

# RELATIONSHIP BETWEEN ANNUAL INCOME AND LOAN BORROWED

Submitted By: -  
Nidhaan Singh Gandhi  
NMIMS NIVERSITY  
FY BBA



## ABSTRACT

This paper studies the increasing trend of people to outsource finance for upcoming projects. People in various occupations and backgrounds apply for loans. These loans may be Personal, Business, Home or even Educational. People apply for loans as per their Annual Income. The Interest rates cause a person to payback more than the actual amount borrowed. So, every individual tries to acquire finance at the lowest rate and also the safest. Similarly, lending Institutions try to charge high amounts of interest to earn good profit. The study is based on Loans Data collected from Lending Club.

## INTRODUCTION

The following data, shows us the loans granted by Lending Club, an online lending website, to people of various occupations, at various interest rates. As the dataset has people with various occupations, the Annual Income of the people also vary.

Here, the dependent variable is the Loan Amount and the dependent variables are Annual Income and Interest Rate charged.

**OBJECTIVES**

- To understand the relationship between Annual Income and Loan Amount borrowed.
- To understand how Loan Amount borrowed is affected by Interest Rate charged by banks.

**SUMMARY**

The data used has been taken from an online data website named Kaggle. This research will help various competing banks, associations and online webpages to study this relationship and in turn alter their offerings to establish a competitive market. This research will also help the General Audience to estimate the Loan Amount and Interest Rates they are offered by the Club according to their Annual Income bracket. Also, they can compare the prevailing interest rates offered by various banks in the market and select the most feasible offer.

**ACTUAL DATA**

	A	B	C	D	E	F	G	H	I	J	K
1	Occupation	Annual Income	Loan Amount	Interest Rate							
2	Chef	55000	2500	13.56							
3	Postmaster	90000	30000	18.94							
4	Administrative	59280	5000	17.97							
5	IT Supervisor	92000	4000	18.94							
6	Mechanic	57250	30000	16.14							
7	Director COE	152500	5550	15.02							
8	Account Manager	51000	2000	17.97							
9	Assistant Director	65000	6000	13.56							
10	Legal Assistant III	53580	5000	17.97							
11	Consultant	70000	28000	11.31							
12	Job Coach Supervisor	65000	11200	8.19							
13	Quality Field Engineer	154000	6500	17.97							
14	Teller	65000	22000	12.98							
15	respiratory therapist	80000	3500	16.14							
16	Worship Director	102500	7000	12.98							
17	Processor	23878	25000	16.91							
18	Neonatal Nurse Practitioner	120000	16000	20.89							
19	Stationary Engineer	75000	13000	14.47							
20	Exhibits director	55000	13000	14.47							
21	driver coordinator	65000	9600	23.4							
22	gas attendant	40000	3500	20.89							
23	Financial Relationship Associate	33000	16000	26.11							
24	Sale Representative	90000	13000	23.4							
25	Operator	68107	23000	20.89							
26	Manager	43000	8000	23.4							
27	Nursing Supervisor	150000	32075	11.8							
28	Material Handler	80000	10000	19.92							
29	Instructional Coordinator	51000	16000	17.97							
30	Program Manager	49700	3300	23.4							
31	Budget & Procurement Specialist	33190	3500	20.89							
32	Foreman shop	74000	10000	20.89							
33	Substance Abuse Counselor	40000	6000	6.46							
34	Business Intelligence Engineer	171000	25000	27.27							
35	FC	74800	20000	12.98							
36	Information technology	53000	25000	17.97							
37	BANKER	52000	1250	23.4							

**DESCRIPTIVE STATISTICS**

Annual Income	Loan Amount	Interest Rate	Annual Income
₹10,000.00	₹1,500.00	13.50	Mean
₹10,000.00	₹1,500.00	13.50	Standard Error
₹10,000.00	₹1,500.00	13.50	Median
₹10,000.00	₹1,500.00	13.50	Mode
₹10,000.00	₹1,500.00	13.50	Standard Deviation
₹10,000.00	₹1,500.00	13.50	Range
₹10,000.00	₹1,500.00	13.50	Minimum
₹10,000.00	₹1,500.00	13.50	Maximum
₹10,000.00	₹1,500.00	13.50	Sum
₹10,000.00	₹1,500.00	13.50	Count
₹10,000.00	₹1,500.00	13.50	Largest(1)
₹10,000.00	₹1,500.00	13.50	Smallest(1)
₹10,000.00	₹1,500.00	13.50	Confidence Level(95.0%)

**ANNUAL INCOME**

On an average the Annual Income of the people taking loan is ₹10,000.  
 Half of the Annual Income of the people taking loan is above or below ₹10,000.  
 Most people with the Annual Income ₹10,000 take loans.  
 The Annual Income of all people are on an average of ₹10,000.00150 away from the mean, that is, ₹10,000.  
 Deviation between the maximum and minimum Annual Income is ₹10,000.  
 The people of a minimum Annual Income ₹1,500 have taken a loan.  
 The Annual Income of the dataset is Positively Skewed, that is, the right tail of the distribution is longer than the left.  
 The Annual Income has Positive Kurtosis, that is, the distribution has darker tails and taller peaks.

Loan Amount
Mean
Standard Error
Median
Mode
Standard Deviation
Range
Minimum
Maximum
Sum
Count
Largest(1)
Smallest(1)
Confidence Level(95.0%)

**LOAN AMOUNT**

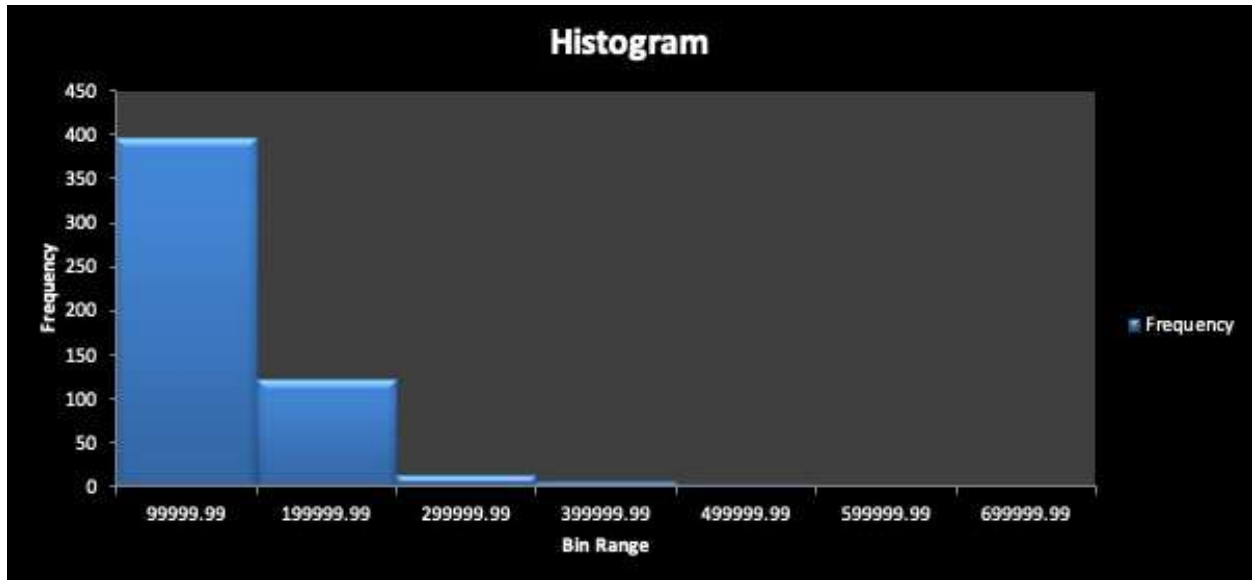
The average Loan Amount given is ₹10,000.  
 Half of the Loan Amount given is above or below ₹10,000.  
 Most minimum Loan Amount given is ₹10,000.  
 All of the Loan Amounts given are on an average ₹10,000 away from the mean, that is, ₹10,000.  
 Deviation between the maximum and minimum Loan Amount given is ₹10,000.  
 The Minimum Loan Amount given is ₹1,000.  
 The Maximum Loan Amount given is ₹10,000.  
 The Loan Amount is Positively Skewed, that is, the right tail of the distribution is longer than the left.  
 The Loan Amount has Negative Kurtosis, that is, the distribution has lighter tails and flatter peaks.

**HISTOGRAM OF ANNUAL INCOME**

Class Interval	Bin Range	Bin Range	Frequency
0-100000	99999.99	99999.99	395
100000-200000	199999.99	199999.99	121
200000-300000	299999.99	299999.99	13
300000-400000	399999.99	399999.99	5
400000-500000	499999.99	499999.99	2
500000-600000	599999.99	599999.99	0
600000-700000	699999.99	699999.99	0
700000-800000	799999.99	799999.99	1
		More	0

**HISTOGRAM**

It is a diagram consisting of rectangles whose area is proportional to the frequency of a variable and whose width is equal to the class interval. In the data, the maximum people with Annual Income between 0 to ₹1,00,000 and minimum people with Annual Income between ₹5,00,000 to ₹7,00,000.

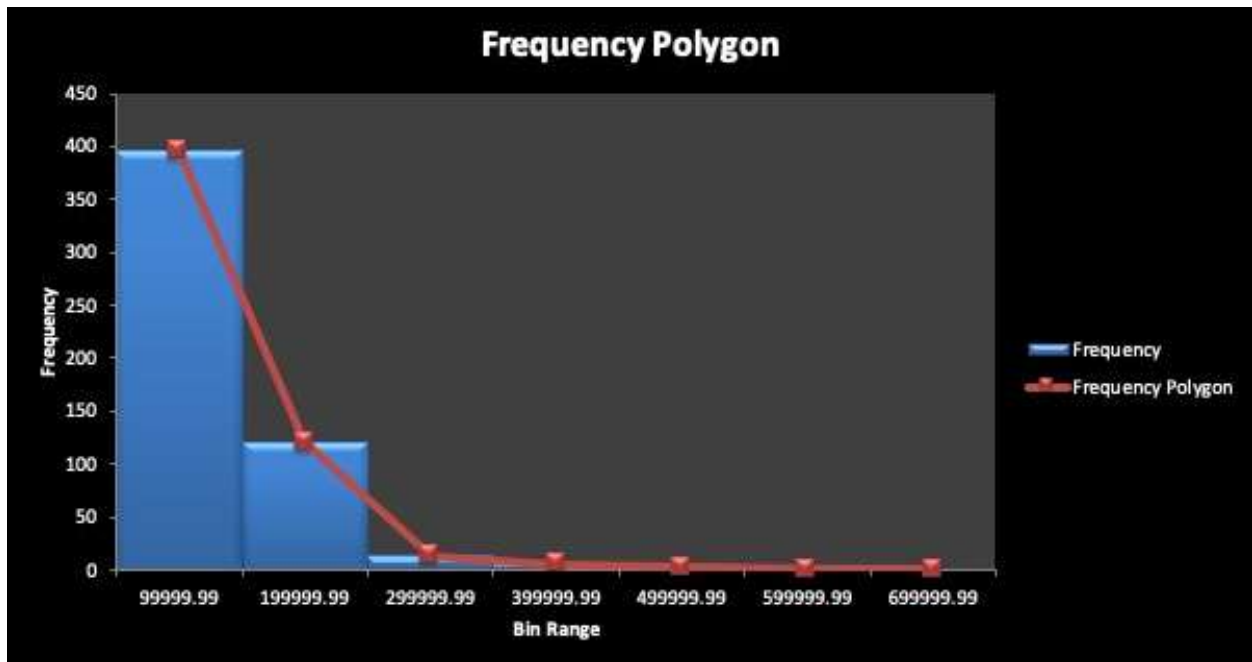


**FREQUENCY POLYGON OF ANNUAL INCOME**

Class Interval	Bin Range	Bin Range	Frequency
0-100000	99999.99	99999.99	395
100000-200000	199999.99	199999.99	121
200000-300000	299999.99	299999.99	13
300000-400000	399999.99	399999.99	5
400000-500000	499999.99	499999.99	2
500000-600000	599999.99	599999.99	0
600000-700000	699999.99	699999.99	0
700000-800000	799999.99	799999.99	1
		More	0

**FREQUENCY POLYGON**

It is a graph constructed by using lines to join the midpoints of each interval, or bin. The heights of the points represent the frequencies. In the data, the average people taking a loan are of a maximum Annual Income of ₹99,999.99 and Annual Income of ₹59,999.99.

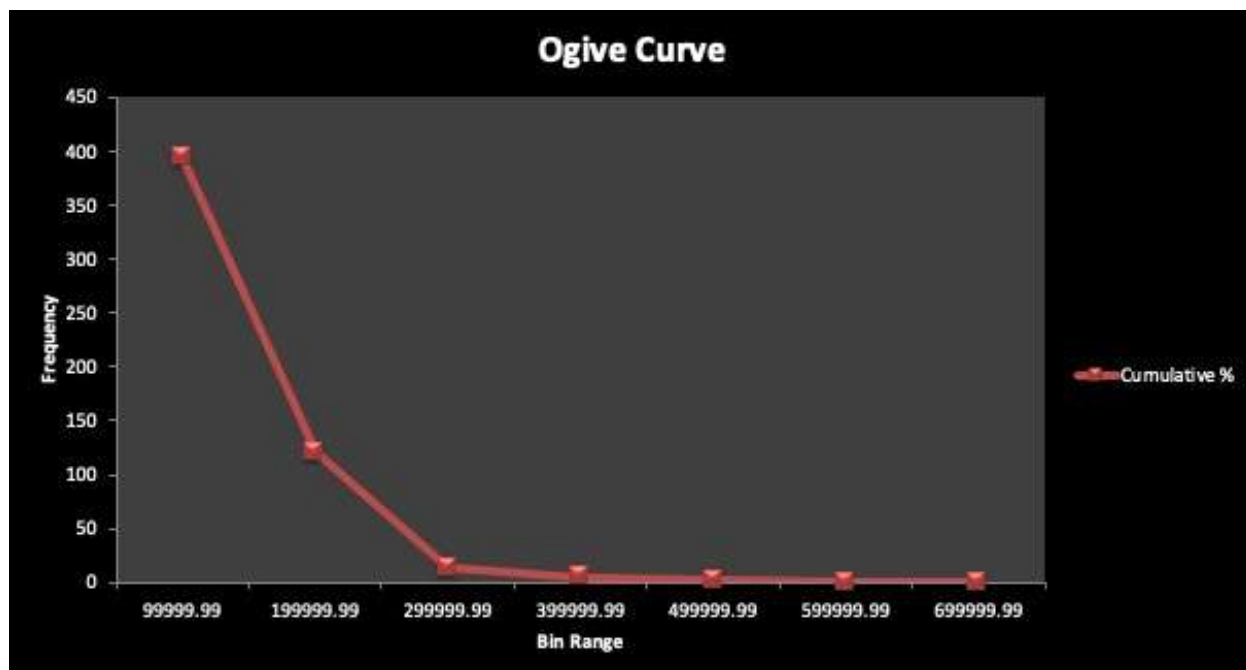


### OGIVE CURVE OF ANNUAL INCOME

Class Interval	Bin Range	Bin Range	Frequency
0-100000	99999.99	99999.99	395
100000-200000	199999.99	199999.99	121
200000-300000	299999.99	299999.99	13
300000-400000	399999.99	399999.99	5
400000-500000	499999.99	499999.99	2
500000-600000	599999.99	599999.99	0
600000-700000	699999.99	699999.99	0
700000-800000	799999.99	799999.99	1
		More	0

#### OGIVE CURVE

It is a graph showing the curve of a cumulative distribution function. The points plotted are the upper class limit and the corresponding cumulative frequency.

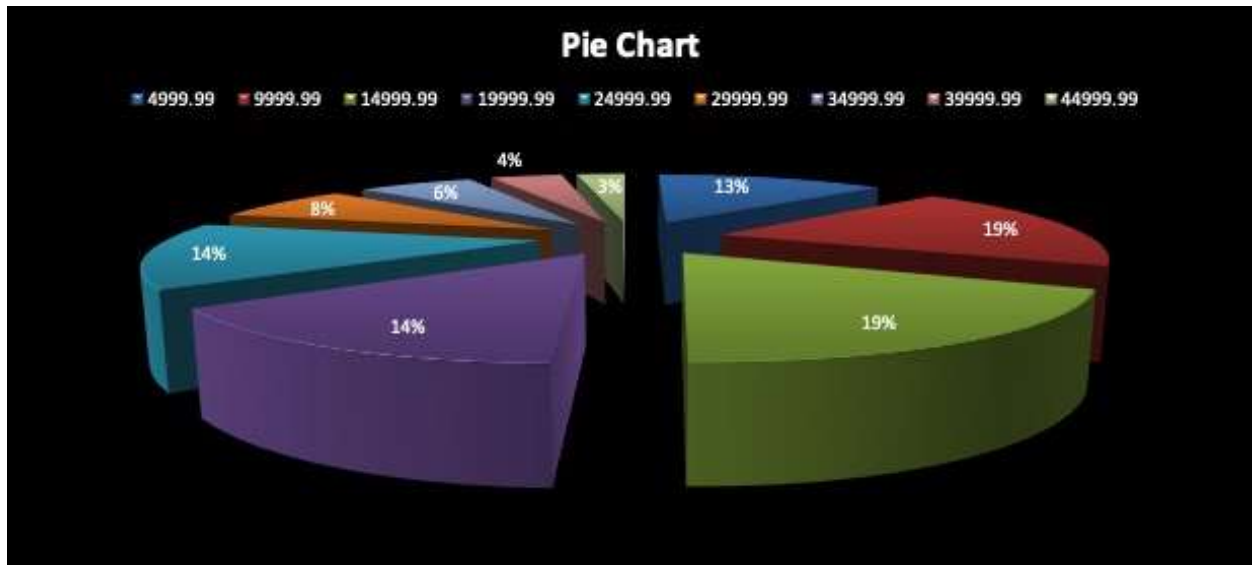


### PIE CHART OF LOAN AMOUNT

Class Interval	Bin Range	Bin Range	Frequency
0-5000	4999.99	4999.99	70
5000-10000	9999.99	9999.99	101
10000-15000	14999.99	14999.99	100
15000-20000	19999.99	19999.99	78
20000-25000	24999.99	24999.99	75
25000-30000	29999.99	29999.99	41
30000-35000	34999.99	34999.99	35
35000-40000	39999.99	39999.99	22
40000-45000	44999.99	44999.99	15
		More	0

#### PIE CHART

It is a type of graph in which a circle is divided into sectors that each represent a proportion of the whole. In the data, maximum people have taken a Loan Amounting between ₹5,000 to ₹10,000 and minimum people have taken a Loan Amounting between ₹40,000 to ₹45,000.

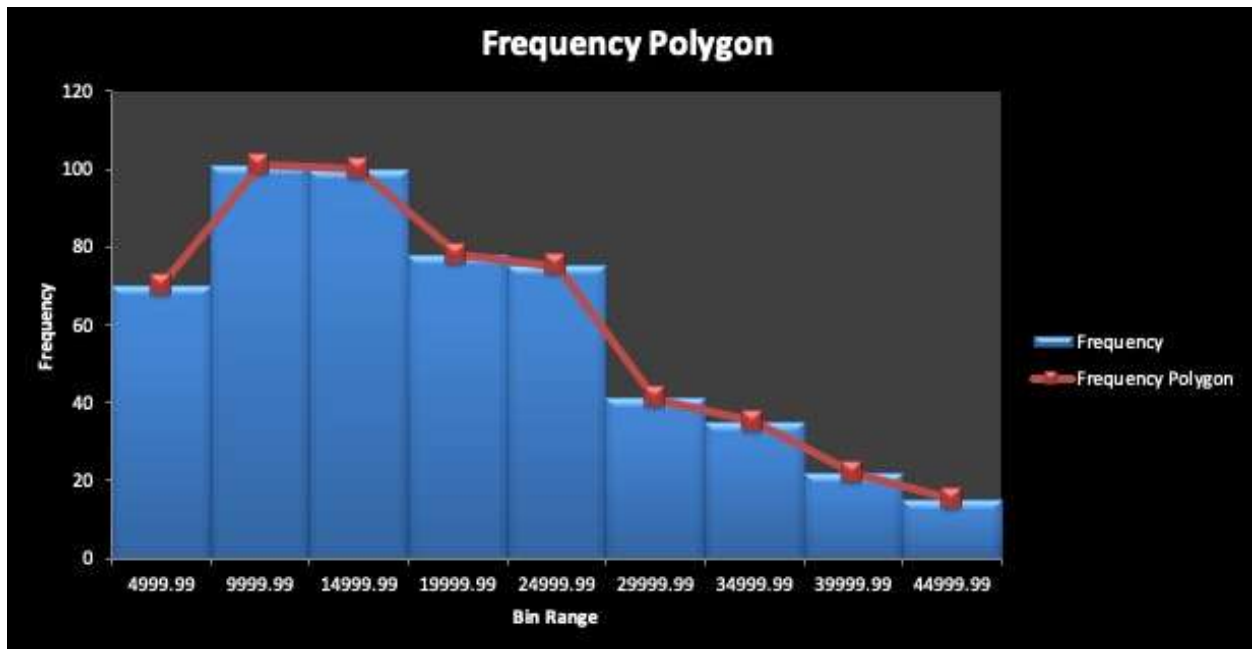


### FREQUENCY POLYGON OF LOAN AMOUNT

Class Interval	Bin Range	Bin Range	Frequency
0-5000	4999.99	4999.99	70
5000-10000	9999.99	9999.99	101
10000-15000	14999.99	14999.99	100
15000-20000	19999.99	19999.99	78
20000-25000	24999.99	24999.99	75
25000-30000	29999.99	29999.99	41
30000-35000	34999.99	34999.99	35
35000-40000	39999.99	39999.99	22
40000-45000	44999.99	44999.99	15
		More	0

#### FREQUENCY POLYGON

It is a graph constructed by using lines to join the midpoints of each interval, or bin. The heights of the points represent the frequencies. In the data, the average maximum people have taken Loan Amounting ₹9999.99 and minimum people have taken Loan Amounting ₹44999.99.



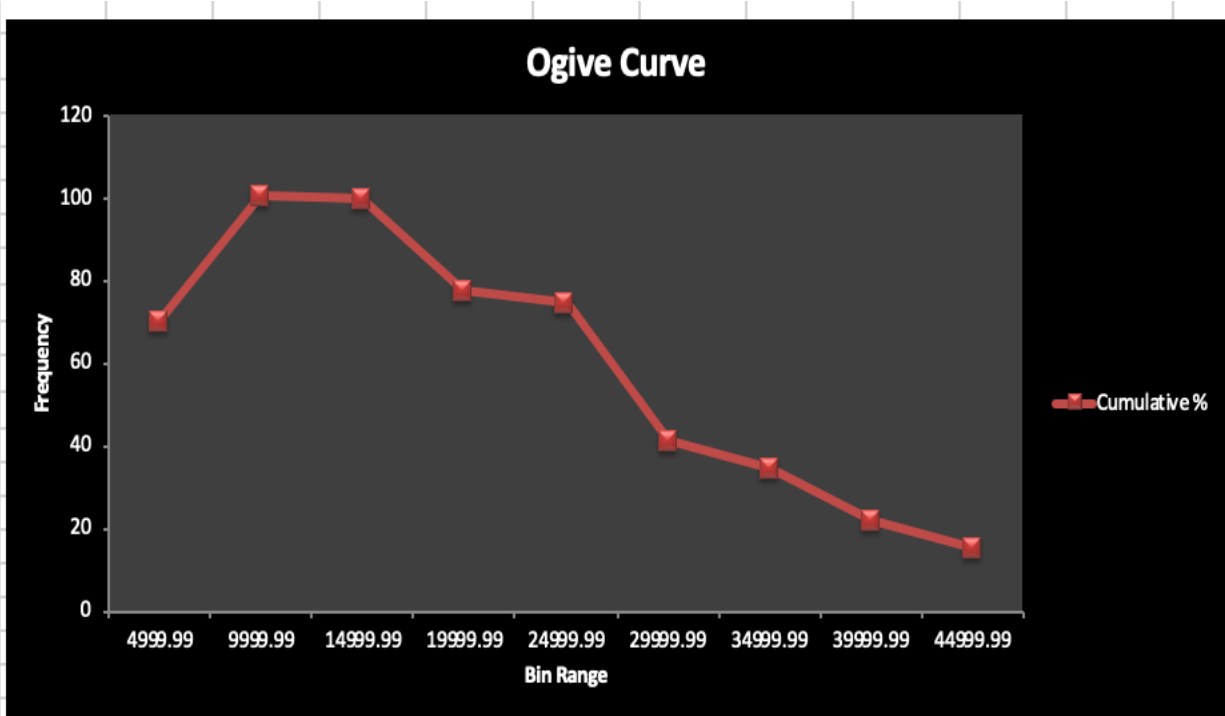
**OGIVE CURVE FOR LOAN AMOUNT**

Class Interval	Bin Range	Bin Range	Frequency
0-5000	4999.99	4999.99	70
5000-10000	9999.99	9999.99	101
10000-15000	14999.99	14999.99	100
15000-20000	19999.99	19999.99	78
20000-25000	24999.99	24999.99	75
25000-30000	29999.99	29999.99	41
30000-35000	34999.99	34999.99	35
35000-40000	39999.99	39999.99	22
40000-45000	44999.99	44999.99	15
		More	0

**OGIVE CURVE**

It is a graph showing the curve of a cumulative distribution function. The points plotted are the upper class limit and the corresponding cumulative frequency.





**SINGLE CORRELATION**

	<i>Annual Income</i>	<i>Loan Amount</i>
<i>Annual Income</i>	1	
<i>Loan Amount</i>	0.30272284	1

**CORRELATION**

It is the measure of the extent to which two or more variables fluctuate together. There exists an Imperfect Positive Correlation between Annual Income and the Loan Amount borrowed. The coefficient of correlation is 0.302722840242261.

If a person's Annual Income increases, then he will borrow more money but less than proportionately.

**MULTIPLE CORRELATION**

	Annual Income	Loan Amount	Interest Rate
Annual Income	1		
Loan Amount	0.30272284	1	
Interest Rate	-0.004779682	0.013594334	1

**MULTIPLE CORRELATION**

It is the measure of the extent to which two or more variables fluctuate together. There exists an Imperfect Positive Correlation between Annual Income and the Loan Amount borrowed and Loan Amount borrowed and Interest Rate charged. The coefficient of correlation is 0.302722840242261 and 0.0135943342133409 respectively.

If a person's Annual Income increases, then he will borrow more money but less than proportionately and as a person's Loan Amount borrowed increases, the Interest Rate charged also increases but less than proportionately.

There also exists an Imperfect Negative Correlation between Annual Income and Interest Rate charged. The coefficient of correlation is -0.0047796824709014. If a person's Annual Income increases, the Interest Rate charged also decreases, but less than proportionately.

**SINGLE REGRESSION**

SUMMARY OUTPUT									
<i>Regression Statistics</i>									
Multiple R	0.30272284								
R Square	0.091641118								
Adjusted R Square	0.089943251								
Standard Error	9630.506427								
Observations	537								
<b>ANOVA</b>									
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>				
Regression	1	5005931985	5005931985	53.97425963	7.63326E-13				
Residual	535	49619459916	92746654.05						
Total	536	54625391902							
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95.0%</i>	<i>Upper 95.0%</i>	
Intercept	11374.85179	719.9480962	15.79954422	1.98133E-46	9960.579976	12789.1236	9960.579976	12789.1236	
Annual Income	0.050238998	0.006838292	7.346717609	7.63326E-13	0.036805803	0.063672192	0.036805803	0.063672192	
<b>Regression Equation</b>									
<b>Y= A+BX</b>									
Y= 11374.85 + 0.05X									
<b>REGRESSION</b>									
It determines the strength of a dependent variable on an independent variable.									
Y is the Dependent Variable, that is, Loan Amount borrowed.									
X is the Independent Variable, that is, Annual Income of the person.									
If the dependent variable, that is Loan Amount Borrowed changes by ₹1, it is because of an increase in 0.05 units of the Independent Variable, that is, Annual Income of the person.									

**MULTIPLE REGRESSION**

SUMMARY OUTPUT								
<b>Regression Statistics</b>								
Multiple R	0.303096292							
R Square	0.091867362							
Adjusted R Square	0.088466117							
Standard Error	9638.319008							
Observations	537							
<b>ANOVA</b>								
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
Regression	2	5018290678	2509145339	27.00991548	6.69664E-12			
Residual	534	49607101224	92897193.3					
Total	536	54625391902						
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95.0%</i>	<i>Upper 95.0%</i>
Intercept	10945.38113	1380.432424	7.928951063	1.30023E-14	8233.637103	13657.12516	8233.637103	13657.12516
Annual Income	0.050250929	0.006843917	7.342422039	7.87747E-13	0.036806626	0.063695232	0.036806626	0.063695232
Interest Rate	31.24981932	85.6766623	0.364741325	0.715448907	-137.0548173	199.554456	-137.0548173	199.554456
<b>Regression Equation</b>								
$Y = A + BX_1 + CX_2$								
$Y = 11374.85 + 0.05X_1 + 31.25X_2$								

**MULTIPLE REGRESSION**

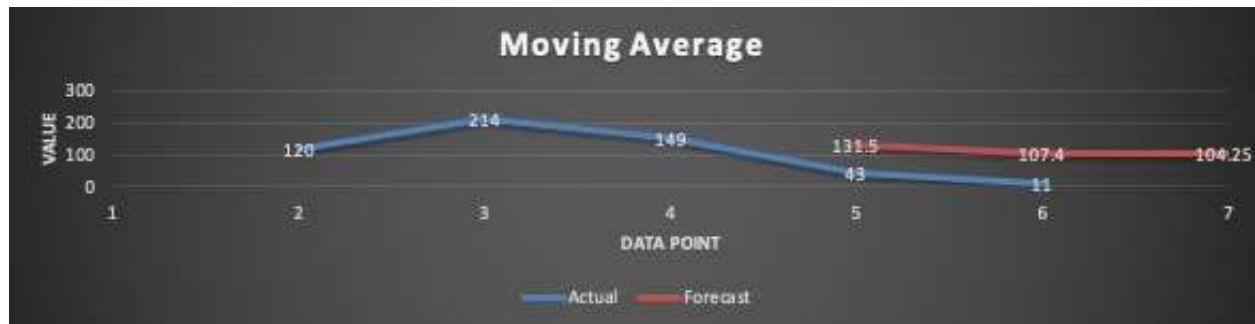
It determines the strength of a dependent variable on more than one independent variables.  
 Y is the Dependent Variable, that is, Loan Amount borrowed.  
 Xn are the Independent Variables, that is, Annual Income of the person(X1), Interest Rate charged (X2).  
 If the dependent variable, that is Loan Ampount Borrowed changes by ₹1, it is because of an increase in 0.05 units of the Independent Variable X1, that is, Annual Income of the person and an increase in 31.24% of the Independant Variable X2, that is, Rate of interest charged.

**MOVING AVERAGE**

Row Labels	Count of Interest Rate	Count of Interest Rate	#N/A
<0 or (blank)			#N/A
5-10	120	120	#N/A
10-15	214	214	#N/A
15-20	149	149	131.5
20-25	43	43	107.4
25-30	11	11	104.25
<b>Grand Total</b>	<b>537</b>		

**Moving Average**

As the interest rates increases, people tend to borrow less.



### ANALYSIS

If a person's Annual Income increases, then he will borrow more money but less than proportionately.

If a person's Annual Income increases, then he will borrow more money but less than proportionately and as a person's Loan Amount borrowed increases, the Interest Rate charged also increases but less than proportionately.

If a person's Annual Income Increases, the Interest Rate charged also decreases, but less than proportionately.

If the dependent variable, that is Loan Amount Borrowed changes by ₹1, it is because of an increase in 0.05 units of the Independent Variable, that is, Annual Income of the person.

If the dependent variable, that is Loan Amount Borrowed changes by ₹1, it is because of an increase in 0.05 units of the Independent Variable X1, that is, Annual Income of the person and an increase in 31.24% of the Independent Variable X2, that is, Rate of Interest charged.

As the interest rates increases, people tend to borrow less.

### CONCLUSION

From the study and analysis, we conclude that:

Amount of Loan Borrowed is affected by Annual Income.

As Annual Income increases, the Interest Rate doesn't affect the Loan Amount borrowed.

As Interest Rate increases, people tend to borrow less.