



Government of India
Ministry Of Skill Development & Entrepreneurship
Directorate General of Training
NATIONAL SKILL TRAINING INSTITUTE TRIVANDRUM
Opp.Block Office, Kazhakuttam, Trivandrum - 695102

Project Documentation

Topic: - Heart Attack Analysis & Prediction

Module: - 5 (Data Analytics)

Trade: -Advanced Diploma in IT Networking & Cloud

Submitted By

ROLL NO	NAME
ADIT/TVM/19/007	ATHIRA P V
ADIT/TVM/19/015	SOORYA S DAS

IBM Faculty

TO / ADIT

AD / Section In charge

ABSTRACT

1. INTRODUCTION

1.1. OBJECTIVE

1.2. PROJECT DESCRIPTION

1.3. SCOPE OF WORK

1.4 FEASIBILITY OF STUDY

2. HARADWARE & SOFTWARE REQUIREMENTS

2.1. HARDWARE REQUIUREMENT

2.2. SOFTWARE REQUIREMENTS

3. BACKGROUND

3.1. BACKGROUND OF THE STUDY

3.2. PROBLEM STATEMENT

4. RESEARCH METHODOLOGY

4.1. RESEARCH DESIGN

4.2. SYSTEM DEVELOPMENT METHODOLOGY

5. DATA ANALYSIS

6. CONCLUSION

7. REFERNCES

ABSTRACT

We are creating a Data Analysis Project on Heart attack analysis & Prediction. This project is an attempt of data analyzing Heart attack analysis & Prediction with the help of data science and data analytics in excel. Heart attacks is one of the biggest causes of morbidity and mortality among the population of the world. Prediction of cardiovascular disease is regarded as one of the most important subjects in the section of clinical data analysis. The amount of data in the healthcare industry is huge. Data mining turns the large collection of raw healthcare data into information that can help to make informed decisions and predictions.

1. INTRODUCTION

Worldwide, heart attack is the primary cause of death, and the possible detection at an earlier stage avoids the attacks. The process of making a decision based on previous information is termed as a biological prediction. Medical specialists produce data with prosperity of hidden information, and it is not accurately being utilized efficiently for predictions. For this reason, the analysis changes the unused data into a dataset for modelling by various data mining (DM) approaches. DM is the procedure of valuable mining from a huge number of databases. According to World Health Organization every year, due to heart problems, more than 12 million deaths happen worldwide. For heart attacks, the factors of pulse rate, cholesterol, and blood pressure are the main reason—some non-modifiable factors such as drinking, smoking also reason for heart attacks.

1.1. OBJECTIVE

The main objective of this research is to develop a heart prediction system. The system can discover and extract hidden knowledge associated with diseases from a historical heart data set. Heart attack analysis & prediction system aims to exploit data mining techniques on medical data set to assist in the prediction of the heart attack analysis & prediction system.

1.2. PROJECT DESCRIPTION

Data analysis proves to be crucial in the medical field. It provides a meaningful base to critical decisions. It helps to create a complete study proposal. One of the most important uses of data analysis is that it helps in keeping human bias away from medical conclusion with the help of proper statistical treatment. By use of data mining for exploratory analysis because of nontrivial information in large volumes of data. The health care industries collect huge amounts of data that contain some hidden information, which is useful for making effective decisions for providing appropriate results and making 2 effective decisions on data, some data mining techniques are used to better the experience and conclusion that have been give.

1.3. SCOPE OF WORK

Here the scope of the project is that integration of clinical decision support with computer-based patient records could reduce medical errors, enhance patient safety, decrease unwanted practice variation, and improve patient outcome. This suggestion is promising as data modeling and analysis tools, e.g., data mining, have the potential to generate a knowledge-rich environment which can help to significantly improve the quality of clinical decisions

1.4. FEASIBILITY OF STUDY

Heart attack analysis and predictor system will use the data mining knowledge to give a user-oriented approach to new and hidden patterns in the data. The knowledge which is implemented can be used by the healthcare experts to get better quality of service and to reduce the extent of adverse medicine effect.

2. HARADWARE & SOFTWARE REQUIREMENTS

2.1. HARADWARE REQUIREMENTS

- Personal Computer or Laptop
- Processor – Intel Core i3
- Hard Disk Capacity – 1 Tb

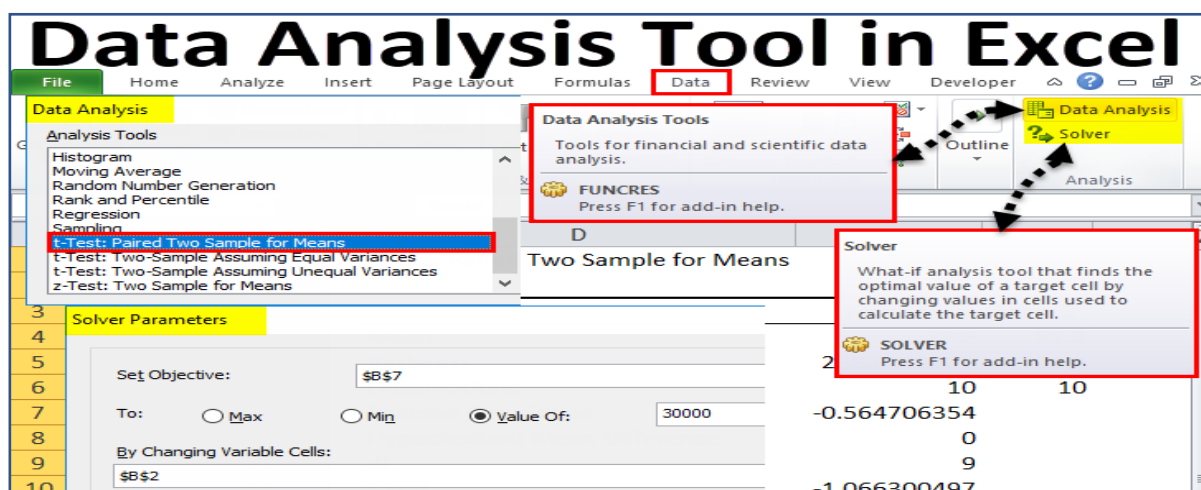
2.2. SOFTWARE REQUIREMENTS

- Operating System – Windows OS (10th Gen)
- Data Analysis Tool – MS Excel (Charts, Graph)

- MS Excel

Microsoft Excel is one of the most widely used tools in any industry. While some enjoy playing with pivotal tables and histograms, others limit themselves to simple pie-charts and conditional formatting.

Some may create an artwork out of the dull monochrome Excel, while others may be satisfied with its data analysis. In this discussion, we will make a deep delving analysis of Microsoft Excel and its utility. We will focus on how to analyse data in Excel Analytics, the various tricks, and techniques for it. The discussion will also explore the various ways to analyse data in Excel.



3. BACKGROUND

Among all fatal disease, heart attacks diseases are considered as the most prevalent. Medical practitioners conduct different surveys on heart attacks and gather information of heart patients, their symptoms and disease progression. Increasingly are reported about patients with common diseases who have typical symptoms. In this fast moving world people want to live a very luxurious life so they work like a machine in order to earn lot of money and live a comfortable life therefore in this race they forget to take care of themselves, because of this there food habits change their entire lifestyle change, in this type of lifestyle they are more tensed they have blood pressure, sugar at a very young age and they don't give enough rest for themselves and eat what they get and they even don't bother about the quality of the food if sick the go for their own medication as a result of all these small negligence it leads to a major threat that is the heart attacks

3.1. BACKGROUND OF THE STUDY

Data analysis proves to be crucial in the medical field. It provides a meaningful base to critical decisions. It helps to create a complete study proposal. One of the most important uses of data analysis is that it helps in keeping human bias away from medical conclusion with the help of proper statistical treatment. By use of data mining for exploratory analysis because of nontrivial information in large volumes of data. The health care industries collect huge amounts of data that contain some hidden information, which is useful for making effective decisions for providing appropriate results and making 2 effective decisions on data, some data mining techniques are used to better the experience and conclusion that have been given. Heart predictor system will use the data mining knowledge to give a user-oriented approach to new and hidden patterns in the data. The knowledge which is implemented can be used by the healthcare experts to get better quality of service and to reduce the extent of adverse medicine effect.

3.2. PROBLEM STATEMENT

Heart attacks can be managed effectively with a combination of lifestyle changes, medicine and, in some cases, surgery. With the right treatment, the symptoms of heart attacks can be reduced and the functioning of the heart improved. The predicted results can be used to prevent and thus reduce cost for surgical treatment and other expensive. The overall objective of

my work will be to predict accurately with few tests and attributes the presence of heart attacks. Attributes considered form the primary basis for tests and give accurate results more or less. Many more input attributes can be taken but our goal is to predict with few attributes and faster efficiency the risk of having heart attacks. Decisions are often made based on doctors' intuition and experience rather than on the knowledge rich data hidden in the data set and databases. This practice leads to unwanted biases, errors and excessive medical costs which affects the quality of service provided to patients. Data mining holds great potential for the healthcare industry to enable health systems to systematically use data and analytics to identify inefficiencies and best practices that improve care and reduce costs. According to (Wurz & Takala, 2006) the opportunities to improve care and reduce costs concurrently could apply to as much as 30% of overall healthcare spending. The successful application of data mining in highly visible fields like e-business, marketing and retail has led to its application in other industries and sectors. Among these sectors just discovering is healthcare. The healthcare environment is still „information rich“ but „knowledge poor“. There is a wealth of data available within the healthcare systems.

4. RESEARCH METHODOLOGY

4.1. RESEARCH DESIGN

I will be using the experimental type of research design. It is a quantitative research method. Basically, it is a research conducted with a scientific approach, where a set of variables are kept constant while other set of variables are being measured as the subject of the experiment. This is more practically while conducting face recognition and detection as it monitors the behaviors and patterns of a subject to be used to acknowledge whether the subject matches all details presented and cross checked with previous data. It is an effect research method as it is time bound and focuses on the relationship between the variables that give actual results.

4.2. SYSTEM DEVELOPMENT METHODOLOGY

The methodology of software development is the method in managing project development. There are many models of the methodology are available such as Waterfall model Incremental model, RAD model, Agile model, Iterative model and Spiral model. However, it still needs to be considered by developer to decide which is will be used in the project. The methodology model is useful to manage the project efficiently and able to help developer from getting any problem during time of development. Also, it helps to achieve the objective and scope of the projects. In order to build the project, it need to understand the stakeholder requirements.

5.DATA ANALYSIS

❖ DATA SET

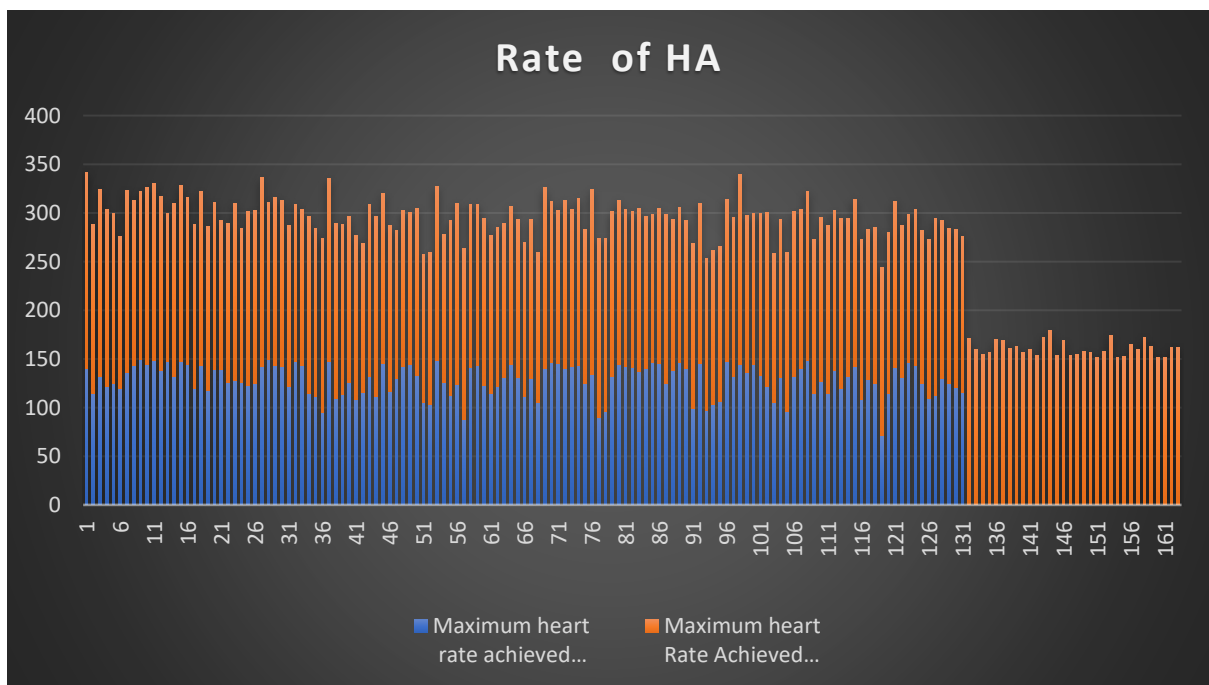
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
	Age of the person	Gender of the person	Chest Pain type	Resting blood pressure (in mm Hg)	cholesterol in mg/dl fetched via BMI sensor	(fasting blood sugar > 120 mg/dl) (1 = true; 0 = false)	resting electrocardiographic results	maximum heart rate achieved	exercise induced angina (1 = yes; 0 = no)	Previous peak	Slope	number of major vessels (0-3)	Thal rate	Target variable	
1															
2	29	1	1	130	204	0	0	202	0	0	2	0	2	1	
3	34	1	3	118	182	0	0	174	0	0	2	0	2	1	
4	34	0	1	118	210	0	1	192	0	0.7	2	0	2	1	
5	35	0	0	138	183	0	1	182	0	1.4	2	0	2	1	
6	35	1	1	122	192	0	1	174	0	0	2	0	2	1	
7	35	1	0	120	198	0	1	130	1	1.6	1	0	3	0	
8	35	1	0	126	282	0	0	156	1	0	2	0	3	0	
9	37	1	2	130	250	0	1	187	0	3.5	0	0	2	1	
10	37	0	2	120	215	0	1	170	0	0	2	0	2	1	
11	38	1	2	138	175	0	1	173	0	0	2	4	2	1	
12	38	1	3	120	231	0	1	182	1	3.8	1	0	3	0	
13	39	1	2	140	321	0	0	182	0	0	2	0	2	1	
14	39	0	2	94	199	0	1	179	0	0	2	0	2	1	
15	39	0	2	138	220	0	1	152	0	0	1	0	2	1	
16	39	1	0	118	219	0	1	140	0	1.2	1	0	3	0	
17	40	1	3	140	199	0	1	178	1	1.4	2	0	3	1	
18	40	1	0	110	167	0	0	114	1	2	1	0	3	0	
19	40	1	0	152	223	0	1	181	0	0	2	0	3	0	
20	41	0	1	130	204	0	0	172	0	1.4	2	0	2	1	
21	41	0	1	105	198	0	1	168	0	0	2	1	2	1	
22	41	1	1	135	203	0	1	132	0	0	1	0	1	1	
23	41	1	2	112	250	0	1	179	0	0	2	0	2	1	
24	41	1	2	130	214	0	0	168	0	2	1	0	2	1	
25	41	0	2	112	268	0	0	172	1	0	2	0	2	1	

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
	Age of the person	Gender of the person	Chest Pain type	Resting blood pressure (in mm Hg)	cholesterol in mg/dl fetched via BMI sensor	(fasting blood sugar > 120 mg/dl) (1 = true; 0 = false)	resting electrocardiographic results	maximum heart rate achieved	exercise induced angina (1 = yes; 0 = no)	Previous peak	Slope	number of major vessels (0-3)	Thal rate	Target variable	
274	66	0	2	146	278	0	0	152	0	0	1	1	2	1	
275	66	1	1	160	246	0	1	120	1	0	1	3	1	0	
276	66	0	0	178	228	1	1	165	1	1	1	2	3	0	
277	66	1	0	112	212	0	0	132	1	0.1	2	1	2	0	
278	67	0	2	115	564	0	0	160	0	1.6	1	0	3	1	
279	67	0	2	152	277	0	1	172	0	0	2	1	2	1	
280	67	0	0	106	223	0	1	142	0	0.3	2	2	2	1	
281	67	1	0	160	286	0	0	108	1	1.5	1	3	2	0	
282	67	1	0	120	229	0	0	129	1	2.6	1	2	3	0	
283	67	1	0	125	254	1	1	163	0	0.2	1	2	3	0	
284	67	1	0	100	299	0	0	125	1	0.9	1	2	2	0	
285	67	1	0	120	237	0	1	71	0	1	1	0	2	0	
286	67	1	2	152	212	0	0	150	0	0.8	1	0	3	0	
287	68	1	2	118	277	0	1	151	0	1	2	1	3	1	
288	68	0	2	120	211	0	0	115	0	1.5	1	0	2	1	
289	68	1	2	180	274	1	0	150	1	1.6	1	0	3	0	
290	68	1	0	144	193	1	1	141	0	3.4	1	2	3	0	
291	69	0	3	140	239	0	1	151	0	1.8	2	2	2	1	
292	69	1	3	160	234	1	0	131	0	0.1	1	1	2	1	
293	69	1	2	140	254	0	0	146	0	2	1	3	3	0	
294	70	1	1	156	245	0	0	143	0	0	2	0	3	1	
295	70	1	0	145	174	0	1	125	1	2.6	0	0	3	1	
296	70	1	0	130	322	0	0	109	0	2.4	1	3	2	0	
297	70	1	2	160	269	0	1	112	1	2.9	1	1	3	0	
298	71	0	1	160	302	0	1	162	0	0.4	2	2	2	1	
299	71	0	2	110	265	1	0	130	0	0	2	1	2	1	
300	71	0	0	112	149	0	1	125	0	1.6	1	0	2	1	
301	74	0	1	120	269	0	0	121	1	0.2	2	1	2	1	
302	76	0	2	140	197	0	2	116	0	1.1	1	0	2	1	

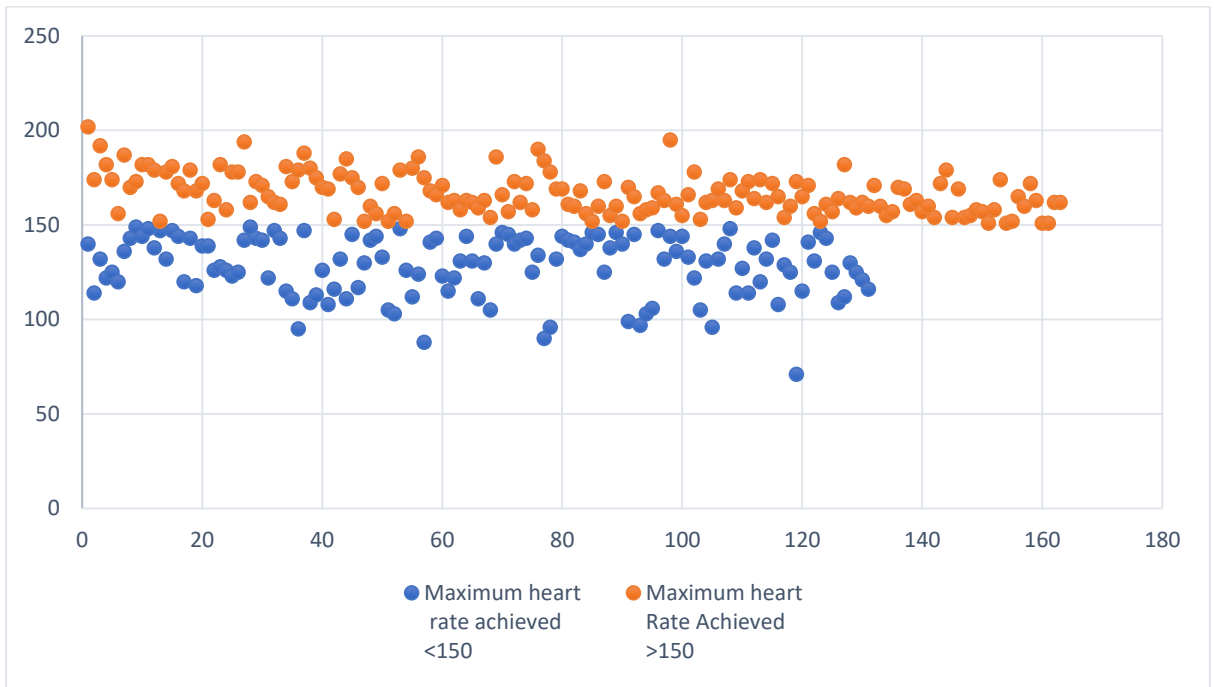
❖ RATE OF HA

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	Maximum heart rate achieved	Maximum heart rate achieved <150	Maximum heart Rate Achieved >150	Max & Min in Heart Rate Achieved															
2	174	140	202	Minimum 71															
3	192	114	174																
4	182	132	192																
5	174	122	182	Maximum 195															
6	130	125	174																
7	156	120	156																
8	187	136	187																
9	170	143	170																
10	173	149	173																
11	182	144	182																
12	182	148	182																
13	179	138	179																
14	152	147	152																
15	140	132	178																
16	178	147	181																
17	114	144	172																
18	181	120	168																
19	172	143	179																
20	168	118	168																
21	132	139	172																
22	179	139	153																
23	168	126	163																
24	172	128	182																
25	153	126	158																
26	163	123	178																

❖ COLUMN CHART



❖ **SCATTERD**



◆ AGE

heart - Excel

FileHomeInsertPage LayoutFormulasDataReviewViewDeveloperHelpTell me what you want to doShare

Clipboard

Font

Alignment

Number

Styles

Cells

Editing

Calibri11

Wrap Text

General

Conditional Formatting

Format as Table

Cell Styles

Insert

Delete

Format

AutoSum

Fill

Sort & Find & Filter

Select

Clear

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1	Age of the person	Age of the person (<50)	Age of the person (>65)	Age of the person (50 -65)																			
2	29	29	66	50																			
3	34	34	66	50																			
4	34	34	66	50																			
5	35	35	66	50																			
6	35	35	66	50																			
7	35	35	66	50																			
8	35	35	66	50																			
9	37	37	67	51																			
10	37	37	67	51																			
11	38	38	67	51																			
12	38	38	67	51																			
13	39	39	67	51																			
14	39	39	67	51																			
15	39	39	67	51																			
16	39	39	67	51																			
17	40	40	67	51																			
18	40	40	68	51																			
19	40	40	68	51																			
20	41	41	68	51																			
21	41	41	68	52																			
22	41	41	69	52																			
23	41	41	69	52																			
24	41	41	69	52																			
25	41	41	70	52																			
26	41	41	70	52																			

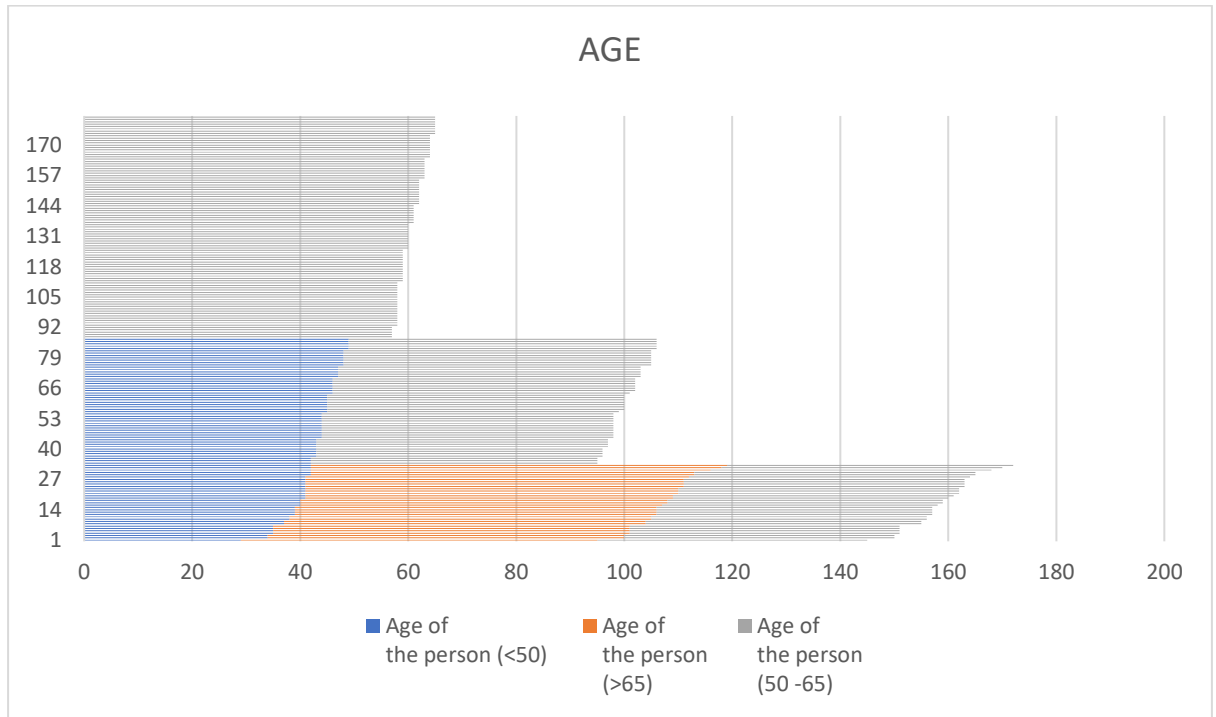
Ready

heartRate of HASEXAGERegression

Type here to search

14:1703-07-2021

❖ COLUMN CHART



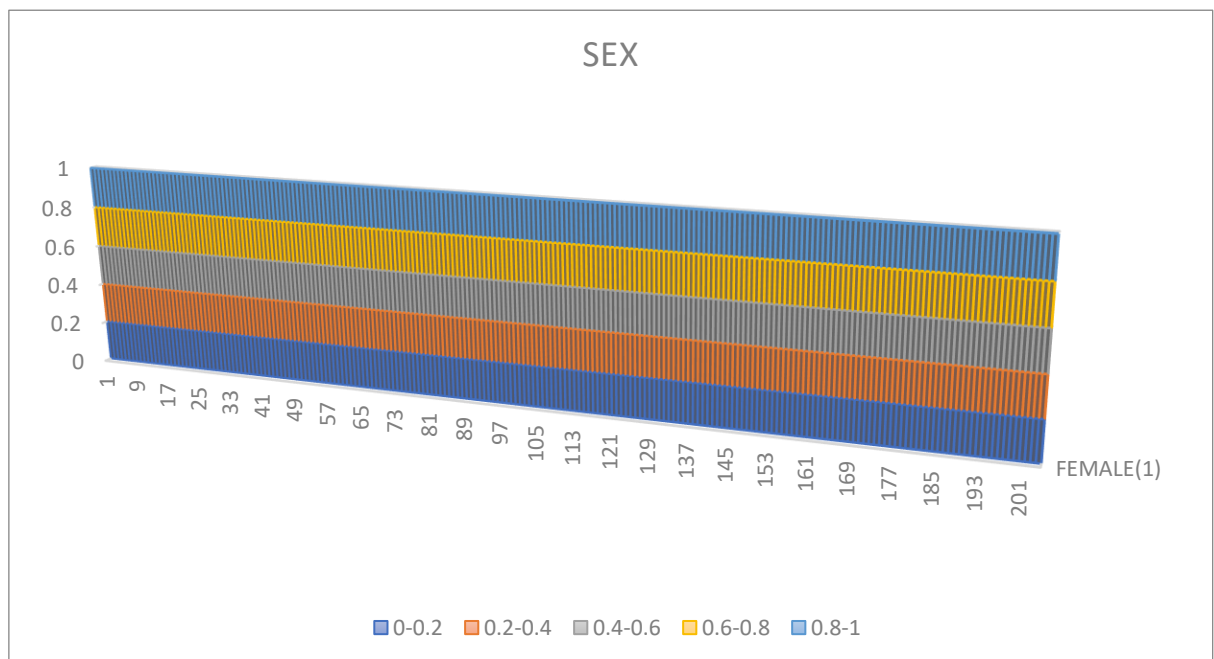
❖ LINE CHART



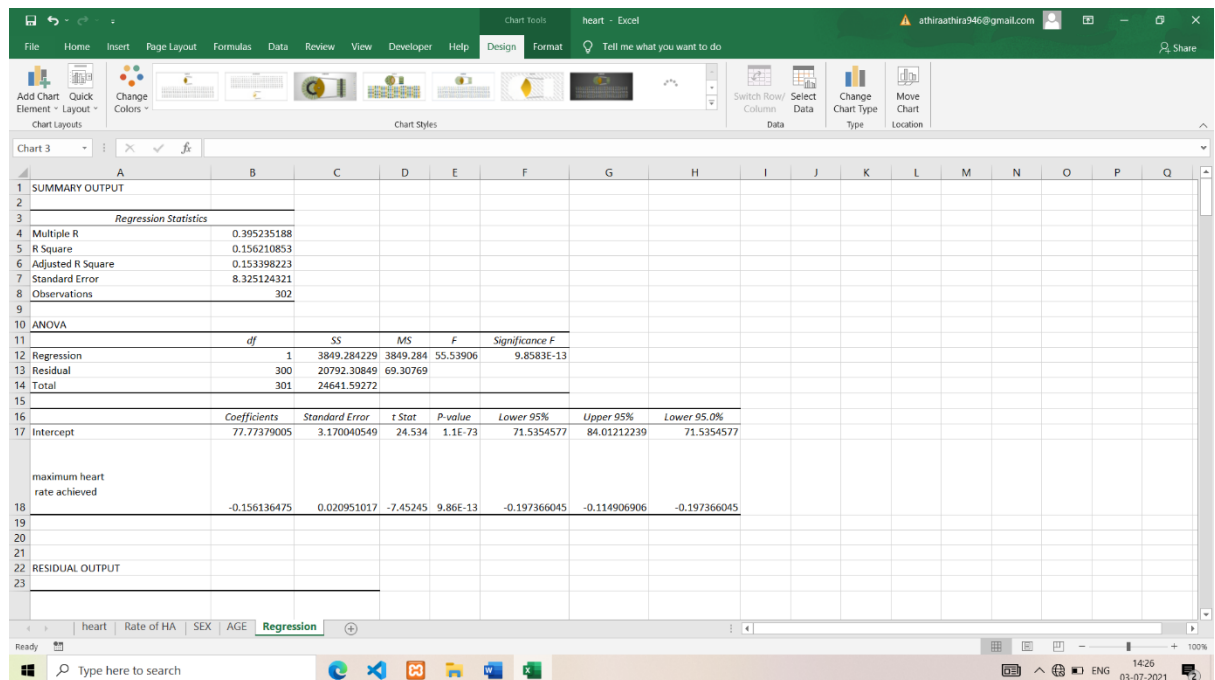
❖ SEX

	Gender of the person	FEMALE(1)	MALE(0)
1			
2	0	1	0
3	0	1	0
4	1	1	0
5	1	1	0
6	1	1	0
7	1	1	0
8	0	1	0
9	1	1	0
10	1	1	0
11	1	1	0
12	0	1	0
13	0	1	0
14	1	1	0
15	1	1	0
16	1	1	0
17	1	1	0
18	0	1	0
19	0	1	0
20	1	1	0
21	1	1	0
22	1	1	0
23	0	1	0
24	1	1	0
25	0	1	0
26	1	1	0
27	1	1	0

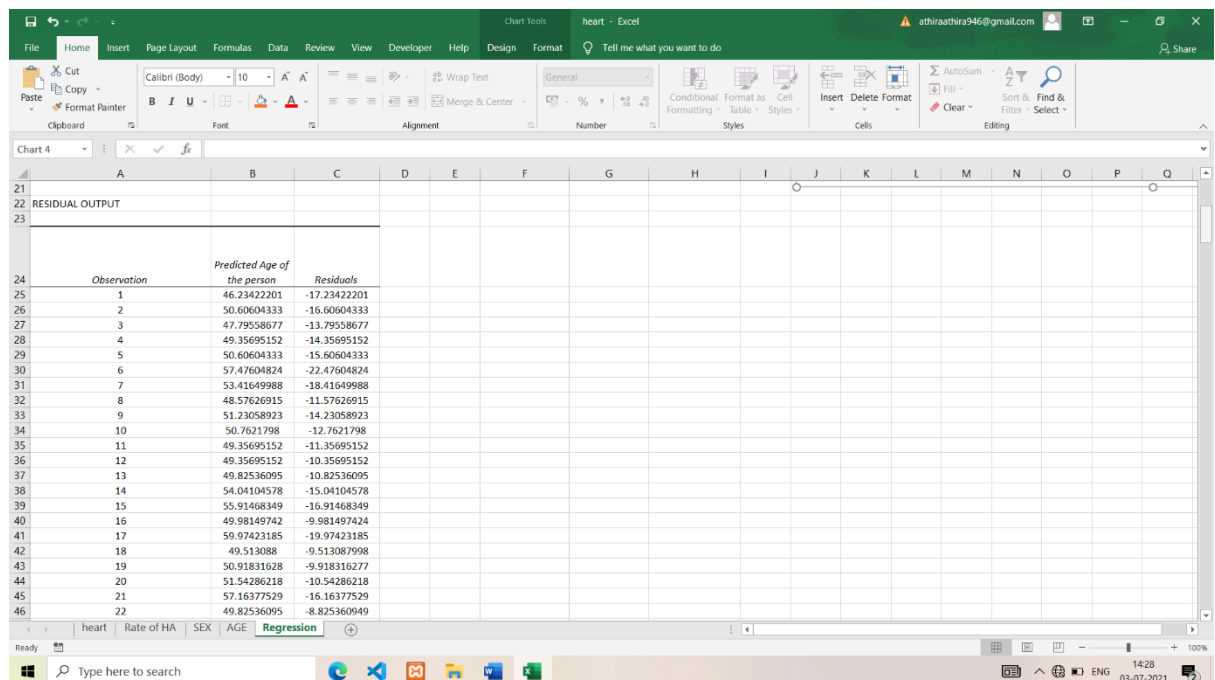
❖ SURFACE



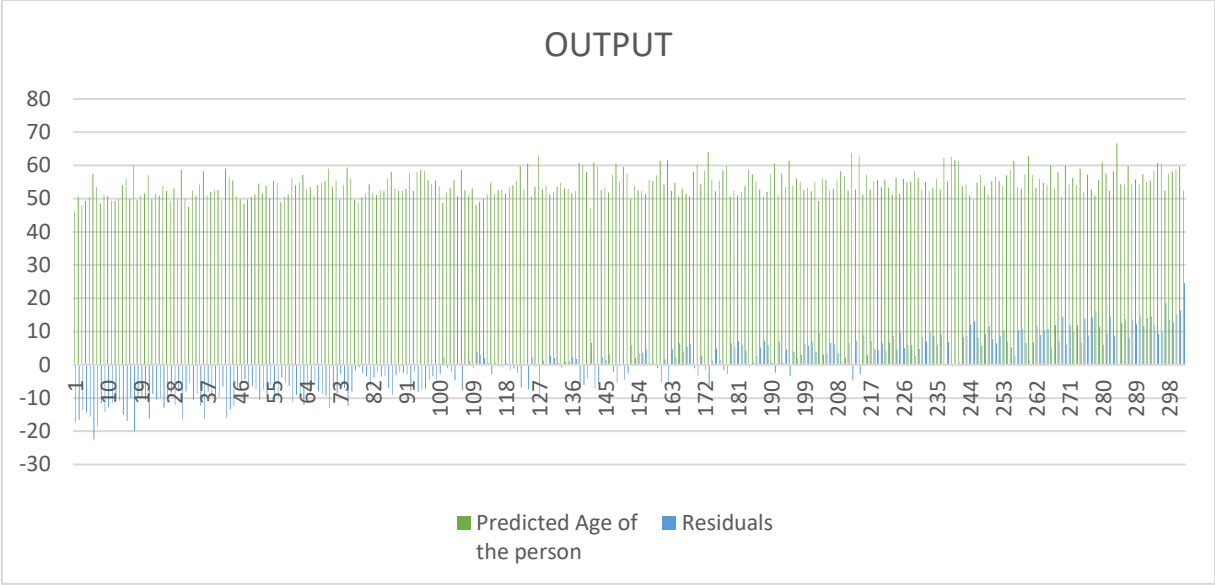
❖ REGRESSION



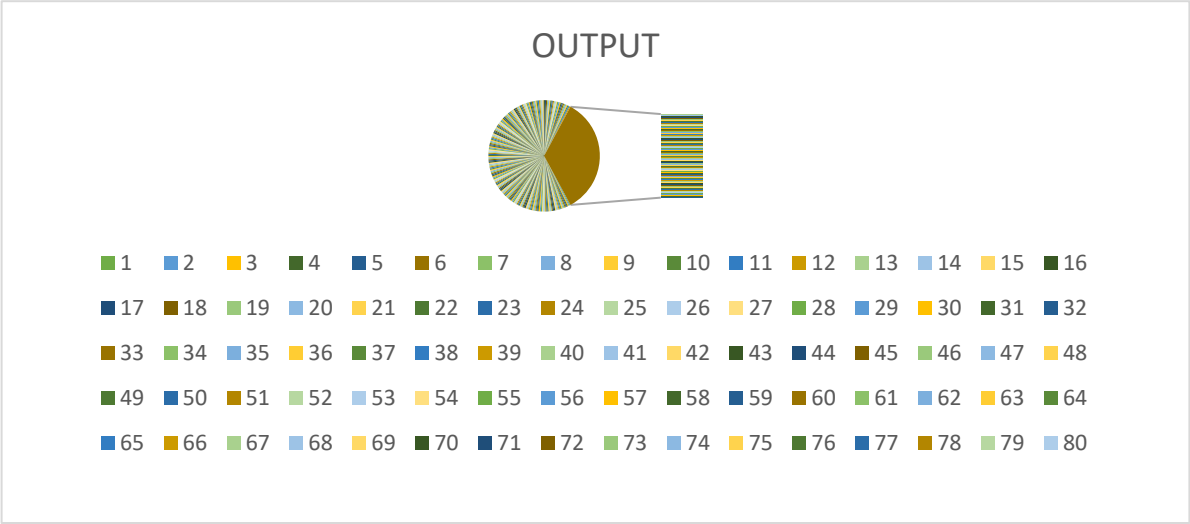
❖ RESDUAL OUTPUT



❖ COLUMN CHART



❖ **PIE CHART**



6. CONCLUSION

Heart attack analysis and prediction one of the leading causes of death worldwide and the early prediction of heart attack analysis and prediction is important. The computer aided heart attack analysis and prediction system helps the physician as a tool for heart attack analysis and prediction diagnosis. Some Heart attack analysis and prediction classification system is reviewed in this paper. From the analysis it is concluded that, data mining plays a major role in heart attack analysis and prediction classification. Neural Network with offline training is a good for disease prediction in early stage and the good performance of the system can be obtained by preprocessed and normalized dataset. The classification accuracy can be improved by reduction in features.

7. REFERNCES

- [Heart Attack Risk Prediction Using Machine Learning | by Amayo Mordecai II | Towards Data Science](#)
- [Heart Disease Prediction. Cleveland Heart Disease\(UCI Repository\)... | by Shubhankar Rawat | Towards Data Science](#)
- [ML | Heart Disease Prediction Using Logistic Regression . - GeeksforGeeks](#)