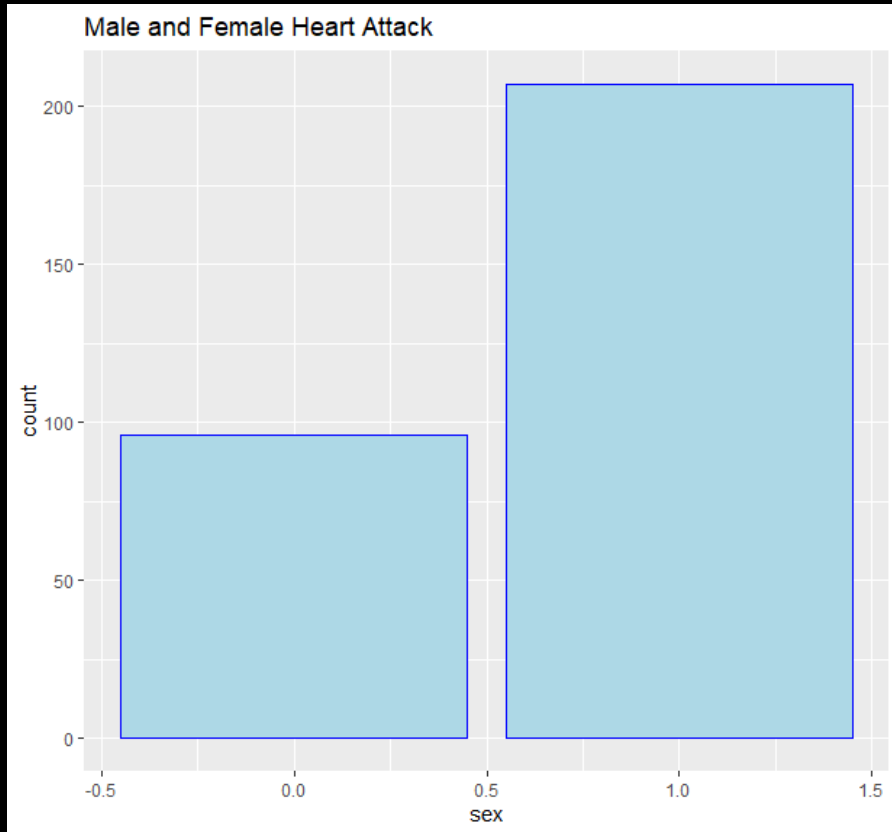
27.017

22.608

19.307

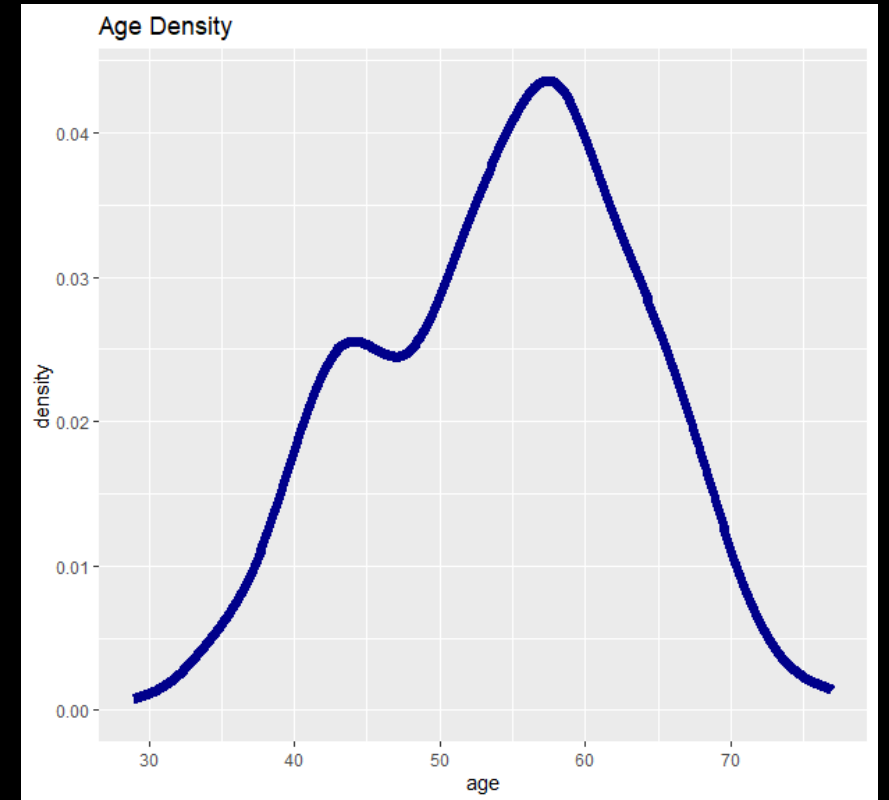
15.822

Male and Female heart attack



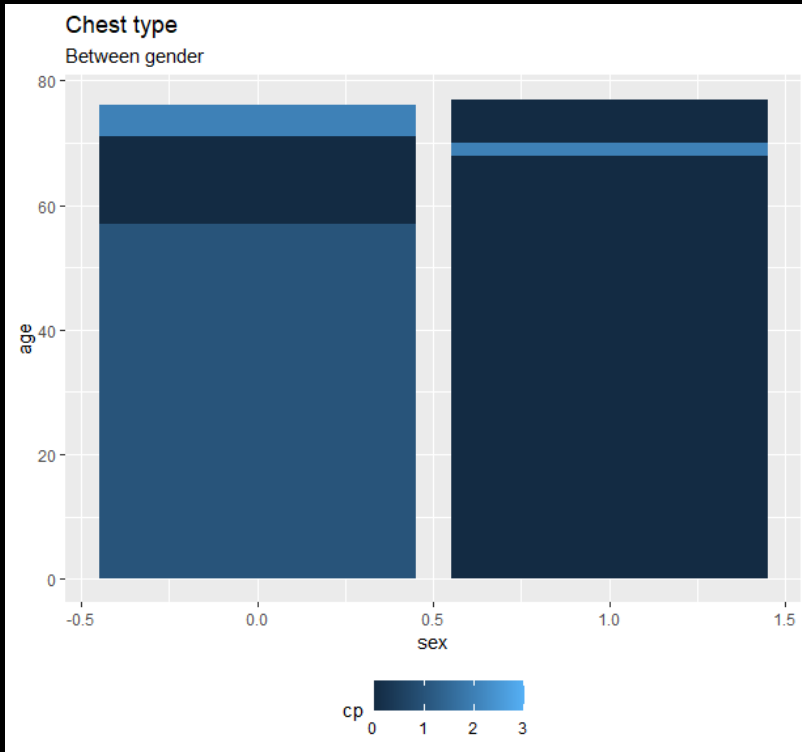
```
#Male and Female heart attack
ggplot(HeartAttack, aes(sex)) + geom_bar(color="blue", fill='light blue' )+
  labs(title= 'Male and Female Heart Attack')+
```

Age Density



```
#Age Density
ggplot(HeartAttack, aes(age)) +
  geom_density(lwd=2.5, col='dark blue')+
  labs(title='Age Density')
```

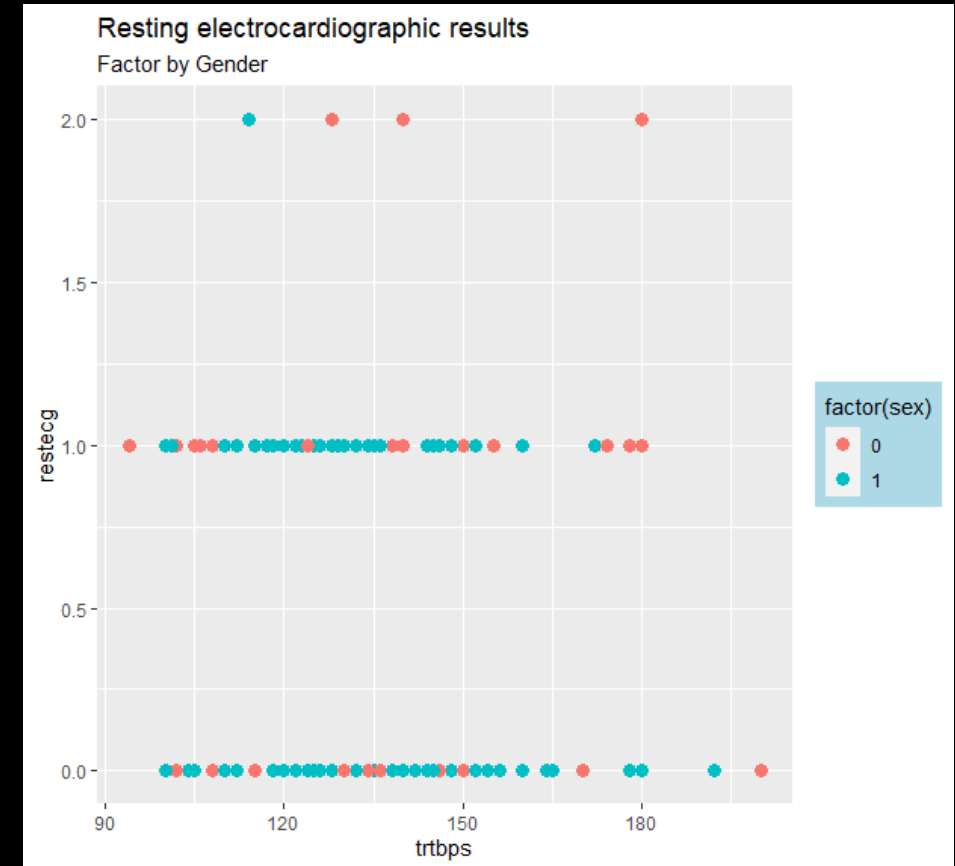
Chest Type Between Gender



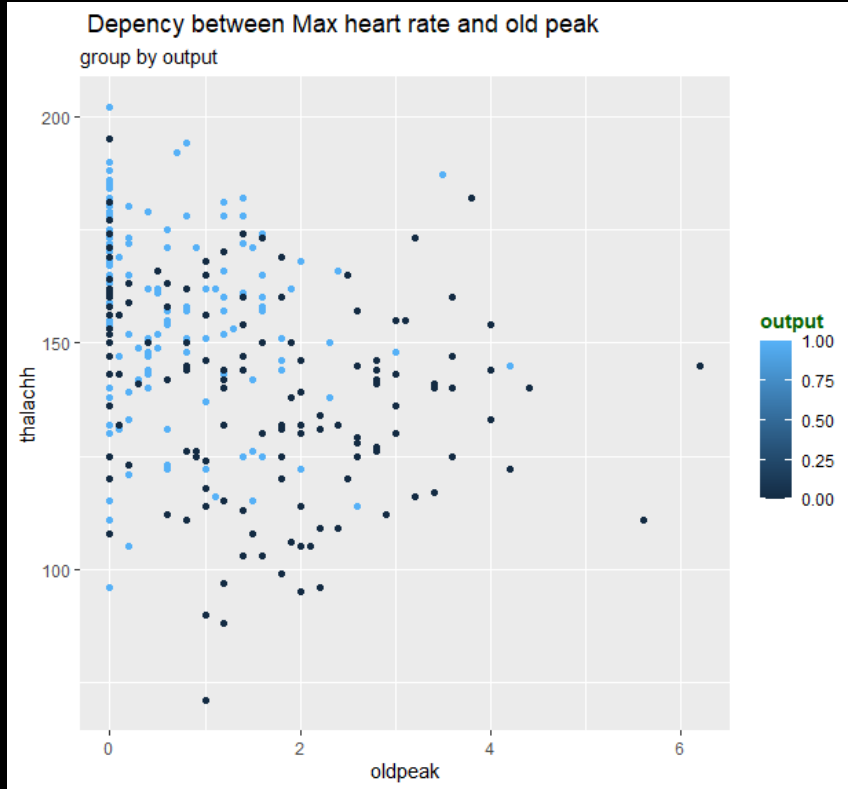
```
#ChestType Between Gender
ggplot(HeartAttack,aes(fill=cp,x=sex,y=age)) +
geom_bar(position="dodge", stat="identity")+
labs(title='Chest type', subtitle='Between gender')+
theme(legend.position = 'bottom')
```

Resting result

```
ggplot(HeartAttack,aes(x=trtbps,y=restecg,color=factor(sex)))+
geom_point(size=3)+
theme(legend.background = element_rect(fill='lightblue',size=0.5, linetype = 'solid'))+
labs(title='Resting electrocardiographic results',subtitle='Factor by Gender')
```



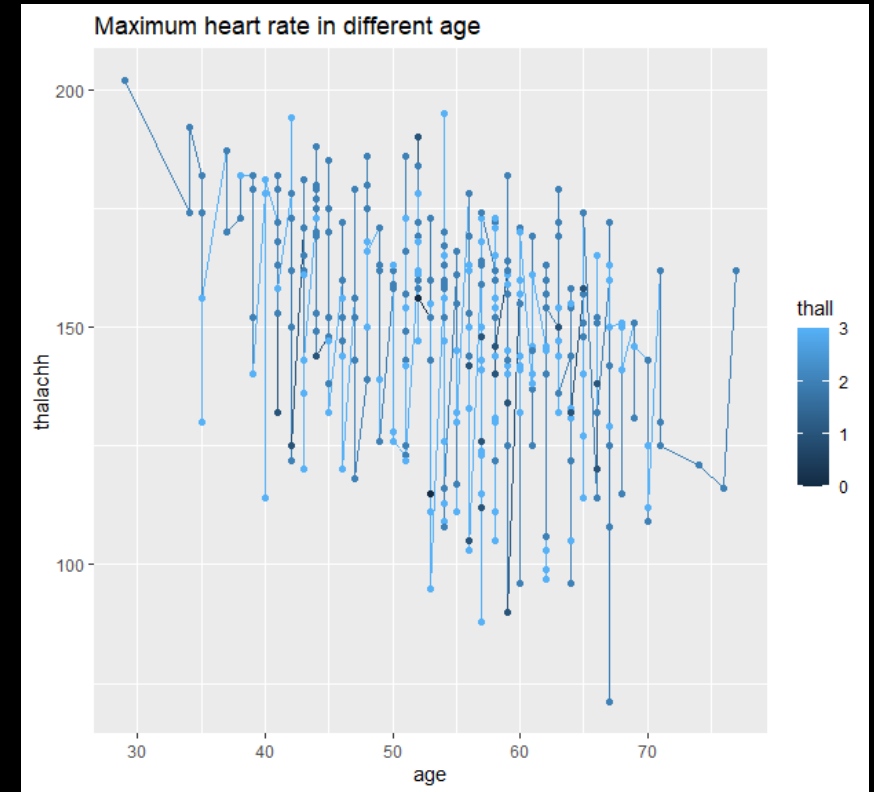
Max heart rate and old peak dependency



Group by output information

```
ggplot(HeartAttack, aes(oldpeak, thalachh, colour = output)) +  
  geom_point() +  
  theme(legend.title = element_text(colour = 'darkgreen', face = 'bold')) +  
  labs(title = 'Dependency between Max heart rate and old peak', subtitle = 'group by output')
```

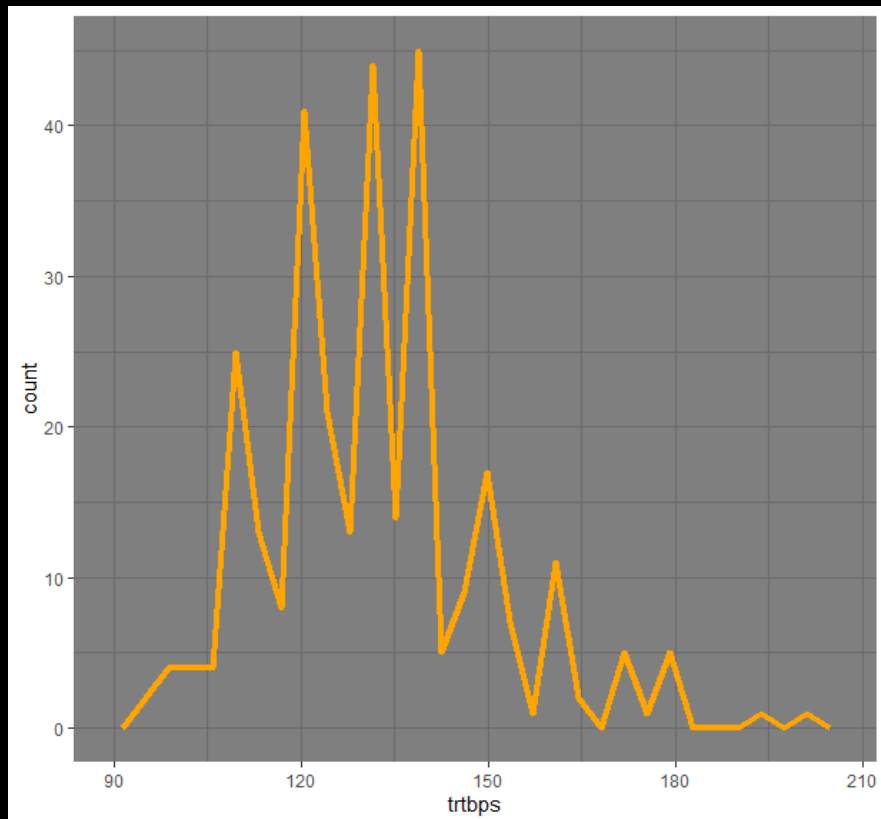
Max heart rate in different age



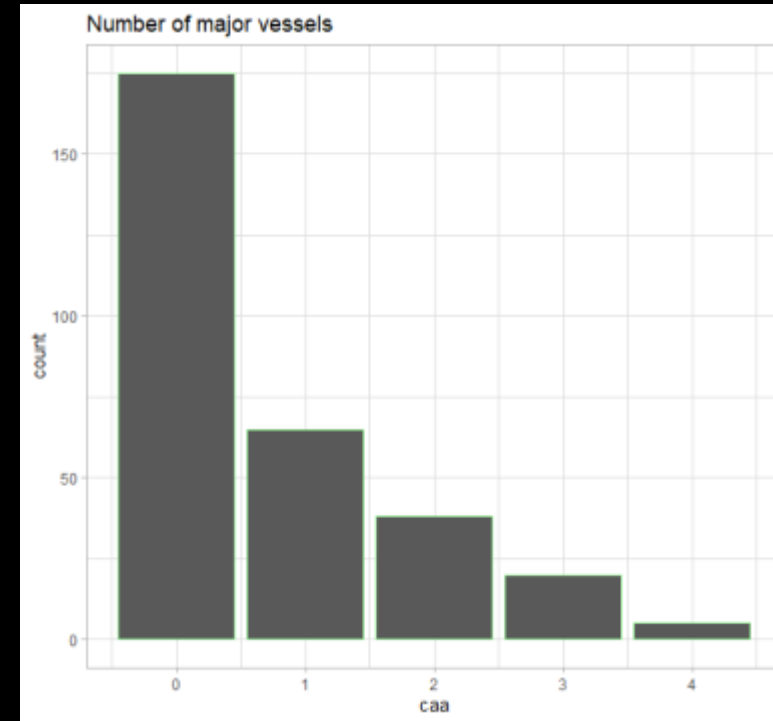
```
heart_rate = ggplot(HeartAttack, aes(age, thalachh, colour = thall))  
|  
heart_rate +  
  geom_line() +  
  geom_point() +  
  labs(title = 'Maximum heart rate in different age')
```


Blood pressure density

```
ggplot(HeartAttack,aes(x=trtbps))+  
  geom_freqpoly(color='orange',size=1.5)+  
  theme_dark()
```



Number of major vessels

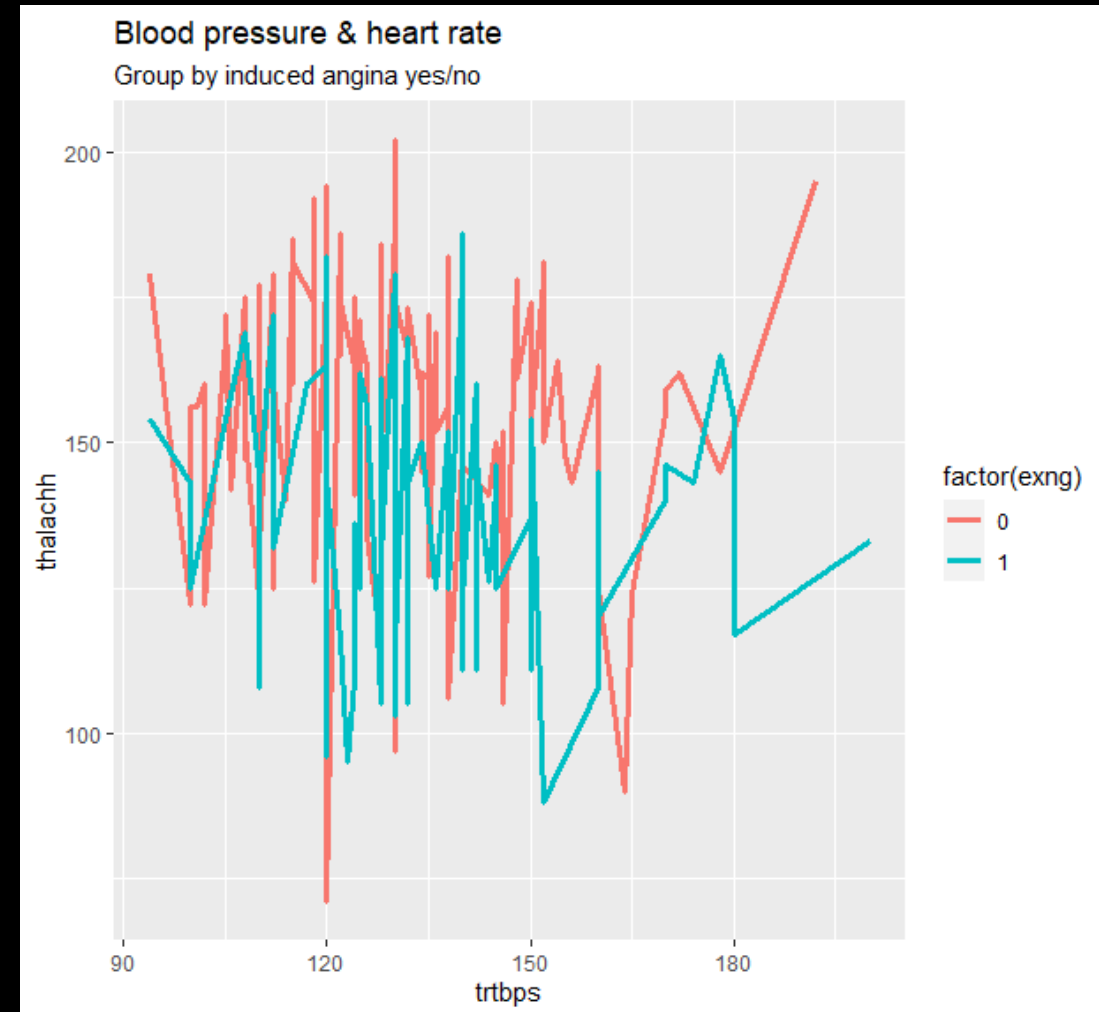


```
ggplot(HeartAttack) +  
  geom_bar(aes(x = caa), position = 'dodge',color='lightgreen') +  
  labs(title='Number of major vessels')+  
  theme_light()
```

Dependency between blood pressure and heart rate

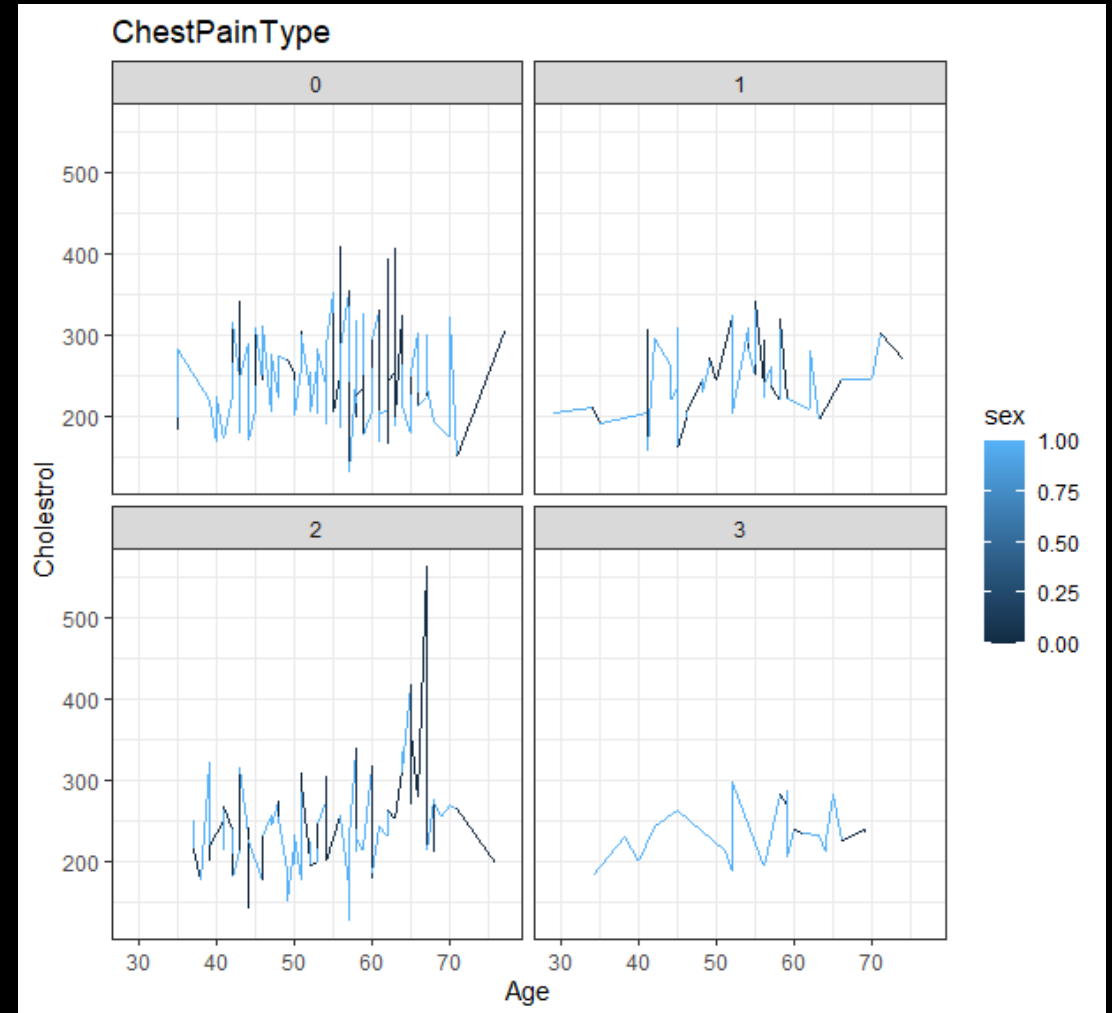
Angina = yes/not

```
ggplot(HeartAttack) +  
  geom_line(aes(x = trtbps, y = thalachh, color = factor(exng)), lwd = 1.25) +  
  labs(title='Blood pressure & heart rate',  
        subtitle = 'Group by induced angina yes/no')
```



Cholesterol by Age group by Chest Pain Type

```
ggplot(HeartAttack, mapping = aes(x = age, y = chol, color = sex)) +  
  geom_line() +  
  facet_wrap(facets = vars(cp)) +  
  labs(title = "ChestPainType",  
        x = "Age",  
        y = "Cholesterol") +  
  theme_bw()
```



House Data



Two Data

1. Heart Attack (from kaggle.com)
2. Home (create by myself)



Every single code in this presentation was written by myself.



	price	rent	area	room	m2	floor	sum.floor	citycentre	items	quality
1	183000	400	Hezi Aslanov	3	130	10	13	0	2	2
2	172900	380	Hezi Aslanov	3	210	17	18	0	1	1
3	92800	450	Mikrorayon	2	72	6	9	0	2	3
4	425000	650	Nerimanov	3	177	4	16	1	2	3
5	560000	800	Nesimi	5	255	0	0	1	2	4
6	642000	1200	Baki	4	215	0	0	1	2	4
7	133000	300	Baki	1	62	0	0	1	0	0
8	375000	600	Genclik	4	207	0	0	1	1	4
9	420000	650	Xetai	3	160	10	20	1	2	4
10	180000	350	8 Noyabr	3	125	6	14	1	1	3
11	115000	300	Mikrorayon	2	68	2	9	0	0	1
12	255000	500	Xetai	3	132	5	7	1	2	3
13	185000	350	Mikrorayon	4	130	1	9	0	1	2
14	90000	400	Yasamal	2	50	12	20	1	1	2
15	132000	380	Nizami r.	2	70	8	19	1	0	0
16	115000	300	Mikrorayon	2	70	2	9	0	0	0
17	350000	550	Mikrorayon	4	200	4	18	0	2	3
18	310000	600	Sahil m.	3	100	2	5	1	1	3
19	165000	480	Nerimanov	2	80	14	17	1	1	1
20	280000	500	Xetai	3	113	5	11	1	2	2

Main points of data

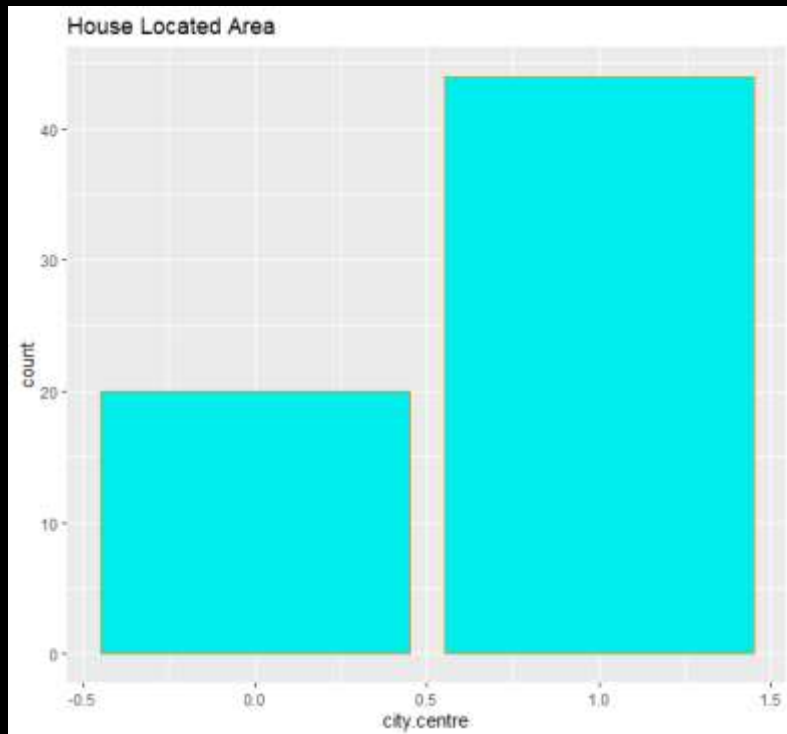
variable Types

Home		64 obs. of 10 variables									
\$ price	: int	183000	172900	92800	425000	560000	642000	...			
\$ rent	: int	400	380	450	650	800	1200	300	600	650	350
\$ area	: chr	"Hezi Aslanov"	"Hezi Aslanov"	"Mikrorayon..."							
\$ room	: int	3	3	2	3	5	4	1	4	3	3
\$ m2	: int	130	210	72	177	255	215	62	207	160	125
\$ floor	: int	10	17	6	4	0	0	0	0	10	6
\$ sum.floor	: int	13	18	9	16	0	0	0	0	20	14
\$ city.centre	: int	0	0	0	1	1	1	1	1	1	1
\$ items	: int	2	1	2	2	2	2	0	1	2	1
\$ quality	: int	2	1	3	3	4	4	0	4	4	3

```
1 Home = read.csv("C:/Users/FX505DT/Desktop/Home.csv", header=TRUE, stringsAsFactors=FALSE)
2 library(ggplot2)
3 #m2- Per Square Metr
4 #CityCentre 0 1
5 #Quality 0 1 2 3 4
6     # 0-Very Bad
7     # 1-Bad
8     # 2-Normal
9     # 3-Good
10    # 4-Very Good
11
12 #Items 0 1 2
13     # 0-Empty
14     # 1-Only Useful Things
15     # 2-Full
16
```

House Loacated Area (City Centre 0/1)

```
ggplot(Home,aes(citycentre))+  
  geom_bar(fill='cyan2',color='darkorange')+  
  labs(title='House Located Area')
```



Price Room counts dependency



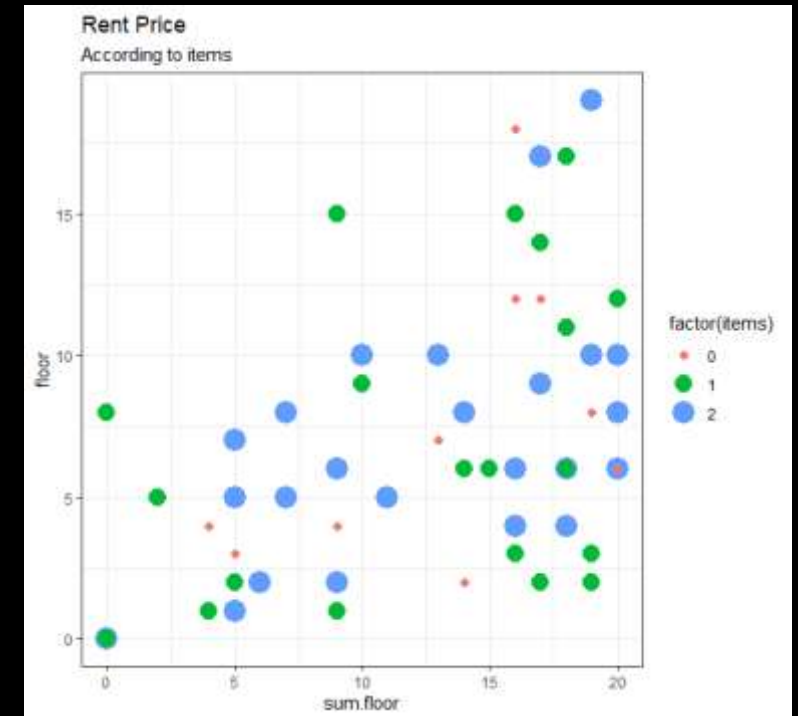
```
ggplot(Home,aes(price,room))+  
  geom_point(aes(shape = factor(room),color=factor(room)),size=3)+  
  geom_smooth()+  
  labs(title='Dependency between price and room')+  
  theme(legend.text = element_text(size=15,face='italic'))
```


Rent Price according to area



```
ggplot(Home,aes(rent,colour=factor(city centre)))+  
  geom_freqpoly(lwd=1,binwidth=50)+  
  labs(title='Rent Price',subtitle='According to area')+  
  theme(legend.background = element_rect(colour='red',fill='cyan4'),  
        legend.title = element_text(face='bold'))
```

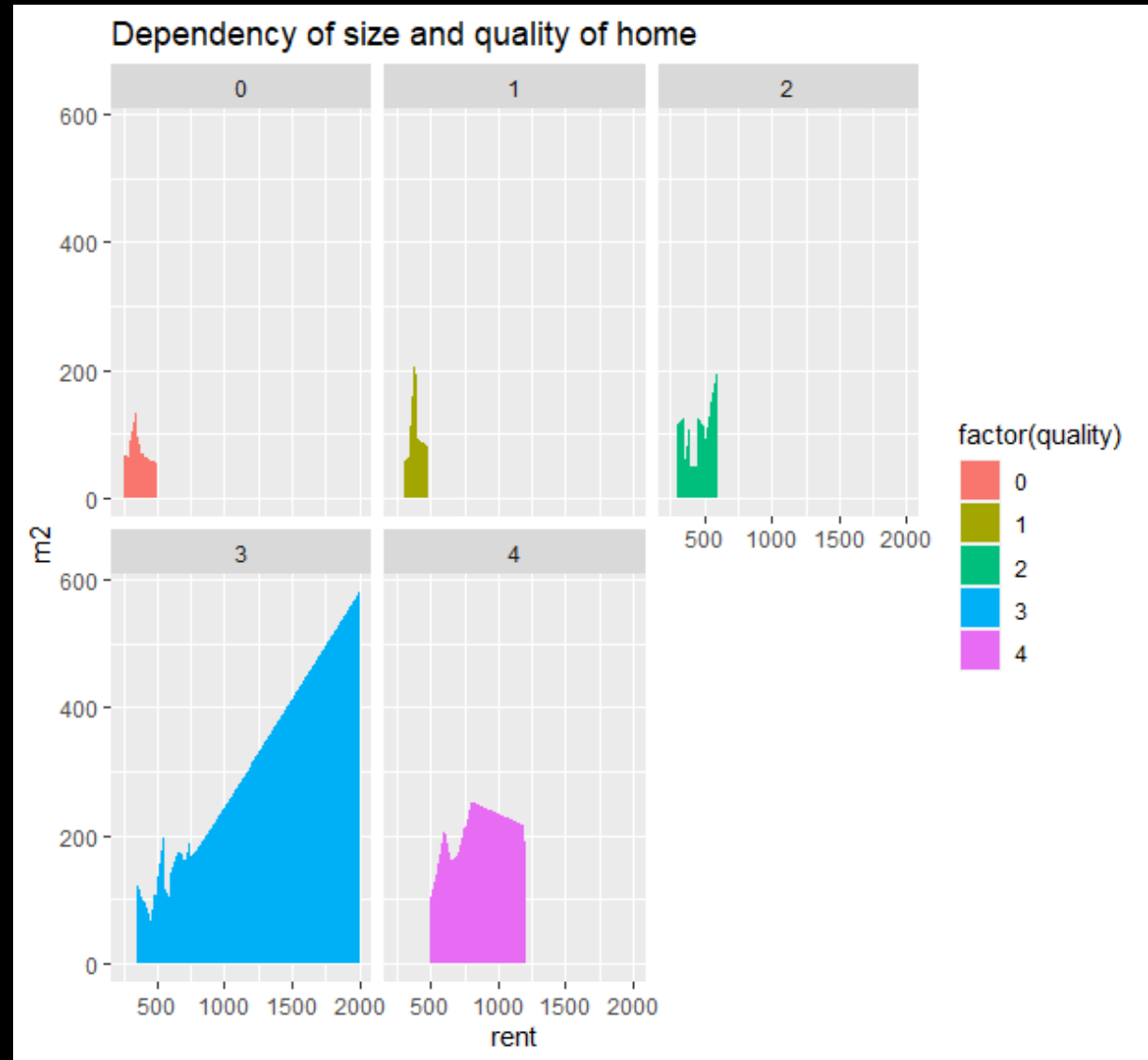
Rent Price according to items



```
ggplot(Home) +  
  geom_point(mapping = aes(x = sum.floor, y = floor,  
                           color = factor(items),  
                           size = factor(items)))+  
  labs(title='Rent Price',subtitle='According to items')+  
  theme_bw()
```

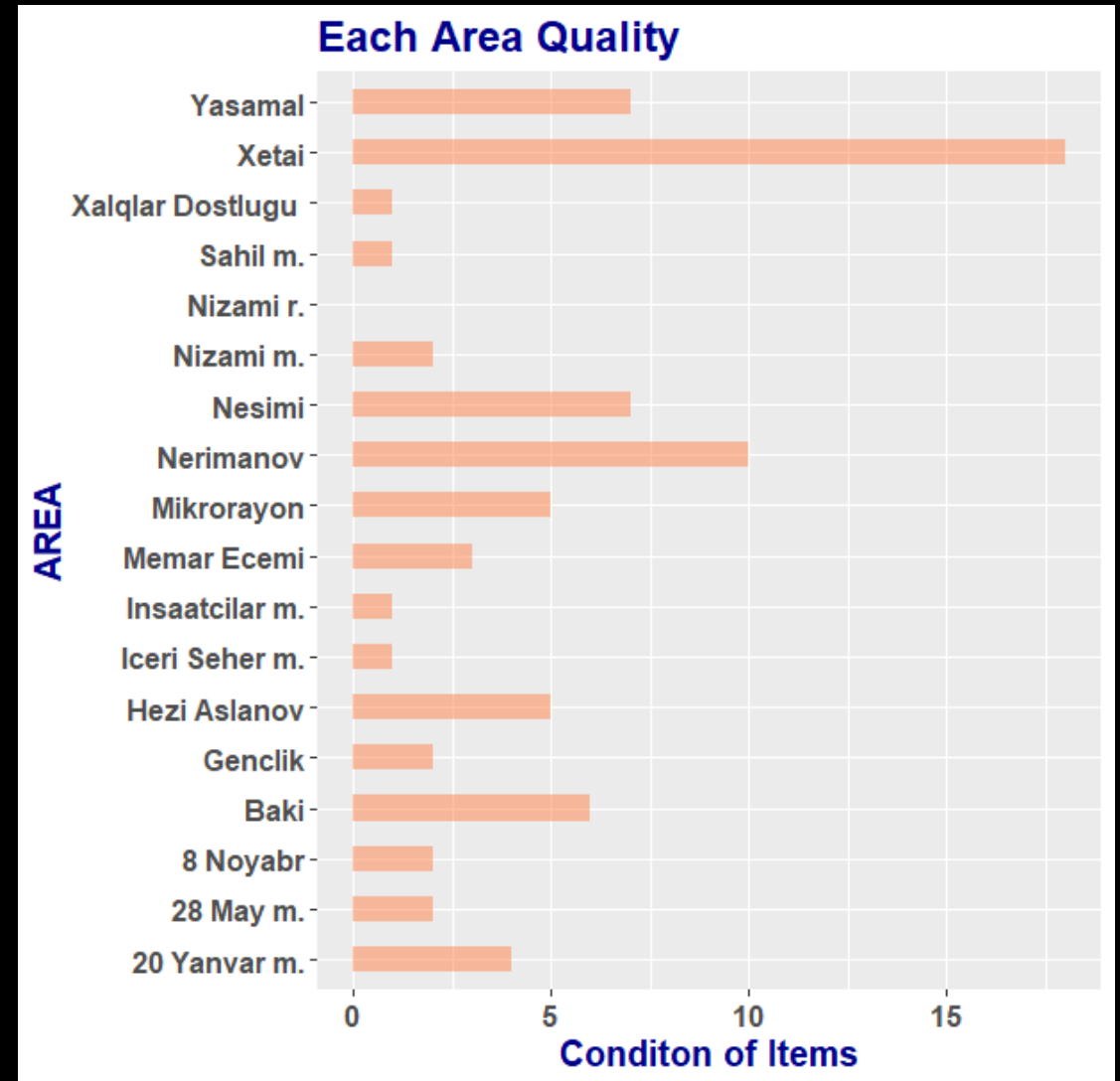
Surface and quality of house

```
ggplot(Home,aes(rent,m2,fill=factor(quality)))+  
  geom_area()+  
  labs(title='Dependency of size and quality of home')+  
  facet_wrap(~quality)
```



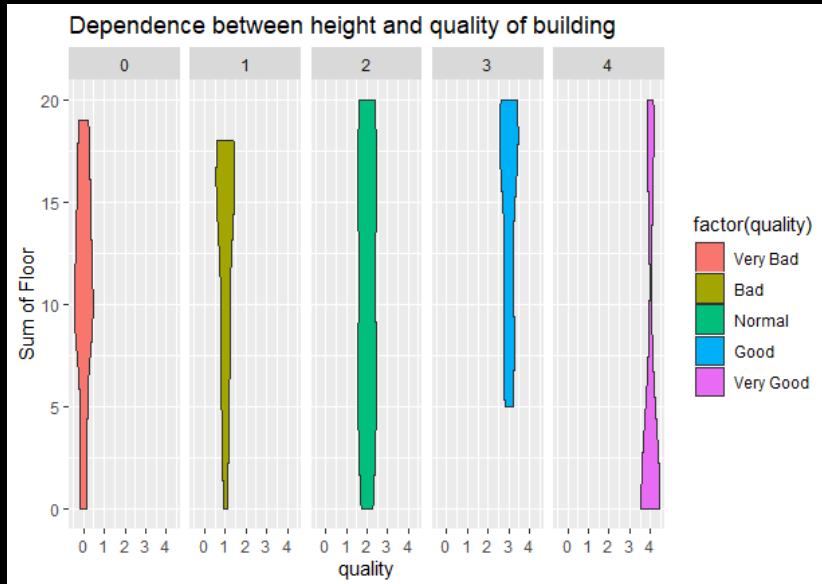
Items condition for each area

```
ggplot(Home, aes(x=area, y=items)) +  
  geom_bar(stat="identity", fill="sienna", alpha=0.5, width=0.5) +  
  coord_flip() +  
  labs(title='Each Area Quality', x='AREA', y='Conditon of Items') +  
  theme(text=element_text(size=15, colour='darkblue', face='bold'))
```



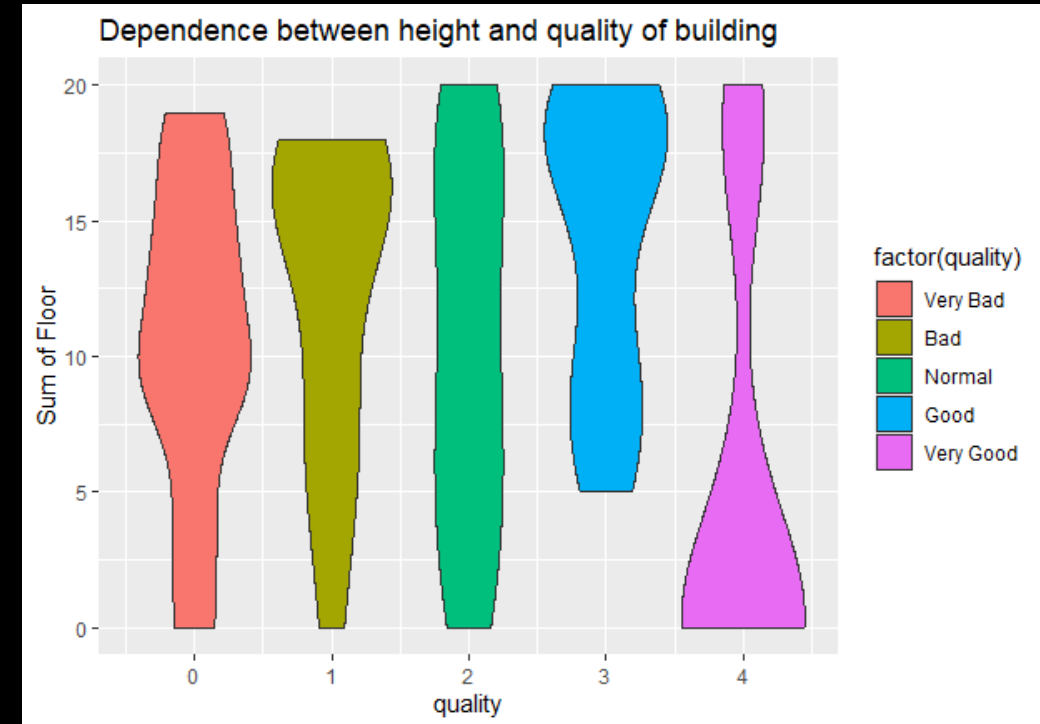
Dependence between height and quality of building

```
ggplot(Home,mapping = aes(x = quality, y = sum.floor, fill = factor(quality)))+  
  geom_violin()+  
  labs(title='Dependence between height and quality of building')+  
  ylab('Sum of Floor')+  
  facet_grid(~quality)+  
  scale_fill_viridis_d()+  
  scale_fill_discrete(labels = c('Very Bad','Bad','Normal','Good','Very Good'))
```



Without facet_grid

With facet_grid



Thank You



For your attention !