Pub Survey Analysis : Helping Pubs To Enhance Their Business

Sanket Dilip Dayama
x18143652
MSc. Data Analytics
National College of Ireland

***Abstract*— A pub is a place where people meet for different purposes, which includes hangout with friends, corporate parties, meet ups and screenings of sports. There are different types of people who enjoy spending time and money in pubs which is a business perspective for many pubs. A lot of factors depend on revenue generation for this type of business as choices of people differ according to their lifestyle and occupation. It is a challenge or a task to retain the loyal customers and also enhance the business simultaneously. A study on the reason for visiting a particular pub by the people is necessary, so that the expectations and behaviour of spending of the customers can be used to implement some strategies to retain and increase the business. This study analyses and tries to identify the parameters which reveal the behaviour of customers. This will generate key insights upon which different strategies can be developed to increase the revenue. The relation between different parameters which can increase the business of pubs has been tried to find out using different techniques like identifying the correlation and significance of the parameters affecting it using IBM SPSS and graphically visualising them into Tableau software. The study shows that there is a relation between males and females on the frequency of going into pubs. Also, some insights are gathered on different relations formed between the parameters studied.**

Keywords : Pubs, business, behaviour of customers and key insights.

I. INTRODUCTION

The term public house in the past which are called pubs now had the main purpose of serving drinks and food as their business. Previously, pubs were present as a networking place in village areas. The sole purpose of people was to meet over drinks and discuss trending news and topics related to daily life. In recent years, party culture and night life has increased all over the world. This made the pub business evolve their establishment into different types of themed environments, sports bars and other amenities[1]. The pubs create a lot of business over people visiting and spending time at these places. The most important phenomenon for this business is to satisfy customers with the right environment, identifying their needs and serving them according to their preferences. There is a need of keeping the people engaged to the place so that the behaviours and trends can be identified. The parameters on where customers prefer to spend should be monitored as it would be a key insight for growth.

An analysis on human engagement to these places will uncover the facts of regular customers and what is it that people prefer the most when entering that place[2]. This business runs on the preference of the people as there is plenty of competition, people seek an active environment where the can spend time and enjoy connecting with others. The primary aim of this study is to analyse the factors which can help to retain the current business and also finding the relation between the parameters which can help the business to increase and make it profitable. This analysis also can help with pubs spending on the breweries or buying alcohol from third parties, the environment and organising events smartly by making avoiding unnecessary losses.

II. HYPOTHESIS

The null hypothesis for this study is the frequency of the people going to pubs remains the same by considering the parameters like types of people on different occasions and their income.

III. LITERATURE REVIEW

Author Andrew Baker discusses about an analysis done on pubs in England. A survey consisting of 56 questions was created and which would determine the attachment of the people with the particular pubs as to find how often customers visit that place. This method was applied on analysing the behaviour in pubs because of the admirable work done by Lansangan who made this approach on finding the attachment of the customers to coffee places[3]. The use of chi square test for independence and ANOVA was used to find the significance and identifying the frequency of visiting the pubs.[4]

Another approach made by author Tutenges was in terms of collecting the ethnographic data from danish cities and to find the behaviour and reasons of drinking alcohol which helps in increase in sales of alcohol in pubs. In order to analyse and for finding the relation, two things were focused on which were the pubs should be youth oriented and consisting of themed and of traditional ambience. An analytical software tool Nvivo was used to analyse the collected data. The results found were pubs with larger venues and dance floors had more sales than the ones having traditional ambience. It was also found that advertisements also helped to increase the sales of the alcohol by implementing schemes like happy hour and organising musical events[3]. The increase sales of alcohol is directly proportional to increase in business of the pubs.

A study on customer behaviour done by Valeria Anttonen on finnish people visiting the pubs gives an idea of reasons to visit pubs or spending behaviour of people of different lifestyle and in different age groups. The data was gathered by doing a survey. The analysis was performed using excel and important insights were gathered for people visiting pubs which resulted in spending good time and watching sports turned out to be the maximum in percentage. The people with working lifestyle than a student spent more on the quality of alcohol than on taking discounted meals.A granular analysis such as type of beers ordered and beverage interests were also monitored [5].

IV. Dataset

The dataset used for this study was created by doing an online survey which included the questions of customer behaviour, psychology and information related to spending and their preferences and was downloaded from kaggle.[6] It consists of different attributes which includes the spending style of the customers and the types of people contributing to the pub revenue. The dataset consists of frequency as a dependent variable and gender as dichotomous variable with number of categorical variables along with continuous variables. The metadata for different attributes present is as follows,

1) Timestamp : It consists of time and date of survey answered.
2) Age : Age of the people visiting pubs.
3) Gender : The gender of people going to pubs.
4) Income : The income of people by occupation.
5) WTS : The willingness of people to spend in pubs.
6) Primary Imp : Primary importance given by people while choosing to go to pub.
7) Sec Imp : The secondary importance of people for visiting a pub.
8) Lifestyle : The lifestyle of the person
9) Occasion : preferred occasion for visiting the pub.
10) stratum : The area where the people come from that is urban, rural and capital.
11) Occupation : Occupation of people visiting pubs

Timestamp	Age	Gender	Income	Occupation	Fav_Pub	WTS	Freq	Prim_Imp	Sec_Imp	Stratum	Lifestyle	Occasions
2017/02/25 10:52:03	19	Male	100000	Student	Station	2000	Several times in a mon	Environment	Menu	Capital	Nightlife	Hang outs with friends
2017/02/25 10:53:19	19	Female	50000	Student	Calumet	2000	rarely (once two week	Music	Pricing	Capital	Adventure/traveling/exploring	Hang outs with friends
2017/02/25 10:54:05	20	Male	100000	Student	Liberty	3000	rarely (once two week	Environment	Music	Capital	Busy(student life, work)	Hang outs with friends
2017/02/25 10:55:09	18	Male	0	Student	Calumet	3000	Several times in a mon	Environment	Music	Capital	Art	Hang outs with friends
2017/02/25 10:55:39	19	Male	90000	Student	Station	2000	Several times a week	Environment	Menu	Capital	Adventure/traveling/exploring	Hang outs with friends
2017/02/25 10:55:41	18	Female	0	Student	Irish	1000	Several times in a mon	Environment	Music	Capital	Art	Hang outs with friends
2017/02/25 10:55:45	19	Female	0	Student	Station	2000	Several times in a mon	Environment	Pricing	Capital	Busy(student life, work)	Hang outs with friends
2017/02/25 10:56:45	19	Male	150000	Student	Irish	3000	Several times in a mon	Environment	Music	Capital	Nightlife	Hang outs with friends
2017/02/25 10:57:00	18	Female	0	Student	Calumet	2000	Several times a week	Environment	Music	Capital	Adventure/traveling/exploring	Hang outs with friends
2017/02/25 10:57:58	19	Male	200000	Working	VOID	5000	Several times in a mon	Environment	Pricing	Capital	Nightlife	Special events/parties
2017/02/25 10:58:12	19	Male	120000	Student	DAS	2000	Several times in a mon	Environment	Pricing	Urban	Art	Hang outs with friends
2017/02/25 10:58:19	17	Female	30000	Student	Station	5000	Several times in a mon	Music	Pricing	Capital	Adventure/traveling/exploring	Hang outs with friends
2017/02/25 10:59:18	18	Male	0	Student	Station	2000	Several times in a mon	Environment	Music	Capital	Sport	Hang outs with friends
2017/02/25 10:59:28	17	Female	0	Student	37 pub	5000	rarely (once two week	Environment	Pricing	Urban	Art	Hang outs with friends
2017/02/25 10:59:59	17	Male	0	Student	37 pub	7000	rarely (once two week	Environment	Music	Capital	Busy(student life, work)	Hang outs with friends
2017/02/25 11:01:22	19	Male	0	Student	Venue	10000	Several times a week	Environment	Music	Capital	Nightlife	Hang outs with friends
2017/02/25 11:02:07	21	Male	130000	Student	37 pub	2000	rarely (once two week	Environment	Pricing	Urban	Sport	Hang outs with friends
2017/02/25 11:02:28	18	Female	50000	Student	Cantaloup	2000	rarely (once two week	Music	Environm	Capital	Adventure/traveling/exploring	Hang outs with friends
2017/02/25 11:05:04	18	Female	40000	Student	Irish	2000	Several times in a mon	Menu	Pricing	Capital	Art	Hang outs with friends
2017/02/25 11:05:10	20	Female	100000	Student	Cantaloup	2000	Several times in a mon	Environment	Menu	Capital	Busy(student life, work)	Hang outs with friends
2017/02/25 11:05:29	18	Female	80000	Student	Irish	5000	rarely (once two week	Environment	Menu	Capital	Adventure/traveling/exploring	Hang outs with friends
2017/02/25 11:05:32	18	Male	25000	Student	Station	3500	rarely (once two week	Menu	Music	Urban	Adventure/traveling/exploring	Special events/parties
2017/02/25 11:06:08	20	Male	60000	Student	Cantaloup	20000	rarely (once two week	Environment	Music	Capital	Art	Hang outs with friends
2017/02/25 11:07:08	19	Female	0	Student	Calumet	4000	rarely (once two week	Environment	Pricing	Capital	Art	Hang outs with friends

Fig. 1: Dataset used

V. Methods and Implementation

A. Data pre-processing

Data pre-processing is an important task and a part of data analysis. The data downloaded or received from any sources are not in the form which can be used directly for analysis. Techniques such as data cleaning and transformation are required to format the data as required for the analysis. The data is prepared for analysis by using R programming language. The data consisted of excel sheet and was imported in RStudio using the library readxl. The data consisted of NAs and irrelevant values were removed by using dplyr library in R.

B. Implementation

IBM-SPSS was to find out the significance of the variables in order to study which variable is responsible as a contributing factor. A correlation matrix was developed so that the significance can be measured. The results of the

Correlation Matrix

	Occupation (1)	Occupation (2)	Occupation (3)	WTS	Prim_Imp(1)	Prim_Imp(2)	Prim_Imp(3)	Stratum(1)	Stratum(2)	Occasions(1)	Occasions(2)	Occasions(3)
Constant	.000	.000	-.690	.059	-.426	-.348	-.407	-.042	.000	-.029	.000	.000
Age	.000	.000	.576	-.157	-.017	-.017	-.032	-.177	.000	-.042	.000	.000
Gender (1)	.000	.000	-.146	.248	.066	.081	-.050	-.088	.000	-.012	.000	.000
Occupation(1)	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
Occupation(2)	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
Occupation(3)	.000	.000	1.000	-.102	.018	-.015	.044	-.075	.000	-.083	.000	.000
WTS	.000	.000	-.102	1.000	-.054	.021	-.074	.016	.000	-.108	.000	.000
Prim_Imp(1)	.000	.000	.018	-.054	1.000	.680	.938	.053	.000	-.118	.000	.000
Prim_Imp(2)	.000	.000	-.015	.021	.680	1.000	.656	.154	.000	-.001	.000	.000
Prim_Imp(3)	.000	.000	.044	-.074	.938	.656	1.000	.077	.000	-.102	.000	.000
Stratum(1)	.000	.000	-.075	.016	.053	.154	.077	1.000	.000	.002	.000	.000
Stratum(2)	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000
Occasions(1)	.000	.000	-.083	-.108	-.118	-.001	-.102	.002	.000	1.000	.000	.000
Occasions(2)	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000
Occasions(3)	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000
Occasions(4)	.000	.000	-.136	.010	-.184	.013	-.167	.052	.000	.532	.000	.000

Fig. 2: correlation matrix

matrix convey that the variables occupation and occasion are the most significant and contributors for frequency of people visiting the pubs. The significance value should of p less than 0.05. The correlation matrix gives an idea of the significance level of the attributes, but in order to study the customer behaviour and getting the key insights from the data visual analysis is the best method for comparing and analysing the data.

C. Visual Analytics

It is a field which targets analytical reasoning from the interactive visual graphs which is an outcome of informative visuals and scientific visuals.

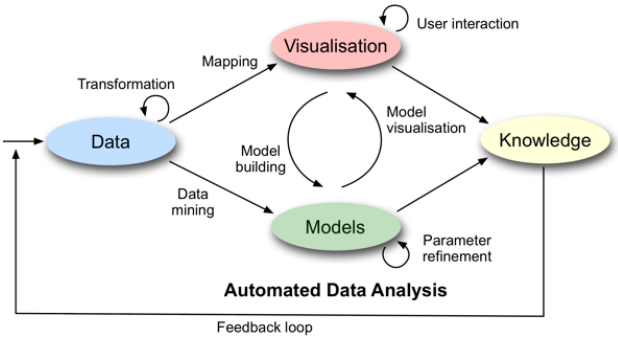

Fig. 3: visual analytics

Data represented visually gives more clear observations and are extremely helpful for gaining knowledge and

understanding any type of customer behaviour. As depicted from Fig.3 the lifecycle of visual analytics describes the transformation of the data according to the requirements as discussed in data pre-processing then comes the mapping of the data columns for creating visuals to extract some meaning from the data which can be user interactive. The impact of visual representation of data is more than any other method as it is understood by all and different inputs can be generated from the people upon their perception of understanding the visualisation. The colours chosen in the visualisations depict some meaning for each parameter, this makes the analysis more easy to understand and identify the pain points. Author Robert Kosara's study on visual analysis of categorical data inspires to use the visual analysis method on the current data set. The categorical data can be termed as categorical dimensions which consist of some meanings. These categorical data termed as dimensions by the author can be visualised and prove extremely helpful in the data sets of surveys to understand the behaviour[8]. For representing the relations between the parameters visual analytics would be the best method as to observe how much they are in relation with each other and key insights can be gathered. According to author Charles Stopler, the primary objective of visual analytics is exploring the data, analysing it and deriving reasons from it. This will help users to understand the complexity of the data and also the hypothesis made can be validated. Interactive graphs help users to understand the data into depth and the knowledge can be extracted with opinions[7]. The data set used in this study consists of a pub survey as mentioned in the data set section. Visual analysis will help the user to identify the relation category wise and can help in understanding the behaviour of the people or customers visiting pub. Tableau software is used for data visualisation. A cleaned excel sheet was imported in the tableau to visualise the data. For visual analysis a split bar graph, horizontal bar graph, an area graph is and a pie chart is used.

VI. Analysis Performed

As mentioned earlier all graphs are created in Tableau software keeping the concept of visual analytics in mind. Some of the columns present as dimensions needed to be converted as measures in order to map the data for creating the graphs. There are 6 visualisations which consists of different types of graphs as mentioned in the section of visual analysis. The first visualisation consists of the survey in terms of gender and their visiting frequency in a pub.

The above graph that is Fig.4 depicts a comparison of customers or people going to pubs by gender. We can observe generally that the female ratio of going to pubs is higher than the males. As we focus on the granularity of data, there is a significant difference between the males and females of rarely going to pubs and several times in a month. But there is a relation in between males and females in going several times a week. The number of people going to pub is

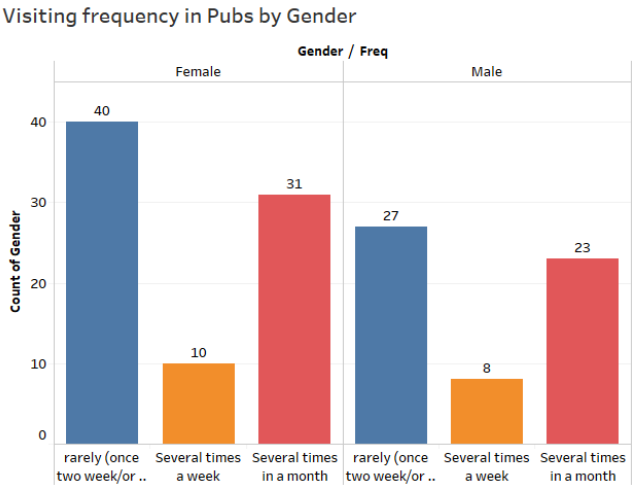

Fig. 4: Visiting frequency in Pub by Gender

almost the same. It is a key insight for the business where is can be increased and revenue can be generated. Different strategies can be made such as happy hours for couples or couple passes or events can be organised which will make the people to spend and will lead to increase in business. Also the couples who visit the pub several times a week, alerts or notifications of discounts can be sent in order to retain the loyal customers.

The second visualisation consists of people visiting the pub on different occasions by gender in an horizontal bar chart.

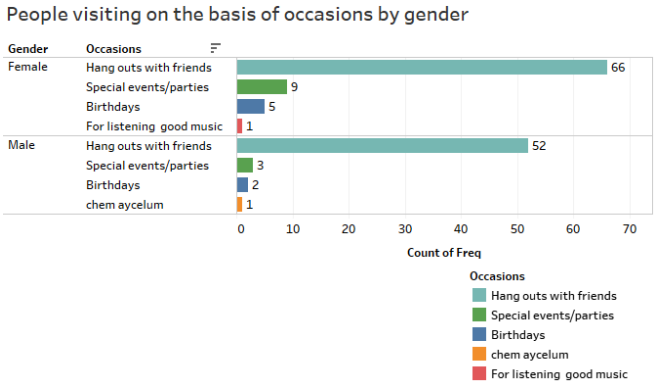

Fig. 5: people visiting pub by occasions and gender

In the above graph we can observe the people going on different occasions in pub and the frequency. The most reason to visit the pub is hanging out with friends, then special events or parties , birthdays and the last for listening the music. It is observed that females lead in going to pub in this analysis too. In the criteria of special events or parties females have a higher contribution than males. This might be the reason of any bachelor parties and events like women's day contribute to this lead. Keeping this in mind the business can be increased by developing strategies like discounts and offers age wise to increase the revenue. As it is said females remember the special days like birthdays more than males their ratio is higher. Schemes can be

developed to retain and improve the sales on these special days. In the category of listening to good music, there needs an improvement as this reason can be developed in to a business requirement and revenue can be generated by a little investment on creating events of different bands and singers.

The third visualisation consists of ratio of people visiting the pub from urban or rural areas. As seen from Fig.6 below, a large number of people visit the pubs from urban areas. It is next to nothing when it comes to comparing against urban for rural areas. The analysis also describes

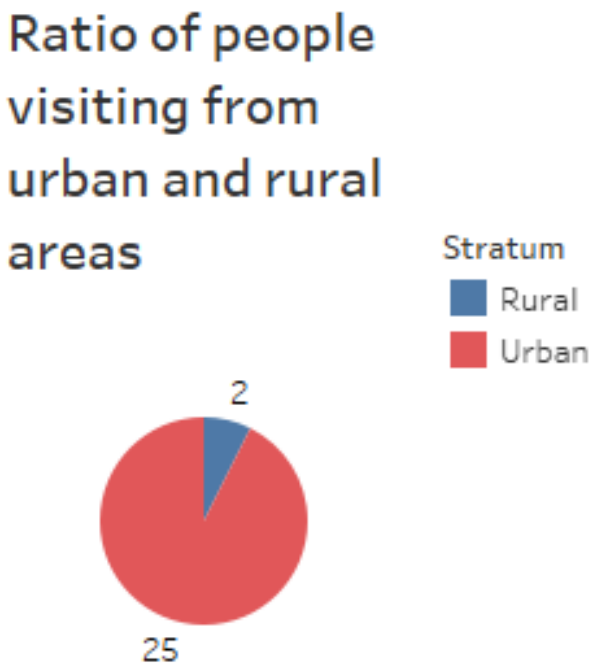

Fig. 6: Ratio of people visiting from urban and rural areas

here that it is very important to retain the urban customers as maximum business is generated by them. Generating revenue from rural areas can be done with some investment by organising traditional festivals and events in villages on their auspicious occasions. The response can be measured and may lie an opportunity to open a chain and increase the business.

The fourth visualisation consists of types of the people visiting pubs and their frequency.
We can observe from the Fig.7 above which represents the types of people visiting pubs and number of times they visit it. It can be observed that the major contributor for the revenue is students for every category present which are visiting rarely once or twice a month, several times a week and several times a month. The students working as part timers are on the secondary level to contribute for the business. These are the customers to be retained and can be given the benefits of loyalty and retain the ongoing business. The another insight observed is the people which can contribute a revenue in large amounts even by visiting

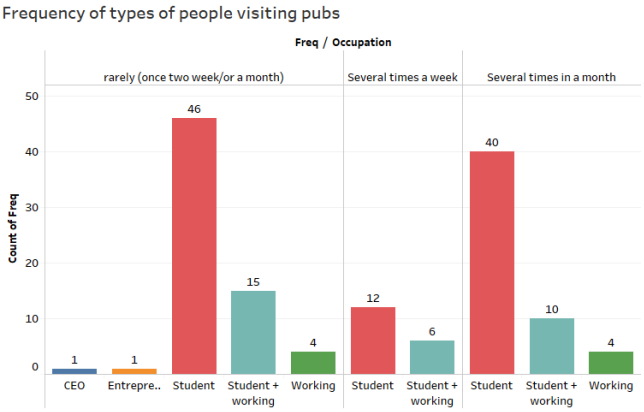

Fig. 7: Frequency of types of people visiting pubs

pubs less times are the entrepreneurs and CEOs. As we can observe they are the least contributors, strategies for companies can be developed by creating alliances and giving contracts to host the parties in the particular pub so that the office parties can generate major revenues even if the frequency is less. The corporate party culture can contribute in large amounts to this business. Along with that the people from working category can also get the benefits after getting passes from their corporate relations with pubs and will be a win-win situation for the customers and most importantly pubs for their business.

The fifth visualisation consists of the income and occupation.

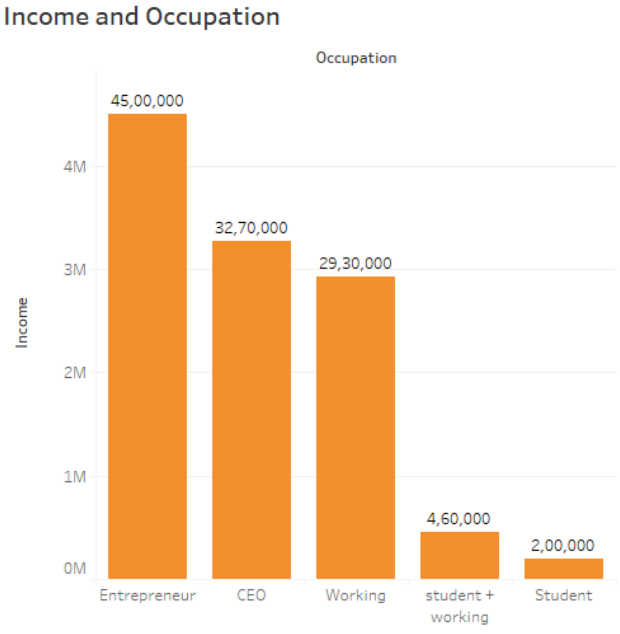

Fig. 8: Income and Occupation

The Fig.8 depicts the income and occupation of the people or customers. This analysis can be related with the previous one of frequency of types of people visiting pubs. This method is called cross comparative analysis where the relationship

between the two factors are observed and insights are gathered. We can observe that the category of entrepreneur and CEOs have the highest income and relating the analysis of strategies for corporate people can yeild more revenue can help here. These people can be recommended the alcohols or food of high price according to meet their standards and chances of buying them are hight because of the high income. The categories of students and students which are working can be suggested with coupons and strategies like happy hours to increase the business.

The sixth visualisation consists of the age groups and how frequently they visit pubs.

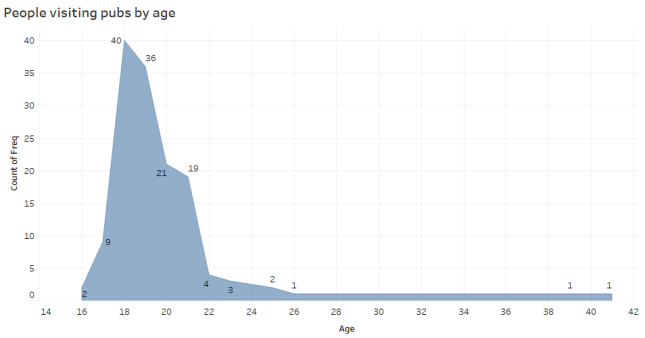

Fig. 9: Frequency of people visiting pubs by age

The Fig.9 above depicts the frequency of people visiting pubs on the factor of age. We can observe that age group from 18 to 21 has the maximum frequency of visiting the pubs. The analysis also describes that it drastically decreases from the age 22 and beyond. Again the concept of cross analysis can be applied for generating key insights.The previous analysis of types of people visiting pubs in Fig.7 can be related to do a deeper analysis by which the behaviour in terms of age and types of people can be measured. The majority of people working lie in the age group of greater than 24. As discussed earlier strategies related to corporate can be developed and the necessary business measures can be implied.

The seventh visualisation consists of the primary importance given according to different occupations by willingness to spend in Pubs.

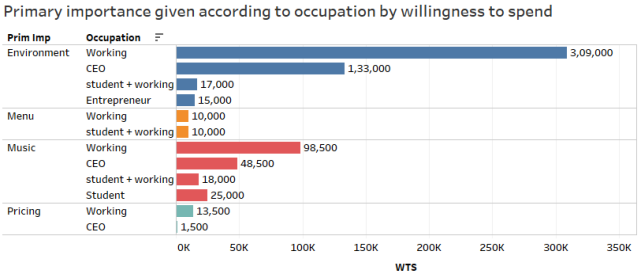

Fig. 10: Primary importance given according to occupation by willingness to spend

The Fig.10 below depicts the primary importance given by the people working in different occupation and their willingness to spend on different types of categories such as environment, menu, music and the pricing. We can observe that the working people give more preference to the different environment of the pub and are willing to spend high on this parameter. Music is also one of the factor which influences people to spend and which is directly proportional to consumption of food and alcohol. It can be seen that the working people have an higher edge in environment and music resulting in contribution of menus too. These are the loyal customers and strategies can be developed in order to retain them. The willingness to spend of other customers which are students should be focused on as they are the customers who have the highest frequency of visiting pubs. Strategies such as happy hours, weekend special discounts or student discounts will help the willingness to spend from this category of customers.

VII. Discussion

As per the analysis performed we have observed what are the factors which can retain a business and the factors responsible for increasing it. The survey data helped us by understanding the behaviour of the people or customers visiting pubs and their preferences.It also helped us by understanding how to increase the business in rural areas and urban areas. These key insights can be used for choosing to open a pub into rural or urban areas. As discussed in the analysis section different strategies can be applied according to the behaviour of the customers and their needs. The results also show that younger people tend to visit the pubs more but the willingness to spend is generated more by the customers who are working. The relation between males and females was also studied which gave significant results of females visiting more on the basis of events. The relation between income of the customers by their occupation and their willingness to spend was also analysed which gathered some knowledge of the behaviour of spending into pubs and how much can these customers spend.Some graphs were analysed using cross analysis technique which inherited key insights from the data analysed.

VIII. Conclusions

This study was performed on a survey data of pubs. Different tools for various purposes like RStudio, IBM SPSS and Tableau were used. The RStudio was used to pre-process the data and prepare the data for analysis. The IBM SPSS tool was used to observe the correlation between the variables and which is more significant and contributing factor for the frequency of customers going in pubs. Lastly, the data was analysed in Tableau creating different graphs. Significant relation between the behaviour of customer of frequency for visiting the pubs considering the factor of different types of people or customers depending on their income was found. The relation differs according to the types of people,their income and their preferences for visiting the pubs to spend the money on it. This rejects our null hypothesis as the

analysis performed gave different results. On the analysis performed students can be considered as the loyal customers and workers have the ability to spend enough. The relation between the gender also describes that females are more fond of visiting the pubs than male on specific events and on a general purpose. This analysis can be used to study the behaviour of customers and strategies can be applied in order to increase the business.

REFERENCES

[1] R. Unless, P. Act, W. Rose, T. If, and W. Rose, " Community cohesion and village pubs in Northern England : an White Rose Research Online URL for this paper : Version : Accepted Version Article : Mount , MP and Cabras , I (2016) Community cohesion and village pubs in Northern England : an econometric study . Regional Studies , 50 (7). pp . 1203-1216 . ISSN 0034-3404 Community Cohesion and Village Pubs in Northern England : An Econometric Study," 2016.

[2] P. J. Sandiford and P. Divers, "The pub as a habitual hub: Place attachment and the Regular customer," *International Journal of Hospitality Management*, no. May, pp. 1–8, 2018.

[3] S. Tutenges and F. Bøhling, "Designing drunkenness: How pubs, bars and nightclubs increase alcohol sales," *International Journal of Drug Policy*, vol. 70, pp. 15–21, 2019.

[4] A. Baker, M. Tolley, and K. M. Hill, "Mooe Thaa a Quiik Piit: IIIestigatiig plaae attaahheet to aa EEglish eeeeatiooal settiig," 2016.

[5] "Marketing Research on Consumer Behaviour and Drinking Habits of People Living in Imatra - Lap- peenranta Region , Finland," 2017.

[6] https://www.kaggle.com/erikhambardzumyan/pubs

[7] C. D. Stolper, A. Perer, and D. Gotz, "Progressive visual analytics: User-driven visual exploration of in-progress analytics," *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, no. 12, pp. 1653–1662, 2014.