

Name : MARIUS MOLDOVEANU

Student Number : 2023028

**Module Title : DATA HANDLING & BUSINESS
INTELLIGENCE**

Assignment Title : Data Analysis & Visualisation

Module Code : LCBB5000

Module Lecturer's Name : Anupam Mazumdar- GP1

Amitave Banik- GP2

Executive Summary

The concerned report identifies the proper methods and techniques of the business and data handling process based on its sufficient data and planning tools. However, it is found in Part 1 which is totally based on product performance and its quality approach through determining the proper usage of excel functions. Moreover, part 2 mainly deals with proper analysis of data and tool through using of SPPSS functions consist of various test. Throughout the test, it is found that there are multiple variables as ice cream, puzzles, and video games as per the choice and significance of students. Thus, the results highlighted that the majority of student like to choose the ice-creams based on various flavors.

Table of Contents

Part 1	4
Introduction.....	4
Body	4
Category-wise product sales	4
Quantity-wise product sales.....	5
Unit price-wise product sales.....	6
Shopping costs-wise product sales	7
Profit-wise product sales	8
Pre-processing data	9
Conclusion	10
Part 2	11
2.1 Identify the number of respondents who like vanilla flavors ice cream?	11
2.2 What are the majority of males and females?	12
2.3 Median and Mean of the respondents who like to choose the strawberry and chocolates flavored ice cream?	13
2.4 variance of K-Mean.....	14
2.5 Data mining.....	15
2.6 Merits and Demerits of Excel and SPSS	16

Part 1

Introduction

Data on sales and profit have to be considered for data regression as well as calculation accurate value of profit and nonprofit period which is supported to take effective division as well as implement strategy in business. This is a statistical method which is considered by different businesses as well as an entity to recycle and interpret past sales and profit data to identify different factors and elements which have affected sales profit.

Body

This part of the report is mainly focused on different factors and elements of data which has affected business revenue as well as the profitability of past business operations. There are different types of variables such as sales units of the sales price as well as units that have affected profit margin and numbers of product sales. Besides that, product category and discount on product has also affected profit margin through increased and decline in profit margin of the product. As stated by Sambrani and Pol (2019, p.161), the data is acquired on sales and interest to buy different category products related to the UK population. The use of excel function as well as other techniques for data interpretation has mainly supported considering different analytical factors and business intelligence for better observation and considering multiple business analytics. Here mainly *If function* has been considered to observe business intelligence as well as steps which have affected business sales and revenue growth. The *If function* has also supported the data analysis process through the use of different graphs and PIVOT tables. Besides that, it is an authentic process for data manipulation as well as various preferences used to obtain positive data such as enhancement in margins and sales unit times.

Category-wise product sales

There are mainly three types of products that have been sold in different three types of products such as furniture, office suppliers and technology. The company has sold nearly **14915600.82** in the last year. The total sales amount of product consists of three types of product such as furniture of **5178590.542** and office suppliers of **3752762.1**. Here furniture has been more sold as compared with office suppliers. The main reason for more sales of furniture is it is a household product and used by different levels of consumers who either have business units or are used for household purposes. On the other hand, the total sales

volume of technology exists at **5984248.182**. The sales value may be affected by different factors such as the requirement and price of the product. As technology is a common need for different levels of individuals as well as it has support to minimise effort as well as costs to perform any operations. The below graphs consist of sales of units through different products and value of three products grand total.

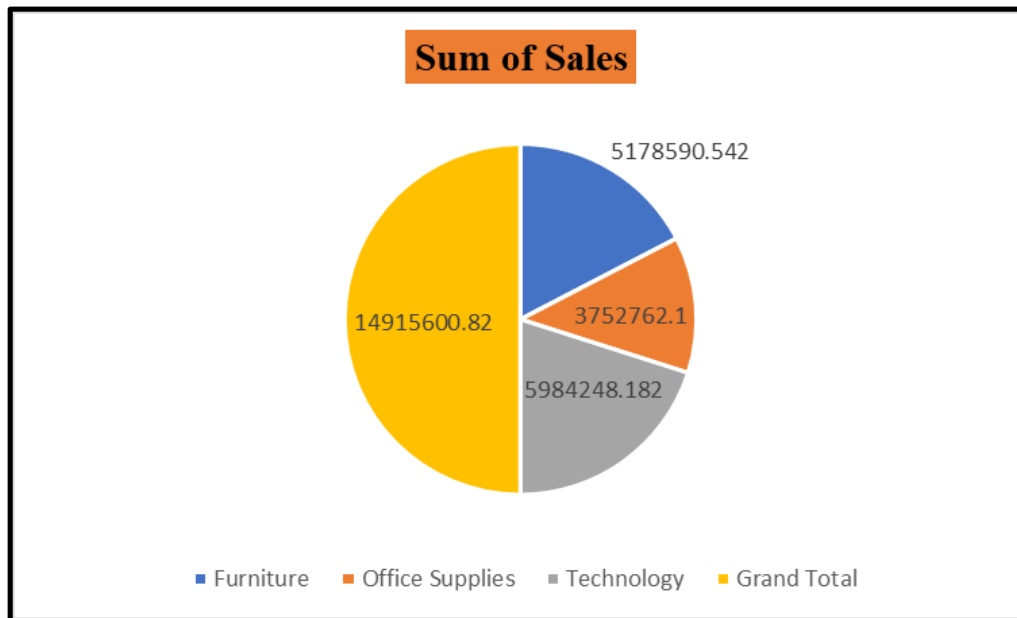


Figure 1: Sales units of three products
(Source: Self-created)

Quantity-wise product sales

Product quantity data has been also manipulated with If as well as other manipulation techniques of excel to identify different elements that affected the sales price of the product. As opined by Singh and Kumari (2020, p.162), here another manipulation technique such as ANOVA and regression has been considered by analysts to reflect different other factors. As argued by Galetsiet *al.* (2020, p.209), these two data handling techniques have also played an important role to execute the mean and median values of sales and costs which has affected business performance as well as product performance. Product performance is mainly affected by mean and median costs and the price of the product nevertheless, the performance of the business has been executed through the use of the total mean and median value of different products.



Figure 2: Sum of order quantity

(Source: Self-created)

Order quality refers to a number of products which are ordered by different levels and types of consumers to the organisation within a specific period of time. Here data which is manipulated to evaluate actual numbers of ordered products has been collected through the company's past records as well as the current demand for products in the market. At the current time, the company handmade sales orders of **214780** quantities from three types of products. Among these products, office suppliers are in higher demand of the product at the volume of **118281**. Besides that, technology has the second highest order quantity at a volume of **52175**. Therefore, the total order quantity of furniture gas exists at **44324**. Thus the demand for the ordered furniture is low in the market as compared with other products which are ordered by different levels of consumers.

Unit price-wise product sales

The unit price has been also considered for the purpose of data manup[ulation as well as implementing If and count If techniques. The unit price of technology is high with a value of **374609** is the highest value among the other two product prices. As argued by Schereskyet al. (2021, p.114), the main reason for increasing the price of technology is the high demand for products in the market as compared with other products. As in current times, technology is more demanding for different purposes such as business as well as household. Besides that,

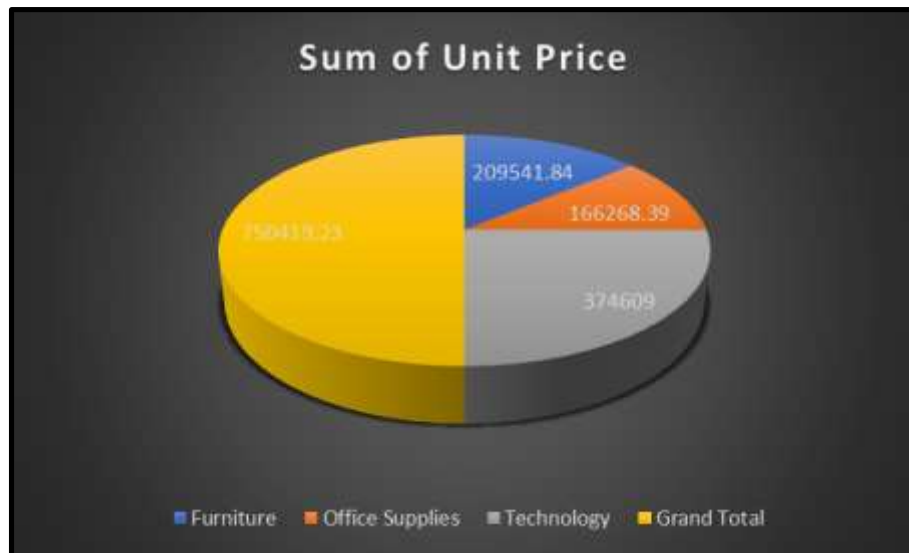


Figure 3: Sum of unit price
(Source: Self-created)

Shopping costs-wise product sales



Figure 4: Sum of shipping costs
(Source: Self-created)

The unit price of the different products has also been manipulated to execute different factors and elements to ensure data technique. The data technique is implemented to execute the price of one product which is affected by another product. As opined by Chen *et al.* (2018, p.315), the price of products has been mainly affected by costs as well as other resources which have been implemented to promote products as well as sales in the market. Among the three products, furniture has the highest costs compared with other products in the market. As shipping costs have also affected product sales. Thus, it is necessary here to consider the shipping costs of product manipulation and interpretation with other techniques to develop

decisions regarding product sales. Shipping costs of furniture exist at **53243.69** which is nearly 50% shipping costs of the rest of the two products. Apart from this, other factors such as office suppliers have total costs at **36095.51** which is 2nd highest in shipping costs among the other two products. As opined by Alambeigiet *al.* (2018, p.261), excel functions such as ***If and Count If*** have played significant roles to execute which product has consumed more shipping costs and selecting alternative sources for transporting products from one location to another location.

Profit-wise product sales

Profit is a key data factor while a business industry wants to take decisions as well as manipulate data to make decisions for launch as well as promotion in the market. As profit has not only a substance product life cycle nevertheless, also sustained business operating in the market. On the other hand, profit of different products as well as aggregated profit is also required for data manipulation and division. As argued by Hoque *et al.* (2020, p.214), it has been found from the data handling process that 38 times the product has been sold at profit. Besides that, **8361 times** have sales at loss. This data suggests that the company should focus on better performance as well as enhance its strategy to promote products in the market to generate profit in each sale.

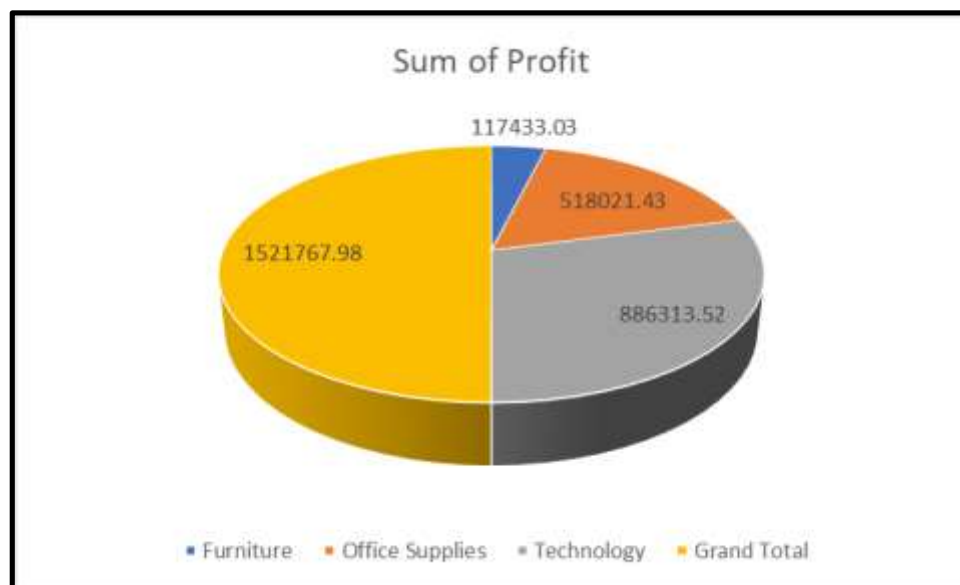


Figure 5: Sum of profit

(Source: Self-created)

The company has generated a total profit of **1521767.98** which is acquired from three products of the company which are sold at different profit percentages as well as prices.

Technology has given the highest profit to companies as compared with other products which are sold in the market. The profit of technology exists at **886313.52** which is nearly 65% of total sales in amount. On the other hand, furniture has given low profit to companies with a value of **117433.03**. Thus, the office supplier has the best profit of the company as it has given an average profit of **518021.43**. This it has executed that, the company has able to d

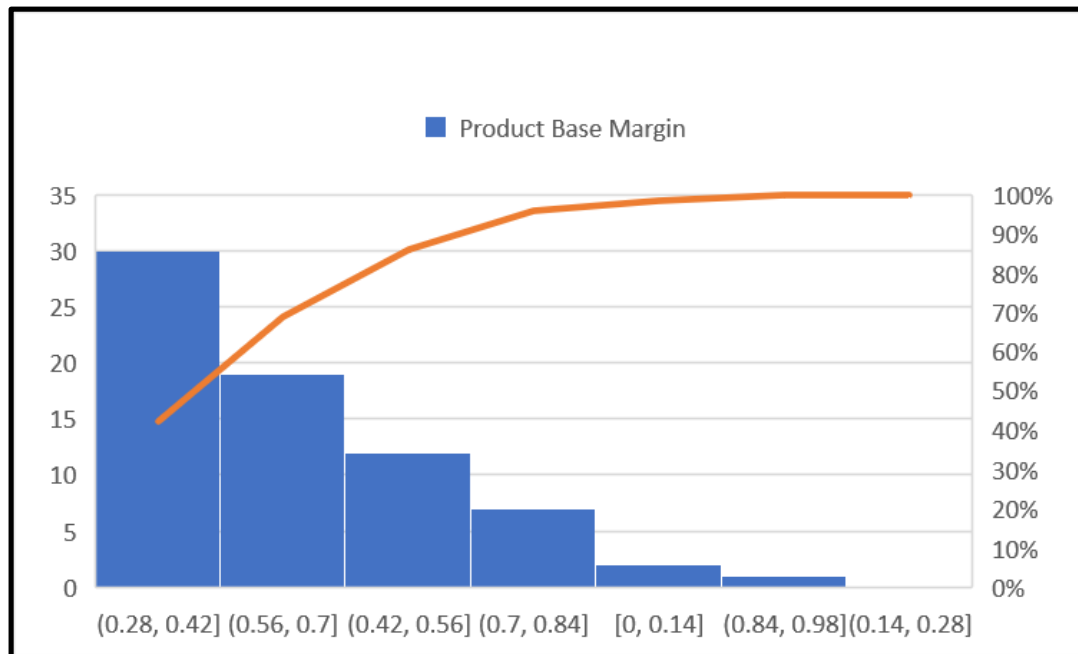


Figure 6: product based margin

(Source: Self-created)

The product-based margin of the different products has declined due to more time companies have sold different types of products at loss. As argued by Sadgaliet *al.* (2019, p.49), the company has sales maximum time with low price in order to enhance its sales quantity. Nevertheless, only furniture products have better order quantity as compared with other products such as technology and sales.

Pre-processing data

Data pre-processing is an innovative tool and technique to determine the profit benefits as it usually aims to transform the business through the transformations of raw data within the business. Thus, the entire data is procured through the superstore product and its value chain creation techniques. Hence, it involves proper order id, quantity and selling units to recognize the profit within the business. In such cases, it can be determined that multiple tools and techniques have been used to evaluate the overall stock and selling data of superstores (Jalil and Hwang, 2019, p.15). Moreover, it can be identified that the proper usage of *the PIVOT*

Table has been used to evaluate the performance of data regarding the products and their benefits.

Performance of products

In the context of product performance, the processing of data and tools that handle the overall business has been evaluated by comparing the representation and possibilities regarding the quantity price. Consecutively, the rate of product and its efficiency is based on entire profit and its sales returns. In such a case the PIVOT table and proper analysis of If functions mainly help to determine whether the product has declined or generates any profits within the business based on each quality product (Castro *et al.* 2019, p.39).

However, the graphical representation identifies the product performance and category-wise sales based on the overall product record. Thus, it can be identified that there has been a huge quantity of products that have been sold based on its numerous categories and product efficiency to generate market returns. Therefore, it needs more evaluation of data through proper usage of excel function rather than the pivot table and if functions. Therefore, the count if function mainly determined the overall profit which is 38 and the declination percentage is 8361. Hence, it needs more effective information through source and gathering data regarding the product and its category of product.

Conclusion

It has concluded that different factors of production such as price as well as costs such as shipping costs have affected product life cycle and performance. Besides that, it has been found from different analytic techniques such as ANOVA and regression stating that furniture has high product sales nevertheless low-profit units. Technology is also a product of a company that has given better returns as compared with other products due to better demand for products in the market.

Part 2

Introduction

Data analysis is an important function that helps business leaders to optimize their performances and measure their financial opportunities. Though, it consists of various business plans and implementations that help the company to reduce its operating cost and effectiveness within the business. This can be done by evaluating the huge amount of data.

Discussion

It can be discussed that the entire survey has been done on the basis of evaluating the basic understanding and interests of the respondents. Hence, it needs the proper evaluation and fundamental approach through using the proper **SPSS** functions and inputs to determine the effectiveness of the respondents (Niño *et al.* 2020, p.41). However, it is more effective through using the multiple variables and input to determine the interest of the respondents regarding video games, ice cream, and puzzles.

2.1 Identify the number of respondents who like vanilla flavors ice cream?

It can be identified through the use of a fundamental approach through the SPSS regarding the majority of students who like vanilla flavors ice cream as more than 90% (Sun *et al.* 2020, p.32). Hence, it can be identified as illustrated below:

ice_cream					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Vanilla	95	47.5	47.5	47.5
	Chocolate	47	23.5	23.5	71.0
	Strawberry	58	29.0	29.0	100.0
	Total	200	100.0	100.0	

Figure 7: frequency of vanilla Flavored ice cream

(Source: created by the learner)

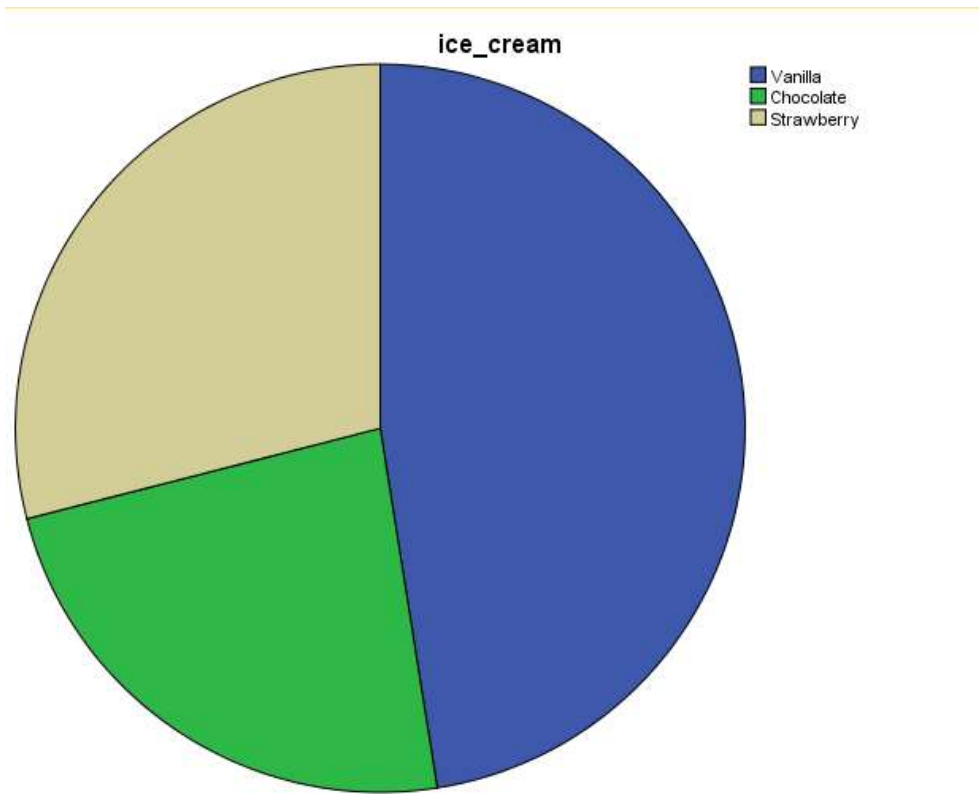


Figure 8: The majority of student who likes the vanilla flavors ice cream

(Source: created by the learner)

Therefore, it can be observed through the above-mentioned representation of *the Pie chart* that the majority of students who like vanilla flavored ice cream are 95 as compared to the other flavored ice cream. However, it can be determined that the majority of ice cream flavors consist of several tastes and choices. Although, the review and rating for the vanilla flavors ice cream are huge as compared to the rest of the flavors.

2.2 What are the majority of males and females?

		Gender			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	91	45.5	45.5	45.5
	Female	109	54.5	54.5	100.0
	Total	200	100.0	100.0	

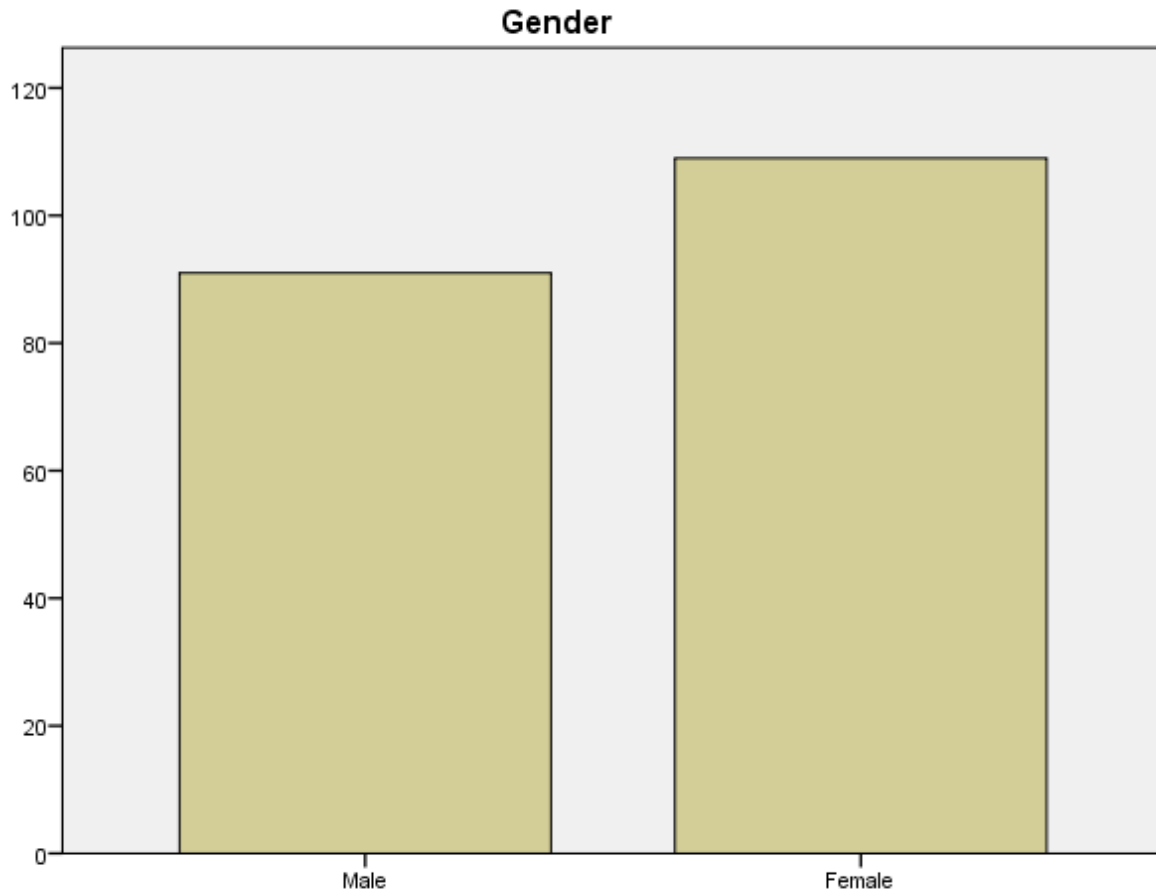


Figure 8: Participation of Gender

(Source: created by the learner)

Gender participation plays a huge role to determine the different choices as per age, sex, and other demographic segments. However, it can be observed through the above table which reflects the majority of male and female respondents (Ahmad *et al.* 2020, p.33). However, the entire survey contains the minimum number of male participants as compared to the female participants. Thus, it can be observed that the female participants are greater than the male participants such as 109. Therefore, it can be identified that the majority of the male is lower than the females because their choice and taste are totally different from the female respondents. Thus, the overall survey has propounded the participation of genders.

2.3 Median and Mean of the respondents who like to choose the strawberry and chocolates flavored ice cream?

Statistics

ice_cream

		Statistic	Bootstrap ^a			
			Bias	Std. Error	95% Confidence Interval	
					Lower	Upper
N	Valid	200	0	0	200	200
	Missing	0	0	0	0	0
Mean		1.82	.00	.06	1.69	1.93
Median		2.00	-.23	.41	1.00	2.00

Figure 9: The majority of respondents who like strawberries and chocolates

(Source: created by the learner)

As per the entire discussion based on the participants who like their respective ice cream flavor's it can be determined by evaluating the mean, and median, which consist of several flavored ice creams. However, it can be identified that the overall discussion has been done through the help of SPSS functions and analysis (Yeung *et al.* 2020, p.120). However, it can be observed through the above-mentioned data in the figure based on the majority of respondents who like the strawberry and chocolates flavored ice cream has been discussed through the rate of frequency and statistics. Hence, it consists of the mean, and median of the students. Thus, it can be determined that the mean null hypothesis value of the respondents has been identified as **1.82 and the median value as 2.00**. Thus, the median components are likely to choose the strawberry flavored ice cream with the overall respondents as more than the 90%.

2.4 variance of K-Mean

Identify the data set

ANOVA						
		Sum of Squares	df	Mean Square	F	Sig.
ice_cream	Between Groups	.027	1	.027	.037	.847
	Within Groups	146.128	198	.738		
	Total	146.155	199			
video	Between Groups	318.337	1	318.337	3.285	.071
	Within Groups	19189.163	198	96.915		
	Total	19507.500	199			
puzzle	Between Groups	62.905	1	62.905	.545	.461
	Within Groups	22873.290	198	115.522		
	Total	22936.195	199			

Figure 10: Test based on ANOVA

(Source: created by the learner)

In the context of cluster analysis, it can be essential in terms of identifying the various applications and their usage is done by the researcher and many marketers. However, it helps to improve the data analysis of the variables, research patterns and development. Moreover, it can be identified through the above-mentioned representation regarding the clustering analysis which is mainly based on respondents and the majority of students (Vallurupalli and Bose, 2018, p.74). This is because of the group-based buying patterns. Hence, it provides an accurate representation of segments in order to achieve more information through the personalization's. Therefore, the figure represents the test based on ANOVA in order to gather the information regarding the majority of constraints and students through the overall respondents. It signifies the mean and median value of the respondents based on frequencies as gender, significance consisting of ***different groups and within groups***. However, in terms of ice creams, it identifies the significant value as .847 which is a positive value and closer to the value of 1. Similarly, in the context of video games, the significance value is determined as .071 and the puzzle signifies as .461 (Zhong *et al.* 2020, p.96). Hence, except for the ice cream, the videos and puzzles are not positively skewed as compared to the null hypothesis. Furthermore, it can be illustrated that the main significance between the various groups and responses evaluate the degree of independent and dependent variables. However, it is totally dependent upon the different variables.

2.5 Data mining

Data mining mainly helps the overall credentials and fraud to analyses the consumer data and financial structure regarding the purchase methods and various modes of transactions. Moreover, data mining is significant in terms of identifying a better understanding between the various tastes and choices of the participants in terms of variables and preferences (Carvalho, 2021, p.12). Thus, it helps to design the new research campaign.

Similarly, it consists of various possibilities and quality resources based on smart technology and initiatives. This is because of gaining the potential outcomes as per the survey or findings of the result. Hence, it operates unpredictably to optimize multiple data techniques through the process of data handling.

Various data mining initiatives and factors have been identified in an affirmative way such as Regression, classification, and Clustering techniques.

Furthermore, the entire data handling and mining techniques mainly identify the proper ways to evaluate the essential source and integrate the positive relationship between the participants and the company operations managers (Yusof *et al.* 2019, p.12). Therefore, an

example has been evaluated as **Mcdonald's** uses an appropriate data mining system to identify the consumer efficiency and target within their business. Moreover, the **Retail sector** is the best source and examples of using data mining techniques within the business to identify opportunities to enhance their productions. However, the concerned companies properly used their data mining techniques to evaluate the knowledge and value principles or information based on data to attract or consume more target market opportunities.

2.6 Merits and Demerits of Excel and SPSS

<i>Functions</i>	<i>Advantages</i>	<i>Disadvantages</i>
SPSS	1. It helps to determine the faster access and quick service to support or acknowledge the specific research findings rather than the excel 2. It identifies an ability to identify through providing more support by the functions in order to evaluate the data. 3. It operates through the help of value proposition techniques as per the information based.	1. SPSS mainly have certain limitations in terms of time consumption. 2. It involves multiple issues and uncertainties regarding the data as dependent and independent variables.
Excel	1. It consists of various calculative features and analyses to gather proper information. 2. It can provide authentic sources and graphs regarding the data.	1. Excel functions include various issues and risks based on cost and capital structure. 2. These functions cannot be well designed by the various work efficiency and collaborative culture.

Table 1: Merits and Demerits of SPSS and Excel

(Source: created by the learner)

Conclusion

The overall presentation of SPSS represents the multiple variables and amount that helps to evaluate the proper data inspections. However, it can be determined that the entirety helps to proper handling of business intelligence. Though, it can be determined that the proper utilization of data interpretation has been done through the help of data sets and business handling techniques such as data mining. Thus, data mining helps the overall study to formulate the techniques and clarification regarding the taste and choice of respondents based on variables. Therefore, it can be more informative and innovative by taking the smart interpretation by adding the graphs, and charts within the entire analysis.

Reference list

Ahmad, S., Miskon, S., Alkanhal, T.A. and Tlili, I., 2020. Modeling of business intelligence systems using the potential determinants and theories with the lens of individual, technological, organizational, and environmental contexts-a systematic literature review. *Applied sciences*, 10(9), p.3208.

Alambeigi, F., Wang, Z., Hegeman, R., Liu, Y.H. and Armand, M., 2018. Autonomous data-driven manipulation of unknown anisotropic deformable tissues using unmodelled continuum manipulators. *IEEE Robotics and Automation Letters*, 4(2), pp.254-261.

Carvalho, M.J.T.B.D., 2021. *Monitoring supplier deliveries and analysis of potential risk factors in supply chain management: a business intelligence and analytics approach* (Doctoral dissertation).

Chen, C., Chen, H., He, Y. and Xia, R., 2018. TBtools, a toolkit for biologists integrating various biological data handling tools with a user-friendly interface. *BioRxiv*, 289660(10.1101), p.289-660.

Cheng, C., Zhong, H. and Cao, L., 2020. Facilitating speed of internationalization: The roles of business intelligence and organizational agility. *Journal of Business Research*, 110, pp.95-103.

Galetsis, P., Katsaliaki, K. and Kumar, S., 2020. Big data analytics in health sector: Theoretical framework, techniques and prospects. *International Journal of Information Management*, 50, pp.206-216.

Hoque, R., Seita, D., Balakrishna, A., Ganapathi, A., Tanwani, A.K., Jamali, N., Yamane, K., Iba, S. and Goldberg, K., 2020. Visuospatial foresight for multi-step, multi-task fabric manipulation. *arXiv preprint arXiv:2003.09044*. pp. 201-218

Jalil, N.A. and Hwang, H.J., 2019. Technological-centric business intelligence: Critical success factors. *International Journal of Innovation, Creativity and Change*, 5(2), p.1499.

Moscoso-Zea, O., Castro, J., Paredes-Gualtor, J. and Luján-Mora, S., 2019. A hybrid infrastructure of enterprise architecture and business intelligence & analytics for knowledge management in education. *IEEE Access*, 7, pp.38778-38788.

Niño, H.A.C., Niño, J.P.C. and Ortega, R.M., 2020. Business intelligence governance framework in a university: Universidad de la costa case study. *International Journal of Information Management*, 50, pp.405-412.

Sadgali, I., Sael, N. and Benabbou, F., 2019. Performance of machine learning techniques in the detection of financial frauds. *Procedia computer science*, 148, pp.45-54.

Sambrani, V.N. and Pol, N., 2019. Percolation of Data—An In Depth Analysis of UNIBIC Sales Dump. *IJRAR-International Journal of Research and Analytical Reviews (IJRAR)*, 6(2), pp.157-162.

Scheresky, G., Wilson, W.W. and Bullock, D.W., 2021. *Spatial and Logistical Competition for Shipments to China from the United States and Brazil* (No. 1186-2022-968).

Singh, J. and Kumari, M., 2020. 9 Agricultural/biomass waste management through “green supply chain way”: Indian “brickfield” perspective. *Supply Chain Sustainability: Modeling and Innovative Research Frameworks*, 2, p.157.

Sun, S., Zheng, X., Villalba-Díez, J. and Ordieres-Meré, J., 2020. Data handling in industry 4.0: Interoperability based on distributed ledger technology. *Sensors*, 20(11), p.3046.

Vallurupalli, V. and Bose, I., 2018. Business intelligence for performance measurement: A case based analysis. *Decision Support Systems*, 111, pp.72-85.

Yiu, L.D., Yeung, A.C. and Jong, A.P., 2020. Business intelligence systems and operational capability: an empirical analysis of high-tech sectors. *Industrial Management & Data Systems*, 120(6), pp.1195-1215.

Yusof, N.I., Zainuddin, N.M.M., Hassan, N.H., Sjarif, N.N.A., Yaacob, S. and Hassan, W.A.W., 2019. A guideline for decision-making on business intelligence and customer relationship management among clinics. *International Journal of Advanced Computer Science and Applications*, 10(8).

