# **CHAPTER FOUR: DATA ANALYSIS**

## 4.1 Introduction

This chapter describes the steps used in setting up the system environment, acquiring the necessary dataset in our local machine, preparing the collected dataset , cleaning up the prepared data.frame ready exploration and data analysis in order to achieve the quality results as expected. BNPL Nigerian dataset(the acquired dataset) provided us with the 371 observations and 11 variables that will be used for the analysis. The information contained in the acquired dataset contains details/track record about customers on how they have been operating in the buy now pay later business. Analysis of the BNPL dataset makes it possible to conduct a comprehensive study of the ways in which knowledge management and data analytics influence the level of success achieved by a business operating under the Buy Now Pay Later (BNPL) model.

In the process, we will apply a correlation plot in order to determine whether or not there were any correlation between the numeric variables that were included in the BNPL dataset. Understanding the relationships between the variables is necessary in order to make quality results.

## 4.2 Dataset Description

"Buy now, pay later" refers to a form of short-term financing that gives customers the opportunity to make purchases and then pay for them at a later date, typically without incurring any interest charges. In the case that we are using as an illustration, we have collected the Nigerian BNPL dataset for the purpose of conducting data analysis so that we can come up with insights to improve the overall performance of the BNPL business. The BNPL dataset features demonstrate how each customer engages with this particular business in terms of the amount of money they have, whether or not they have paid, and so on. Each of the 11 variables is a characteristic of our dataset that will be considered during the analysis and will contribute to our overall comprehension of the data. In total, there are eleven variables, four of which are categorical variables, two of which are numeric variables, and one of which is a character type variable. The dataset is saved in a Comma Separated Values (csv) format.

## 4.2.1 Research Objectives

In order to get insights from the BNPL dataset, which is being used as a data source, an activity of this nature requires research objectives to be formulated. We were able to analyze the existing level of knowledge and generate predictions about the future using research objectives statements chosen from our topic by using the BNPL dataset. These statements were based on our topic. By selecting facts that are capable of being proved or refuted, we are able to conduct an analysis of the study's objective statements and assumptions and arrive at acceptable conclusions. The primary emphasis was placed on research objectives statements in order to gather the essential insights that were required to enhance BNPL's business performance in Nigeria. The following is a list of the research objectives considered in our research project during data analysis that were developed to guarantee that the research will be effective.

1. To determine the effect of using data analytics in buy now pay later on businesses’ performance in Nigeria.
2. To evaluate how the data analytics are used in the buy now pay later businesses to influence the knowledge management practices in Nigeria.
3. To analyze the impact of the buy now pay later business strategy on customer satisfaction and business performance.

## 4.3 Analysis

### **4.3.1 System Environment**

System environment provides a platform where the research project analysis will take place. System environment is considered so that it would be simpler and ease to perform the reproducibility of this project. This is done so that the user can be presented with information regarding the different software versions, settings, and packages that are available. If we do this, we will be able to determine which packages in our system environment are required and which ones aren't required to run it. In addition to that, the parameters of our system environment will also be displayed here. The head of the project is now fully aware of the subsequent steps that need to be taken because they have identified the appropriate packages that are present and those that are missing and have finished the processes required to get the environment ready for the analysis of our dataset. It is crucial that we are aware of the degree to which our application is compatible with those of other companies' because this will enable us to determine whether or not certain packages or software have an older version that is incompatible with the current offers on which we want to run our project on.

The contents of the.csv file is read in and then formatted into data frames by the following tools; of R (an open-source statistical computer language) and the RStudio (free IDE for the R language). We are going to make use of a process in R that is referred to as "data munging" to make certain that the data we are working with is free of faults and simple to comprehend. As a component of the process of preparing the data, it was possible to add missing variables and remove the unnecessary them entirely, and the variable headers were renamed to be more understandable It considerably took less time to finish this step given that the dataset that was used for data analysis contained a relatively small amount of "poor data." In addition, a variety of R packages was utilized throughout the analysis. In order to get a deeper level of comprehension of our data, we make use of a wide range of graphical and numerical representations.

### **4.3.2 Data Preparation**

The preparation of the data constitutes the second step of our project. At this point, we concentrated on making a new folder in our personal computer that will be used to store the dataset information. The datasets are saved in a well-known folder so that users of our system can access them in a prompt manner whenever they are needed. We now have a perspective of our dataset because we have presented the first five rows and the last five rows after it has been loaded into memory as shown below.

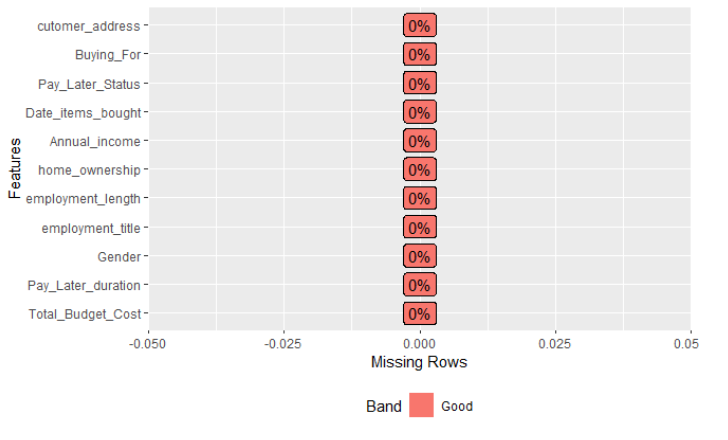
In addition to this, we investigated the variables that are included in our dataset and the kinds of data that are contained within each variable. This is made feasible thanks to a package in the R programming language called Glimpse. This package lays out the variables and the types that correspond to them (which are the things that we are actually looking for in our dataset) (what we are actually looking for in our dataset). Because of this information, we are now aware of the total number of observations that are present in our dataset, as well as the data types that are utilized the most frequently and those that are present in our dataset. This also sheds light on some of the missing values, presuming those values were ever a part of the system to begin with. If the dataset is discovered using any of these approaches or tools, the next automatic step is to clean up our dataset, which cleans up our dataset. This activity is accomplished by utilizing the "tidyverse" R package. We were quite fortunate in that there were no missing values in our circumstance. All of these processes are important, and we need to carry them out in order for our exploratory data analysis to be finished in a straightforward manner.



*Figure 1: The figure above illustrates the description of each variable found in the BNPL Nigerian dataset*

### **4.3.3 Data Cleaning**

After the data has been prepared, the following step is to clean the data. If you want to get the most out of your data exploration, cleaning it up beforehand is the best method to do so. At this point in the process, the data has already been extracted from its primary source and prepared; the next step is to perform an analysis on it. At this level, it is harder to identify and correct errors once they have been made. In order to remedy these problems and render the data suitable for analysis, the data will first require cleaning and then will need to be formatted. Because our dataset contained no missing values, we were required to conduct a thorough examination of the data types. As a result of performing a cross-check on all of the variables, we are able to determine what values are stored in each variable and whether or not the data type that has been assigned to that variable corresponds to the values that are stored in it. Because each variable in our dataset contained categorical data, we classified them as character variables rather than factor variables. These variables included "Gender," "Environment Length," "Home Ownership," and "Pay Later Status." It was necessary to do cleaning on these variables in order to change them from being of the character type to being of the factor type, which is the correct type of data.



# Figure 2: Above is a figure showing cleaned dataset with no missing values

### **4.3.4 Data Statistics**

After cleaning and tidying our data we went on to have a look at the overall dataset statistics. We start by displaying the structure of our dataset. Thereafter, we find the statistics in our dataset using the “describe()” method. This method is found in the “Hmisc” R package. Therefore, we start by loading the “Hmisc” package in order to be able to use the describe function. The results are then found/generated by this function. It is significant to perform descriptive statistics on our dataset. This is because descriptive statistics enables us to present the data in a more meaningful way, which allows simpler interpretation of the data. In the context of statistical analysis, the term "descriptive statistics" refers to a certain style of analysis that contributes to the effective communication of the significance of the data in a manner that is both clear and succinct. For instance, patterns might start to form. On the other hand, descriptive statistics don't let us draw any conclusions about the data we've looked at or about any assumptions we could have had. This is because descriptive statistics don't provide enough information. There is no problem with them that is rooted in their nature.

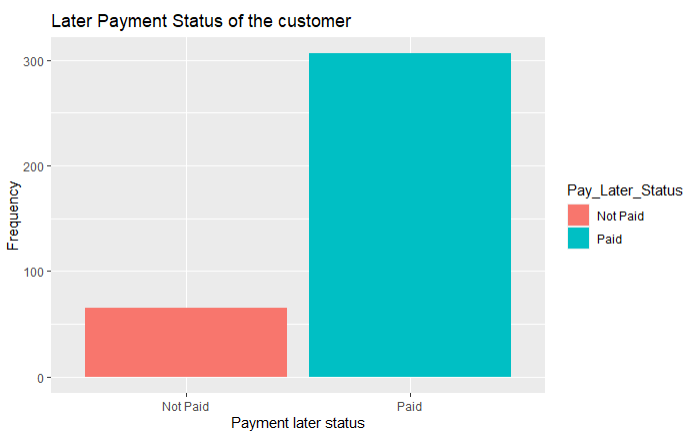
If we presented the data without the assistance of descriptive statistics, it would be quite difficult to comprehend what the data are trying to communicate. This is especially important to keep in mind when there is a substantial amount of data. Using descriptive statistics, the data can be presented in a manner that is easier to comprehend, which ultimately makes data interpretation much simpler. Utilizing descriptive statistics will allow us to have a better understanding of how our data variables are distributed.

# **CHAPTER FIVE: RESULTS AND FINDINGS**

The purpose of this part was to investigate unpaid external data sources (the internet), which may be put to use for the purpose of developing a deeper comprehension of a company model's customer base. More information from a wide variety of sources has been compiled in order to achieve a more in-depth comprehension of the characteristics of the purchase now pay later business model while interacting with customers. Credibility, reliability, relevance, and accessibility were the four factors taken into consideration during the evaluation of the additional data sources (data format). The dataset was made possible by the contribution of a wide array of components. This buy now pay later dataset is being collected for a number of different purposes, the most important of which are research aims, research problems, and research hypotheses. Because I already had the knowledge I required, figuring out which model would be most appropriate for addressing our problem statement was made much less difficult for me.

## 5.1 Research Objective 1

The analysis of the influence of the buy now pay later business strategy on the level of customer satisfaction as well as the success of the company was the primary focus. The "total budget cost" of the clients is taken into consideration here as one of the variables. This occurs when the customers collect their buy now goods from the business store. The "Annual income" of each customer is the second variable that is taken into consideration. In addition, once a customer has made a purchase in the shop, they are given the option to pay for it at a later date. The variables that have been discussed above have a significant impact on the level to which one can determine the impact that a purchase now pay later business plan has on the level of customer happiness, as well as the overall performance of the business. According to the findings that were discovered by applying our dataset, a large number of customers have a history of making payments precisely as intended, in contrast to the smaller number of customers who have not completed their payments precisely as intended. The frequency of the late payment status is represented by the figure below, which offers a clear indication of the number of customers who have paid for the things they purchased and the number of customers who have not yet made their payments.



### Figure 3: Above shows frequency of Later Payment Status of the customer

In a similar manner, we moved on and performed analysis using the variables "Annual income" and "total budget cost." The fact that yearly income and total budget cost have been discovered to have a correlation suggests that the customer's annual income has an effect on the total budget cost that the consumer incurs when shopping at the buy now pay later business. The graph gives the impression that the majority of clients do not spend more than twenty thousand dollars as their total cost budget while making purchases of goods from a store that offers "buy now, pay later" payment plans. In addition, the vast majority of customers who engage in the tactic of "buy now, pay later" have an annual income of approximately fifty thousand dollars (50,000), whereas only a minority of customers have a greater annual income of more than fifty thousand dollars (50,000). When compared to the number of male clients, the number of female customers is slightly higher. The following chart compares the total annual income of each consumer to the total budget cost that was spent on purchasing the goods.

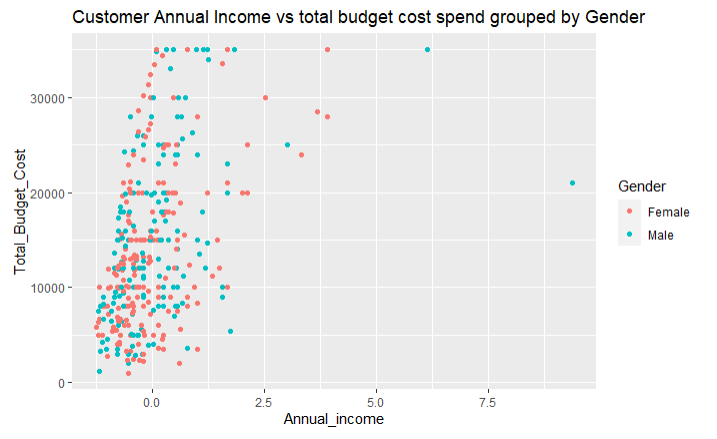


Figure 4: Above illustrates the customer Annual Income vs total budget cost spend by the customer grouped by Gender.

## 5.2 Research Objective 2

The purpose of this objective was to examine how the implementation of data analytics in purchase now pay later models will influence the success of businesses in Nigeria. This objective takes into consideration all of the dataset variables as well as how well they have done in our dataset. In our example, we conducted a number of analyses on a wide variety of variables in order to determine how well they performed in relation to the whole dataset. It was necessary to employ variables such as the combination of the "Buying For" variable and the "Gender" variables. The management of the company is able to determine which categories of products are preferred by which genders by carefully examining the graph that has been supplied to them. In comparison to male consumers, the female gender of the buy now pay later business's clients has a higher rate of satisfaction with all of the company's products and services, as shown by the graph. However, management is also able to determine which of the products now available for purchase performs better than the others. The values from our "Buying For" variable are depicted in the following picture, which compares them to "Gender" variables of various types.

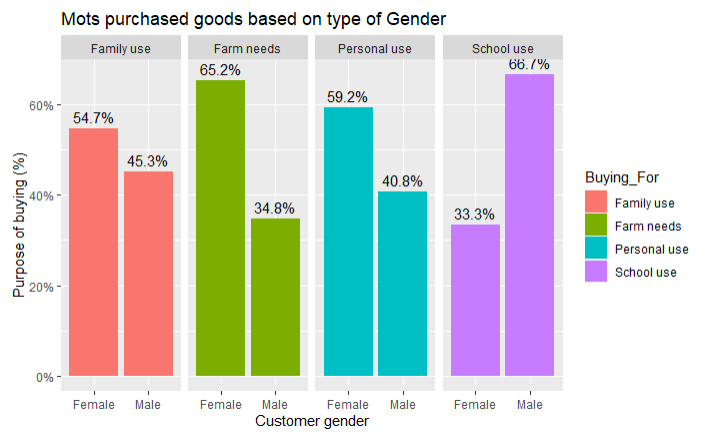
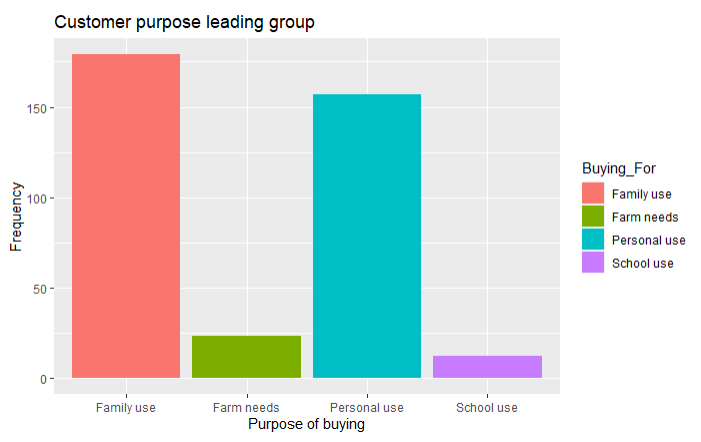


Figure 5: Above shows the “Buying\_For” vs “Gender” type of customers.

## 5.3 Research Objective 3

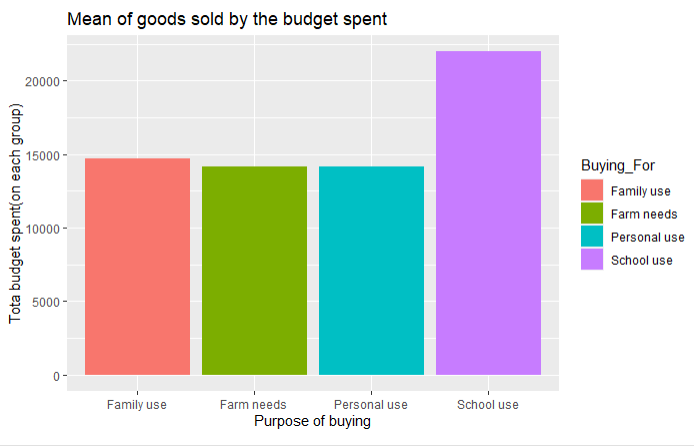
The purpose of this objective was to assess the manner in which buy now pay later companies in Nigeria make use of data analytics to shape their customers' attitudes and behaviors regarding knowledge management. In the field of data analytics, the practice of studying raw data in order to extract conclusions from it is known as "data mining." Data analytics can be thought of as both a field of study and a set of applicable abilities. In order to carry out data analytics, there must first be raw data consisting of relevant variables that are connected to one another and that contain a wealth of information. When an analysis is done on a business organization, having this kind of rich information that is related will lead to results that are relevant to the company organization. Because of this, it is clear that the company has a responsibility to make certain that the data records they collect contain all of the information that is pertinent and that can be required for carrying out data analytics in order to derive patterns that can assist in making the company more successful. Having said all of this, it is abundantly evident that it is the responsibility of the company to guarantee that well-planned and well-executed management procedures are utilized in the general business operations in order to accomplish all of this. When we look at the dataset for the Buy Now Pay Later business, we can see quite clearly some of the procedures that the company has put into place in order to improve data analytics. Variables that are good indicators of variables that play a significant role in the company data include things like the "employment title" of a particular client, the "annual income" of a customer, the "employment length" of each and every customer, and so on. There are obvious signs on how these variables considerably effect the primary goal of this business model (which is to earn enormous cost/revenue from the items sold) when analytics are done on these variables. The main objective of this business model is to make huge profit from the sale of goods. As a result of employing the graphs and plots generated from the company information by using data analytics, our method of management of the business has changed as a business organization. The graphs that were produced as a consequence of data analytics can be simply used by management in order to change the management techniques that they now use.

In case one, using the below diagram, one can tell the best performing goods form the one sold to the customers. Therefore, as a management, new strategies are set to improve on the same product as well as finding an alternative on how to improve the least performing goods.



### Figure 6: .The customer purpose goods leading group

In the second case, using the below diagram, one can tell the top goods that generate high revenue/income to the buy now pay later business after they have been sold to the customers. Therefore, as a management, new strategies are set also to improve on the same product as well as finding an alternative on how to improve the least goods that make less returns into the business.



### Figure 7: The above graph shows the Mean of goods sold to the customers against the budget spent on each group of goods.

# **CHAPTER SIX: DISCUSSION**

This section provides a clear discussion on the results outlined in the previous chapter in detailed. From the discussion made, suggestions and recommendations as well are provided in order to help the Buy Now Pay Later business organization make much informed business decisions. This overall paper provides us with three research objectives which we are aiming at. These research objectives complement the research hypothesis which we are also testing in order to accept or reject them.

## 6.1 Research Objective 1

The first research objective states that we have to analyze the impact of the buy now pay later business strategy on customer satisfaction and business performance. Specific dataset variables were considered in order to attain this objective. With the specific variables in use results were found. According to the results generated, It is evident that many customers who get services from the business organization of buy now pay later were very satisfied. What proves this is the ability of over 90% to customers to pay for the goods they bought later and remaining with only few customers who have not managed to pay back. Also, this is a good indicator of a developed trust between the customers and the business organization with its goods and services offered. The thing here is that, customers have liked the business model of buy now pay later. Its is evident that this business strategy can bring in more customers on board thus the business ends up making revenue out of it in the later payment days.

Further, by following the revenue generated as a result of payments made from the customers, its evident that almost all customers spend more. Goods taken by customers sums up to high budget cost. This means that, customers when using the buy now pay later strategy, find its easy to collect as many goods as possible which in turn sums up to a good amount for each customer.

To conclude on this objective, its clear that, by looking at how each and every customer responds back to the business through the way they make their payments later and the total amount of money they spend in this business as per individual shows that customers are well satisfied with this kind of business model.

## 6.2 Research Objective 2

According to the second research aim, one of our responsibilities is to analyze how the implementation of data analytics in buy now pay later affects the overall performance of businesses in Nigeria. After getting the dataset for buy now pay later in the right location, we carried out the analysis on it. We are able to form opinions based on the representations of plots and graphs, which are the same as those illustrated in the chapter before this one (chapter 4). The findings unequivocally demonstrate that the availability of data analytics makes the buy now, pay later business model an excellent choice. This is due to the fact that the purchase now pay later management, by making use of the graphs that are produced as a consequence of performing analytics on the dataset of their firm, are able to easily come to conclusions and make decisions that are well informed. To provide a bit more context, the visualizations that have been supplied highlight the essential information that is desperately need. With this knowledge, it is possible to determine whether the buy now pay later business model is reasonable in light of the expectations. In our particular instance, one of the several visualizations that were created showed that the girl was the better performer when compared to the guy (the most buying gender as well as the most paying gender). The management is now in a position to determine which gender is more successful in the buy now pay later business as a result of considering this information. As a result of this information, they are in a better position to formulate a strategy to bring in more alluring products that cater to the preferences of the gender that they have determined makes the most purchases. This is particularly useful in the business of "buy now, pay later," which creates a significant quantity of client and sales data on a daily basis. Without the assistance of data analytics in place, the management absolutely cannot be in a position to be able to make all of these adjustments. Therefore, in order for the purchase now pay later business to ensure the smooth operation of their own company, it is necessary for them to collect more pertinent data and incorporate data analytics.

## 6.3 Research Objective 3

According to the third and final research objective, we are required to conduct an analysis of the ways in which data analytics are applied in Nigerian buy now, pay later businesses in order to determine their impact on knowledge management practices. It should be obvious that data analytics refers to both the science and the abilities that are required to evaluate raw data in order to draw conclusions about the information. This indicates that data analytics can only be performed in environments where there is a presence of data, and not just any data but data that is relevant to the business in order to generate accurate results that the management of the business can use to improve the operation of the business smoothly. With this in mind, it is the responsibility of the businesses that offer buy now, pay later options to compile the necessary dataset whenever transactions take place so that it can be used in the step that involves data analytics. After then, data analytics will be performed on the information that was gathered from the firms that provide buy now, pay later options. As a consequence of this, crystal-clear visualizations with accurate labels will be produced. The management is responsible for going over these outcomes, which were obtained as a result of carrying out data analyses. Because the information is presented in a visual format, it is much simpler to comprehend what is going on in the company in general. For instance, in our case, a great deal of graphic content has been produced. One of them was a graphic representation of the group of products that were purchased the most frequently by our clients. The products that fell under the family category were the ones that were purchased the most. On the other hand, there was not a significant difference between it and the other items. When an organization has access to the kind of information described above, it is simple to generate opinions. As a result, we as a team are kept abreast of the most recent developments. As a result, we create fresh understanding regarding the management of the firm.

## 6.4 Limitations

In spite of the fact that I was able to successfully analyze my dataset for the Nigerian company BSPL, I was limited in a number of other respects. For this reason, the study could only make use of historical data provided by Nigeria's BSPL business company. I'm basing what I'm doing on the historical records of the Nigerian company BSPL, which is in the business of oil and gas. Even though this is not a negative development, the execution of this data analysis project could have been enhanced if we had used data that was collected in real time. In order to conduct data analysis, we make use of a particular kind of data that is continuously produced by the BSPL business company in Nigeria.

## 6.5 Conclusion

In conclusion, buy now and pay later businesses generates large volume of dataset each day whenever transactions take place with its customers. As a result, there is need to make this large volume of dataset useful and effective back to the business itself. Thus, here comes in data analytics, a mechanism that helps in finding patterns out of the raw data in order to help the business make informed decisions. This means that data analytics and the buy now and pay later businesses depend on each other. If data is generated without data analytics, it becomes useless to the business. Therefore, it’s evident that, there exist a positive relationship between data analysis and the buy now and pay later businesses. Thus, we accept the null hypothesis.

Further, data analytics provides clear and informed visualizations to as a result from the raw data back to the management. The management of the business are now in a position to go through the visualizations produced and by doing this, new patterns are realized form the results generated. These patterns play a vital role by inserting new knowledge in the management team which later come up with new and different strategies in order to improve and make necessary adjustments so that revenue goes high. With this kind of information, its evident to say that there is a positive correlation between data analytics and the knowledge management practices.

Finally, knowledge plays a vital role in the whole process of managing the buy now pay later business. As we have seen form the visualizations displayed in the results section, we get knowledge by looking at them. With the knowledge that we get, we become creative by coming up with new strategies on how to make our business effective and more profitable than before. Knowledge helps the management team to realize faults, weakness and strengths in the business operation, thus vital. With the knowledge found, we can know how to prevent/mitigate the faults, strength the weakness found and maintain and improve the strength realized. Therefore, there exists a positive relationship between the knowledge management practices and the performance of the buy now pay later businesses in Nigeria.

## 6.6 Recommendations

I do recommend that, buy now pay later business can thrive best whenever data analytics mechanism is applied in the business operation. Therefore, business operating under buy now pay later business model should consider integrating data analytics in the business in order to increase the business income whenever later payment comes from the customers. Knowledge plays a vital role in enhancing the effectiveness of the business. The fact that data analytics is the best place to generate knowledge required, thus indicates that the business still should consider data analytics again and again as the source of thriving their business. Therefore, data analytics is the way to go. I also recommend that, buy now pay later business should consider collecting data using the relevant variables. Variables that have no relation when used during data collection will generate raw data that cannot derive accurate patterns to thrive this business since it is the relation of these variables that leads generate the patterns that is needed in the business operation