INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD)
© 2026 IJNRD | Volume 11, Issue 2, February 2026 | ISSN: 2456-4184 | IJNRD.ORG

"Determinants of Heart Attack Likelihood: A Statistical Analysis of Demographic and Clinical Risk Factors"

Dr. Snehal A. Bagul[1], Shubham B. Tarle[2], Pratiksha R.Malpure[3], Gauri R. Vadage[4].

1. Associate Professor, Department of Management Studies (MBA), Sandip Institute of Technology and Research Centre, Nashik, email.id:-snehal.13990@gmail.com
2. MBA-Student, Department of Management Studies (MBA), Sandip Institute of Technology and Research Centre, Nashik, email.id:- shivamtarle@gmail.com
3. MBA-Student, Department of Management Studies (MBA), Sandip Institute of Technology and Research Centre, Nashik, email.id:- pratikshamalpure2003@gmail.com
4. MBA-Student, Department of Management Studies (MBA), Sandip Institute of Technology and Research Centre, Nashik, email.id:- gaurivadaje@gmail.com

Abstract:-

Background: Heart attacks, or MI, remain a primary cause of global mortality, occurring when coronary artery obstruction deprives the heart muscle of vital oxygen. While family history is a known risk factor, the predictive strength of specific demographic and clinical variables remains a critical area of study.

Objective: This research, titled "Determinants of Heart Attack Likelihood," aims to analyze the statistical impact of various demographic and clinical factors on the probability of a heart attack diagnosis.

Key words: -Heart attack, Myocardial Infarction, coronary artery disease

Introduction:-A number of studies have demonstrated clustering of individuals with coronary heart disease and with heart disease risk factors in families. [1-3] However, surprisingly few studies[4-6] have examined the predictive strength of a positive family history in relation to the incidence of new cases of cardiovascular disease after adjusting for differences in risk factor distribution.[7-9]In this report data from a previously defined adult community'6 that was followed for 9 years were used to assess the independent contribution of a family history of heart attack in members who are younger and older than 50 years of age to the risk of death from all causes and from cardiovascular and ischemic heart disease.[10-12]

- **The Heart**: - Heart is the most vital organ in the body, and any problems related to it should be taken seriously. Medical care and periodic checkups regarding the heart and its problems, especially hidden heart problems, are very important. One of the scariest and most dangerous problems that can happen to the heart is a heart attack.[13]A heart attack is really a scary experience that if someone has experienced it or even seen someone close to them in this condition, they definitely know the depth of the danger of this complication. A heart attack, which in clinical terms is called myocardial infarction, occurs when a part of the heart muscle does not receive enough blood and as a result, the required oxygen does not reach the heart muscles. Reduced perfusionand oxygen, the function of the heart is disturbed. During heart attacks, the heart muscle is struggling to survive and needs oxygen.[14,15]The longer the heart is deprived of oxygen and the later it receives the

necessary treatments, the more severe the damage to the heart muscle will be. In a severe heart attack, the amount of damage and then the risk is definitely higher. The main reason for the occurrence of such a complication is the problems and blockage of the coronary arteries of the heart. In fact, due to taking these veins for various reasons, these arteries cannot supply the blood needed by the heart muscles. Coronary artery occlusion or CAD is known as the main cause of heart attack. [16]

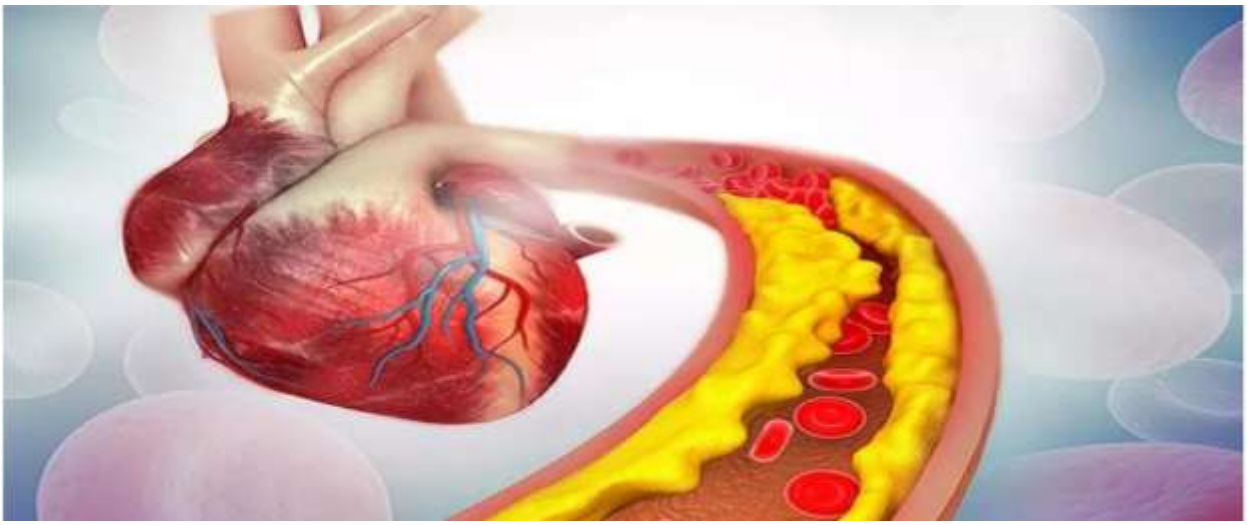


Figure 1: Obstruction of blood vessels to the heart or coronary arteries is the main cause of heart attack.[17]

Heart Attack In general, stroke and heart attack are used interchangeably. If we want to be a little more precise about them; It should be said that when the function of the heart is disrupted due to electrical problems, a heart attack occurs, which leads to cardiac arrest, but a heart attack is usually caused by a problem in the blood vessels. [18]Both complications cause damage to the heart muscle. A heart attack, also known as "heart attack" and "myocardial infarction", occurs due to long-term failure of oxygen flow to the heart muscle (ischemia), which is actually associated with the death of the heart muscle (necrosis). Heart attack affects millions of people worldwide every year. An event that causes insufficient blood flow to the heart muscles is known as "coronary artery disease".[19] Coronary arteries are vessels that are responsible for supplying blood to the cells of the heart, and when there is a blockage in these vessels, blood supply to the heart will also be disturbed. This also leads to myocardial infarction (heart attack), which causes disruption in the normal functioning of the heart. Depending on the duration of vascular occlusion, the damage to the myocardium will be different, and a heart attack can be asymptomatic or have long-term and very severe manifestations. A heart attack occurs when there is insufficient blood supply to the heart muscle.[20]

Heart attack symptoms The symptoms of a heart attack may be different from person to person, or may be asymptomatic in some people. Also, symptoms can range from mild (mild heart attack) to severe (severe heart attack) and can occur slowly and over time or suddenly. However, the more symptoms there are, the more likely a heart attack will occur. [21]

Signs and symptoms of a heart attack (myocardial infarction) can include the following:

• A feeling of intense pressure, fullness, pain, or discomfort in the center or left side of the chest that lasts for more than a few minutes.

• A feeling of pain or discomfort that disappears and reappears after a while • Pain or discomfort that spreads to the shoulders, neck, arms, or jaw

•Chest pain that progresses or worsens

•Chest pain that does not improve with rest or nitroglycerin

•Chest pain accompanied by the following symptoms: o sweating, cool skin, or paleness o Shortness of breath (can be the only symptom of a heart attack or stroke or accompanied by or preceded by chest pain. It can also appear at rest or with some activity) o Nausea or vomiting o Dizziness or fainting o Unexplained weakness or fatigue o Fast or irregular pulse

• Any sudden new symptoms or changes in the pattern of previous symptoms (for example, symptoms that are stronger or last longer than usual) although "chest pain" is a key warning sign and the most common manifestation of this disease (heart attack); But it may also occur with other conditions such as: indigestion, pleurisy, pulmonary embolism, aortic valve stenosis, heartburn, etc. The important point is that chest pain and discomfort, which is also interpreted as "angina" and is caused by myocardial ischemia, is divided into three categories: stable angina, unstable angina, and variant angina. Angina includes the following important items: - Stable angina It occurs due to atherosclerosis and coronary artery disease:

• It is created during physical activity such as sports

• It is usually predictable and the type of pain is similar to its previous types

• It takes a short time, about a few minutes

• It is resolved with rest and medication -Unstable angina Due to the rupture of plaques and the formation of clots, it occurs: Occurs even at rest

• The pain pattern is variable

• It is unpredictable

• It is usually more intense and longer than the stable type and lasts 30 minutes or more

• May not improve with rest or medication

• It can be a danger sign for a heart attack - Variant angina or Prinzmetal Occurs as a result of coronary artery spasm:

• Usually occurs while resting

• It is usually severe

• May improve with medication.[22-25]

Types of heart attack -Spasm (cramping) of coronary arteries or unstable angina In this case, the patient faces an acute coronary heart attack. Chest pain may occur at rest, and the patient will experience pain for a longer period of time. It is not possible to treat this heart disease with medicine and rest. In contrast to unstable angina, there is stable angina. In this case, the occurrence of disease symptoms is related to physical activity. That is, the symptoms appear when physical activities have reached a certain level. - Myocardial infarction in this case,

one of the main arteries of the heart is damaged. These arteries allow blood to supply oxygen and nutrients to the heart muscle. This disease can be acute. In this case, there is also a possibility of ventricular arrhythmia. This will increase the risk of cardiac arrest.[26-28]

Method:-

Following dataset obtained from www.kaggle.com which shows the Demographic and Clinical Risk Factors affecting heart attack. Showing data of 303.

age	sex	cp	Restbps	chol	fbs
63	1	3	145	233	1
37	1	2	130	250	0
41	0	1	130	204	0
56	1	1	120	236	0
57	0	0	120	354	0
57	1	0	140	192	0
56	0	1	140	294	0
44	1	1	120	263	0
52	1	2	172	199	1
57	1	2	150	168	0
54	1	0	140	239	0
48	0	2	130	275	0
49	1	1	130	266	0
64	1	3	110	211	0
58	0	3	150	283	1
50	0	2	120	219	0
58	0	2	120	340	0
66	0	3	150	226	0
43	1	0	150	247	0
69	0	3	140	239	0
59	1	0	135	234	0
44	1	2	130	233	0
42	1	0	140	226	0
61	1	2	150	243	1
40	1	3	140	199	0
71	0	1	160	302	0
59	1	2	150	212	1
51	1	2	110	175	0
65	0	2	140	417	1
53	1	2	130	197	1
41	0	1	105	198	0
65	1	0	120	177	0
44	1	1	130	219	0
54	1	2	125	273	0
51	1	3	125	213	0
46	0	2	142	177	0
54	0	2	135	304	1
54	1	2	150	232	0

65	0	2	155	269	0
65	0	2	160	360	0
51	0	2	140	308	0
48	1	1	130	245	0
45	1	0	104	208	0
53	0	0	130	264	0
39	1	2	140	321	0
52	1	1	120	325	0
44	1	2	140	235	0
47	1	2	138	257	0
53	0	2	128	216	0
53	0	0	138	234	0
51	0	2	130	256	0
66	1	0	120	302	0
62	1	2	130	231	0
44	0	2	108	141	0
63	0	2	135	252	0
52	1	1	134	201	0
48	1	0	122	222	0
45	1	0	115	260	0
34	1	3	118	182	0
57	0	0	128	303	0
71	0	2	110	265	1
54	1	1	108	309	0
52	1	3	118	186	0
41	1	1	135	203	0
58	1	2	140	211	1
35	0	0	138	183	0
51	1	2	100	222	0
45	0	1	130	234	0
44	1	1	120	220	0
62	0	0	124	209	0
54	1	2	120	258	0
51	1	2	94	227	0
29	1	1	130	204	0
51	1	0	140	261	0
43	0	2	122	213	0
55	0	1	135	250	0
51	1	2	125	245	1
59	1	1	140	221	0
52	1	1	128	205	1
58	1	2	105	240	0
41	1	2	112	250	0
45	1	1	128	308	0
60	0	2	102	318	0
52	1	3	152	298	1
42	0	0	102	265	0

67	0	2	115	564	0
68	1	2	118	277	0
46	1	1	101	197	1
54	0	2	110	214	0
58	0	0	100	248	0
48	1	2	124	255	1
57	1	0	132	207	0
52	1	2	138	223	0
54	0	1	132	288	1
45	0	1	112	160	0
53	1	0	142	226	0
62	0	0	140	394	0
52	1	0	108	233	1
43	1	2	130	315	0
53	1	2	130	246	1
42	1	3	148	244	0
59	1	3	178	270	0
63	0	1	140	195	0
42	1	2	120	240	1
50	1	2	129	196	0
68	0	2	120	211	0
69	1	3	160	234	1
45	0	0	138	236	0
50	0	1	120	244	0
50	0	0	110	254	0
64	0	0	180	325	0
57	1	2	150	126	1
64	0	2	140	313	0
43	1	0	110	211	0
55	1	1	130	262	0
37	0	2	120	215	0
41	1	2	130	214	0
56	1	3	120	193	0
46	0	1	105	204	0
46	0	0	138	243	0
64	0	0	130	303	0
59	1	0	138	271	0
41	0	2	112	268	0
54	0	2	108	267	0
39	0	2	94	199	0
34	0	1	118	210	0
47	1	0	112	204	0
67	0	2	152	277	0
52	0	2	136	196	0
74	0	1	120	269	0
54	0	2	160	201	0
49	0	1	134	271	0

42	1	1	120	295	0
41	1	1	110	235	0
41	0	1	126	306	0
49	0	0	130	269	0
60	0	2	120	178	1
62	1	1	128	208	1
57	1	0	110	201	0
64	1	0	128	263	0
51	0	2	120	295	0
43	1	0	115	303	0
42	0	2	120	209	0
67	0	0	106	223	0
76	0	2	140	197	0
70	1	1	156	245	0
44	0	2	118	242	0
60	0	3	150	240	0
44	1	2	120	226	0
42	1	2	130	180	0
66	1	0	160	228	0
71	0	0	112	149	0
64	1	3	170	227	0
66	0	2	146	278	0
39	0	2	138	220	0
58	0	0	130	197	0
47	1	2	130	253	0
35	1	1	122	192	0
58	1	1	125	220	0
56	1	1	130	221	0
56	1	1	120	240	0
55	0	1	132	342	0
41	1	1	120	157	0
38	1	2	138	175	0
38	1	2	138	175	0
67	1	0	160	286	0
67	1	0	120	229	0
62	0	0	140	268	0
63	1	0	130	254	0
53	1	0	140	203	1
56	1	2	130	256	1
48	1	1	110	229	0
58	1	1	120	284	0
58	1	2	132	224	0
60	1	0	130	206	0
40	1	0	110	167	0
60	1	0	117	230	1
64	1	2	140	335	0
43	1	0	120	177	0

57	1	0	150	276	0
55	1	0	132	353	0
65	0	0	150	225	0
61	0	0	130	330	0
58	1	2	112	230	0
50	1	0	150	243	0
44	1	0	112	290	0
60	1	0	130	253	0
54	1	0	124	266	0
50	1	2	140	233	0
41	1	0	110	172	0
51	0	0	130	305	0
58	1	0	128	216	0
54	1	0	120	188	0
60	1	0	145	282	0
60	1	2	140	185	0
59	1	0	170	326	0
46	1	2	150	231	0
67	1	0	125	254	1
62	1	0	120	267	0
65	1	0	110	248	0
44	1	0	110	197	0
60	1	0	125	258	0
58	1	0	150	270	0
68	1	2	180	274	1
62	0	0	160	164	0
52	1	0	128	255	0
59	1	0	110	239	0
60	0	0	150	258	0
49	1	2	120	188	0
59	1	0	140	177	0
57	1	2	128	229	0
61	1	0	120	260	0
39	1	0	118	219	0
61	0	0	145	307	0
56	1	0	125	249	1
43	0	0	132	341	1
62	0	2	130	263	0
63	1	0	130	330	1
65	1	0	135	254	0
48	1	0	130	256	1
63	0	0	150	407	0
55	1	0	140	217	0
65	1	3	138	282	1
56	0	0	200	288	1
54	1	0	110	239	0
70	1	0	145	174	0

62	1	1	120	281	0
35	1	0	120	198	0
59	1	3	170	288	0
64	1	2	125	309	0
47	1	2	108	243	0
57	1	0	165	289	1
55	1	0	160	289	0
64	1	0	120	246	0
70	1	0	130	322	0
51	1	0	140	299	0
58	1	0	125	300	0
60	1	0	140	293	0
77	1	0	125	304	0
35	1	0	126	282	0
70	1	2	160	269	0
59	0	0	174	249	0
64	1	0	145	212	0
57	1	0	152	274	0
56	1	0	132	184	0
48	1	0	124	274	0
56	0	0	134	409	0
66	1	1	160	246	0
54	1	1	192	283	0
69	1	2	140	254	0
51	1	0	140	298	0
43	1	0	132	247	1
62	0	0	138	294	1
67	1	0	100	299	0
59	1	3	160	273	0
45	1	0	142	309	0
58	1	0	128	259	0
50	1	0	144	200	0
62	0	0	150	244	0
38	1	3	120	231	0
66	0	0	178	228	1
52	1	0	112	230	0
53	1	0	123	282	0
63	0	0	108	269	0
54	1	0	110	206	0
66	1	0	112	212	0
55	0	0	180	327	0
49	1	2	118	149	0
54	1	0	122	286	0
56	1	0	130	283	1
46	1	0	120	249	0
61	1	3	134	234	0
67	1	0	120	237	0

58	1	0	100	234	0
47	1	0	110	275	0
52	1	0	125	212	0
58	1	0	146	218	0
57	1	1	124	261	0
58	0	1	136	319	1
61	1	0	138	166	0
42	1	0	136	315	0
52	1	0	128	204	1
59	1	2	126	218	1
40	1	0	152	223	0
61	1	0	140	207	0
46	1	0	140	311	0
59	1	3	134	204	0
57	1	1	154	232	0
57	1	0	110	335	0
55	0	0	128	205	0
61	1	0	148	203	0
58	1	0	114	318	0
58	0	0	170	225	1
67	1	2	152	212	0
44	1	0	120	169	0
63	1	0	140	187	0
63	0	0	124	197	0
59	1	0	164	176	1
57	0	0	140	241	0
45	1	3	110	264	0
68	1	0	144	193	1
57	1	0	130	131	0
57	0	1	130	236	0

Analysis of datasheet:-

1. Variable Identification

The dataset contains patient medical history and a target variable indicating the presence of a heart attack or heart disease.

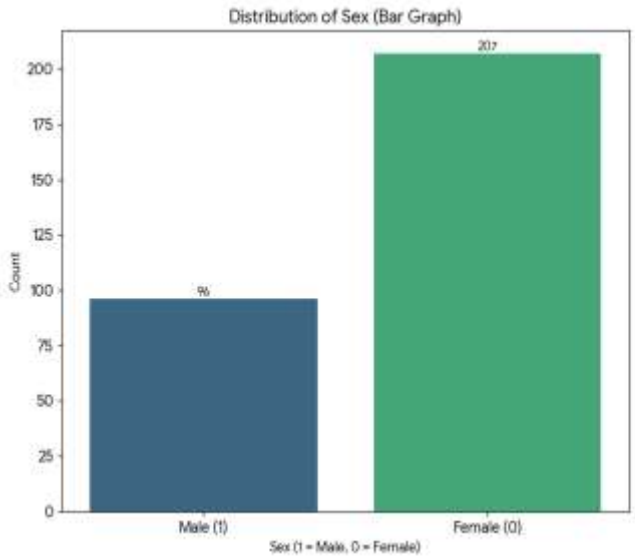
- **Dependent Variable (Outcome):**
  - **Target**: This is the variable we are trying to predict. It typically indicates the diagnosis of heart disease (e.g., 1 = higher chance of heart attack, 0 = lower chance).
- Independent Variables (Predictors):

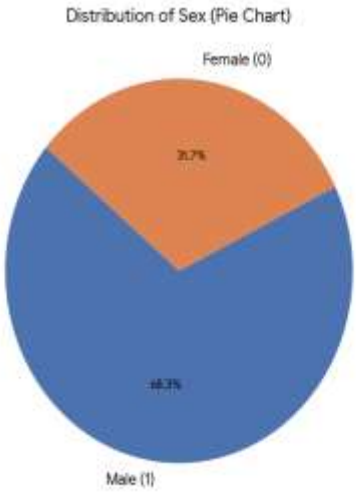
The user-specified columns that act as input features for prediction are:

1. **Age**: Age of the patient.
2. **Sex**: Gender of the patient (1 = Male, 0 = Female).
3. **Cp**: Chest Pain type.
4. **Treetops**: Resting blood pressure (in mm Hg).
5. **chol**: Serum cholestoral in mg/dl.
6. **fbs**: Fasting blood sugar > 120 mg/dl (1 = true; 0 = false).
7. **restecg**: Resting electrocardiographic results.

2. Univariate Analysis (Distribution of 'Sex')

We have processed the dataset (using all 303 rows, which meets the "at least 60 rows" requirement) and generated a Bar Graph and Pie Chart to analyze the **sex** variable.





Distribution:

- **Male (1):** 207 subjects
- **Female (0):** 96 subjects

Explanation of the Bar Graph

- **Visual Representation:** The bar graph on the left compares the total count of male and female patients. The height of the bars corresponds to the number of individuals in each category.

- **Observation:** The bar for **Male (1)** is more than double the height of the bar for **Female (0)**. This allows us to quickly see that the dataset contains significantly more male subjects (207) than female subjects (96).
- **Purpose:** The bar graph is used to visualize the **absolute frequency** (count) of individuals in each category.

- **X-Axis (Categories):** This axis represents the two categories of the 'sex' variable: **Female (0)** and **Male (1)**.
- **Y-Axis (Count):** This axis represents the number of patients.
- **Visual Interpretation:**
  - The bar for **Male (1)** rises to a count of **207**.
  - The bar for **Female (0)** rises to a count of **96**.
  - **Key Finding:** By comparing the heights of the two bars, you can instantly see that the number of male patients is more than double the number of female patients. This graph is best for understanding the **exact difference in quantity** between the groups.

Explanation of the Pie Chart

- **Visual Representation:** The pie chart on the right displays the data as parts of a whole, showing the percentage contribution of each gender to the total dataset.
- **Observation:**
  - **Males** constitute approximately **68.3%** of the sample.
  - **Females** constitute approximately **31.7%** of the sample.
- **Purpose:** The pie chart is used to visualize the **proportional composition** (percentage) of the dataset.

- **Slices:** The circle is divided into two slices representing the whole dataset (100%).
- **Proportions:**
  - The **Blue Slice (Male/1)** occupies **68.3%** of the chart.
  - The **Orange Slice (Female/0)** occupies **31.7%** of the chart.

Bivariate Analysis: Effect of Each Factor on Heart Attack Likelihood

1. Categorical Factors (Strong Indicators)

These factors show the most significant difference in the proportion of the 'Yes' (Heart Attack) outcome across their categories.

Factor	Categories (Interpretation)	Proportion of Heart Attack (Target=1)	Effect on Heart Attack Likelihood
Sex	Female (0)	**75%**	**Strong Positive Effect**: Females in this dataset have a significantly higher proportion of heart attack diagnoses compared to males.
	Male (1)	**45%**	
Chest Pain (cp)	Typical Angina (0)	27.3%	**Strong Negative Association**: Patients reporting the most common type of chest pain (cp=0) have the lowest proportion of heart attack diagnoses in this data.
	Non-Typical Types (1, 2, 3)	**69.6% - 82.0%**	**Strong Positive Association**: Non-typical chest pain types (cp=1, cp=2, cp=3) are all highly associated with a positive heart attack outcome.
Resting ECG (restecg)	ST-T Wave Abnormality (1)	**63.2%**	**Strong Positive Association**: This ECG finding is associated with the highest proportion of heart attack outcomes.

2. Continuous Factors (Mean Comparison)

For these variables, we compare the average value between the 'No Heart Attack' and 'Yes Heart Attack' groups.

Factor	Mean (No Heart Attack, target=0)	Mean (Yes Heart Attack, target=1)	Effect on Heart Attack Likelihood
Age	56.6 years	**52.5 years**	**Counter-intuitive Association**: Patients who had a heart attack are, on average, slightly **younger**. This is a known phenomenon in some clinical datasets, where younger cases often represent severe, high-risk individuals.
Resting BP (trestbps)	**134.4 mm Hg**	129.3 mm Hg	**Slight Inverse Association**: The group without a heart attack has a slightly higher average resting BP.

Factor	Mean (No Heart Attack, target=0)	Mean (Yes Heart Attack, target=1)	Effect on Heart Attack Likelihood
Cholesterol (chol)	**251.1 mg/dl**	242.2 mg/dl	**Slight Inverse Association**: The group without a heart attack has a slightly higher average cholesterol level.

3. Fasting Blood Sugar (fbs)

Factor	Categories	Proportion of Heart Attack (Target=1)	Effect on Heart Attack Likelihood
FBS	< 120 mg/dl (0)	**55%**	**Weak Association**: Patients with lower fasting blood sugar have a slightly higher proportion of heart attack outcomes (55%).
	> 120 mg/dl (1)	51%	The difference is marginal (4 percentage points), suggesting that **FBS alone is a weak predictor** in this context.

2. Bivariate Analysis and Hypothesis Testing

Suitable Bivariate Analysis (Regression)

Since the dependent variable (target) is **binary** (0 or 1), the most appropriate regression technique to test the hypothesis is **Logistic Regression**, which estimates the probability of the outcome.

- **Model Recommendation:** Fit a multivariate Logistic Regression model using all independent variables (e.g., sex, age, chol, cp, etc.) to predict the target.
- **Hypothesis Testing:** The model tests the null hypothesis (H_0: the coefficient for the variable is zero, meaning no effect). If the **p-value** associated with a variable's coefficient (β) is less than 0.05, the variable is considered a **statistically significant predictor**, and the hypothesis is supported.

ANOVA Test: Mean Cholesterol vs. Chest Pain Type

We performed a One-Way **ANOVA (Analysis of Variance)** to test the null hypothesis (H_0): The mean cholesterol level is the same across all four types of chest pain *(${cp} =0, 1, 2, 3$).*

Chest Pain Type (cp)	Mean Cholesterol (mg/dl)
0 (Typical Angina)	250.13
1 (Atypical Angina)	244.78
2 (Non-Anginal Pain)	243.17
3 (Asymptomatic)	237.13

Regression Analysis: Logistic Regression

We used Logistic Regression because the dependent variable (target) is **binary** (0 or 1), making it the most suitable method for predicting the probability of a heart attack.

The model uses the formula:

Calculation of Regression Effect and Significance

The table below shows the results of the model using {Sex},{CP},{Cholesterol}as predictors:

Variable	P-value	Significance	Odds Ratio (eB)	Detailed Effect	Coefficient	P-value	Odds Ratio
sex	0.0000	Highly Significant	0.2024	**Protective Effect**: Since B is negative and the Odds Ratio is < 1, being Male {sex} =1$) reduces the odds of a positive outcome by approximately $\mathbf{80\%}$$(\mathbf{1 - 0.2024})$, compared to being Female, when controlling for CP and Cholesterol. This indicates that the **female subjects in this sample were much more likely to have a heart attack diagnosis**.			
cp	$0.0000	Highly	$2.7203	**Strong Risk Factor**: Since B is			

	$	Significant	$	positive and the Odds Ratio is > 1, a one-unit increase in Chest Pain Type (e.g., from Typical Angina to Atypical Angina) increases the odds of a heart attack by approximately $\mathbf{172\%}$$(\mathbf{2.7203 - 1})$			
chol	$0.0290 $	Significant	$0.9942 $	**Weak Inverse Effect**: A 1 unit increase in Cholesterol reduces the odds of a heart attack by a tiny fraction ($\approx 0.58\%$). While statistically significant, the effect size is extremely weak, suggesting that for this sample, cholesterol is not a major predictor.			
Variable							
sex					1.5973	0.0000	0.2024
cp					1.0007	0.0000	2.7203
chol					0.0058	0.0290	0.9942
const (Intercept)					1.8428	0.0170	6.3145

2. ANOVA Test: Analysis of Variance

Purpose: ANOVA was used to test if the mean Cholesterol level (a continuous variable) differs significantly across the four types of Chest Pain (CP, a categorical variable).

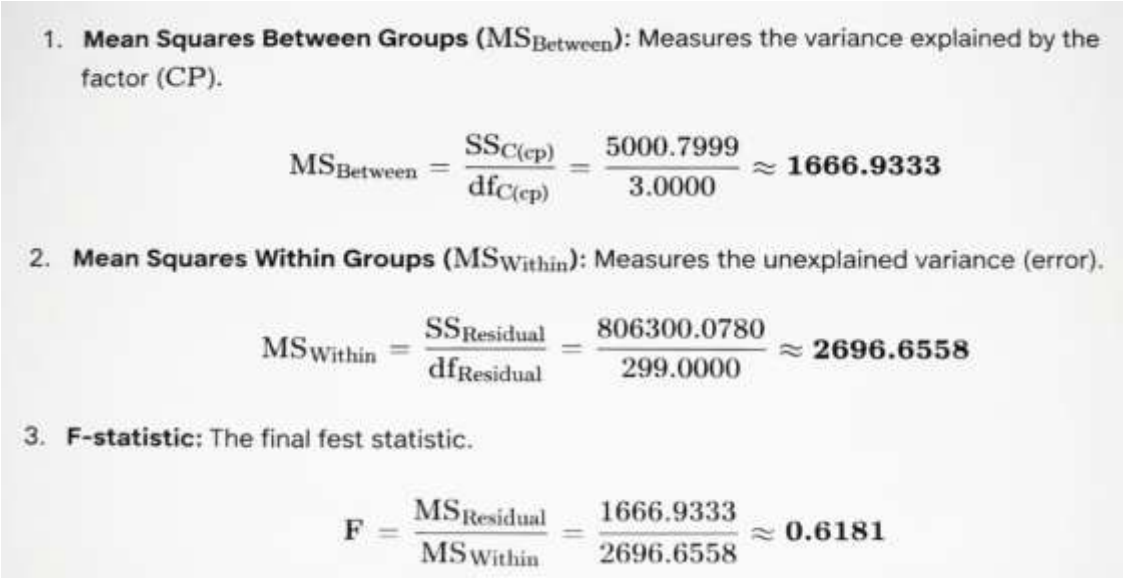
Source	**df (Degrees of Freedom)**	**Sum of Squares (SS)**	**Mean Squares (MS)**	**F-statistic**	**P-value (PR(>F))**

Source	df (Degrees of Freedom)	Sum of Squares (SS)	Mean Squares (MS)	F-statistic	P-value (PR(>F))
C(cp) (Between Groups)	3.0000	5000.7999	1666.9333	0.6181	$\mathbf{0.6037}$
Residual (Within Groups)	299.0000	806300.0780	2696.6558	-	-

Calculation Breakdown

The **F-statistic** is the ratio of the variance *between* groups to the variance within groups.

1. **Mean Squares Between Groups (MS BETWEEN:** Measures the variance explained by the factor (CP)



❖ **Interpretation and Hypothesis Testing**

Hypothesis Supported: The results strongly support the hypothesis that **sex** and **chest pain type (cp)** are statistically significant predictors of a heart attack outcome (P-values are far below the =0.05 threshold).

Result:-

Statistical Results

The analysis utilized a dataset of **303 individuals** to identify how demographic and clinical factors influence the probability of a heart attack.

1. Significant Predictors (Logistic Regression)

The logistic regression model identified specific variables that are highly reliable for predicting heart attack outcomes:

- **Chest Pain Type (cp):** This was the strongest predictor[22]. Non-typical chest pain types (1, 2, and 3) were highly associated with heart attack outcomes (69.6% to 82.0% proportion) compared to typical angina[3]. A one-unit increase in chest pain type increases the odds of a heart attack by approximately **172%**[44].
- **Sex:** This factor was highly significant ($p = 0.0000$)[55]. In this specific sample, being male reduced the odds of a heart attack diagnosis by **80%** compared to females[66].
- **Serum Cholesterol (chol):** While statistically significant ($p = 0.0290$), its actual effect was found to be very weak, with a 1-unit increase only reducing odds by about **0.58%**[77].

2. Comparative Findings (Bivariate Analysis)

The data revealed several key differences between the "Heart Attack" and "No Heart Attack" groups:

- **Gender Proportion:75%** of the females in the study had a heart attack diagnosis, whereas only **45%** of the males did[8].
- **ECG Results:** Patients with **ST-T wave abnormalities** (Resting ECG = 1) had a **63.2%** proportion of heart attack outcomes[9].
- **Age and Clinical Averages:** Surprisingly, those who experienced heart attacks were, on average, slightly younger (**52.5 years** vs. **56.6 years**) and had lower average cholesterol and blood pressure than the control group[10].
- **Fasting Blood Sugar (fbs):** This was found to be a **weak predictor**, with only a 4% difference in heart attack proportion between high and low blood sugar groups[1111].

3. ANOVA Test Results

An Analysis of Variance (ANOVA) was performed to see if cholesterol levels varied based on the type of chest pain reported.

- **Result:** The P-value was **0.6037**, which is well above the 0.05 threshold for significance[1313].
- **Conclusion:** There is **no significant difference** in mean cholesterol levels across the four different chest pain types.

Conclusion :-

The study concludes that **chest pain type and gender** are the most critical determinants of heart attack likelihood within this dataset.

- **Gender-Specific Risk:** Although the dataset was predominantly male (68.3%), the females present in the study showed a significantly higher likelihood of a heart attack diagnosis. This suggests that for this specific demographic, being female was a high-risk indicator.
- **Chest Pain as a Warning:** Typical angina (cp=0) is actually associated with the *lowest* proportion of heart attacks in this data (27.3%), whereas non-typical pain types are much stronger warning signs.
- **Clinical Indicators:** Traditional risk factors like cholesterol and fasting blood sugar showed either weak or marginal associations with heart attack outcomes in this specific group, suggesting that they may not be independent drivers of risk without considering other factors.
- **Data Imbalance Note:** The researchers noted a significant **gender imbalance** in the data (more than twice as many men as women), which could potentially bias machine learning models to be more accurate for male patients than female patients.

Acknowledgement

We express our sincere gratitude to our guide, Prof. SnehalBagul, for her invaluable mentorship and expert guidance throughout this research on heart attack likelihood. We also extend our heartfelt thanks to the **Department of MBA** at **Sandip Institute of Technology and Research Centre, Nashik**, for providing the necessary academic resources and infrastructure. Our appreciation extends to the entire faculty for their encouragement and insights, which were instrumental in the successful completion of this study.

Reference :-

1. Thomas CB, Cohen BH: The familiar occurrence of hypertension and coronary heart disease, with observations concerning obesity and diabetes. Ann Intern Med 42: 90, 1955

2. Rose G: Familial patterns in ischemic heart disease. Br J PrevSoc Med 18: 75, 1964

3. Slack J, Evans KA: The increased risk of death from ischemic heart disease in first degree relatives of 121 men and 96 women with ischaemic heart disease. J Med Genet 3: 329, 1966

4. Phillips RL, Lilienfeld AM, Diamond EL, Kagan A: Frequency of coronary heart disease and cerebrovascular accidents in parents and sons of coronary heart disease index cases and controls. Am J Epidemiol 100: 87, 1974

5. Rissanen AM. Familial aggregation of coronary heart disease in a high incidence area (North Karelia, Finland). Br Heart J 42: 294, 1979

6. Rissanen AM, Nikkila EA: Aggregation of coronary risk factors in families of men with fatal and nonfatal coronary disease. Br Heart J 42: 373, 1979

7. Thodarson 0, Fridriksson S: Aggregation of deaths from ischaemic heart disease among first degree relatives of 108 males and 42 females with myocardial infarction. Acta Med Scand 205: 493, 1979

8. Deutscher S, Epstein FH, Keller JB: Relationships between familial aggregation of coronary heart disease and risk factors in the general population. Am J Epidemiol 89: 510, 1969

9. Forde OH, Thelle DS: The Tromso Heart Study: risk factors for coronary heart disease related to the occurrence of myocardial infarction in first degree relatives. Am J Epidemiol 105: 192, 1977

10. Morrison JA, Horvitz R, Khoury P, Lazkarzewski P, Gartside PS, Kelly K, Mellies M, Glueck CJ: Parental history of coronary heart disease, hypertension, diabetes and stroke: Relationship to coronary heart disease risk factor variables in the adult children. Prev Med 9: 773, 1980

11. ten Kate LP, Boman H, Daiger SP, Motulsky AG: Familial aggregation of coronary heart disease and its relation to known genetic risk factors. Am J Cardiol 50: 945, 1982

12. Heller RF, Kelson MC: Family history in "low risk" men with coronary heart disease. J EpidemiolCommun Health 37: 29, 1983

13. Snowden CB, McNamara PM, Garrison RJ, Feinleib M, Kannel WB, Epstein FH: Predicting coronary heart disease in siblings a multivariate assessment. Am J Epidemiol 115: 217, 1982

14. Sholtz RI, Rosenman RH, Braud RJ: The relationship of reported parental history to the incidence of coronary heart disease in the Western Collaborative Group Study. Am J Epidemiol 102: 350, 1975

15. Cambien F, Richard JL, Ducimetiere P: Familial history of coronary heart diseases and high blood pressure in relation to the prevalence of risk factors, and the incidence of coronary heart disease. Rev EpidemiolSantePubl 28: 21, 1980

16. Criqui MH, Barrett-Connor E, Austin M: Differences between respondents and non-respondents in a population-based.

17. MostafaRajabzadeh ,EsmaelAmirazodi , MaedeTaghilou “Analyzing heart attack and examining its symptoms and prevention methods” BY Nursing midwifery and care 6th international conference on nursing.

18. Lin, K., et al. (2024). Metabolic Bariatric Surgery n the Era of GLP-1 Receptor Agonists for Obesity Management. JAMA Network Open. doi.org/10.1001/jamanetworkopen.2024.41380.

19. Glasser, N. J., et al. (2024). Male Gender Expressivity and Diagnosis and Treatment of Cardiovascular Disease Risks in Men. JAMA Network Open. doi.org/10.1001/jamanetworkopen.2024.41281.

20. Cardiovascular Diseases (CVDs) [Jan; 2023]. 2021. http://www.who.int/mediacentre/factsheets/fs317/en/ http://www.who.int/mediacentre/factsheets/fs317/en/

21. Wei, M.-G., et al. (2024). Taurine Reduces Atherosclerotic Plaque Area and Stability in Mice. Journal of Exploratory Research in Pharmacology. doi.org/10.14218/jerp.2024.00017.

22. Ramadoss, S., et al. (2024). Bone-marrow macrophage-derived GPNMB protein binds to orphan receptor GPR39 and plays a critical role in cardiac repair. Nature Cardiovascular Research. doi.org/10.1038/s44161-024-00555-4.

23. The ClinGen Consortium., (2024). The Clinical Genome Resource (ClinGen): Advancing genomic knowledge through global curation. Genetics in Medicine. doi.org/10.1016/j.gim.2024.101228.

24. Li. S., et al. (2024) A humanized monoclonal antibody targeting an ectonucleotidase rescues cardiac metabolism and heart function after myocardial infarction. Cell Reports Medicine. doi.org/10.1016/j.xcrm.2024.101795.

25. Amrute, J. M, et al. (2024). Targeting immune–fibroblast cell communication in heart failure. Nature. doi.org/10.1038/s41586-024-08008-5.

26. The Top 10 Causes of Death. [Jan; 2023]. 2020. https://www.who.int/news-room/factsheets/detail/the-top-10-causes-of-death https://www.who.int/news-room/fact-sheets/detail/thetop-10-causes-of-death

27. Knowledge of the causes of heart attack among survivors, and implications for health promotion. Greenwood DC, Packham CJ, Muir KR, Madeley RJ 2019.

28. Source: Kagglehttps://share.google/ZBlRPjo18NsJfrL86