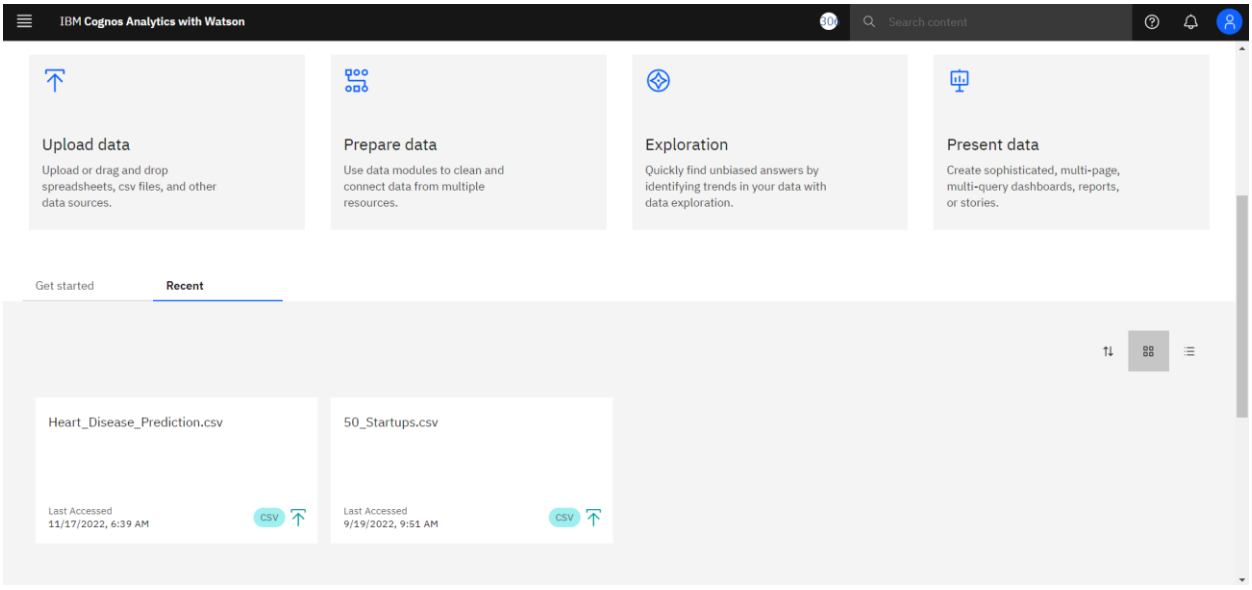
Project Development Phase

Sprint 1

Date	30 October 2022
Team id	PNT2022TMID21213
Project name	VISUALIZING AND PREDICTING HEART DISEASEWITH AN INTERACTIVE DASHBOARD

LOADING DATASET

First we upload the given dataset “Heart disease Prediction Dataset” in IBM Cognos:

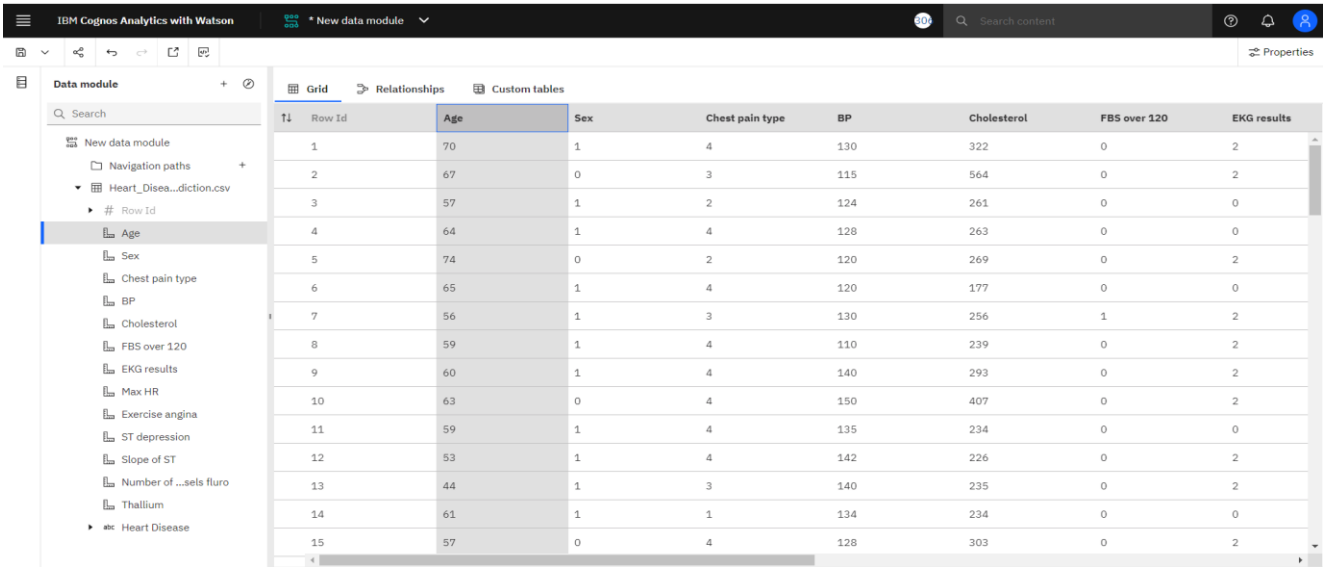


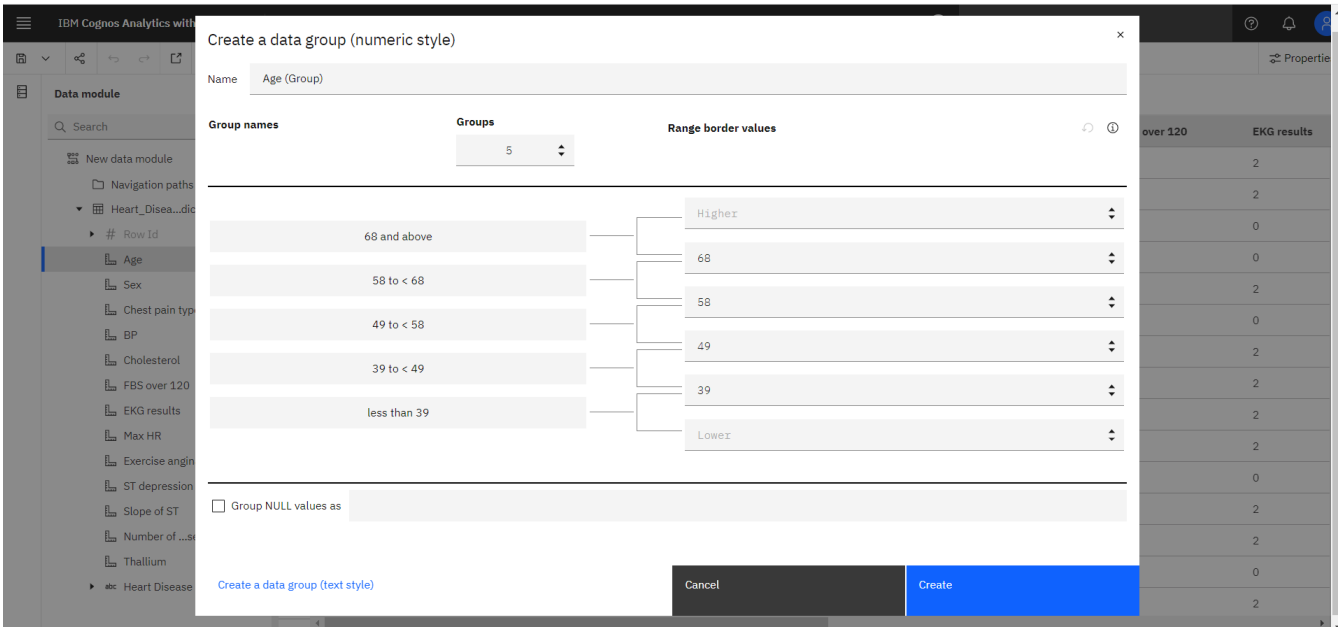
PREPARATION OF DATA

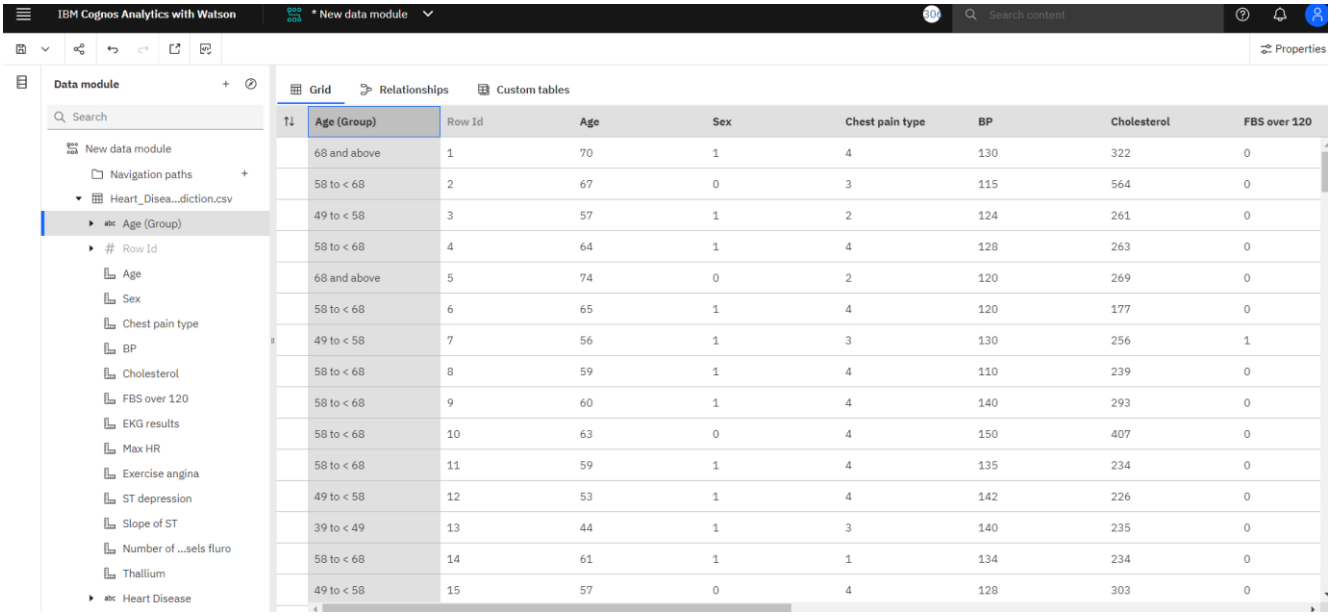
Now we work to process the dataset
In IBM Cognos there is option of preparation of dataset, with which the basic preprocessing is done easily.
Normally this has to be done in python , with codes via jupyter notebook/colab.
But here the same preprocessing can be done just by GUI also the transit and combination of data as tables is easier.

Here we group the age which contains various values according to the user/patient.
For easier visualization and prediction we combine them into groups.

AGE GROUPS



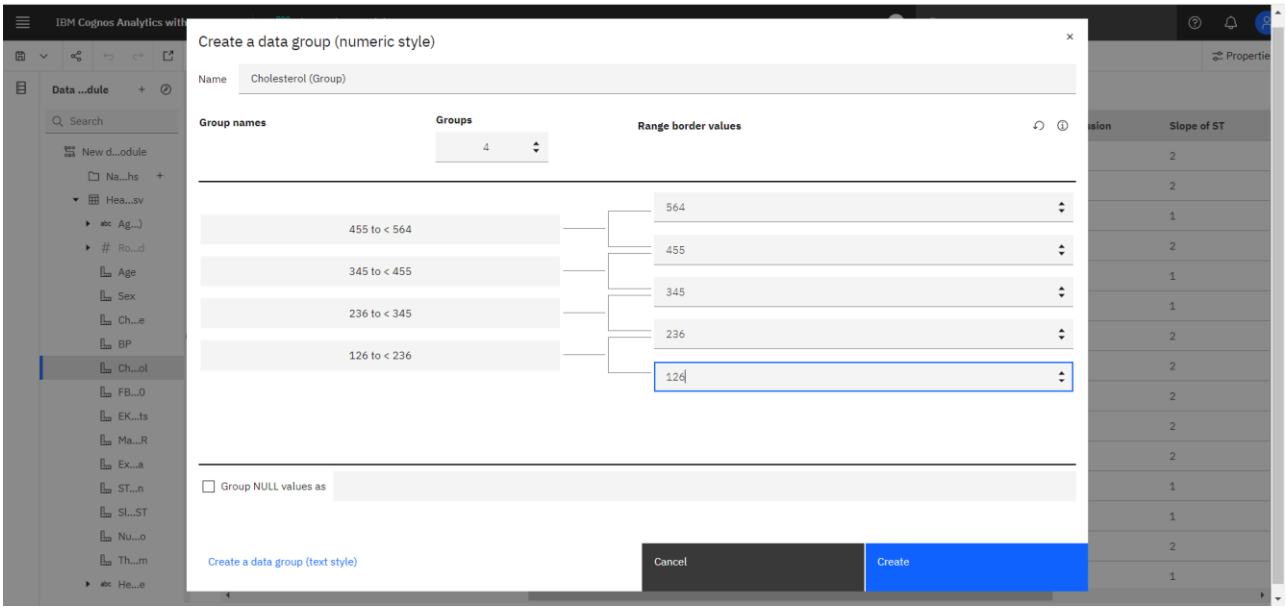


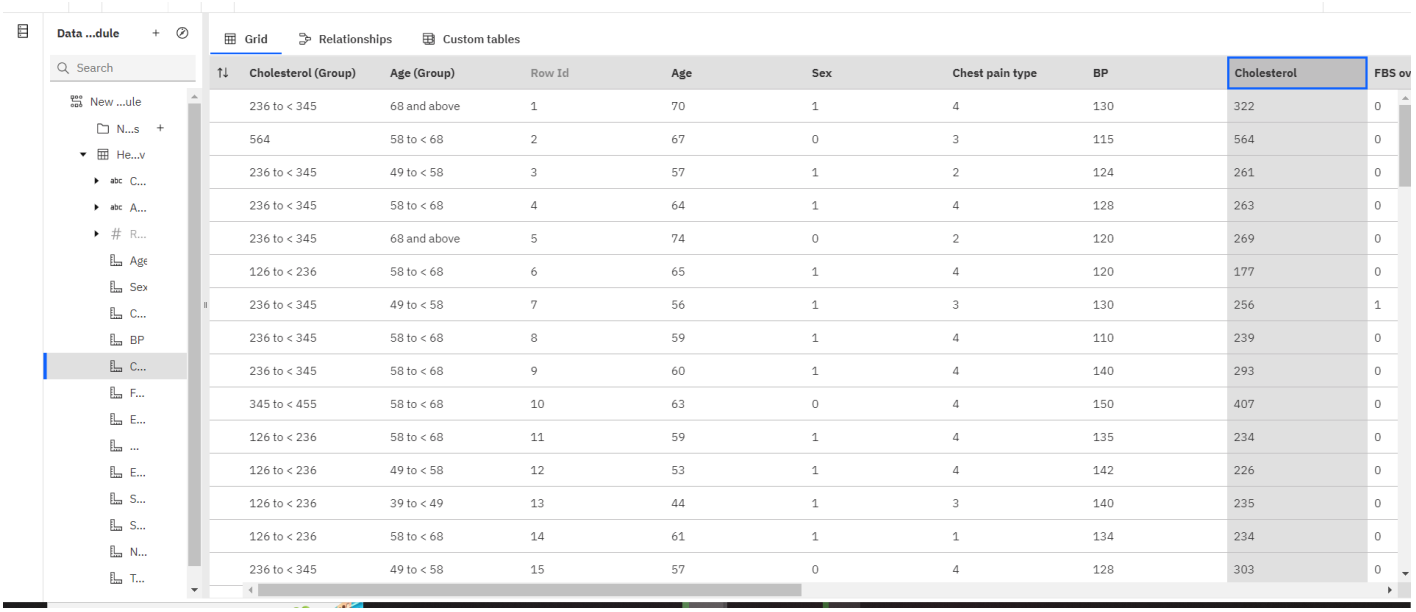


Age (Group)	Row Id	Age	Sex	Chest pain type	BP	Cholesterol	FBS over 120
68 and above	1	70	1	4	130	322	0
58 to < 68	2	67	0	3	115	564	0
49 to < 58	3	57	1	2	124	261	0
58 to < 68	4	64	1	4	128	263	0
68 and above	5	74	0	2	120	269	0
58 to < 68	6	65	1	4	120	177	0
49 to < 58	7	56	1	3	130	256	1
58 to < 68	8	59	1	4	110	239	0
58 to < 68	9	60	1	4	140	293	0
58 to < 68	10	63	0	4	150	407	0
58 to < 68	11	59	1	4	135	234	0
49 to < 58	12	53	1	4	142	226	0
39 to < 49	13	44	1	3	140	235	0
58 to < 68	14	61	1	1	134	234	0
49 to < 58	15	57	0	4	128	303	0

Now the cholesterol of the user is grouped so that it can be assigned variables in range.

CHOLESTEROL USER GROUP

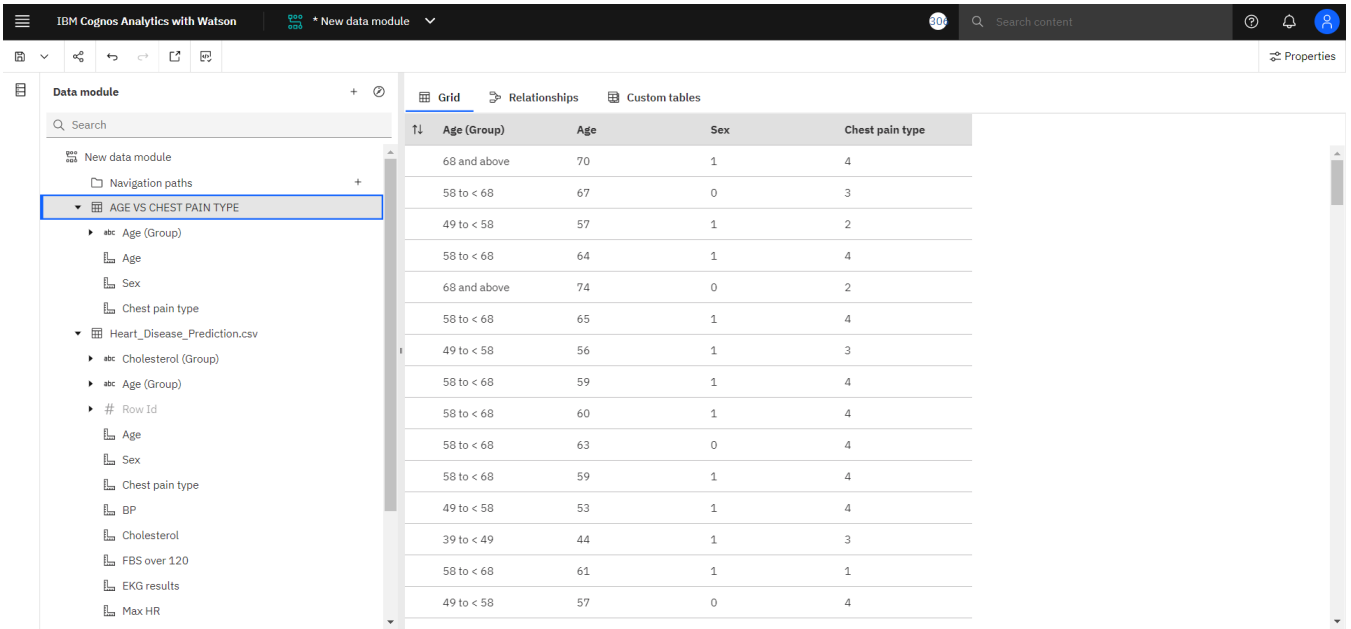




Cholesterol (Group)	Age (Group)	Row Id	Age	Sex	Chest pain type	BP	Cholesterol	FBS ov
236 to < 345	68 and above	1	70	1	4	130	322	0
564	58 to < 68	2	67	0	3	115	564	0
236 to < 345	49 to < 58	3	57	1	2	124	261	0
236 to < 345	58 to < 68	4	64	1	4	128	263	0
236 to < 345	68 and above	5	74	0	2	120	269	0
126 to < 236	58 to < 68	6	65	1	4	120	177	0
236 to < 345	49 to < 58	7	56	1	3	130	256	1
236 to < 345	58 to < 68	8	59	1	4	110	239	0
236 to < 345	58 to < 68	9	60	1	4	140	293	0
345 to < 455	58 to < 68	10	63	0	4	150	407	0
126 to < 236	58 to < 68	11	59	1	4	135	234	0
126 to < 236	49 to < 58	12	53	1	4	142	226	0
126 to < 236	39 to < 49	13	44	1	3	140	235	0
126 to < 236	58 to < 68	14	61	1	1	134	234	0
236 to < 345	49 to < 58	15	57	0	4	128	303	0

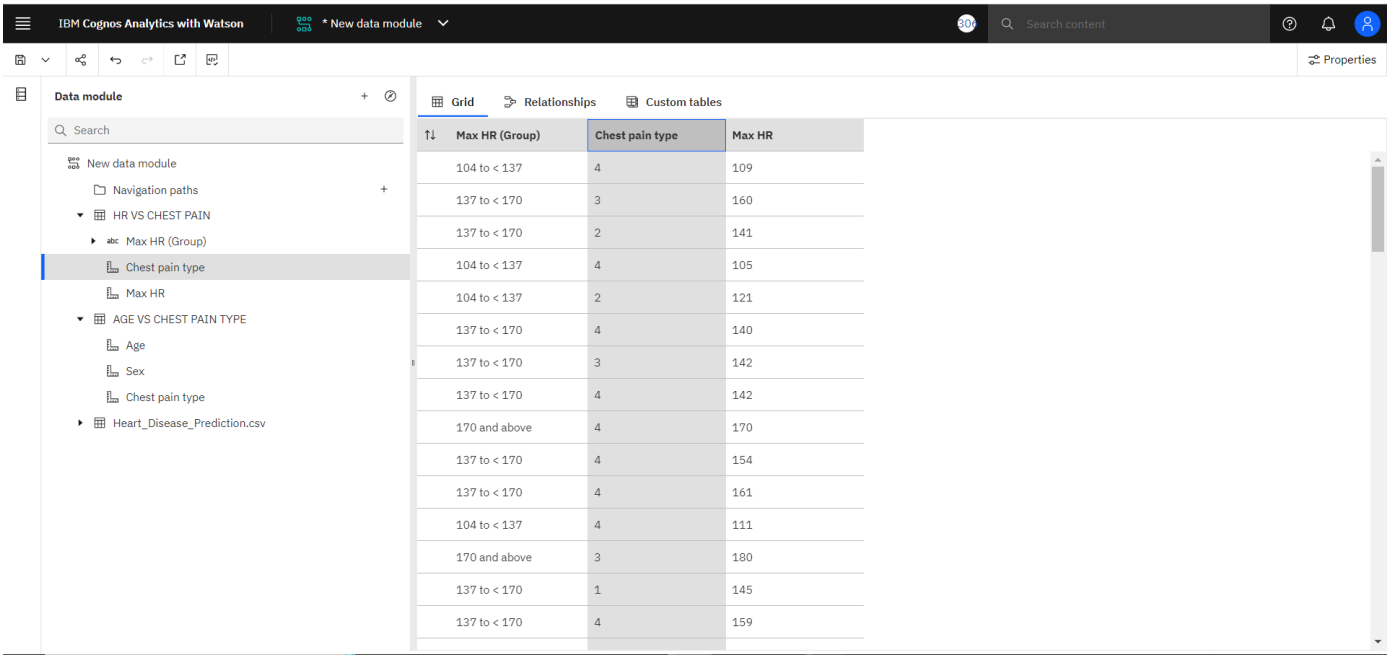
To compare how there is relationship between various variables wrt to independent and dependent variables we create different table copies. Also we join them in order to find the correct relationship.

NEW TABLE FOR AGE VS CHEST PAIN TYPE



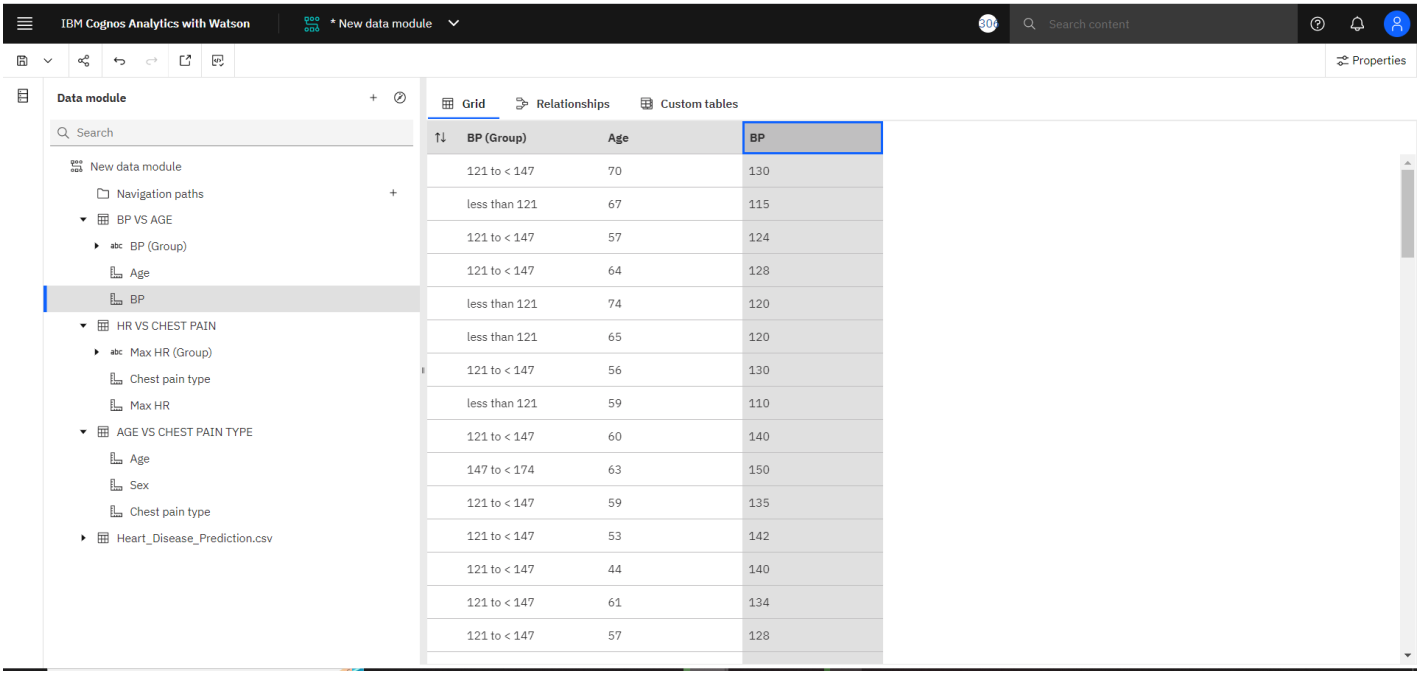
Age (Group)	Age	Sex	Chest pain type
68 and above	70	1	4
58 to < 68	67	0	3
49 to < 58	57	1	2
58 to < 68	64	1	4
68 and above	74	0	2
58 to < 68	65	1	4
49 to < 58	56	1	3
58 to < 68	59	1	4
58 to < 68	60	1	4
58 to < 68	63	0	4
58 to < 68	59	1	4
49 to < 58	53	1	4
39 to < 49	44	1	3
58 to < 68	61	1	1
49 to < 58	57	0	4

TABLE FOR MAX HR VS CHEST PAIN



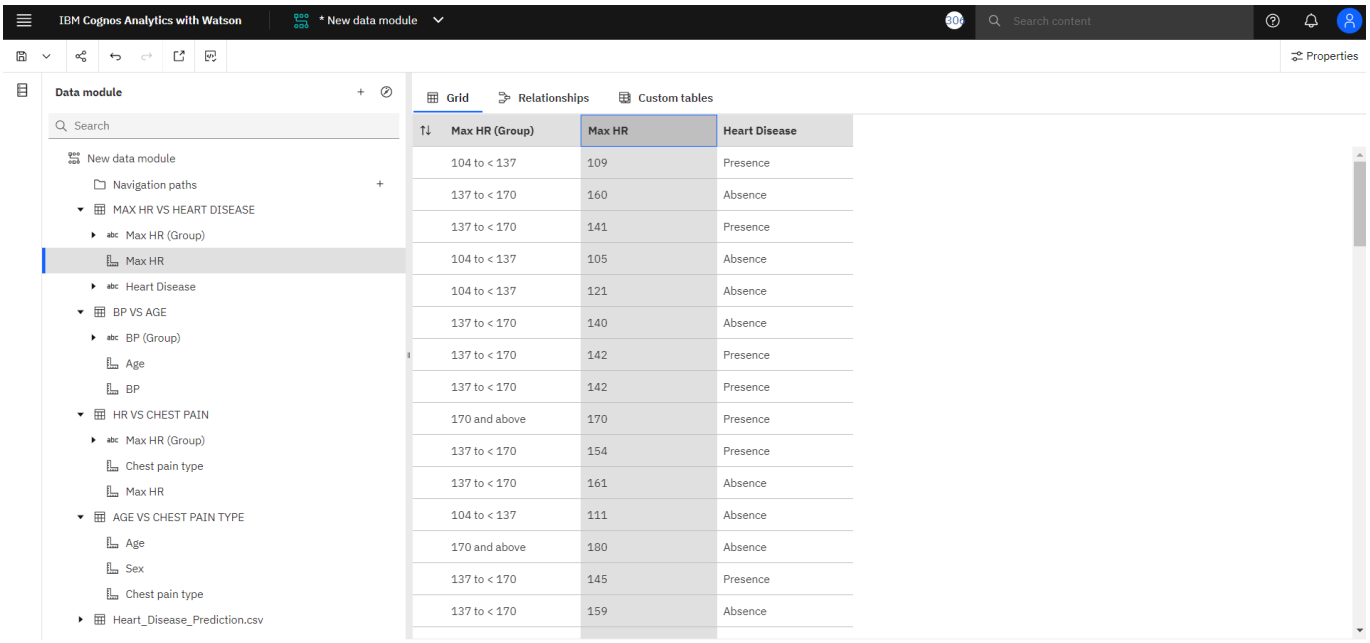
Max HR (Group)	Chest pain type	Max HR
104 to < 137	4	109
137 to < 170	3	160
137 to < 170	2	141
104 to < 137	4	105
104 to < 137	2	121
137 to < 170	4	140
137 to < 170	3	142
137 to < 170	4	142
170 and above	4	170
137 to < 170	4	154
137 to < 170	4	161
104 to < 137	4	111
170 and above	3	180
137 to < 170	1	145
137 to < 170	4	159

TABLE FOR BP VS AGE



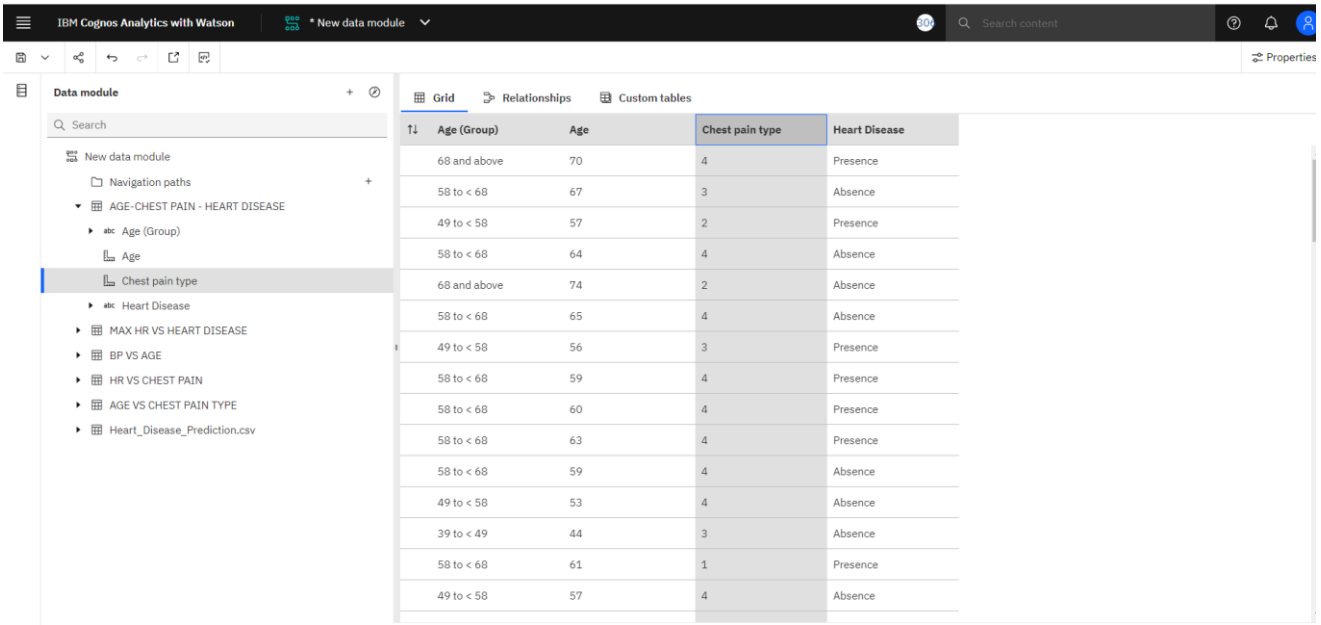
BP (Group)	Age	BP
121 to < 147	70	130
less than 121	67	115
121 to < 147	57	124
121 to < 147	64	128
less than 121	74	120
less than 121	65	120
121 to < 147	56	130
less than 121	59	110
121 to < 147	60	140
147 to < 174	63	150
121 to < 147	59	135
121 to < 147	53	142
121 to < 147	44	140
121 to < 147	61	134
121 to < 147	57	128

TABLE FOR MAX HR VS EXISTING HEART DISEASE



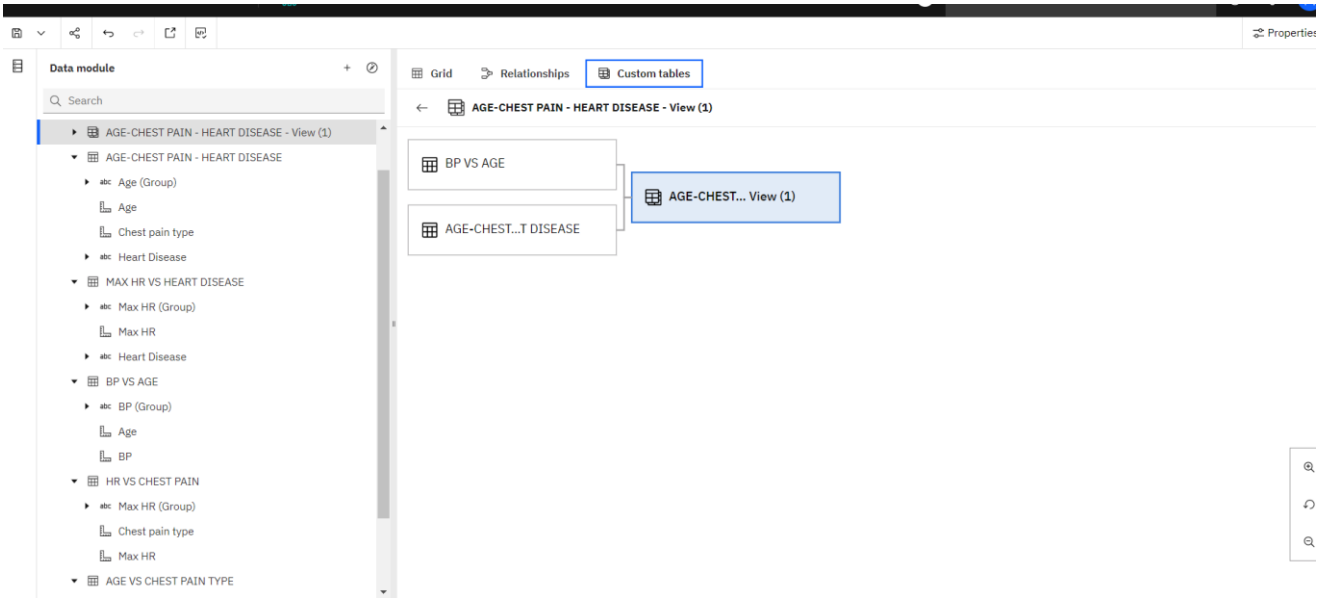
Max HR (Group)	Max HR	Heart Disease
104 to < 137	109	Presence
137 to < 170	160	Absence
137 to < 170	141	Presence
104 to < 137	105	Absence
104 to < 137	121	Absence
137 to < 170	140	Absence
137 to < 170	142	Presence
137 to < 170	142	Presence
170 and above	170	Presence
137 to < 170	154	Presence
137 to < 170	161	Absence
104 to < 137	111	Absence
170 and above	180	Absence
137 to < 170	145	Presence
137 to < 170	159	Absence

TABLE FOR AGE – CHEST PAIN VS EXISTING HEART DISEASE

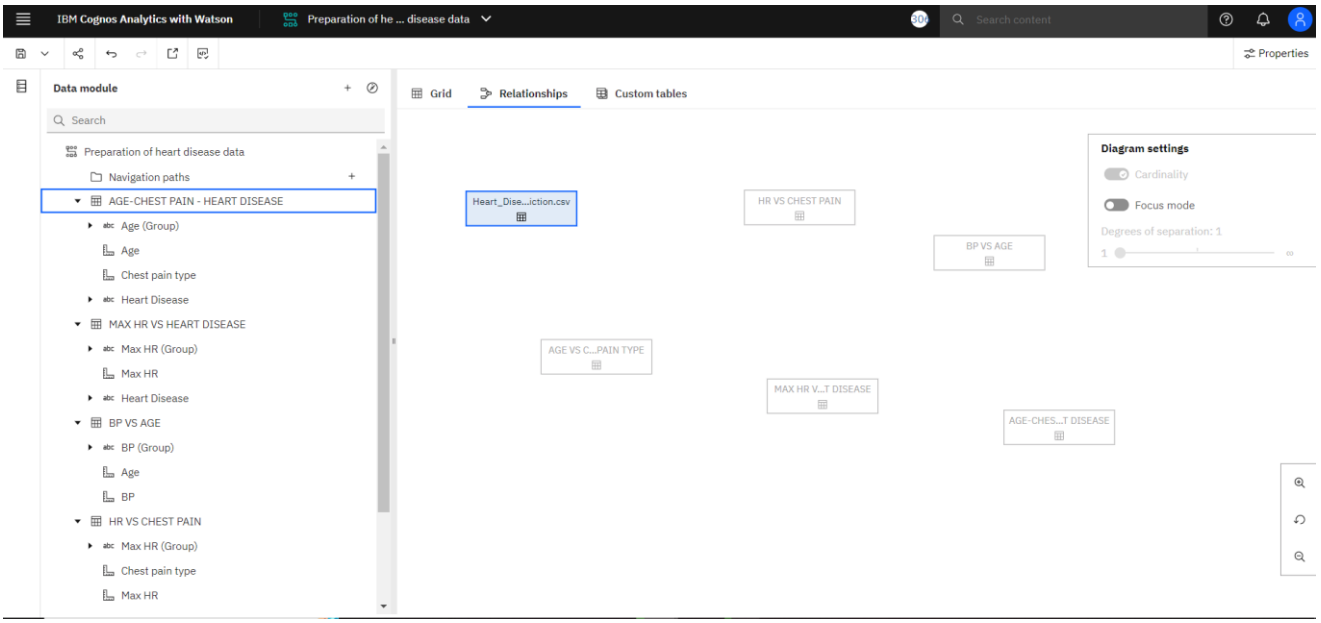


Age (Group)	Age	Chest pain type	Heart Disease
68 and above	70	4	Presence
58 to < 68	67	3	Absence
49 to < 58	57	2	Presence
58 to < 68	64	4	Absence
68 and above	74	2	Absence
58 to < 68	65	4	Absence
49 to < 58	56	3	Presence
58 to < 68	59	4	Presence
58 to < 68	60	4	Presence
58 to < 68	63	4	Presence
58 to < 68	59	4	Absence
49 to < 58	53	4	Absence
39 to < 49	44	3	Absence
58 to < 68	61	1	Presence
49 to < 58	57	4	Absence

COMBINATION OF TABLES



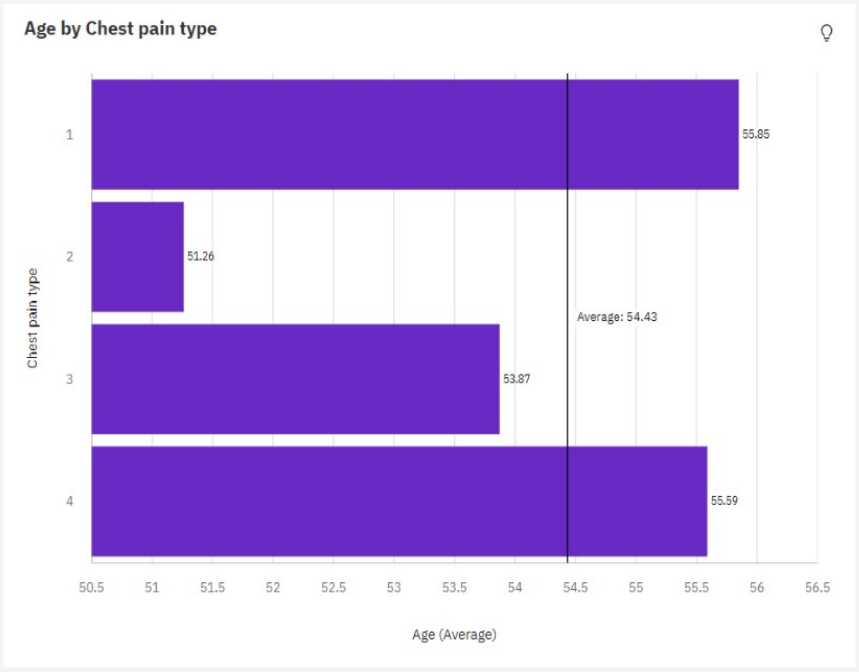
OVERALL DATASET



Now with the above prepared new tables and explored processed datasets we do the data exploration to find the required relationship in form for graphs

EXPLORATION:

Exploration of AGE BY CHEST PAIN TYPE



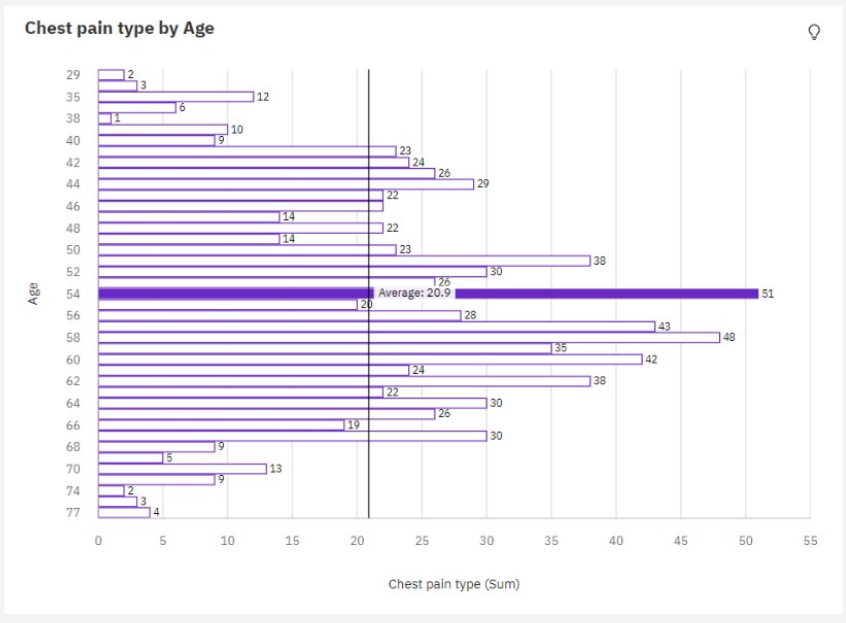
Details

Over all **chest pain types**, the average of **Age** is 54.43.

The average values of **Age** range from 51.26, occurring when **Chest pain type** is 2, to 55.85, when **Chest pain type** is 1.

The most common value of **Chest pain type** is 4, occurring 129 times, which is 47.8 % of the total.

Exploration of CHEST PAIN BY AGE



Details

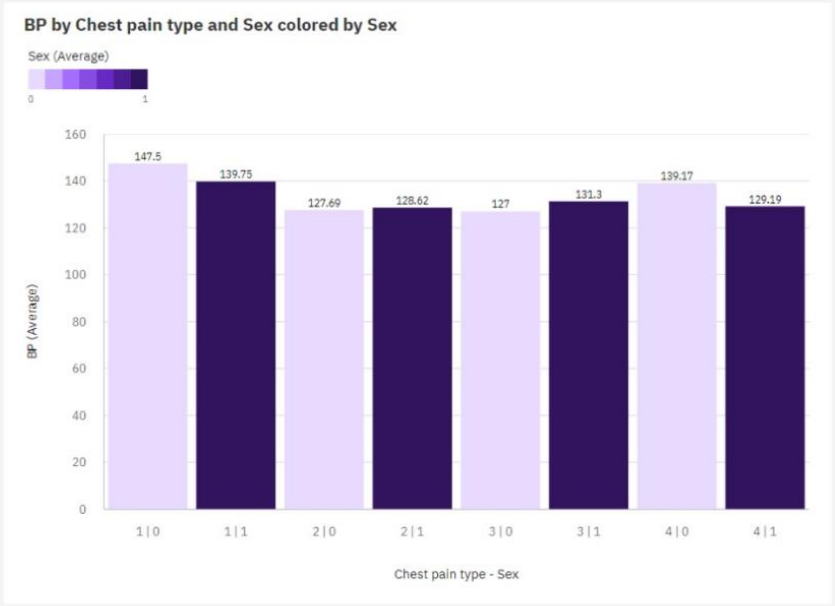
Over all **ages**, the sum of **Chest pain type** is 857.

For **Chest pain type**, the most significant values of **Age** are 54 and 58, whose respective **Chest pain type** values add up to 99, or 11.6 % of the total.

Chest pain type ranges from 1, when **Age** is 38, to 51, when **Age** is 54.

Chest pain type is unusually high when **Age** is 54.

Exploration of BP VS CHEST PAIN TYPE AND GENDER



Details

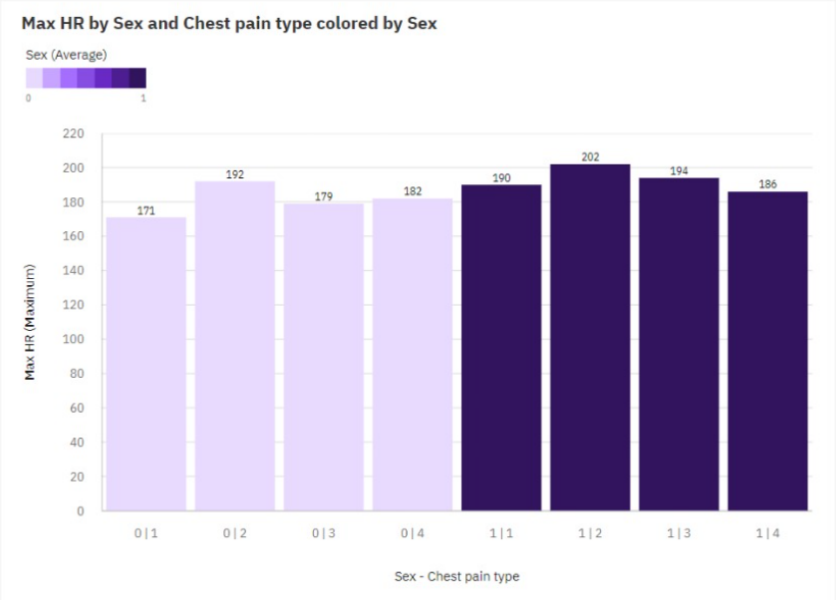
The most common value of **Chest pain type - Sex** is 2|1, occurring 129 times, which is 47.8 % of the total.

The total number of results for **Sex** is 270.

Over all **chest pain type - sexes**, the average of **BP** is 0.6778.

The average values of **BP** range from 0, occurring when **Chest pain type - Sex** is 1|0, to 1, when **Chest pain type - Sex** is 1|0.

Exploration of HEART RATE during CHEST PAIN



Details

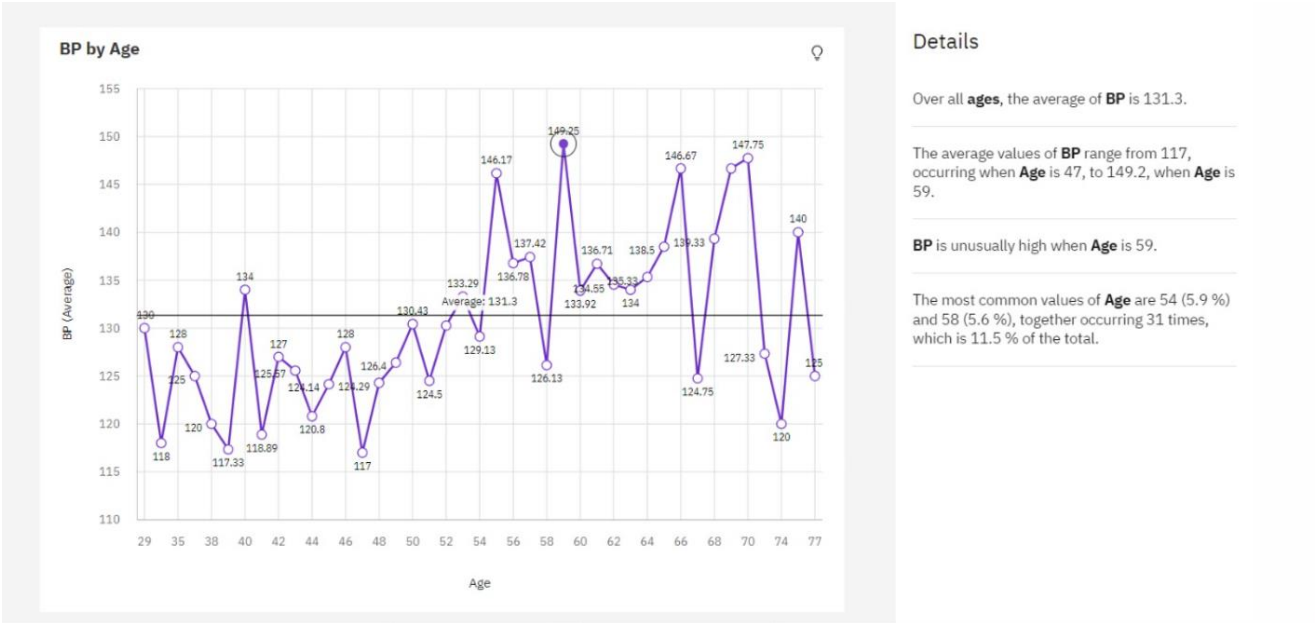
The total number of results for **Max HR** is 270.

The total number of results for **Sex** is 270.

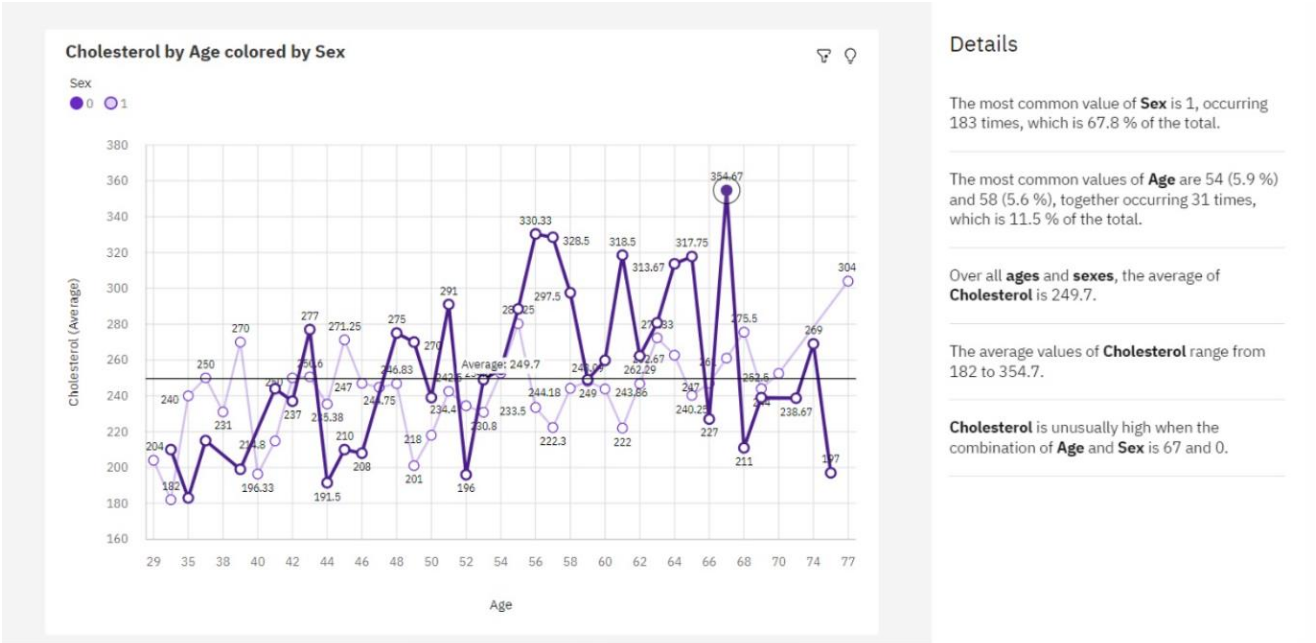
The most common value of **Sex - Chest pain type** is 0|2, occurring 183 times, which is 67.8 % of the total.

The largest value of **Max HR** is 3, occurring when **Sex - Chest pain type** is 0|1.

Exploration of BP by AGE



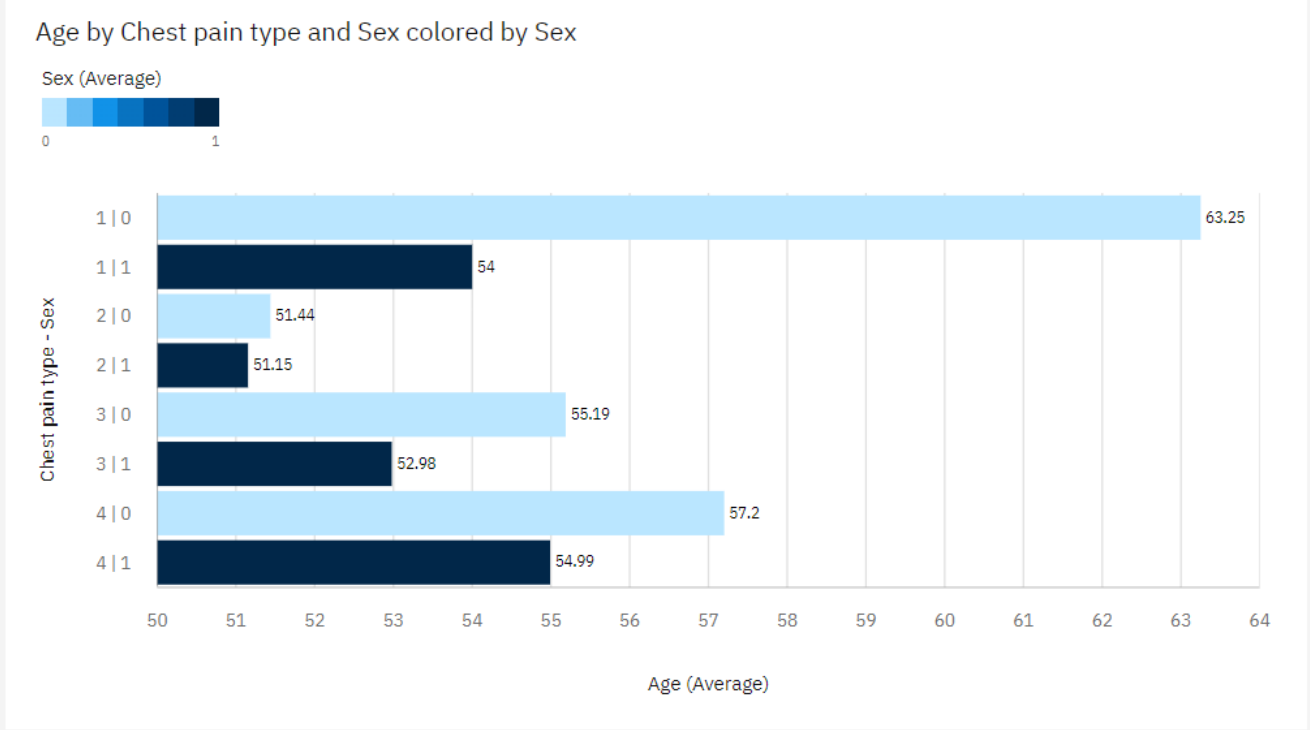
Exploration of CHOLESTEROL by AGE and GENDER



Now after exploration we know the required relationship between the dependent and independent variables, hence we try to visualize the relations in other forms.

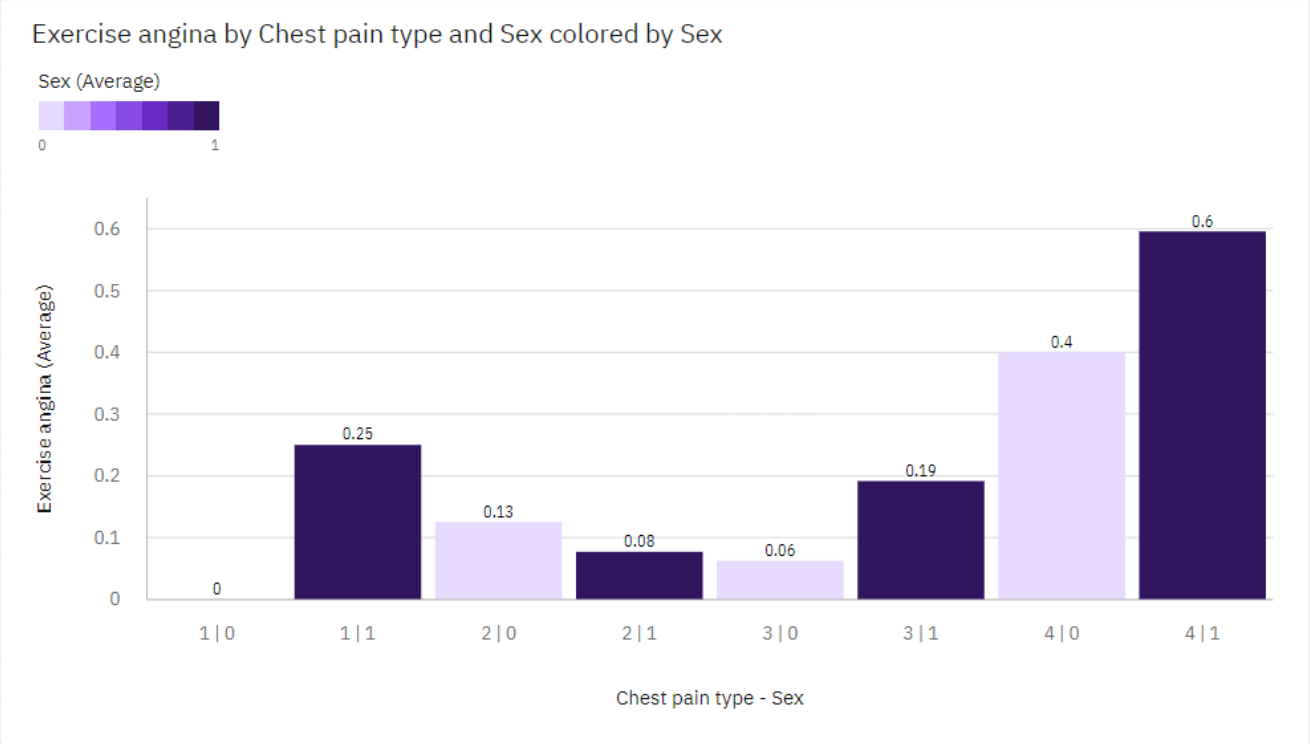
DATA VISUALIZATON

Average Age For Different Chest Pain Types:



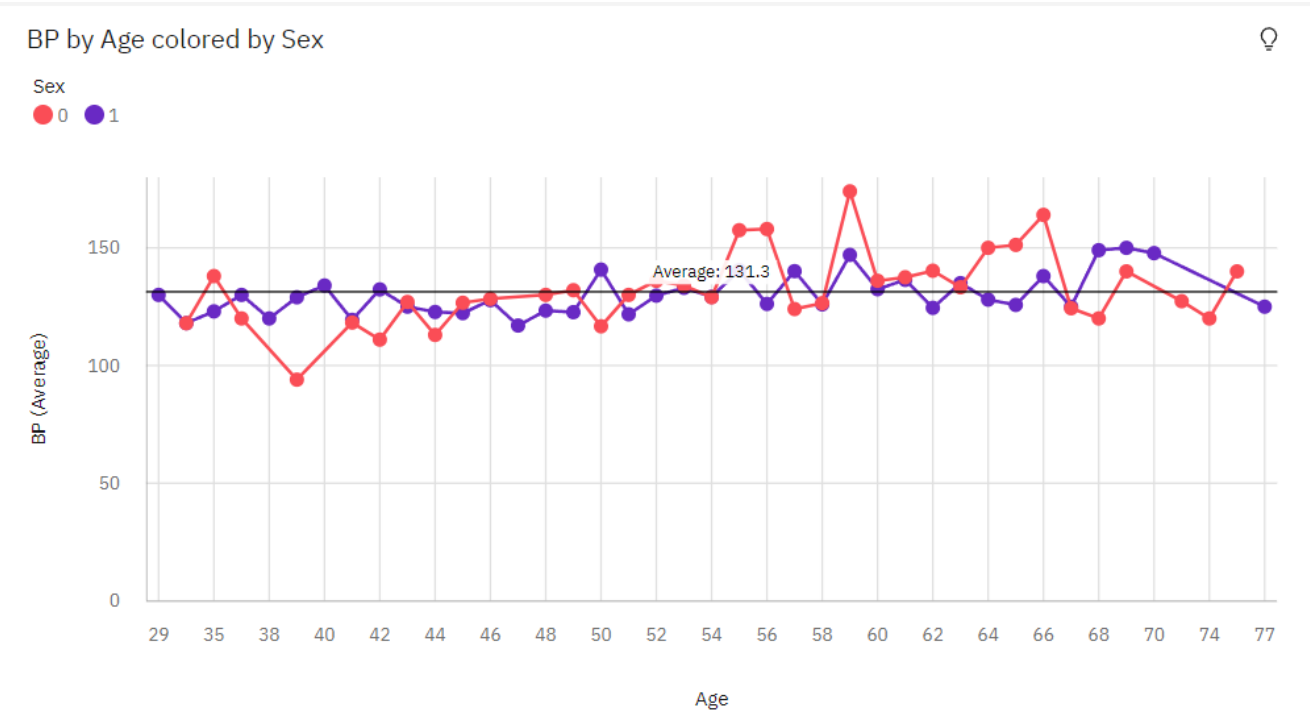
Here, a horizontal bar graph is developed between Age (Average Values) and Chest Pain, where Chest Pain is differentiated by Gender.

Average Exercise Angina During Chest Pain:



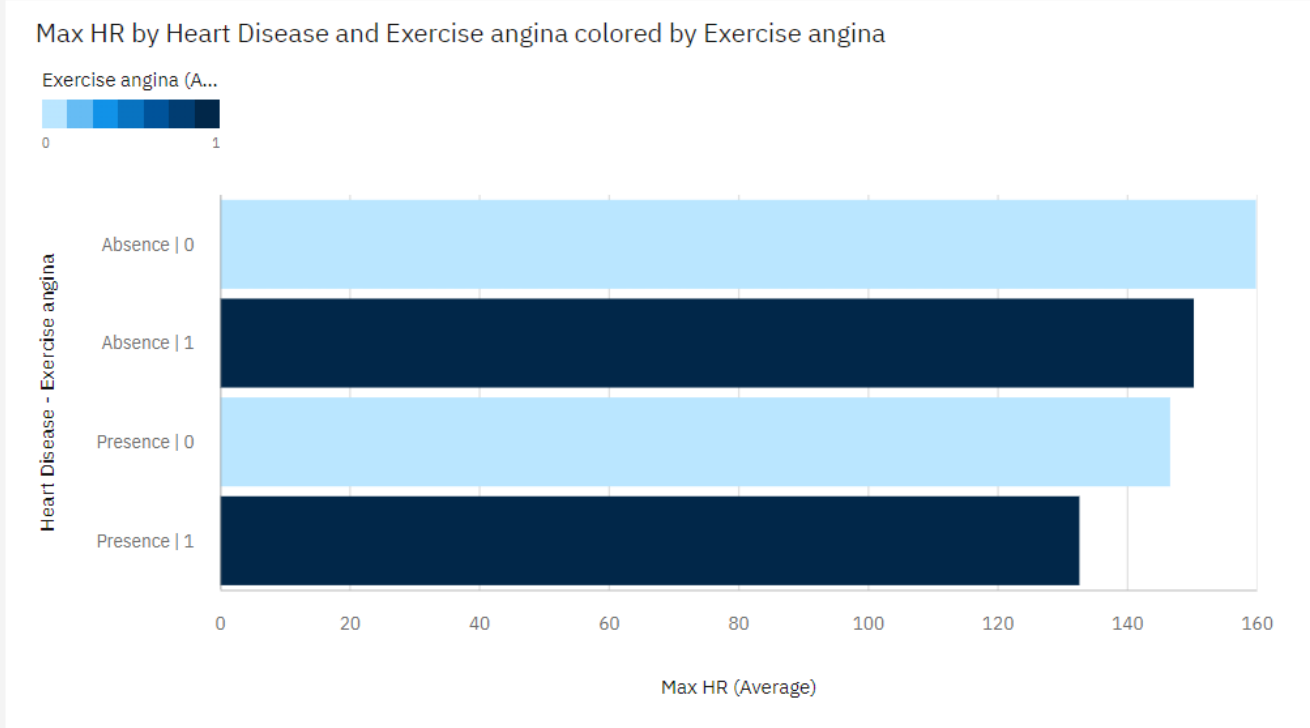
Here, a vertical bar graph is developed between Chest Pain, where Chest Pain is differentiated by Gender and Exercise Angima (Average).

BP Variation With Respect To Age:



Here, a line graph is developed between Age and BP (Average), where BP is differentiated by Gender.

Effect Of Existing Heart Disease On Average Of Exercise Angina:



Here, a horizontal bar graph is developed between Maximum Heart Rate (Average) and Heart Disease, where Heart Disease is differentiated by Exercise Angima.

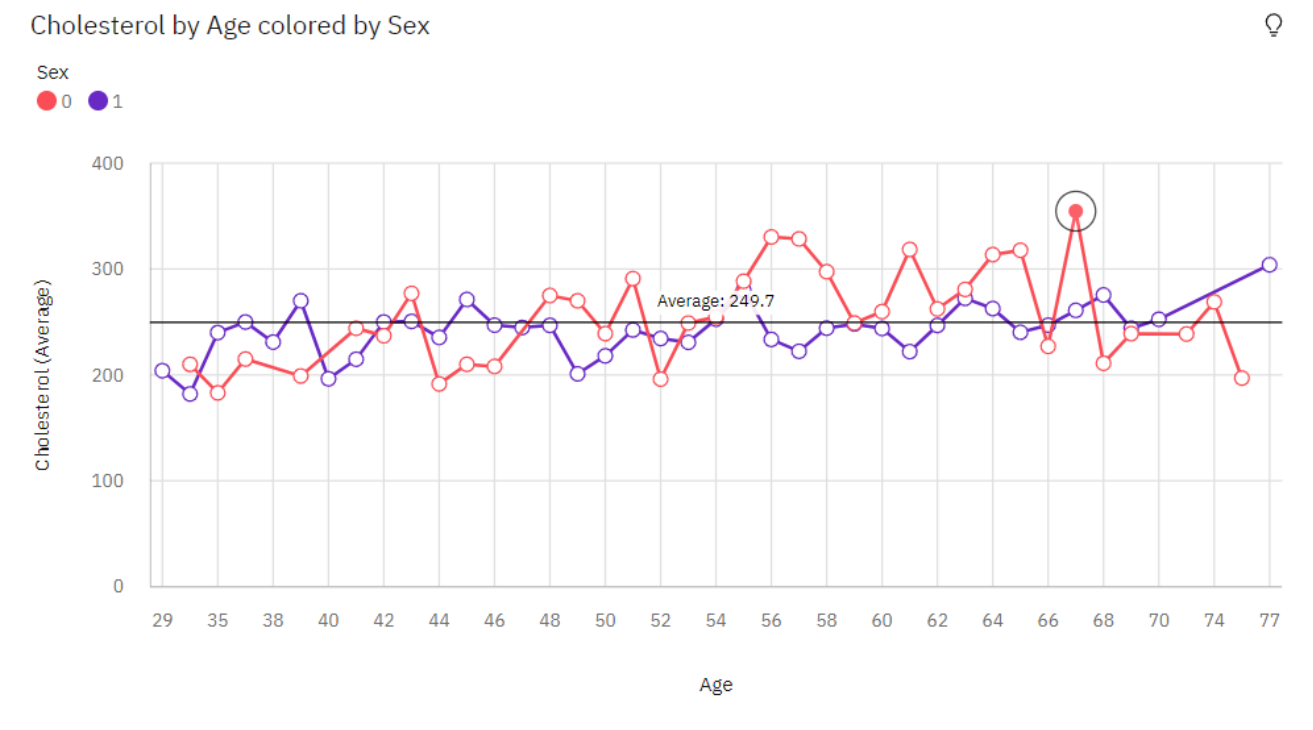
Average Age For Different Types Of Chest Pain in Existing Heart Diseases:

Heart Disease for Chest pain type and Sex

Heart Disease	1	2	3	4	Summary
0	4	16	32	35	**87**
1	16	26	47	94	**183**
Summary	**20**	**42**	**79**	**129**	**270**

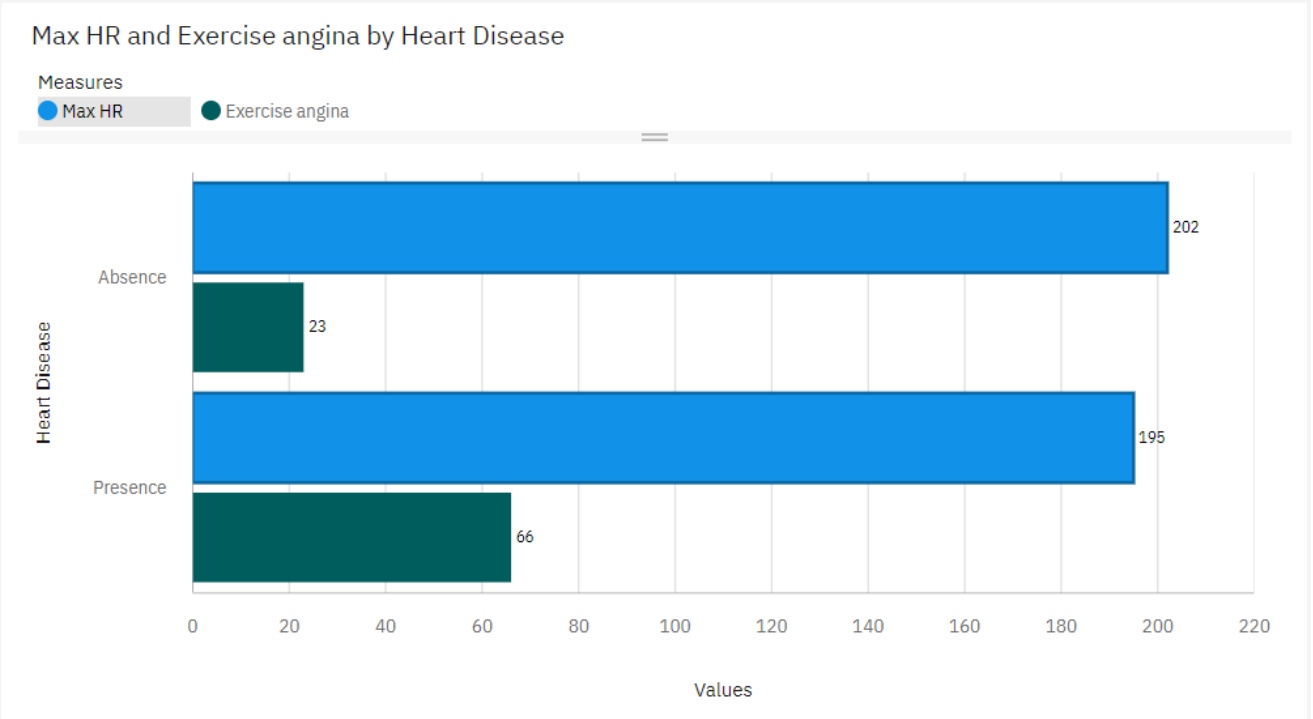
Here, a table is developed between Heart Disease and Gender.

Serum Cholesterol Levels Vs Age:



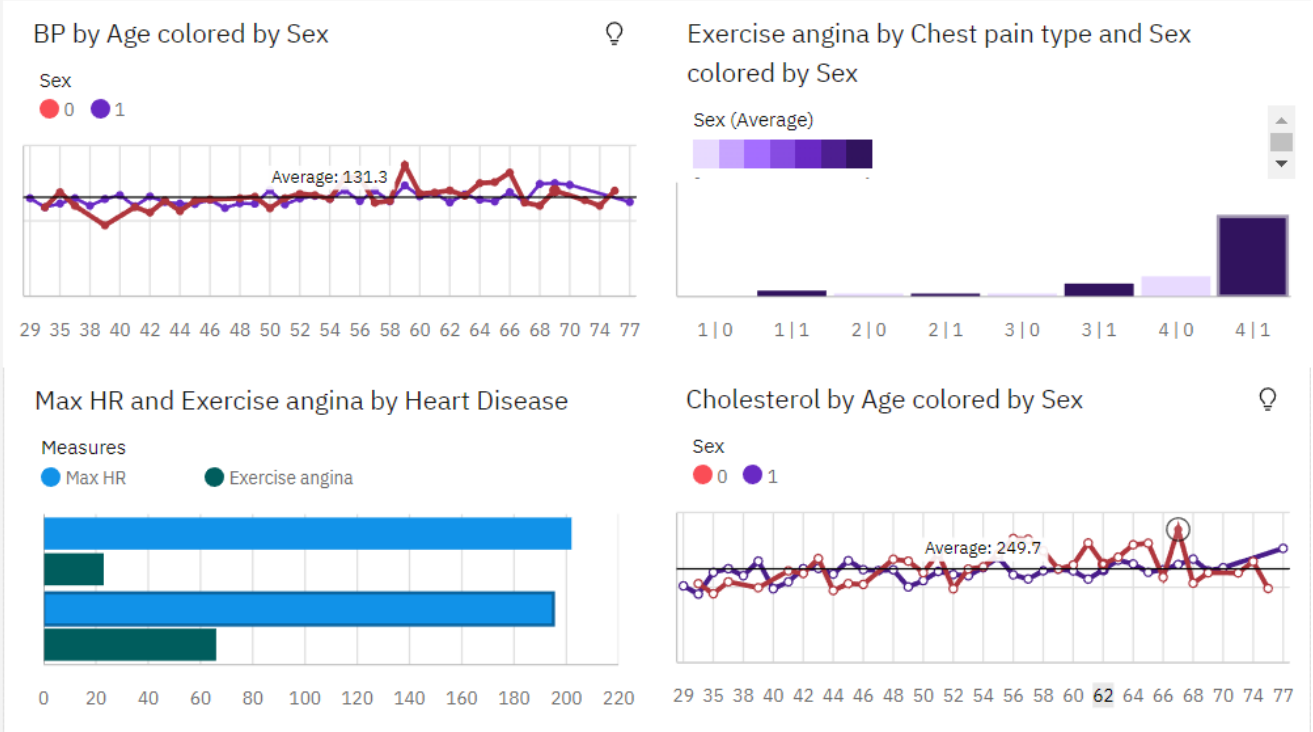
Here, a line graph is developed between Age and Cholestrol (Average), where it is differentiated by Gender.

Maxium Heart Rate In Existing Heart Disease By Exercise Angina:



Here, a horizontal bar graph is developed with Maximum Heart Rate (Average) and Heart Disease.

Dashboard Showing Different Types Of Visuals:



Here, different visualizations are combined into a single dashboard.