Customer Behavior

Nathaniel del Rosario

Objectives

Goals: Find any trends in different characteristics of orders at a restaurant to see if there are any useful findings that can help the business

More specifically, find correlations between parameters such as customer age and their spending habits

Hypothesis

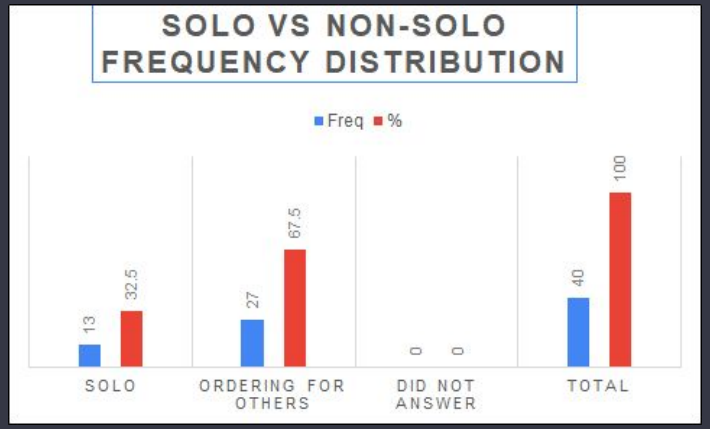
- People of older ages will spend more and will visit more frequently because of higher disposable income and more free time (retired individuals) / needing a meal after work.
- In addition, they may have children who they are ordering for, so their orders will have more items and be more expensive.
- This means that we believe there are positive correlations between age, times visited, order total & items ordered.

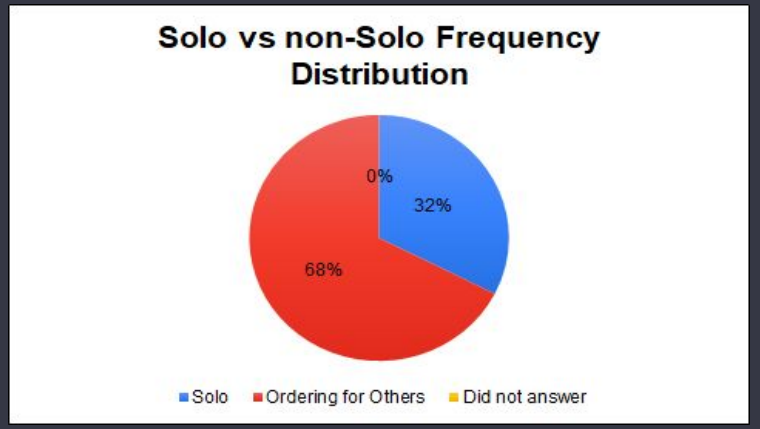
Methodology

- I collected data on an in person survey basis
- The population is the customer base of a restaurant I work at
- Sample size is 40, so n = 40, df = 39
- Used random sampling to ensure each customer had an equal likelihood of being surveyed; this meant equally offering surveys to pick up orders, for-here orders, and walk-ins.

Summary of Data:

Ordering Solo vs Ordering for Others





Customer Age vs Order Total

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.51075299
R Square	0.26086861
Adjusted R Square	0.24141779
Standard Error	1.07669366
Observations	40

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	15.54776931	15.54776931	13.41169845	0.000758434
Residual	38	44.05223069	1.159269229		
Total	39	59.6			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	2.1398259	0.433538362	4.935724458	1.62252E-05	1.262173367	3.017478428	1.262173367	3.017478428
X Variable 1	0.41131665	0.11231413	3.662198581	0.000758434	0.183948579	0.638684718	0.183948579	0.638684718

The correlation is positive with a magnitude of .5107, making it statistically significant. ($F < .05$)

This indicates that if your age is higher, you are more likely to have a higher order total ($)

Items Ordered vs Customer Age:

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.506965							
R Square	0.257014							
Adjusted F	0.237461							
Standard E	0.727919							
Observatio	40							
ANOVA								
	df	*SS*	*MS*	*F*	*ignificance F*			
Regressior	1	6.965071	6.965071	13.14495	0.000843			
Residual	38	20.13493	0.529867					
Total	39	27.1						
	Coefficients	*tandard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*pper 95.0%*
Intercept	1.372688	0.293102	4.683312	3.55E-05	0.779334	1.966041	0.779334	1.966041
X Variable	0.275299	0.075932	3.625597	0.000843	0.121583	0.429016	0.121583	0.429016

The correlation is positive and with a magnitude of .501 statistically significant
(F = .0008 < .05)

This indicates that as your age increases, you are more likely to order more items off the menu.

Customer Visits vs Order Total

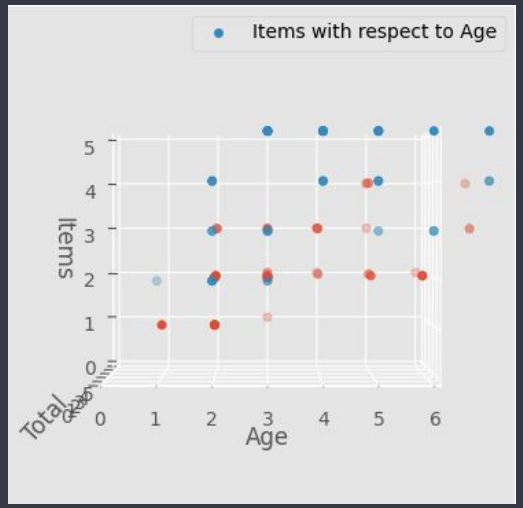
SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.214352							
R Square	0.045947							
Adjusted F	0.02084							
Standard E	1.223257							
Observatio	40							
ANOVA								
	df	*SS*	*MS*	*F*	*ignificance F*			
Regressior	1	2.738422	2.738422	1.830059	0.184115			
Residual	38	56.86158	1.496357					
Total	39	59.6						
	Coefficients	*tandard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*Upper 95.0%*
Intercept	3.062096	0.442169	6.925166	3.11E-08	2.166971	3.957221	2.166971	3.957221
X Variable	0.217335	0.160656	1.352797	0.184115	-0.1079	0.542566	-0.1079	0.542566

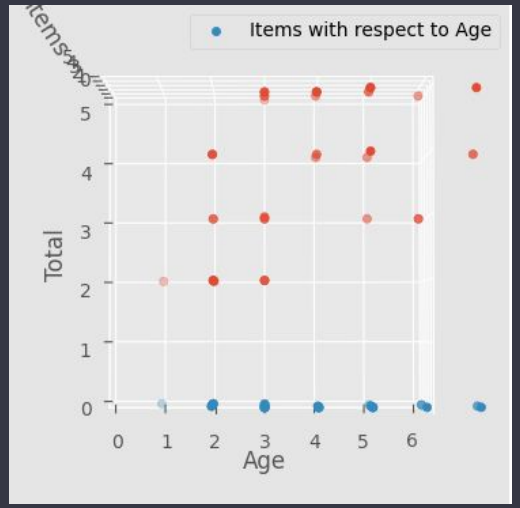
There is **no correlation** as the magnitude is .21 and the Sig. F value is .18 which is greater than the threshold of .05.

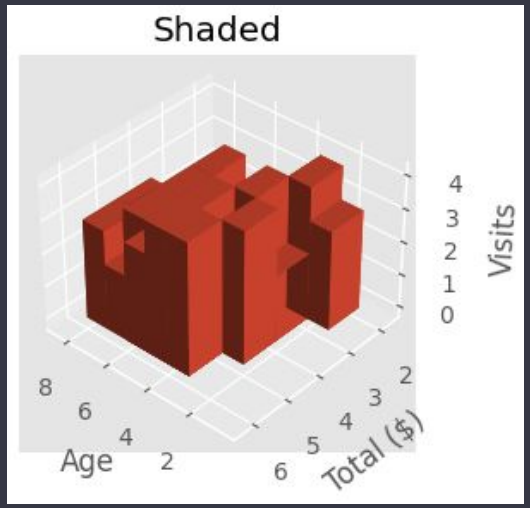
Thus we can not infer that as customers visit more they will spend more

Visualizing Results Using Matplotlib In Python

Plotting these independent variables on a 3 dimensional plane, Total, Items, and Visits, we see that the scatter plots help to support our findings







*Note the clear linear correlations between Items & Age / Total & Age

*Note how “Visits” does not show any correlation with “Total”

Conclusions

- Proved that there is a correlation between customer age and order total as well as customer age and items ordered in the sample obtained
- Did not prove that there is any correlation between customer age and times visited in the sample obtained
- Showed that a majority (68%) of the collected sample ordered for more than one person
- Shortcomings:
    - Could have surveyed a larger sample for more realistic results

Further Applications

From a business standpoint, these results allow new decisions to be made about catering more to the audience, our customers

Looking forward, we can use machine learning to train, classify, and predict on a larger sample of a population

If future researchers were to elaborate on this study, I would request that they classify types of customers, which would make it easier to predict customer behavior based on their demographic/behavior, and more importantly implement deals catered to those customers

These new offers could range from **senior citizen discount**, **returning customer discounts**, **to incentives to buy more items in one visit**