Journal of Theoretical and Applied Information Technology
15th June 2023. Vol.101. No 11

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

BUY NOW PAY LATER SERVICES ON GENERATION Z: EXPLORATORY DATA ANALYSIS USING MACHINE LEARNING

YOSY ARISANDY [1] , YOSZA BIN DASRIL [1], SHAHRUL NIZAM BIN SALAHUDIN[1], MUCH AZIZ MUSLIM[1], ARISMAN ADNAN[2], GOH KHANG WEN[3]

[1]Faculty of Technology Management and Business, Universiti Tun Husein Onn Malaysia, 86400 Johor, Malaysia

[2]Faculty Mathematics and Natural Sciences, Universitas Riau, Pekanbaru, 28293 Indonesia

[3]Faculty of Data Science and Information Technology, INTI International University, 71800 Nilai, N.Sembilan, Malaysia

E-mail: [1]gp200002@student.uthm.edu.my, [1]yosza@uthm.edu.my, [1]shahrulns@uthm.edu.my, [1]gp200017@student.uthm.edu.my [2]arisman.adnan@lecturer.unri.ac.id, [3]khangwen.goh@newinti.edu.my

ABSTRACT

The buy now, pay later (BNPL) business model is an innovative approach to installment loans. It allows customers to take immediate possession of their purchase, with or without a down payment. Furthermore, the majority of BNPL loans are set up to require four payments. However, this type of loan comes with its own set of risks and challenges. This article examines the risk of BNPL as a product for consumers known as Generation Z. The data used is secondary data provided by Kaggle in csv format (loan data.csv) contains 159,584 postpaid customer records and 28 features analyzed through descriptive and Exploratory Data Analysis (EDA). The results show that the majority of pay later clients are married and known as millennials are the ones who used pay later services the most (52.10%). Generation Z has the greatest rate of loan defaults which is about 34.16% with the time employee is about 0-8 months (35.8%). Furthermore, the results indicated that the unemployed generation Z has the highest default percentage of 32.16%. This Exploration data analytic is viewed as a step towards gaining a better understanding of consumers so that predictions, suggestions, and recommendations can be made for potential customers and market paylater segmentation to find the right target market, thereby positively impacting company profits.

Keywords: *Buy Now Pay Later, Risky, Generation Z, Exploratory Data Analysis, Machine Learning*

1. INTRODUCTION

Nowadays, technology and innovations have an active impact on many aspects of human existence, sparking digital transformation in business and society. Digital technology has grown and expanded rapidly during the past 20 years. The widespread availability of various types of gadgets, the development of digital platforms, and the advent of social media and applications resulted in the birth of a generation of millennials, or individuals born at the turn of the century and growing up during this period [1] . Millennials, also known as Generation Y (Gen Y), are the demographic cohort that comes after Generation X (Gen X) and before Generation Z (Gen Z). Gen Z, also known as the internet generation, is the demographic cohort that follows Millennials and precedes Generation Alpha. Also, refers to those individuals who were born between the years 1995 and 2012. [2], makes reference to individuals who were born digitally and are heavily immersed with technology [3], full of confidence, loving self-expression, adventure, and more eager to achieve their values [4].

The BNPL business model, on the other hand, is a novel approach to instalment loans. It enables customers to take possession of their purchase immediately, with or without a down payment. Furthermore, the bulk of BNPL loans are structured to require four installments. However, this form of financing comes with its own set of risks and challenges. Therefore, age is the most significant issue in today's digital world [5] As a result, there exists disparities between different consumer groups

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

(e.g., Gen Y and Gen Z) and their consumer expectations [6]. Consumers of today and the foreseeable future are members of Gen Z[7] .

E-fastest commerce's expansion has occurred between 2020 and the present. Customers' preferences and shopping habits are shifting due to the current pandemic situation, which lends credence to this trend [8]. Gen Z is challenging since they seem to behave differently from previous generations, which can influence customer behavior [9]. They desire efficient transaction processing, simple access to pertinent information, enhanced functionality, and a broader selection of financing products and services that can be personalized [10].

Ernst & Young stated that Gen Z is most likely to shop online for efficiency [9]. Its conducted a recent study [11] to determine how the two groups of young people differ. The firm questioned 1,000 adults and 400 adolescents. Gen Z shops online for "efficiency" 63 percent of Gen Z respondents said they shop online because it "saves [them] time," whereas only 55 percent of millennials said the same. 53% of Gen Z respondents to a survey agreed that "the selection is easier online," compared to 44% of millennial respondents. In comparison to teens, just 41% of millennials agreed with the statement "Prices are lower online." Compared to teens, only 21% of millennials agreed with the statement that "products are categorized [online] in a way that makes purchasing easy." Bilal find, if a customer believes an online store has high-quality offerings and is easy to use, they are more likely to make a purchase there [12]

The most ardent supporters of Buy Now Pay Later (BNPL) payments are members of the younger generations, specifically millennials and Gen Z[13]. And it's easy to see why: these individuals have limited credit histories but ambitious goals [14]. Therefore, they do not like conventional payment options such as credit cards or loans [15]. They require a rapid, easy, and minimally administrative procedure[16]. Previously, it took months or even years to establish financial credibility using conventional methods [17]. However, with real-time data, all of this is now possible.

Cassandra Napoli, senior strategist at trend-forecasting firm WGSN Insight, explains, "Gen Z is one of the most avid adopters of BNPL apps, but it's not always clear how the platforms behave, and they frequently differ from one another." First-time users of these services who are young and impressionable may be unaware of the financial commitment they are making [18].

2. RESEARCH SCOPE

This research examines Gen Z's BNPL risk. Exploratory Data Analysis (EDA) for customer BNPL dataset. This work uses exploratory data analysis to improve machine learning processing.

2.1 Data selection

Starting from importing libraries and BNPL dataset

2.2 Finding Data type

In the dataset, there are three sorts of categories: object, int64 and float64.

2.3 Data clean

By finding missing value and handle, identify duplicate entries/row and drop it, feature scaling.

2.4 Exploratory Data Analysis

Evaluated and compared the research on EDA and paylater with other studies, as shown in Table 1.

Table 1. Evaluate of EDA and generation z in previous research

No	Paper	Aims	Evaluate
1	Exploratory Big Data Statistical Analysis The Impact Of People Life's Characteristics On Their Educational Level[19]	Analyze the educational data in Egypt census (2017), and discuss the relations between educational data features as indicators of educational levels of Egyptians, by regression model used as a traditional statistical method to categorical data analysis.	The current research is also about EDA, but the methodology employed blends statistical methods with matrix visualization.
2	Dynamic graph exploration by interactively linked node-link diagrams and matrix visualizations[20]	a visually and algorithmically scalable method for interactively linking node-link and adjacency matrix visualizations of graphs	In previous studies, the MV approach was used in the field of graph visualization studywithout being integrated with statistical methods. This study, meanwhile, employs financial data, especially consumer data on paylater services.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3	Fintech and responsibility: Buy-now-pay-later arrangements[17]	This study analyse the preference young adults for BNPL relative to credit cards and the role of financial literacy and traits including propensity to plan and save	The findings suggest that financial literacy reduces perceived BNPL benefits and that lower financial literacy is associated with more benefits and less risks. Although we conduct the same study on BNPL services, our focus is on the risk of BNPL as a product for Generation Z consumers.
4	Closer together or further apart? Values of hero generations Y and Z during crisis[21]	Comparing the basic human values of Generation Z and Generation Y,	Generation Y is twice as worried about the economy and finances than Generation Z, according to study. Y prioritizes financial security more than Z. Generation Z worries about future uncertainty.
5	Generation Z's perceptions and attitudes toward debt: a case study of young consumers in rural Michigan, USA[22]	The majority of Generation Z consumers are aware that immediate credit is detrimental. Moreover, these young clients stated that being in debt is never a pleasant experience.	Related with this study, but the current study was conducted to explore loan default data on gen z

3. MATERIALS AND METHOD

The dataset used in this study is secondary data originating from the Pay later dataset available on the Kaggle website. Kaggle [23] is a website for discussing ideas, obtaining inspiration, competing with other data scientists, learning new information and coding methodologies, and watching various examples of real-world applications for data science. This study uses secondary data in the form of the pay later loan dataset. The Kaggle-provided dataset in csv format (loan data.csv). The dataset used consists of 159,584 pay later customer data with 28 features.

The data processing using Exploratory Data Analysis (EDA). EDA is part of data science process [24]. Exploratory data analysis and descriptive statistics, according to Döring and Bortz [25], are used to find prospective and possibly unexpected impacts and trends in the data. Descriptive analysis are commonly used in exploratory data analysis. Summaries of a set of observations or a subset of data are provided by descriptive statistics. These numbers and visualizations could be numerical, like the mean, mode, median, percentiles, max, min, etc., or visual, like graphs and plots. Univariate, bivariate, and multivariate analysis are all possible within the realm of descriptive statistics [19], [26], [27].

The analysis is done using Python programming. EDA in machine learning utilizing multiple Python programming libraries generates graphic representations that provide information about processed datasets, including boxplot graphs, histograms, lineplots, and correlation graphs using the seaborn library. Data processing Machine learning and the construction tool of this machine learning project is using Google Collaboratory, also called Colab. Colab is a free environment for Jupyter notebooks that runs in the cloud and stores notebooks on Google Drive.

4. RESULTS

4.1 Importing Libraries and Dataset

Import CSV files format, loading data into a Python environment is the first step in data analysis. This is the only format in which pandas can import a local dataset to Python for pre-processing. Import the NumPy library to manage vectors and matrices, SciPy for time series and tabular data structures is an additional frequently utilized library of note, Matplotlib, a cross-platform data visualization and graphical charting package for Python and NumPy. Import Seaborn to provide a sophisticated interface for creating visually appealing and instructive statistical visuals.

4.2 Finding data type

There are 3 types of categories in dataset:

4.2.1. Object

The format of the object means the variable is categorical. The categorical variables in this dataset are client gender, client marital status, client loan purpose, client residential status, client time at employer, repaid date month.

4.2.2. Int64

This represents an integer variable of features. int64: income verified, client age, loan number, loan amount, loan term, max amount taken, max tenor taken, settle days, first payment default, loan default, application date month, approval date month, disbursement date month, first payment due date month, due date month.

4.2.3. Float64

This represents a variable that has multiple decimal values involved. They are also

Journal of Theoretical and Applied Information Technology
15th June 2023. Vol.101. No 11

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

numeric variables. float64: client income, client number phone contacts, client average calls per day, interest rate, payment ratio, loan income ratio.

Data types are an important concept because statistical methods can only be used with certain data types. It is necessary to analyze continuous data differently from categorical data, otherwise it will result in incorrect analysis. Therefore, knowing the type of data you are dealing with allows you to choose the correct analysis method.

4.3 Finding missing value and handle

Detecting missing values numerically and applying the Missingno package to perform a graphical analysis of the dataset. Missing values and outliers are examples of problematic data that can impede analysis [28]. This dataset has no missing values.

4.4 Identify duplicate entries and drop it

Any duplicate, blank, or otherwise abnormal data has an effect on the analysis process [29].To determine if duplicates exist in a data set, one can utilize the duplicated() function. After deleting the 12 duplicates, there are 159,584 unique customer records.

4.5 Feature scaling

Feature scaling is a technique used to normalize the range of independent variables or data components. It is also known as data normalization and is often performed during the data pre-processing step. There are two data features with an anomalous distribution.

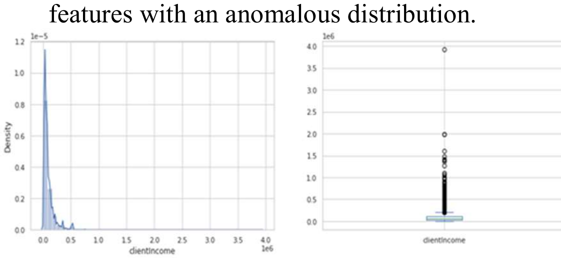

Figure 1. Pre-log transformation client income

Figure 1 describe Most of the data in the applicant's income distribution and loan amount Figure 3 are to the left, which means that it is not normally distributed. The distribution is skewed to the right (positive skewness). Handle to normalize it in the next section because the algorithm works better when the data is normally distributed. Because of these outliers, most of the data in loan amounts is on the left and the right tail is longer. This is called right slope (or positive slope). One way to get rid of skewness is to perform a log transformation. Take the log transformation, it doesn't affect the smaller values much, but subtracts the larger values. So, we get a distribution like the normal distribution Figure 2 and Figure 4.

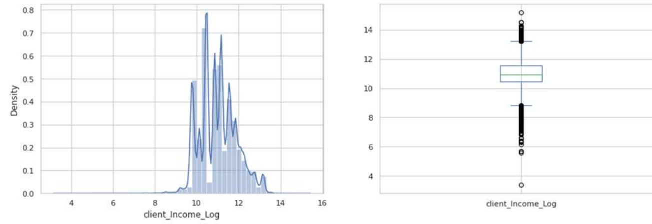

Figure 2. Loan amount before log transformation

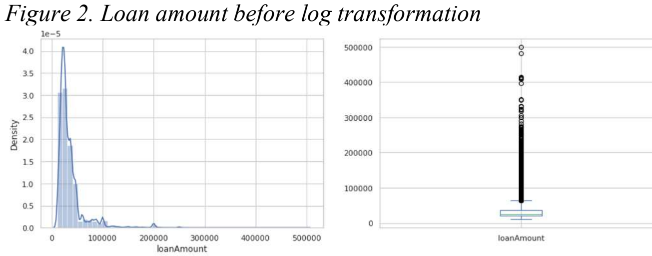

Figure 3. Loan amount before log transformation

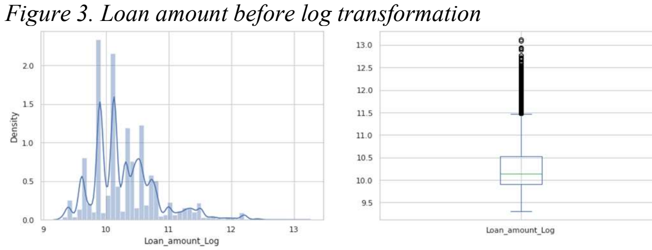

Figure 4. Loan amounts illustrate the effect of log transformation.

4.6 Exploratory Data Analysis (EDA)

It is possible to perform an analysis of some of the essential details contained in the dataset by utilizing EDA.

4.6.1. Uni variate analysis

One of the most basic methods of data description is univariate analysis. Uni comes from the Latin word for "one," therefore the analysis focuses on only one. Accordingly, the primary goal of Uni-variate analysis is to characterize a single feature, rather than conventional connections among features.

From the full dataset, Figure 5 feasible to examine information about customers of pay later services, including the fact that the majority of pay later clients are married. Consequently, using the greatest pay later is for business purposes. Men are more likely than women to use paylater services, according to client gender data.

To acquire data about pay later users in the gen z age category, the age data will be divided into four categories: gen z (18-25), millennial (26-35), gen x (36-45), baby

Journal of Theoretical and Applied Information Technology
15th June 2023. Vol.101. No 11

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

boomer (46-55) and 56 above by using bins algorithm. The data is again displayed as a chart Figure 6 to figure out the percent of each age range that grouped with.

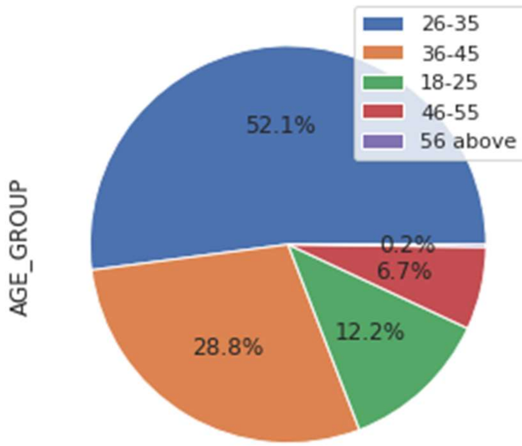

Figure 6. Customers pay Later based on generation category.

The generation of consumers known as millennials are the ones who used pay later services the most (52.10%), followed by the generation known as Gen X (28.80%), and then Gen Z (12.20%).

4.6.2. Bivariate analysis

Bivariate analysis is one of the most straightforward types of descriptive statistics. It refers to the side-by-side comparison of two features for possible correlations. As in Figure 7 comparing between age and default data

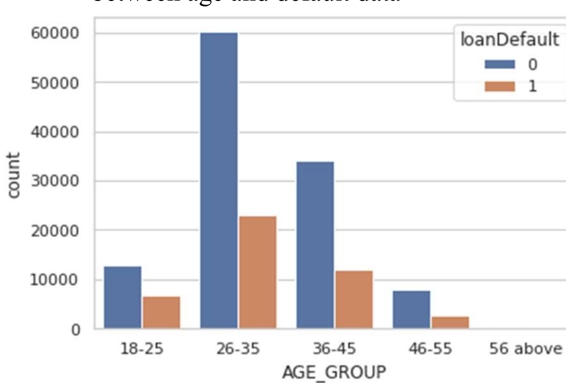

Figure 7. Loan Default Percentage By Age Group

Table 1. Age-related loan default rates

index	AGEZ_GROUP	Not Default	default	Total	Default Percentage
0	18-25	12810	6647	19457	34.16
1	26-35	60177	22905	83082	27.57
2	36-45	34089	11883	45972	25.85
3	46-55	7991	2741	10732	25.54
4	56 above	251	87	338	25.74
5	All	115318	44263	159581	27.74

According to Table 1 Gen Z has the greatest rate of loan defaults at about 34.16%.

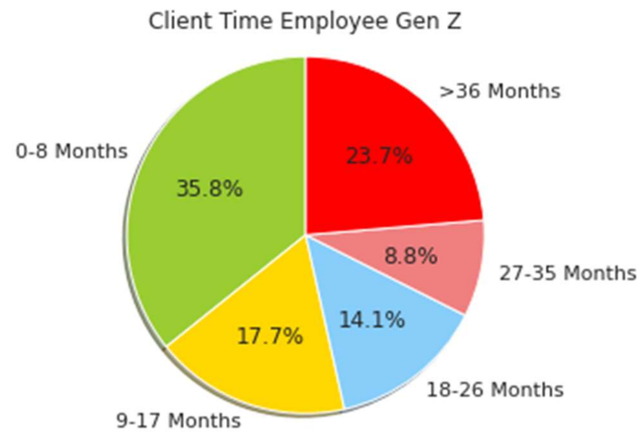

Figure 8. Client Time Employee Gen Z

Figure 8 show that Gen Z has the highest proportion (35.80%) of 0-8 month employment periods.

4.6.3. Multivariate Analysis

Aiming to decipher how the various fields in the dataset interact with one another or discovering interactions between variables with more than two possible levels. Table 2 and Figure 9 illustrate the correlation between customer time at employer, age, and payment default.

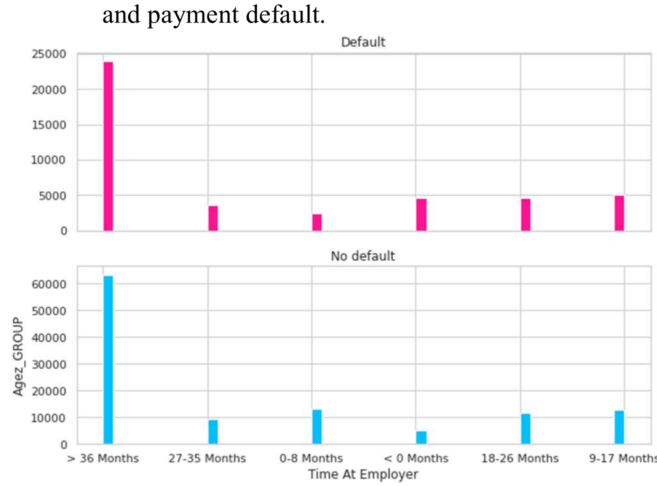

Figure 9. The Correlation Between Customer Time At Employer, Age, And Payment Default

Table 2. The correlation between employer duration and payment default of Gen Z

No	client Time At Employer	Not Default	default	Total	Default Percentage
0	0-8 Months	13052	4610	17662	26.101234
1	18-26 Months	11515	4565	16080	28.389303
2	27-35 Months	9390	3660	13050	28.045977
3	9-17 Months	12906	5033	17939	28.05619
4	< 0 Months	5136	2435	7571	32.162198
5	> 36 Months	63322	23960	87282	27.45125
6	All	115321	44263	159584	27.73649

Figure 9 and table 2 show that the Gen Z with a working period of less than 0 months has the highest default percentage of 32.16%

4.6.4. Matrix Visualization

Matrix visualization (MV) for visualizing and clustering symbolic data using interval-valued conceptual data as an example; it is by far the most prevalent symbolic data type in both the research and in practice[30].

Using the Pearson correlation coefficient, It can determine column pairwise correlations using the corr() function. Figure 10 illustrate the correlations with a heat map. Heat Map Colorize data for a graphical representation. Variables with deeper colors demonstrate a stronger association.

Table 3. Interpretation of the correlation coefficient[31], [32]

Size of Correlation	Interpretation
.90 to 1.00 (−.90 to −1.00)	Very high positive (negative) correlation
.70 to .90 (−.70 to −.90)	High positive (negative) correlation
.50 to .70 (−.50 to −.70)	Moderate positive (negative) correlation
.30 to .50 (−.30 to −.50)	Low positive (negative) correlation
.00 to .30 (.00 to −.30)	Almost no correlation

Table 3 interprets the value coefficient derived from a simple Pearson correlation analysis, also known as the Pearson Product Moment. The correlation value (r) spans from 1 to -1, with values closer to 1 or -1 indicating that the association between the two variables is strengthening, and values closer to 0 indicating that the relationship is weakening. Positive numbers imply a unidirectional (X increases, Y increases) relationship, whereas negative values indicate an inverse (X declines, Y increases) relationship (X increases, Y decreases).

Table 4. Correlation value between test variables

Var1	Var2	Correlation
approvalDateMonth	applicationDateMonth	1
FirstPaymentDueDateMonth	disbursementDateMonth	1
FirstPaymentDueDateMonth	approvalDateMonth	1
disbursementDateMonth	applicationDateMonth	1
disbursementDateMonth	approvalDateMonth	1
FirstPaymentDueDateMonth	applicationDateMonth	1
dueDateMonth	disbursementDateMonth	0.89
dueDateMonth	approvalDateMonth	0.89
dueDateMonth	applicationDateMonth	0.89
dueDateMonth	FirstPaymentDueDateMonth	0.89
paymentRatio	settleDays	0.85
Loan_amount_Log	interestRate	0.78
Loan_amount_Log	loanTerm	0.75
loanTerm	interestRate	0.74
firstPaymentDefault	settleDays	0.65
loanDefault	settleDays	0.63
loanDefault	firstPaymentDefault	0.62
loanDefault	paymentRatio	0.57
firstPaymentDefault	paymentRatio	0.55
clientMaritalStatus	clientAge	0.53
repaidDateMonth	paymentRatio	0.51
repaidDateMonth	settleDays	0.47
maxTenorTaken	maxAmountTaken	0.45
interestRate	incomeVerified	0.42
loanTerm	incomeVerified	0.38
Loan_amount_Log	incomeVerified	0.37
repaidDateMonth	loanDefault	0.33
clientTimeAtEmployer	clientAge	0.32
maxTenorTaken	incomeVerified	0.32
dueDateMonth	loanTerm	0.32

Table 4 illustrate that Age, income, and loan amount are three characteristics that have a negative impact on loan default, as shown by the graph. For instance, the younger the customer, the higher the likelihood of loan default. The connection between employee clients and age is r = 0.32 (poor positive correlation). This data reveals that employees spend less time with younger customers. According to the report, 32% of Gen Z still does not have a job with clients. Less than zero months and zero to eight months.

5. DISCUSSION

The analysis performed yielded some information regarding the utilization of pay-later services. Based on gender, males utilize pay-later services more frequently than women. In line with the findings of a survey performed by the data insight center on consumer e-commerce behavior, 62% of pay later service customers are male and 38% are female. One of the reasons why men favor pay-later services is their practicality.

Gen Z is not the most common consumer of pay later services. The millennial generation, aged 26 to 35, is the generation most likely to use service pay later. This finding is consistent with the findings of a survey conducted by the Katadata Insight Center in 2021 regarding the online shopping habits of Indonesian consumers. The poll reveals that the proportion of transactions made by members of the millennial generation is the largest, at 48%, compared to the Gen Z at about 29%.

Gen Z has the greatest default rate on pay later services. Compared to earlier generations, Gen Z devotes the majority of their income to online shopping, according to survey results. Gen Z has the highest proportion of spending relative to income compared to older generations, who dedicate less money to online shopping. But Gen Z are not financially independent and do not have a steady income. According to the graph, when compared to other generations based on length of time employed, Gen Z has the highest number of people who have worked for 0-8 months. The multivariant analysis demonstrates that consumers with a year of service of 0 or less months or who are unemployed have the largest number of defaults.

6. CONCLUSION

In contrast with prior studies [21]comparing the basic human values of Generation Z and Generation Y, it was discovered that Generation Y is twice as concerned about the economy and financial stability as Generation Z. Generation Y values financial security more than Generation Z. While Generation Z is more concerned about future unpredictability. The majority of consumers from Generation Z are aware that using instant credit is not good[22]. In addition, these young customers said that being in debt is never a comfortable situation.

When a default occurs, Gen Z is exposed to credit risk. Credit risk is the possibility that a loan may not be repaid by the borrower. According to the EDA, the unstable income element is a result of Gen Z's lack of employment and short employment tenure, which is one of the sources of credit risk for Gen Z.

The deployment of data analysis exploration methodologies on pay later data is expected to benefit service providers by making it possible to make predictions, suggestions, and recommendations to identify potential clients and pay later market segmentation, with a favorable effect on company profitability. However, additional research is required to identify other characteristics outside employment that contribute to credit risk in Gen Z.

ACKNOWLEDGMENT

Communication of this research is made possible through monetary assistance by Universiti Tun Hussein Onn Malaysia and the UTHM Publisher's office via publication fund E15216

REFERENCES:

[1] I. Kovshova, "Digital Transformation: The Age of Innovations in Business and Society," in *Digital Transformation*, A. Petrillo, F. De Felice, M. V. Achim, and N. Mirza, Eds. Rijeka: IntechOpen, 2022. doi: 10.5772/intechopen.102797.

[2] M. Seemiller, C., & Grace, *Generation Z: A Century in the Making*, 1st ed. Routledge, 2018. doi: https://doi.org/10.4324/9780429442476.

[3] J. MacKenzie, R. McGuire, and S. Hartwell, "The First Generation of the Twenty-First Century," 2016. [Online]. Available: http://magid.com/sites/default/files/pdf/MagidPluralistGenerationWhitepaper.pdf

[4] J. Liang, J. Li, and Q. Lei, "Exploring the Influence of Environmental Values on Green Consumption Behavior of Apparel: A Chain Multiple Mediation Model among Chinese Generation Z," *Sustain.*, vol. 14, no. 19, 2022, doi: 10.3390/su141912850.

[5] J. yup Lee, "Contesting the digital economy and culture: Digital technologies and the transformation of popular music in Korea," *Inter-Asia Cult. Stud.*, vol. 10, no. 4, pp. 489–506, 2009, doi: 10.1080/14649370903166143.

[6] C. V. Priporas, N. Stylos, and A. K. Fotiadis, "Generation Z consumers' expectations of interactions in smart retailing: A future agenda," *Comput. Human Behav.*, vol. 77, pp. 374–381,

2017, doi: 10.1016/j.chb.2017.01.058.

[7] S. Ö. and A. Okumuş, "The birth of Omni-channel marketing and new dynamics of consumers approach to Retail channels," in *Marketing Management in Turkey*, S. Ozturkcan and E. Y. Okan, Eds. Emerald Publishing Limited, 2018, pp. 247–272. doi: 10.1108/978-1-78714-557-320181029.

[8] C. I. Ratnapuri, M. Karmagatri, D. Kurnianingrum, I. D. Utama, and A. Darisman, "Users Opinion Mining of Tiktok Shop Social Media Commerce To Find Business Opportunities for Small Businesses," *J. Theor. Appl. Inf. Technol.*, vol. 101, no. 1, pp. 214–222, 2023.

[9] M. Schlossberg, "Teen Generation Z is being called 'millennials on steroids,' and that could be terrifying for retailers," *Insider*, 2016. https://www.businessinsider.com/millennials-vs-gen-z-2016-2 (accessed Jan. 19, 2023).

[10] S. Kaabachi, S. Ben Mrad, and T. Barreto, "Reshaping the bank experience for GEN Z in France," *J. Mark. Anal.*, vol. 10, no. 3, pp. 219–231, 2022, doi: 10.1057/s41270-022-00173-8.

[11] M. Merriman, "Rise of Gen Z: new challenge for retailers," *Ernst Young LLP*, pp. 1–12, 2015, [Online]. Available: http://www.ey.com/Publication/vwLUAssets/EY-rise-of-gen-znew-challenge-for-retailers/$FILE/EY-rise-of-gen-znew-challenge-for-retailers.pdf

[12] B. Eneizan, A. Alsaad, A. Alkhawaldeh, H. N. Rawash, and O. Enaizan, "E-WOM, trust, usefulness, ease of use, and online shopping via websites: The moderating role of online shopping experience," *J. Theor. Appl. Inf. Technol.*, vol. 98, no. 13, pp. 2554–2565, 2020.

[13] M. Backman and J. Caporal, "Study: Buy Now, Pay Later Services Grow in Popularity," *The Ascent*, 2022. https://www.fool.com/the-ascent/research/buy-now-pay-later-statistics/ (accessed Jan. 18, 2023).

[14] T. Payments, "The Latest 'Buy Now Pay Later' Consumer Insights for Your Business," *www.trustpayments.com*, 2022. https://www.trustpayments.com/author/trustpayments/ (accessed Jan. 18, 2023).

[15] B. Novendra and sarah S. Aulianisa, "Konsep Dan Perbandingan Buy Now, Pay Later Dengan Kredit Perbankan Di Indonesia: Sebuah Keniscayaan Di Era Digital Dan Teknologi," *J. Rechts Vinding Media Pembin. Huk. Nas.*, vol. 9, no. 2, p. 183, 2020, doi: 10.33331/rechtsvinding.v9i2.444.

[16] D. A. Z. Lia and S. L. Natswa, "Buy-now-pay-later (bnpl): generation z's dilemma on impulsive buying and overconsumption intention," *Proc. BISTIC Bus. Innov. Sustain. Technol. Int. Conf. (BISTIC 2021)*, vol. 193, no. Bistic, pp. 130–137, 2021.

[17] P. Gerrans, D. G. Baur, and S. Lavagna-Slater, "Fintech and responsibility: Buy-now-pay-later arrangements," *Aust. J. Manag.*, vol. 47, no. 3, pp. 474–502, 2022, doi: 10.1177/03128962211032448.

[18] Andrea cheng, "Buy Now, Pay Later: How Gen Z Is Falling Into Debt," *elle.com*, 2022. https://www.elle.com/fashion/shopping/a40154656/gen-z-buy-now-pay-later-debt/ (accessed Jan. 18, 2023).

[19] S. Z. A. Elhady, N. I. Ghali, A. Abo-Elfetoh, and A. M. Idrees, "Exploratory Big Data Statistical Analysis the Impact of People Life'S Characteristics on Their Educational Level," *J. Theor. Appl. Inf. Technol.*, vol. 100, no. 5, pp. 1495–1509, 2022.

[20] M. Burch, K. B. ten Brinke, A. Castella, G. K. S. Peters, V. Shteriyanov, and R. Vlasvinkel, "Dynamic graph exploration by interactively linked node-link diagrams and matrix visualizations," *Vis. Comput. Ind. Biomed. Art*, vol. 4, no. 1, 2021, doi: 10.1186/s42492-021-00088-8.

[21] S. Azimi, Y. Andonova, and C. Schewe, "Closer together or further apart? Values of hero generations Y and Z during crisis," *Young Consum.*, vol. 23, no. 2, pp. 179–196, Jan. 2022, doi: 10.1108/YC-03-2021-1300.

[22] S. Squires and H. W. L. Ho, "Generation Z's perceptions and attitudes toward debt: a case study of young consumers in rural Michigan, USA," *Young Consum.*, vol. 24, no. 2, pp. 133–148, Jan. 2023, doi: 10.1108/YC-07-2022-1567.

[23] Kaggle, "Kaggle," *https://www.kaggle.com/datasets/ibikunlegabriel/loan-data*, 2010. https://www.kaggle.com/datasets/ibikunlegabriel/loan-data (accessed Oct. 06, 2022).

[24] K. Sorg and H. Khobzi, "A decade of the Swiss electronic vaccination Record: Some insights based on an exploratory data analysis," *Int. J. Med. Inform.*, vol. 158, p. 104660, 2022, doi: 10.1016/j.ijmedinf.2021.104660.

[25] N. Döring and J. Bortz, "Forschungsmethoden und evaluation," *Wiesbad. Springerverlag*, 2016.

[26] A. Khademi, "Applied Univariate, Bivariate, and Multivariate Statistics," *J. Stat. Softw.*, vol.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

72, no. Book Review 2, pp. 1–4, 2016, doi: 10.18637/jss.v072.b02.

[27] Julien i.e. hoffman, *BASIC BIOSTATISTICS FOR MEDICAL AND BIOMEDICAL PRACTITIONERS*, 2nd Editio. Stacy Masucci, 2019.

[28] M. A. Muslim *et al.*, “New model combination meta-learner to improve accuracy prediction P2P lending with stacking ensemble learning,” *Intell. Syst. with Appl.*, vol. 18, no. December 2022, p. 200204, 2023, doi: 10.1016/j.iswa.2023.200204.

[29] M. A. Muslim, Y. Dasril, A. Alamsyah, and T. Mustaqim, “Bank predictions for prospective long-term deposit investors using machine learning LightGBM and SMOTE,” *J. Phys. Conf. Ser.*, vol. 1918, no. 4, 2021, doi: 10.1088/1742-6596/1918/4/042143.

[30] C. H. Kao *et al.*, “Exploratory data analysis of interval-valued symbolic data with matrix visualization,” *Comput. Stat. Data Anal.*, vol. 79, pp. 14–29, 2014, doi: 10.1016/j.csda.2014.04.012.

[31] D. E. Hinkle, W. Wiersma, and S. G. Jurs, *Applied statistics for the behavioral sciences*, vol. 663. Houghton Mifflin college division, 2003.

[32] M. M. Mukaka, “Statistics corner: A guide to appropriate use of correlation coefficient in medical research.,” *Malawi Med. J.*, vol. 24, no. 3, pp. 69–71, Sep. 2012.

Journal of Theoretical and Applied Information Technology
15th June 2023. Vol.101. No 11

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

APPENDIX

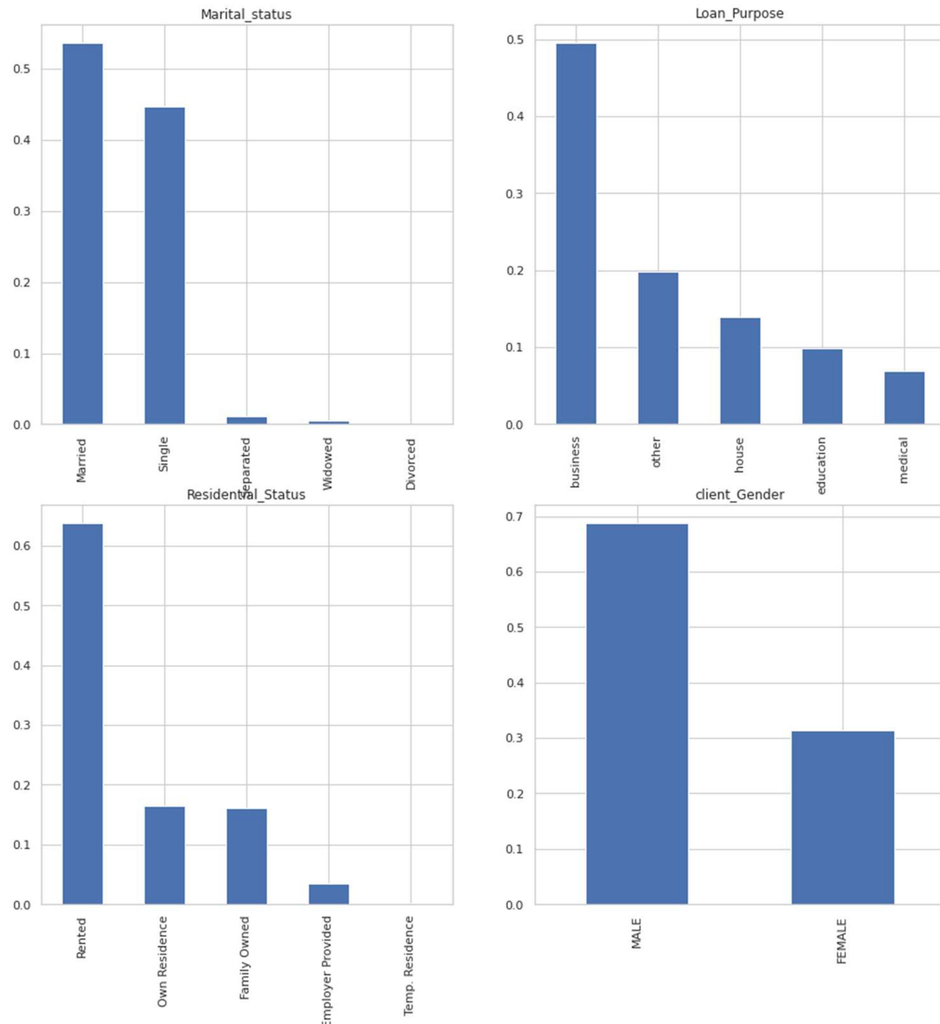

Figure 5. Basic Information about Customer Pay Later

Journal of Theoretical and Applied Information Technology
15th June 2023. Vol.101. No 11

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

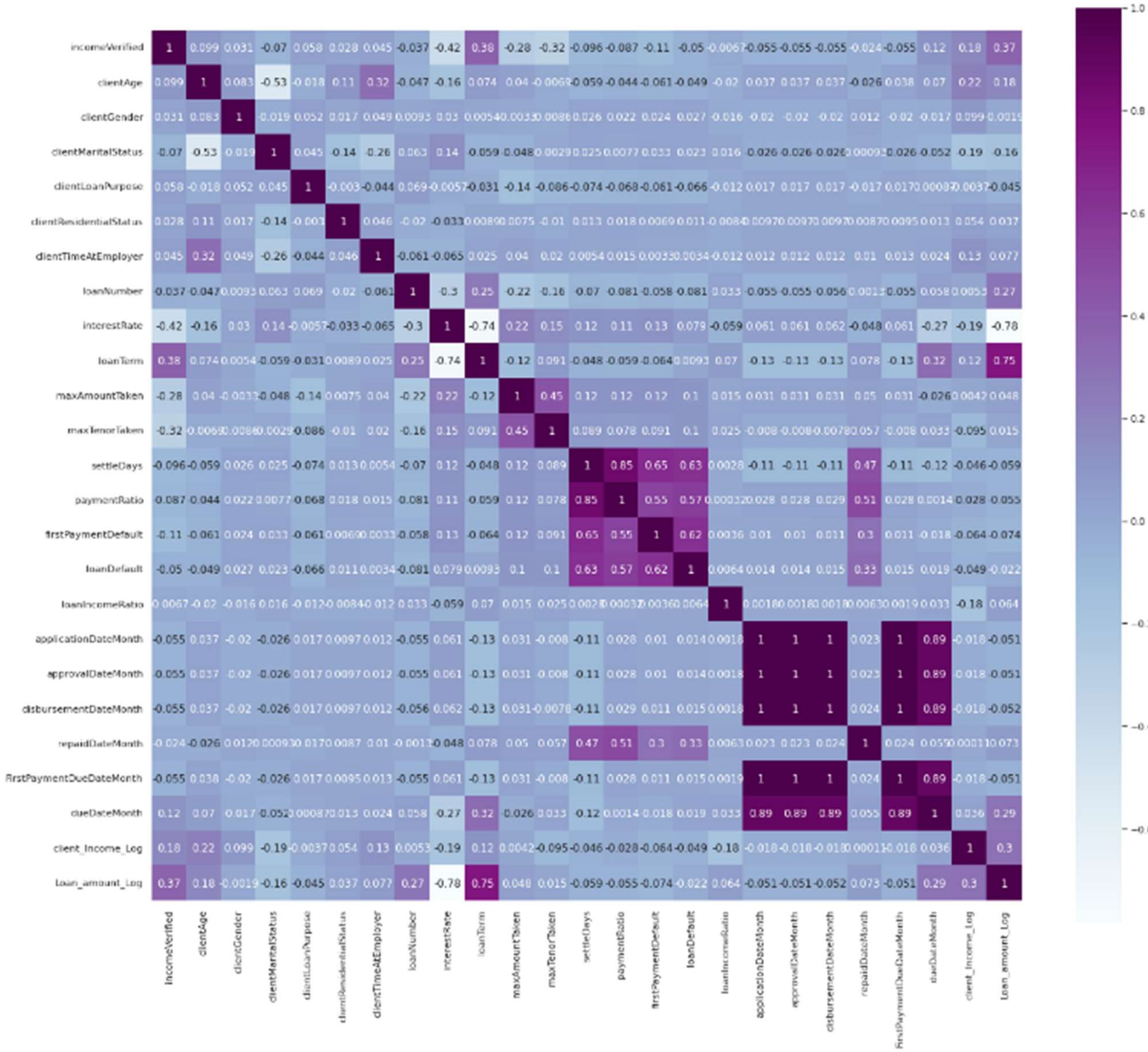

Figure 10. Visualization of Metrix Correlation