International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

A Case Study on Bank Loan by using MS Excel

Pooja Kumawat [1], Mayuree Katara[2]

[1,2]**Department of Computer Science & Engineering Vivekananda institute of technology Jaipur**
pooja.k@vitj.ac.in[1], **Katara.mayuree@vitj.ac.in**[2]

ABSTRACT

Data is being produced very quickly due to increase in information in everyday life. Vast volume of data generated from various organization. That is very difficult to analyze .Data created by different organizations such as healthcare database, banks, sales data, government databases etc. are example of large data processing. Analyzing, Processing and communicating such large amount of data are a challenge. It is very difficult task for analyzing and visualizing such large amount of data. Therefore some tools are required for analyzing such large amount of data. This paper focuses on data analysis and visualizing techniques. This paper shows a case study of bank loan.

Index Terms— Analysis, EDA, visualization, MS Excel

1. Introduction:

We are all always thinking about the future and what is expected to happen in the coming weeks, months, and even years, and in order to do so, a look into the past is required. Businesses must be able to see their development and the factors influencing their sales [1]. In this technological era of large-scale data, organizations must reconsider current techniques to better understand clients in order to achieve a competitive advantage in the market. Data is useless if it cannot be analyzed, comprehended, and applied in context [2].The primary goal here is to read and analyze the available datasets.

I.A. WHAT IS DATA ANALYSIS?

Data analysis is a process that involves creating dataset, studying it, cleaning it and by removing null values, outliers and converting it to generate a useful result. By analyzing previous data helps you to make better decision in the future.

I.B. Data visualization

Data visualization is a process which aims to communicate data effectively and clearly to the user through graphical representation. Effective and efficient data visualization is the key part of the discovery process. It is the intermediate between the human intuition and quantitative context of the data, thus an essential component of the scientific path from data into knowledge and understanding. It is a powerful new technology having a great potential to help researchers as well as companies for building revenue decision [3].

I. C. Visualization Tools:

Data visualization tools organize and present information naturally. For simple visualization of data MS excel can be used many formats like bubble chart, line plots, histograms, tree map can be created. For plotting and graphing Matplotlib was used which has many packages for statistics, clustering and plotting. For creating graphics, processing was used widely. It allows uploading a data set and exploring it with pre build techniques. Spotfire is the best web client and analytical functionality tool. Qliktech may be the best visualization product having interactive drill-down capabilities. Tableau and Power bi has the excellent skill to interact with OLAP cubes. [18].

I.D Data Visualization Techniques

Different data visualization techniques are used for data visualization.

- Box plots
- Histograms
- Heat maps

- Charts
- Tree maps
- Word Cloud/Network diagram

Box Plots

Box plot is a graph that shows how the values in your data are spread. It is a standard way presenting data based on this five attributes ("minimum", first quartile(Q1),median, third quartile(Q3) and "maximum).It can tell you about your outliers and their values. It can also show your data if it is symmetrical, how tightly your data is grouped and how your data is skewed.

Summary of Box Plot

Minimum	Q1 -1.5*IQR
First quartile (Q1/25th Percentile)":	The middle number between the smallest number (not the "minimum") and the median of the dataset
Median (Q2/50th Percentile)":	The middle value of the dataset
Third quartile (Q3/75th Percentile)":	The middle value between the median and the highest value (not the "maximum") of the dataset.
Maximum"	Q3 + 1.5*IQR
interquartile range (IQR)	25th to the 75th percentile.

Histograms

A histogram is a graphical representation of data that uses bars of various heights. Each bar in a histogram divides numbers into ranges. More data falls inside that range, as shown by taller bars. The form and distribution of continuous sample data are shown in a histogram.

It is a plot that enables you to identify and display the frequency distribution (shape) underneath a set of continuous data. This enables the examination of the data for outliers, skewness, and other factors such as outliers and underlying distribution (such as the normal distribution). It relates just one variable and accurately depicts the distribution of numerical data. Count the number of values that fall within each interval, which is achieved by dividing the whole range of values into bins or buckets.

Line Chart

Line plot is used to plot the connection or dependence of one variable on another. To plot the relationship between the two variables, we can simply use the plot function.

Bar Charts

Bar charts are used for comparing different categories or groups. Values of categories can be showed with vertical or horizontal bars. Each bar represents the value length and height.

Pie Chart

This statistical chart is circular and uses slices to show numerical proportion. Each slide's arc length is inversely correlated with the quantity it represents in this diagram. Typically used to compare the components of a whole, they work best with few components and when text and percentages are utilized to explain the contents. However, because the human eye struggles to estimate areas and contrast visual angles, they can be challenging to comprehend.

Scatter Charts

A scatter plot, which is a two-dimensional graphic that represents the joint variation of two data items, is another popular visualization tool. Each marker (such as a dot, square, or plus sign) denotes a particular observation. The value for each observation is shown by the marker location. A scatter plot matrix, which is a set of scatter plots depicting every possible pairing of the measurements that are assigned to the visualization, is created when you assign more than two measures. Scatter plots are used to analyze the relationship between the variable X and Y.

Bubble Charts

In this visualization technique data points are replaced with bubbles and dimension of data is represented in the size of bubbles.

Timeline Charts

Timeline charts show events in chronological sequence, in whichever unit of time the data was collected, such as week, month, year, or quarter, for example the development of a project, an advertising campaign, or the acquisition process. On a timeframe, it displays the order in which past or future events occurred.

Tree Map

For hierarchical data, a tree map is a sophisticated, area-based data visualization that can be challenging to exactly analyze. Bar charts and other less complex visualizations are frequently used.

II. LITERATURE REVIEW

The term visualization is an evolving study area, where many researchers have contributed from the last few decades. Various authors have proposed different techniques and technologies to support data visualization. This section elaborates about how the flow of research has been carried out by the authors and researchers from reputed journals and conferences.

In [4] the author has proposed a Sensor: Network based approach for storing, sharing, visualizing and analyzing data from multiple devices and to interact with each other and with the end user through an open REST- based API. The author has visualized the geographical location of the data stream which when clicked pops up a tabbed window containing different associated information.

In [5] the author has proposed a virtual reality platform for scientific data visualization, a tool for multi-dimensional data visualization using which scientist can interact with the data and their colleagues in the same space. The author has mapped data parameters in different data points, shapes, size, colors, XYZ axis and many more. The author has used iViz a visualization tool which can be run as a standalone application or in a web browser. The author has discussed about a framework of financial time series delivery and visualization which can be used in viewing the historical price movement of a stock [6]. Specialized binary tree (SB- tree) is used for representing the financial time series. Time series data server, SB-tree server and web service contains is the three major components which are distributed on different machines. The system can reduce data volume and can capture the critical points.

In [7] the author has proposed a dashboard for displaying data used for communicating and finding trends in laboratory operation. System is based on .NET scripts, SQL repository. The author depicts that data is collected from the multiple sources like admin, internet and user portal and are stored in database using XML layer, Adobe flash, Action Script, etc. Data is being visualized which is used for laboratory and staff management.

In [8] the author has used a concept of visual web mining for analyzing the web data. A tool named WET is been used for visualization which provides a set of visual metaphor that represents the structure of the websites. The websites exploration tool is used for exploring the websites and for giving the feedback to the website owner for the betterment of the website.

In [9] the author has used a concept for analyzing data for examining the trend and evaluating the eco-environment impact of three gorges project. VC.NET and ArcIMs is the development platform for information system. ArcSDE and oracle 10g are used for management and use of spatial data. The author introduces method and processing and storing the data generated from cross-region, cross-department. Visualization helps in enhancing the data analysis and data mining.

In [10] the author has discussed the problem in compliance management which becomes an obstacle for decision making for effective and efficient monitoring. The person should be provided with compliance software which will help in getting high level information about overall compliance status and low level problem regarding possible problems. The author has designed a dashboard for watching the compliance which avoids the obstacle and decision can be made effectively. In [11] the author has introduced a tool named SECONDA which is used for analyzing both individual and grouped evolution of projects and develops belonging to a software ecosystem, Visualization is implemented in java using JFREECHART libraries. The author has used GNOME ecosystem for studying, under SECONDA. It offers a dashboard for fast visual analysis of local and global matrixes that can be extracted from information stored in the repositories.

In [12] the author has proposed a system for monitoring the user exercising progress and presenting exercise parameters in relation to prescribed targets. This system can be used for monitoring the intensity of the levels recommended by the patients care provider. It uses a miniature wireless 3-axis acceleration tied on the wrist of the patient that transmits acceleration data. The dashboard allows graphical visualization of exercise progress in real time. The author introduces a system where the huge amount of data generated from the collaborative software development tool during the lifecycle of a project can be used to analyze the performance of the individual member, or a team or manager.

[13] They can analyze from different perspectives across different dimensions and visualized in different ways.

In [14] the author has proposed a dashboard which is an integration, validation and visualizing tool for natural language processing. The system helps the system integration team to integrate and validate the system; developers to profile each module and researchers to evaluate and compare the module with the earlier versions. It also supports execution of modules on heterogeneous platform with an easy to use graphical interface developed using eclipse RCP.

III. Proposed Methodology

The primary goals of data visualizations are to analyze the data and communicate it to the user. Visualization's primary objective is to use graphics to connect information in a clear and effective way. We are putting forth a system that will examine and display sales data. The data will be graphed using various parameters to reflect various viewpoints. To find trends for future forecasts, data mining will be used. For analysis and visualizations, a data set from kaggle is used. We have taken the data set of bank loan application

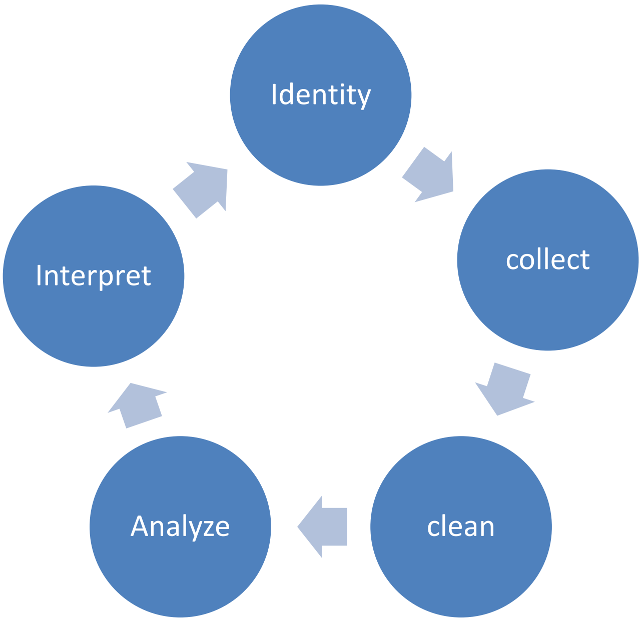


Fig. III. a Data analysis process

III. A In this analysis following questions are solved.

- Present the overall approach of the analysis. Mention the problem statement and the analysis approach briefly
- Identify the missing data and use appropriate method to deal with it. (Remove columns/or replace it with an appropriate value)
- Identify if there are outliers in the dataset. Also, mention why you think it is an outlier. Again, remember that for this exercise, it is not necessary to remove any data points.
- Identify if there is data imbalance in the data. Find the ratio of data imbalance.
- Explain the results of univariate**,** segmented univariate**,** bivariate analysis, etc**.** in business terms.
- Find the top 10 correlation for the Client with payment difficulties and all other cases (Target variable). Note that you have to find the top correlation by segmenting the data frame w.r.t to the target variable and then find the top correlation for each of the segmented data and find if any insight is there. Say, there are 5+1(target) variables in a dataset: Var1, Var2, Var3, Var4, Var5, and Target. And if you have to find top 3 correlation, it can be: Var1 & Var2, Var2 & Var3, Var1 & Var3. Target variable will not feature in this correlation as it is a categorical variable and not a continuous variable which is increasing or decreasing.
- Include visualizations and summarize the most important results in the presentation.

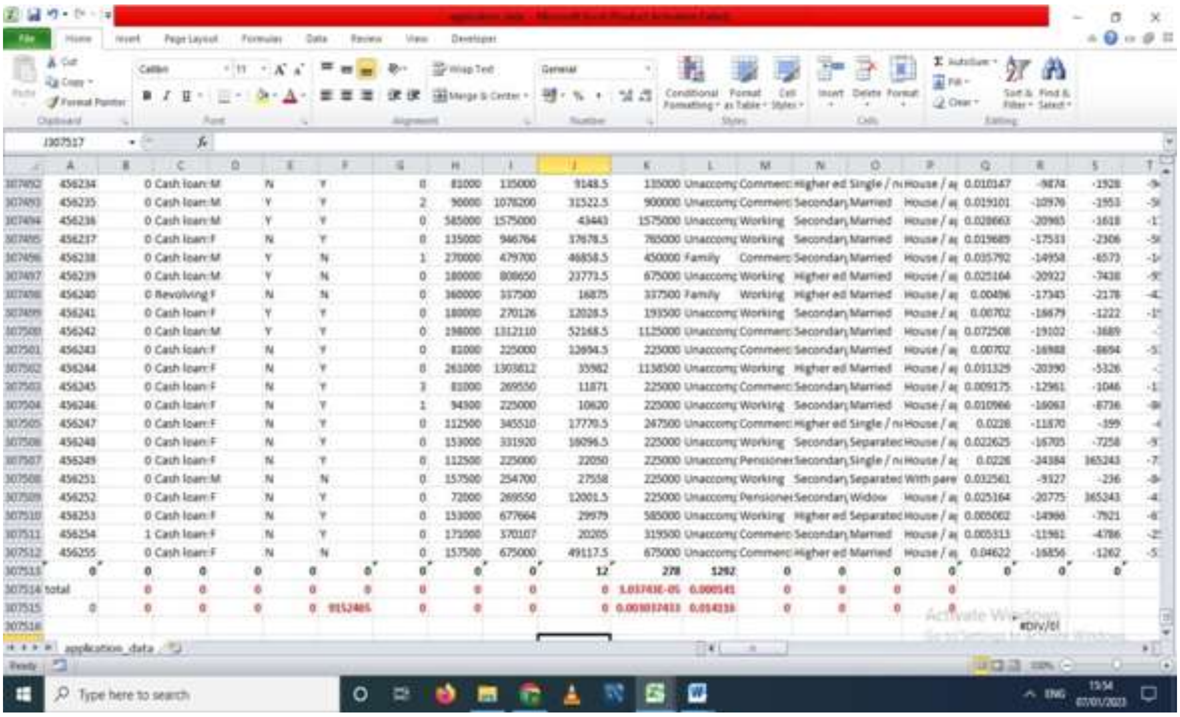
IV. Tech Stack Used

In order to execute this project Microsoft Excel is used. Once the data record is downloaded. After that various formulas and graphs are used to represent the data.

V. Results

We can count the no of blank cell by using

Count blank (A3:A307512)

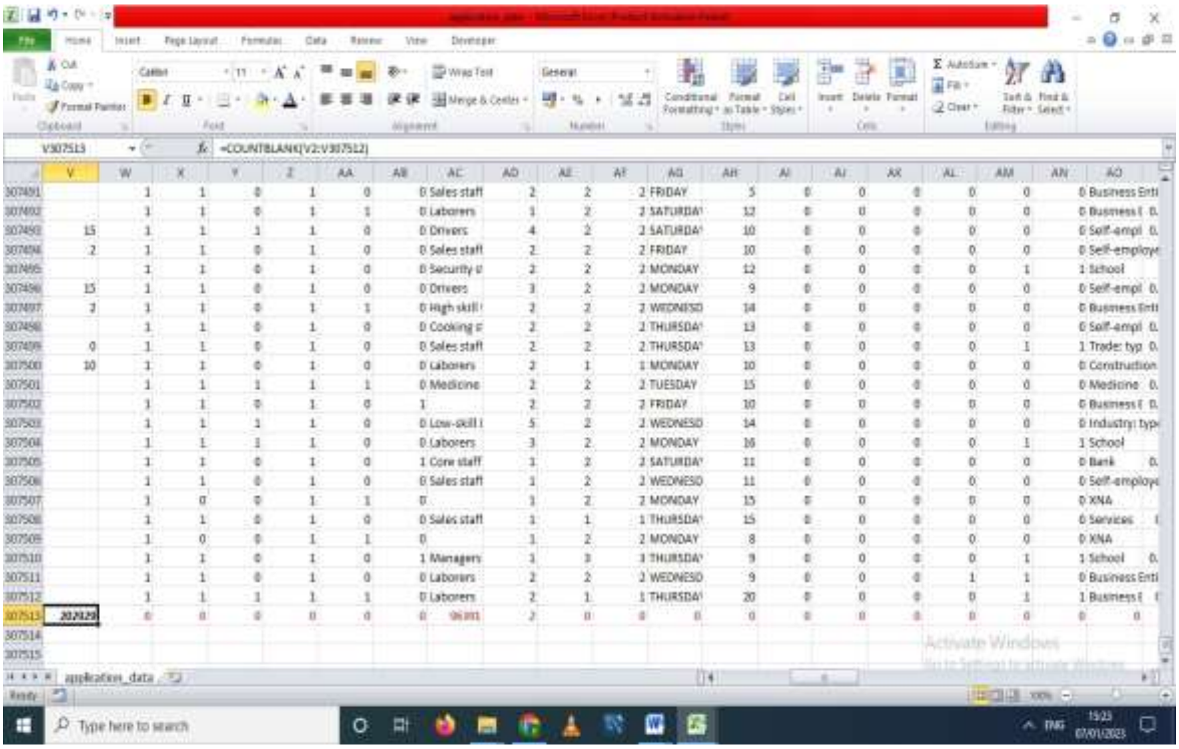


J307517

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
307492	456234	0	Cash loan	M	N	Y	0	81000	135000	9148.5	135000	Unaccomp	Commerc	Higher ed	Single / n	House / a	0.010147	-9874	-1928	-9
307493	456235	0	Cash loan	M	Y	Y	2	90000	1078200	31522.5	900000	Unaccomp	Commerc	Secondary	Married	House / a	0.019101	-10976	-1953	-5
307494	456236	0	Cash loan	M	Y	Y	0	585000	1575000	43443	1575000	Unaccomp	Working	Secondary	Married	House / a	0.028663	-20985	-1618	-1
307495	456237	0	Cash loan	F	N	Y	0	135000	946764	37678.5	765000	Unaccomp	Working	Secondary	Married	House / a	0.019689	-17533	-2306	-5
307496	456238	0	Cash loan	M	Y	N	1	270000	479700	46858.5	450000	Family	Commerc	Secondary	Married	House / a	0.035792	-14958	-6573	-1
307497	456239	0	Cash loan	M	Y	N	0	180000	808650	23773.5	675000	Unaccomp	Working	Higher ed	Married	House / a	0.025164	-20922	-7438	-9
307498	456240	0	Revolving	F	N	N	0	360000	337500	16875	337500	Family	Working	Higher ed	Married	House / a	0.00496	-17345	-2178	-4
307499	456241	0	Cash loan	F	Y	Y	0	180000	270126	12028.5	193500	Unaccomp	Working	Secondary	Married	House / a	0.00702	-16679	-1222	-1
307500	456242	0	Cash loan	M	Y	Y	0	198000	1312110	52168.5	1125000	Unaccomp	Commerc	Secondary	Married	House / a	0.072508	-19102	-3689	-
307501	456243	0	Cash loan	F	N	Y	0	81000	225000	12694.5	225000	Unaccomp	Commerc	Secondary	Married	House / a	0.00702	-16988	-8694	-5
307502	456244	0	Cash loan	F	N	Y	0	261000	1303812	35982	1138500	Unaccomp	Working	Higher ed	Married	House / a	0.031329	-20390	-5326	-
307503	456245	0	Cash loan	F	N	Y	3	81000	269550	11871	225000	Unaccomp	Commerc	Secondary	Married	House / a	0.009175	-12961	-1046	-1
307504	456246	0	Cash loan	F	N	Y	1	94500	225000	10620	225000	Unaccomp	Working	Secondary	Married	House / a	0.010966	-16063	-8736	-8
307505	456247	0	Cash loan	F	N	Y	0	112500	345510	17770.5	247500	Unaccomp	Commerc	Higher ed	Single / n	House / a	0.0228	-11870	-399	-4
307506	456248	0	Cash loan	F	N	Y	0	153000	331920	16096.5	225000	Unaccomp	Working	Secondary	Separated	House / a	0.022625	-16705	-7258	-9
307507	456249	0	Cash loan	F	N	Y	0	112500	225000	22050	225000	Unaccomp	Pensioner	Secondary	Single / n	House / a	0.0228	-24384	365243	-7
307508	456251	0	Cash loan	M	N	N	0	157500	254700	27558	225000	Unaccomp	Working	Secondary	Separated	With pare	0.032561	-9327	-236	-8
307509	456252	0	Cash loan	F	N	Y	0	72000	269550	12001.5	225000	Unaccomp	Pensioner	Secondary	Widow	House / a	0.025164	-20775	365243	-4
307510	456253	0	Cash loan	F	N	Y	0	153000	677664	29979	585000	Unaccomp	Working	Higher ed	Separated	House / a	0.005002	-14966	-7921	-6
307511	456254	1	Cash loan	F	N	Y	0	171000	370107	20205	319500	Unaccomp	Commerc	Secondary	Married	House / a	0.005313	-11961	-4786	-2
307512	456255	0	Cash loan	F	N	N	0	157500	675000	49117.5	675000	Unaccomp	Commerc	Higher ed	Married	House / a	0.04622	-16856	-1262	-5
307513	0	0	0	0	0	0	0	0	0	12	278	1292	0	0	0	0	0	0	0	
307514	total	0	0	0	0	0	0	0	0	0	1.01743E-05	0.000141	0	0	0	0				
307515	0	0	0	0	0	9152465	0	0	0	0	0.003037411	0.014116	0	0	0	0		#DIV/0!		
307516																				

application_data

Fig. V.1 Counting no. of blank cells.



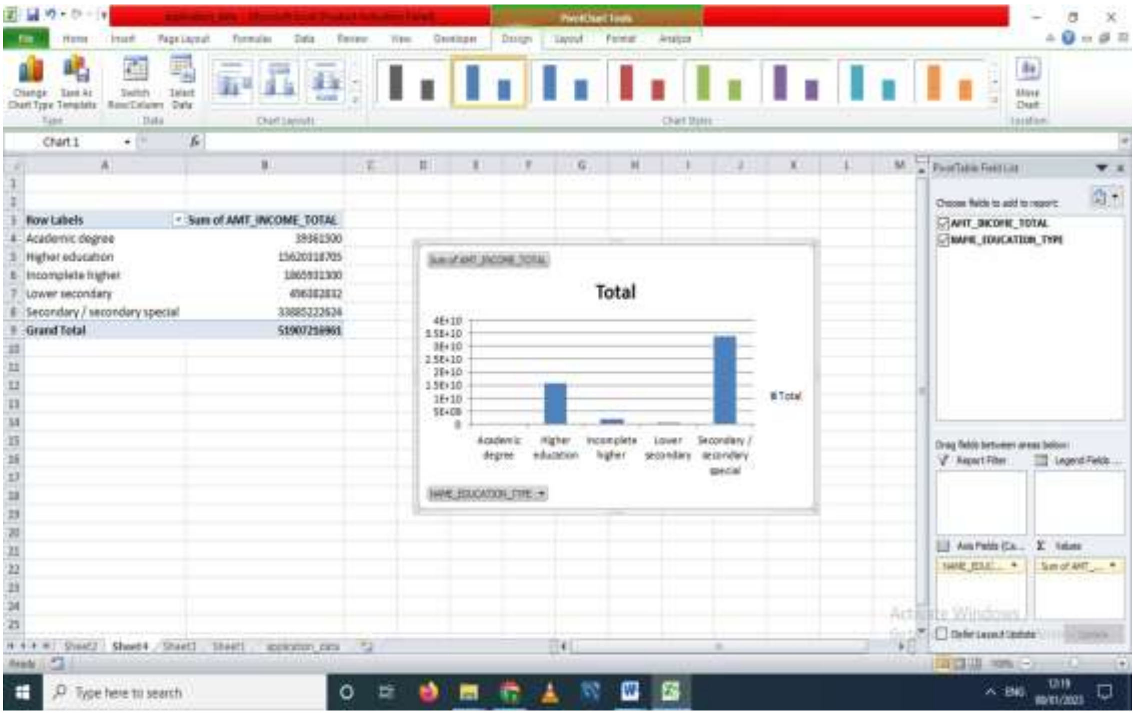
V307513 =COUNTBLANK(V2:V307512)

	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM	AN	AO
307491		1	1	0	1	0	0	Sales staff	2	2	2	FRIDAY	5	0	0	0	0	0	0	Business Enti
307492		1	1	0	1	1	0	Laborers	1	2	2	SATURDA	12	0	0	0	0	0	0	Business I 0.
307493	15	1	1	1	1	0	0	Drivers	4	2	2	SATURDA	10	0	0	0	0	0	0	Self-empl 0.
307494	2	1	1	0	1	0	0	Sales staff	2	2	2	FRIDAY	10	0	0	0	0	0	0	Self-employe
307495		1	1	0	1	0	0	Security st	2	2	2	MONDAY	12	0	0	0	0	1	1	School
307496	15	1	1	0	1	0	0	Drivers	3	2	2	MONDAY	9	0	0	0	0	0	0	Self-empl 0.
307497	2	1	1	0	1	1	0	High skill t	2	2	2	WEDNESD	14	0	0	0	0	0	0	Business Enti
307498		1	1	0	1	0	0	Cooking st	2	2	2	THURSDA	13	0	0	0	0	0	0	Self-empl 0.
307499	0	1	1	0	1	0	0	Sales staff	2	2	2	THURSDA	13	0	0	0	0	1	1	Trade: typ 0.
307500	10	1	1	0	1	0	0	Laborers	2	1	1	MONDAY	10	0	0	0	0	0	0	Construction
307501		1	1	1	1	1	0	Medicine	2	2	2	TUESDAY	15	0	0	0	0	0	0	Medicine 0.
307502		1	1	0	1	0	1		2	2	2	FRIDAY	10	0	0	0	0	0	0	Business I 0.
307503		1	1	1	1	0	0	Low-skill l	5	2	2	WEDNESD	14	0	0	0	0	0	0	Industry: type
307504		1	1	1	1	0	0	Laborers	3	2	2	MONDAY	16	0	0	0	0	1	1	School
307505		1	1	0	1	0	1	Core staff	1	2	2	SATURDA	11	0	0	0	0	0	0	Bank 0.
307506		1	1	0	1	0	0	Sales staff	1	2	2	WEDNESD	11	0	0	0	0	0	0	Self-employe
307507		1	0	0	1	1	0		1	2	2	MONDAY	15	0	0	0	0	0	0	XNA
307508		1	1	0	1	0	0	Sales staff	1	1	1	THURSDA	15	0	0	0	0	0	0	Services (
307509		1	0	0	1	1	0		1	2	2	MONDAY	8	0	0	0	0	0	0	XNA
307510		1	1	0	1	0	1	Managers	1	3	3	THURSDA	9	0	0	0	0	1	1	School 0.
307511		1	1	0	1	0	0	Laborers	2	2	2	WEDNESD	9	0	0	0	1	1	0	Business Enti
307512		1	1	1	1	1	0	Laborers	2	1	1	THURSDA	20	0	0	0	0	1	1	Business I (
307513	202929	0	0	0	0	0	0	96391	2	0	0	0	0	0	0	0	0	0	0	0
307514																				
307515																				

application_data

Total blank cell: sum(=K307513/F307515)*100

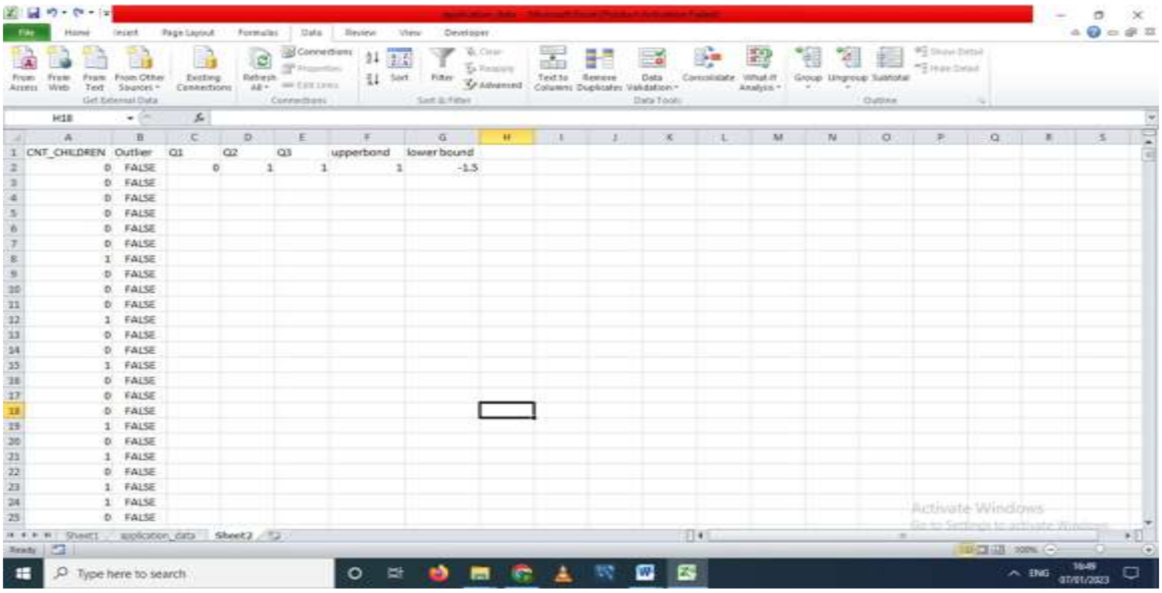
Fig. V.II Counting the no. of blank cells.



Row Labels	Sum of AMT_INCOME_TOTAL
Academic degree	39361300
Higher education	15620118705
Incomplete higher	1865931300
Lower secondary	496382832
Secondary / secondary special	33885222624
Grand Total	51907216961

V. B Find the quartile using quartile function.

In this fig we have calculated the lower bound and upper bound by using quartile function.



	A	B	C	D	E	F	G
1	CNT_CHILDREN	Outlier	Q1	Q2	Q3	upperbond	lower bound
2	0	FALSE	0	1	1	1	-1.5
3	0	FALSE					
4	0	FALSE					
5	0	FALSE					
6	0	FALSE					
7	0	FALSE					
8	1	FALSE					
9	0	FALSE					
10	0	FALSE					
11	0	FALSE					
12	1	FALSE					
13	0	FALSE					
14	0	FALSE					
15	1	FALSE					
16	0	FALSE					
17	0	FALSE					
18	0	FALSE					
19	1	FALSE					
20	0	FALSE					
21	1	FALSE					
22	0	FALSE					
23	1	FALSE					
24	1	FALSE					
25	0	FALSE					

Fig V.II Finding Outlier.

V. C Data Imbalance

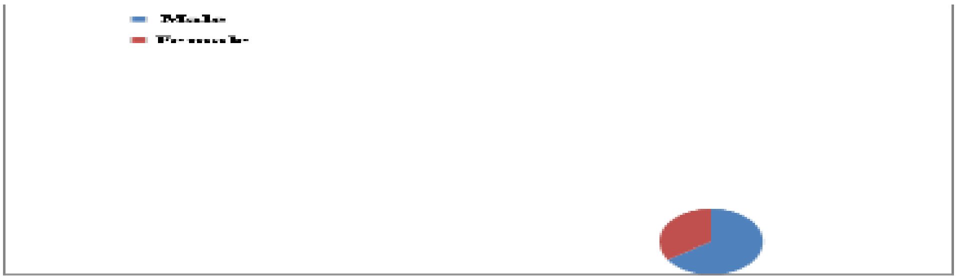


Fig. V.1 Data Imbalance

V.D Univarte Analysis

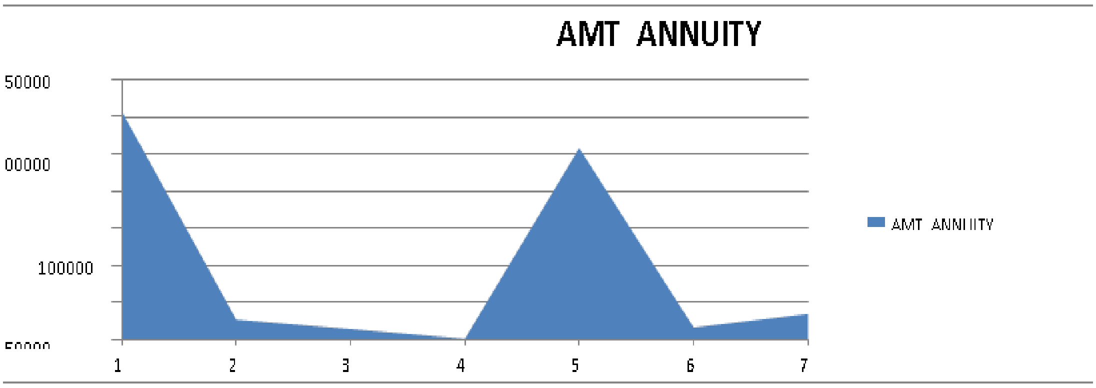
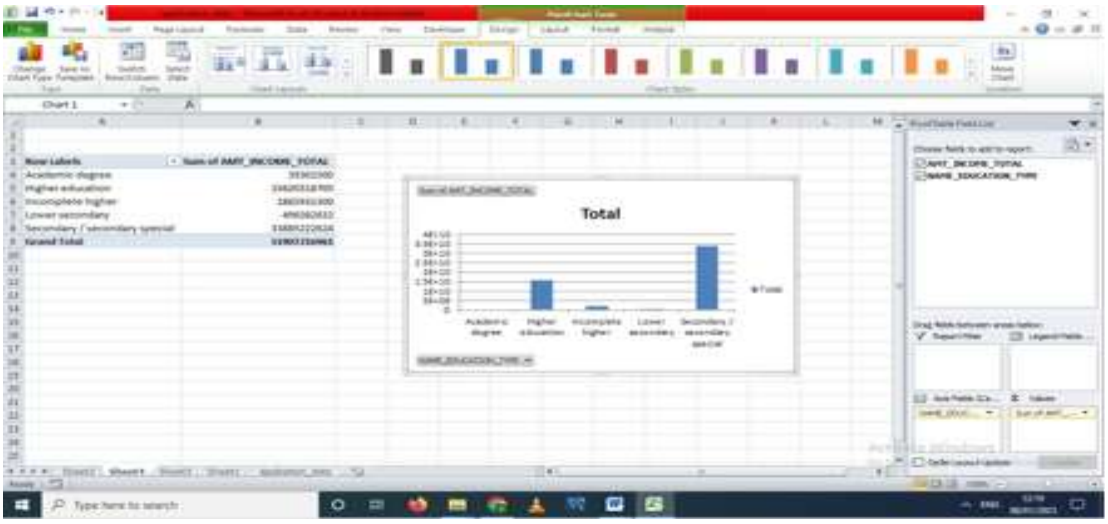


Fig V.IV Univarte analysis

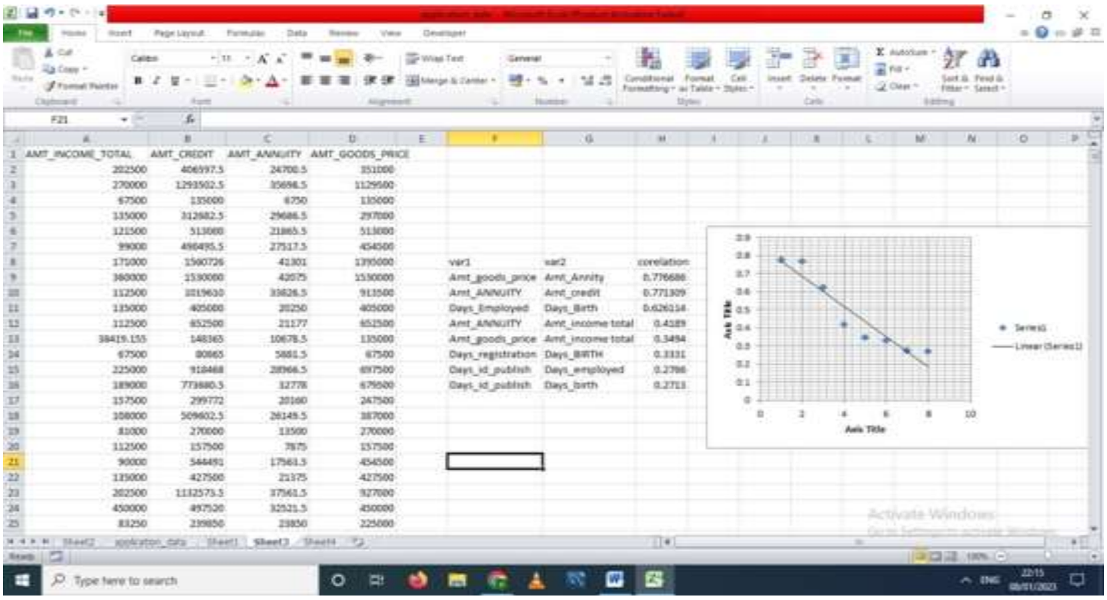
V.D Bivarte Analysis:



Row Labels	Sum of AMT_INCOME_TOTAL
Academic degree	[illegible]
Higher education	[illegible]
Incomplete higher	[illegible]
Lower secondary	[illegible]
Secondary / secondary special	[illegible]
Grand Total	[illegible]

Fig. V.1 Bivarte Analysis

V.E Corelation Between two Variable. = **=CORREL(A2:A47532, B2:B725)**



	AMT_INCOME_TOTAL	AMT_CREDIT	AMT_ANNUITY	AMT_GOODS_PRICE
2	202500	406597.5	24700.5	351000
3	270000	1293502.5	35698.5	1129500
4	67500	135000	6750	135000
5	135000	312682.5	29686.5	297000
6	121500	513000	21865.5	513000
7	99000	490495.5	27517.5	454500
8	171000	1560726	41301	1395000
9	360000	1530000	42075	1530000
10	112500	1019610	33826.5	913500
11	135000	405000	20250	405000
12	112500	652500	21177	652500
13	38419.155	148365	10678.5	135000
14	67500	80865	5881.5	67500
15	225000	918468	28966.5	697500
16	189000	773680.5	32778	679500
17	157500	299772	20160	247500
18	108000	509602.5	26149.5	387000
19	81000	270000	13500	270000
20	112500	157500	7875	157500
21	90000	544491	17563.5	454500
22	135000	427500	21375	427500
23	202500	1132573.5	37561.5	927000
24	450000	497520	32521.5	450000
25	83250	239850	23850	225000

var1	var2	corelation
Amt_goods_price	Amt_Annity	0.776686
Amt_ANNUITY	Amt_credit	0.771309
Days_Employed	Days_Birth	0.626114
Amt_ANNUITY	Amt_income total	0.4189
Amt_goods_price	Amt_income total	0.3494
Days_registration	Days_BIRTH	0.3331
Days_id_publish	Days_employed	0.2786
Days_id_publish	Days_birth	0.2713

Fig. V.1 Correlation between two variable

VI. Conclusions

We have examined various strategies, techniques, and tools in this study, all of which have drawbacks of their own. We have discussed numerous papers that gave us a general understanding of the kind of system needed in the modern era for the analysis and visualizations of data so that business owners and investors may make wise decisions and produce profit. We have suggested a technique where data is processed before being imported and saved in a database. With varied factors and dimensions, this data is visualized. with the help of which the we can easily filter the data . We have suggested a method for importing and storing data.

VII. Future Scope

In our upcoming work, we'll employ a variety of cutting-edge visualization approaches and compile all the graphs and charts onto a single dashboard to aid users in quickly making decisions and producing income.

REFERENCES

[1] Jiawei Han, Micheline Kamber and Jian Pei, “Data Mining Concepts and Techniques”,Third edition, MK Publications, 2009.

[2] M. Tennekes and E. de Jonge, “Top-down Data Analysis with Treemaps,” in Proceedings of the International Conference on Information Visualization Theory and Applications (IVAPP'11), pp. 236–241, March 2011.

[3] P. Hoek, “Parallel Arc Diagrams: Visualizing Temporal Interactions,” Journal of Social Structure, vol. 12, 2011.

[4] Vipul Gupta, Arshan Porsohi and Poornaprajna Udupi, “Sensor Network: An Open Data Exchange for the Web of Things,” in Proceedings of 8th IEEE Conference on Pervasive Computing and Communication Workshop (PERCOM), 2010.

[5] Ciro Donalek, S.G. Djorgovski, Alex cioc, Anwell Wang, Jerry Zhang, Elizabeth Lawler, Stacy Yeh, Ashish Mahbal, Matthew Graham, Andrew, Drake Scott Davidoff and Jeffrey S. Norris, “Immersive and Collaboration Data visualization using Virtual Reality Platforms,” in Proceedings of IEEE International Conference on Big Data, 2014.

[6] Tak-chung Fu, Fu-lai Chung and Chun-fai Lam, “Adaptive Data Delivery Framework for Financial Time Series Visualization,” in Proceedings of IEEE International Conference on Mobile Business (ICMB), pp. 267-273, 2005.

[7] Eric Martin and Vincenzo Di Bernardo, “Enterprise Dashboard Tools for Management of Share-use University Laboratory,” in Proceedings of University Conveinment Industry Micro (UCIM), 2008.

[8] Victor Pascual-cid, “An Information Visualization System for the Understanding of Web Data,” in Proceedings of IEEE Symposium on Information Visualization (INFOVIS), 2008.

[9] Liang Zhu, Bing-Fang, Yue-min Zhou, Xin-hui ma and Leidong yang, “Researches on Eco-environment Data Visualization for Three Gorgeous Project,” in Proceedings of IEEE International Symposium on Information Engineering Electronic Commerce (IEEC), 2009.

[10] Thorben sander, Malthias kehlenbeeck and Michael H. Breithner, “Visualization of Automated Compliance Monitoring and Repairing,” in Proceedings of IEEE Database and Expert System Applications (DEXA), 2012.

[11] Javier Perez, Romuald Deshayes, Methieu Goeminne and Tom Mens, “SECONDA: Software Ecosystem Analysis Dashboard,” in International Conference on Software Maintenance and Reengineering (CSMR), 2010.

[12] Cheol Jeong and Joseph Finkelstein, “Computer-Assisted Upper Extremity Training Using Interactive Biking Exercise (iBikE) Platform,” in Proceedings of IEEE Conference on Engineering in Medicine and Biological Society (EMBC), 2012.

[13] Fleni Stroulia, Isaac Matichuk, Fabio Rocha, Ken Baver, “Interactive Exploration of collaboration software development data,” in IEEE International Conference on Software Maintenance (ICSM), 2013.

[14] Pavan Kumar, Rashid Ahmad, B.D. Chaudary and Mukul K. Sinha, “Enriched dashboard:-An Integration and Visualization Tool for Distributed NLP System on Heterogeneous Platform,” in International Conference on Computer Science and its Applications (ICCSA), 2013.

[15] Ma Xin-hui, Wu Bing-fang, Zhu Liang, etc. “Study on Methods of Hydro-environment Data Visualization for Three Gorges Project” [J], Remote sensing informatics, 2007, pp. 28-31.

[16] M. Lungu, M. Lanza, T. GˆÕrba, and R. Robbes, “The small project observatory: Visualizing software ecosystems,” Science of Computer Programming, vol. 75, no. 4, pp. 264 – 275, 2010.

[17]R. Chang, M. Ghoniem, R. Kosara, W. Ribarsky, J. Yang, E. Suma, C. Ziemkiewicz, D. Kern, and A. Sudjianto, Wire-vis: Visualization of categorical, time-varying data from financial transactions. In IEEE Symposium on Visual Analytics Science and Technology, pages 155–162, 2007.

[18] “The 37 Best Tools for Visualization,” http://www.creativebloq.com/design-tools/data-visualization712402.Harjeet Kaur, Varsha Sahni, and Dr. Manju Bala, “A survey of reactive, proactive, and hybrid routing protocols in MANET: a review”, International Journal of Computer Science and Information Technologies, vol. 4, no. 3, pp. 498-500, 2013