Working with Dataset

Date	11/11/2022
Team ID	PNT2022TMID28411
Project Name	Analytics for Hospitals Health-Care Data

Understanding the Dataset:

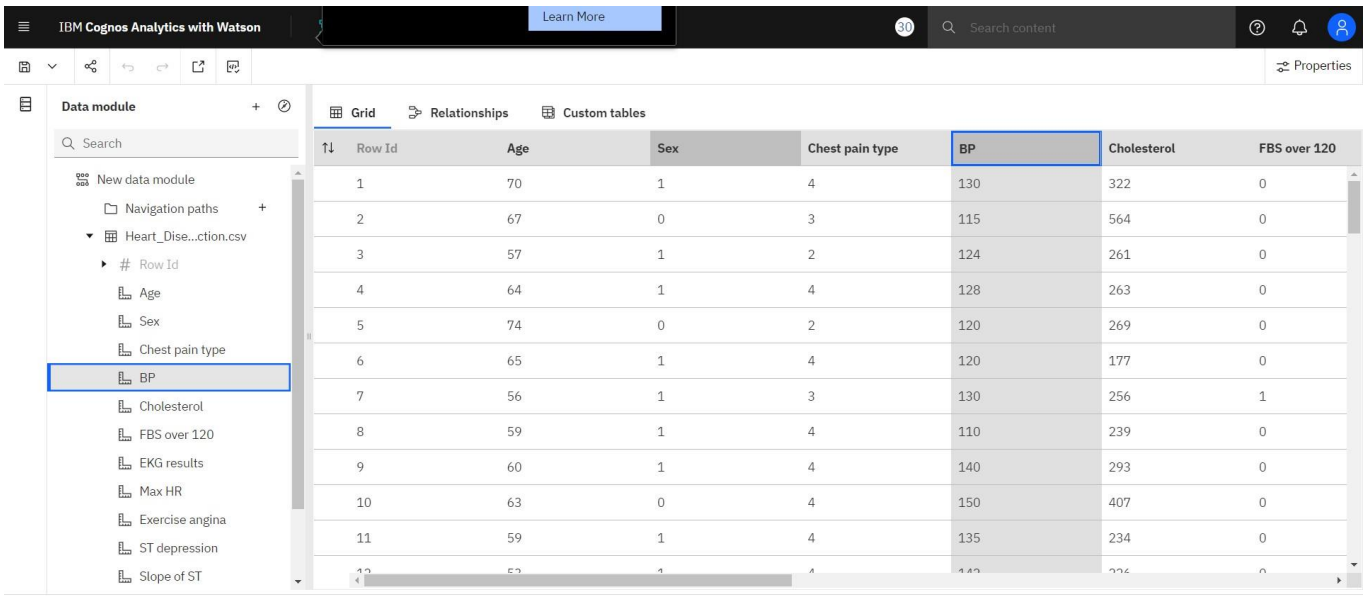
This database contains of 14 fields. The "goal" field refers to the presence of heart disease in the patient. It is integer valued from 0 (no presence) to 4.

The data can be downloaded from the following: https://www.kaggle.com/datasets/rishidamarla/heart-disease-prediction The data dictionary is as follows:

Sno	Field Name		
1	Age	13	Thallium
2	Sex	14	heart disease
3	Chest pain type		
4	BP		
5	Cholesterol		
6	FBS over 120		
7	EKG results		
8	Max HR		
9	Exercise angina		
10	ST depression		
11	Slope of ST		

Loading the Dataset:

This dataset is loaded with all the components that is being listed and these details are stored already in the dataset.



Row Id	Age	Sex	Chest pain type	BP	Cholesterol	FBS over 120
1	70	1	4	130	322	0
2	67	0	3	115	564	0
3	57	1	2	124	261	0
4	64	1	4	128	263	0
5	74	0	2	120	269	0
6	65	1	4	120	177	0
7	56	1	3	130	256	1
8	59	1	4	110	239	0
9	60	1	4	140	293	0
10	63	0	4	150	407	0
11	59	1	4	135	234	0

Exploration of Data:

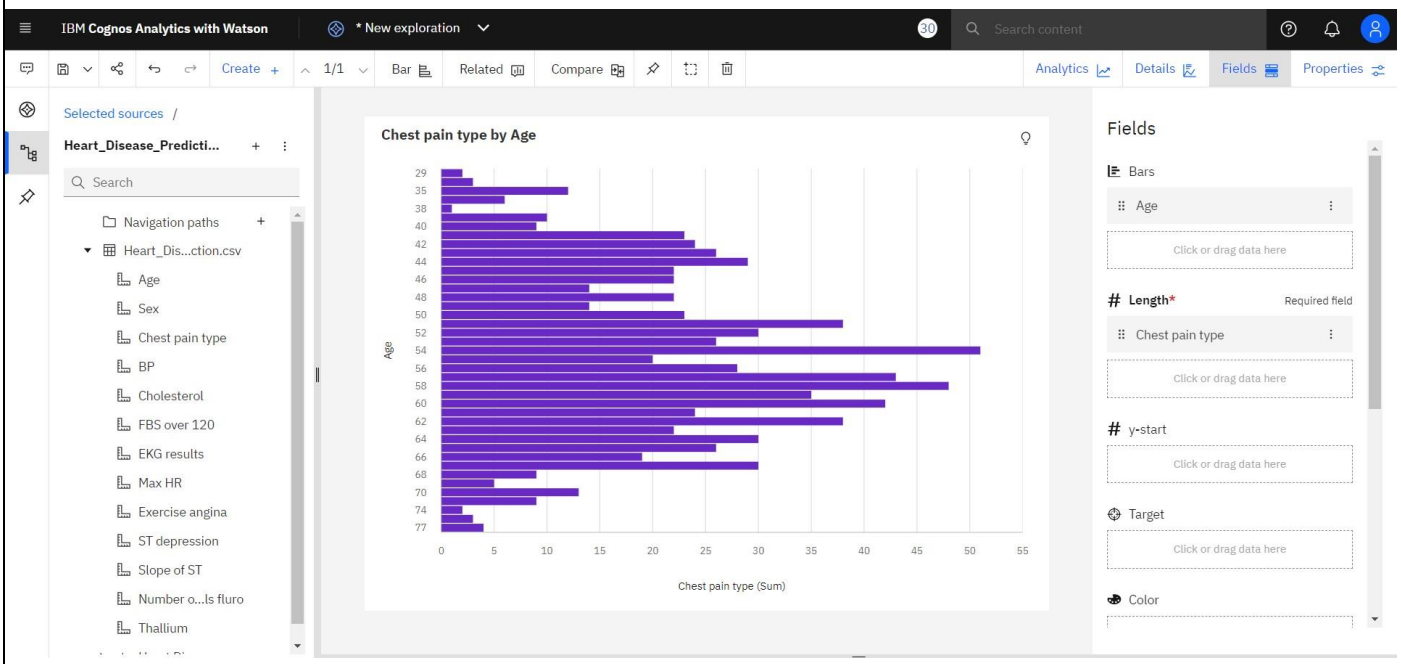
Data exploration is the first step of data analysis used to explore and visualize data to uncover insights from the start or identify areas or patterns to dig into more.

Exploration of Data relationships among the values is presented along with plotting of Average Age for different Chest Pain Types. For visualizing it, we will require the following data:

* Sex

* Age

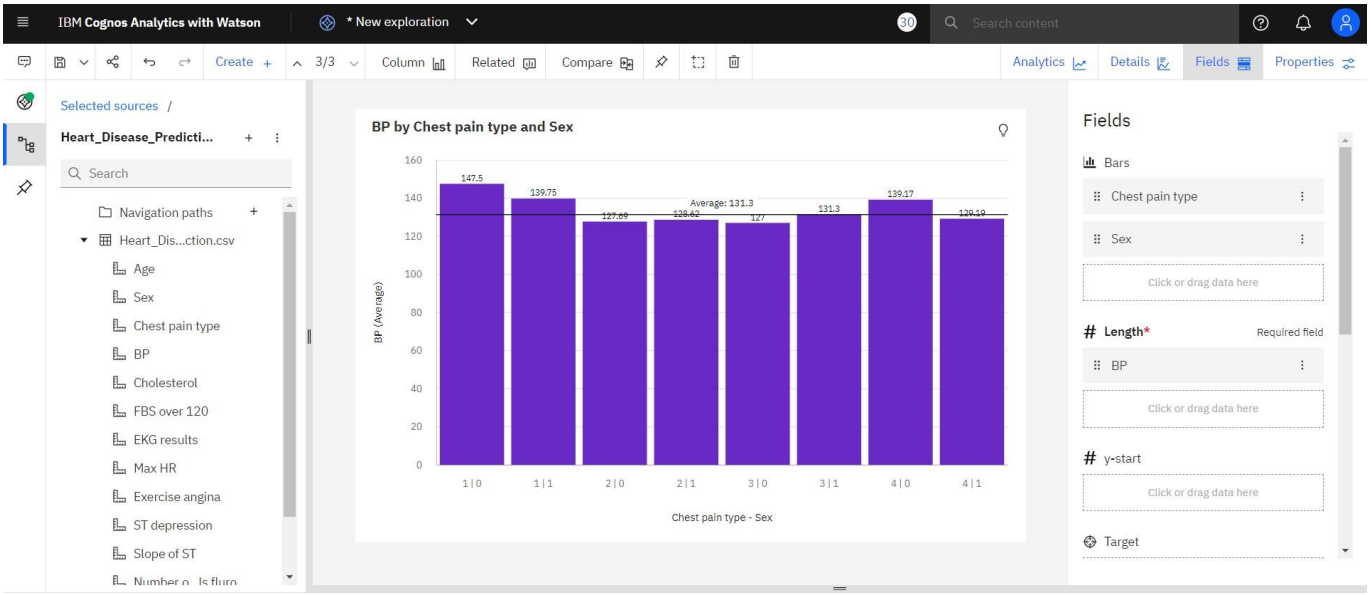
* Chest Pain Type.



Exploration of BPvsChestPainType and Gender:

Data Exploration done in the particular step is done to identify the pain type with respect to the gender and identify the correct remedy for it.

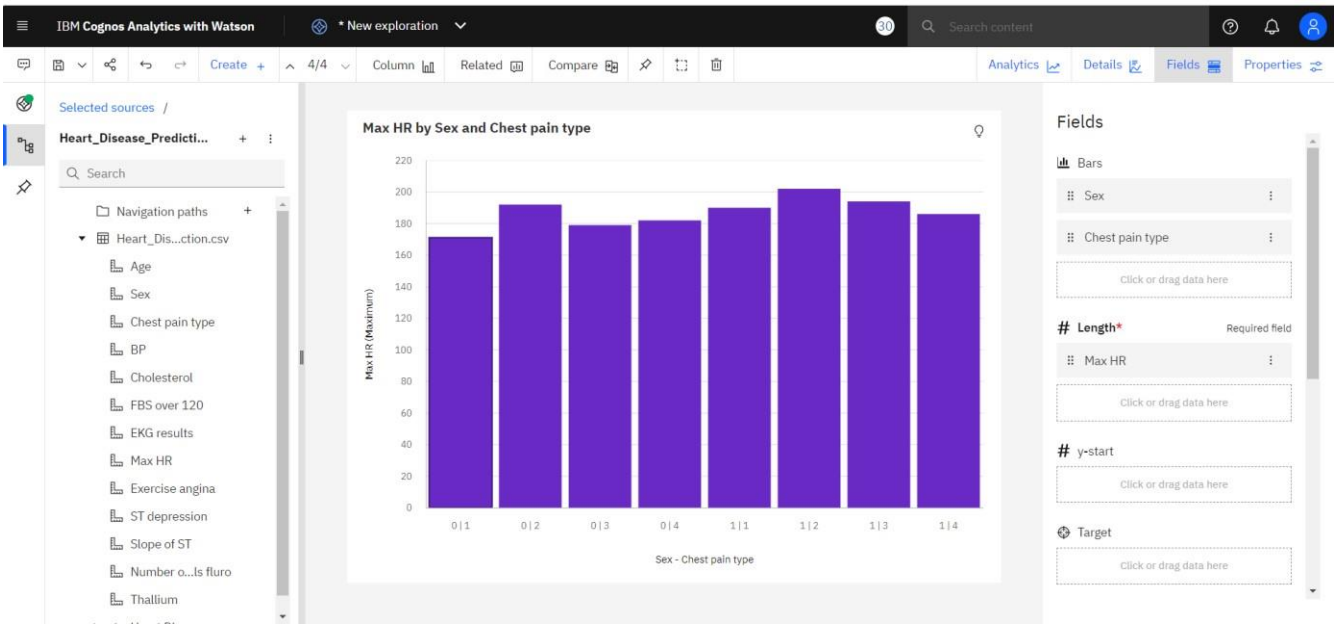
We are going to plot average BP recorded for Male and Female based on Gender during the Chest pain recorded.



Exploration of Max Heartrate During the Chest Pain:

Data exploration in the particular process is getting the details of the heartrate and its highest values during a chest pain.

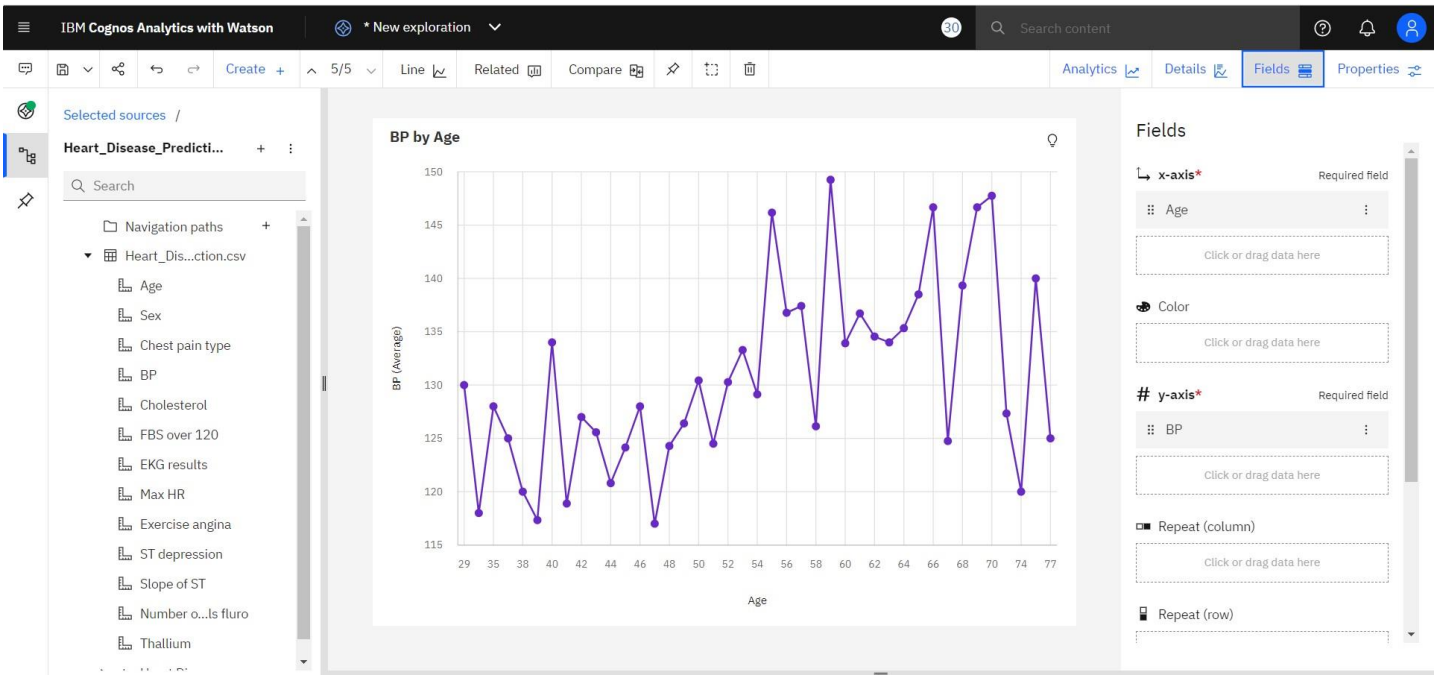
Here we are plotting the average Max Heartbeats recorded for a person based on Gender and Chest Pain Type.



Exploration of BP By Age:

Data exploration in the particular process is getting the details of the Blood Pressure of the user with respect to their age.

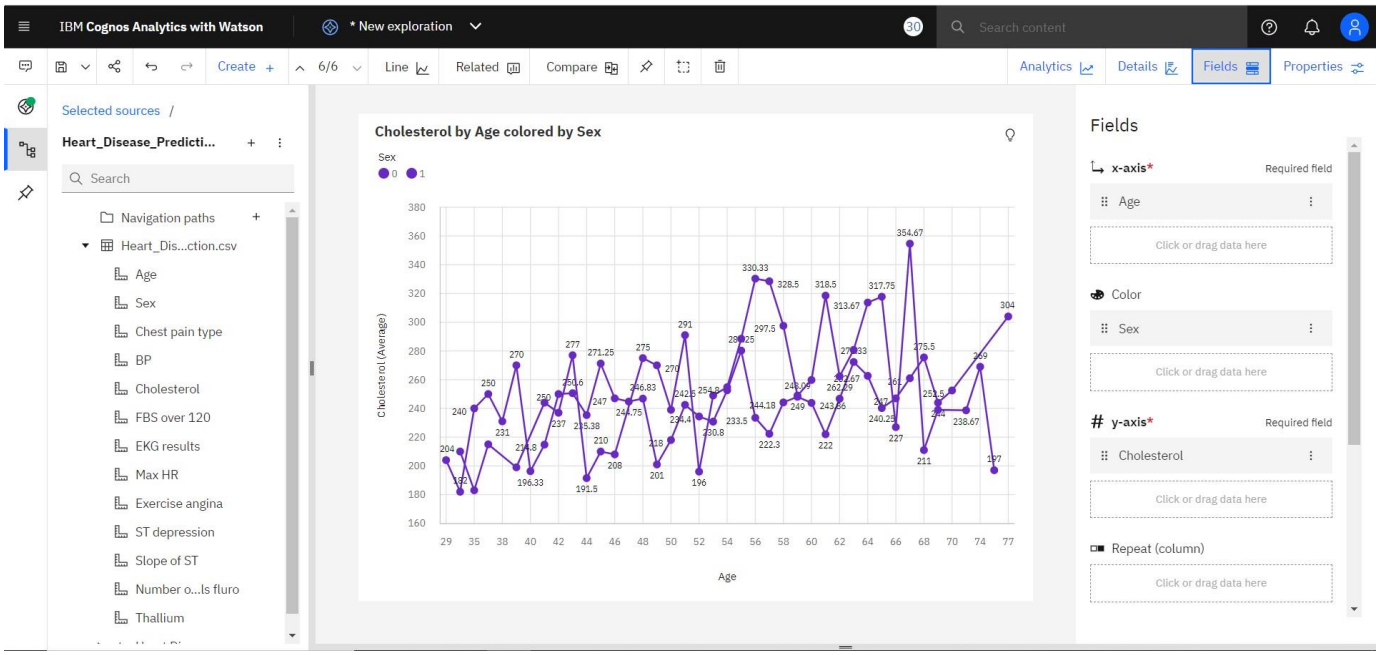
Here we need to consider the age as a dimension, because we want to plot the BP values against it. Changing the age from a measure to dimension will allow us to plot all values from BP present in the dataset corresponding to their age.



Exploration of Cholesterol by Age and Gender:

Data exploration in the particular process is getting the details of the Cholesterol of the user with respect to their age and the respective Gender of the user.

Here we will be exploring the Serum Cholesterol of people recorded with respect to to their age.


IBM Cognos Analytics with Watson
* New exploration
Search content
Create
6/6
Line
Related
Compare
Analytics
Details
Fields
Properties
Selected sources /
Heart_Disease_Predicti...
Search
Navigation paths
Heart_Dis...ction.csv
Age
Sex
Chest pain type
BP
Cholesterol
FBS over 120
EKG results
Max HR
Exercise angina
ST depression
Slope of ST
Number o...ls fluro
Thallium
Cholesterol by Age colored by Sex
Sex
0
1
Cholesterol (Average)
Age
Fields
x-axis*
Required field
Age
Click or drag data here
Color
Sex
Click or drag data here
y-axis*
Required field
Cholesterol
Click or drag data here
Repeat (column)
Click or drag data here