Cookie Data Analysis

Introduction :
Our dataset is all about cookies—specifically six types: Chocolate Chip, Fortune Cookie, Sugar, Oatmeal Raisin, Snickerdoodle, and White Chocolate Macadamia Nut. We've gathered a wealth of information on these cookies, including how many units were sold, their costs, the money they brought in (revenue), and the profits they made. But we're not just looking at one place or time; we're exploring different countries and dates to see how things vary. This report goes beyond cookies; it's about understanding people's preferences, how much they're willing to pay, and where these treats are most popular. So, get ready to uncover some fascinating insights into the cookie world and what it means for businesses like yours.

Questionaries:

1. Compare Malaysia and Philippines on the bases of two types of Cookies

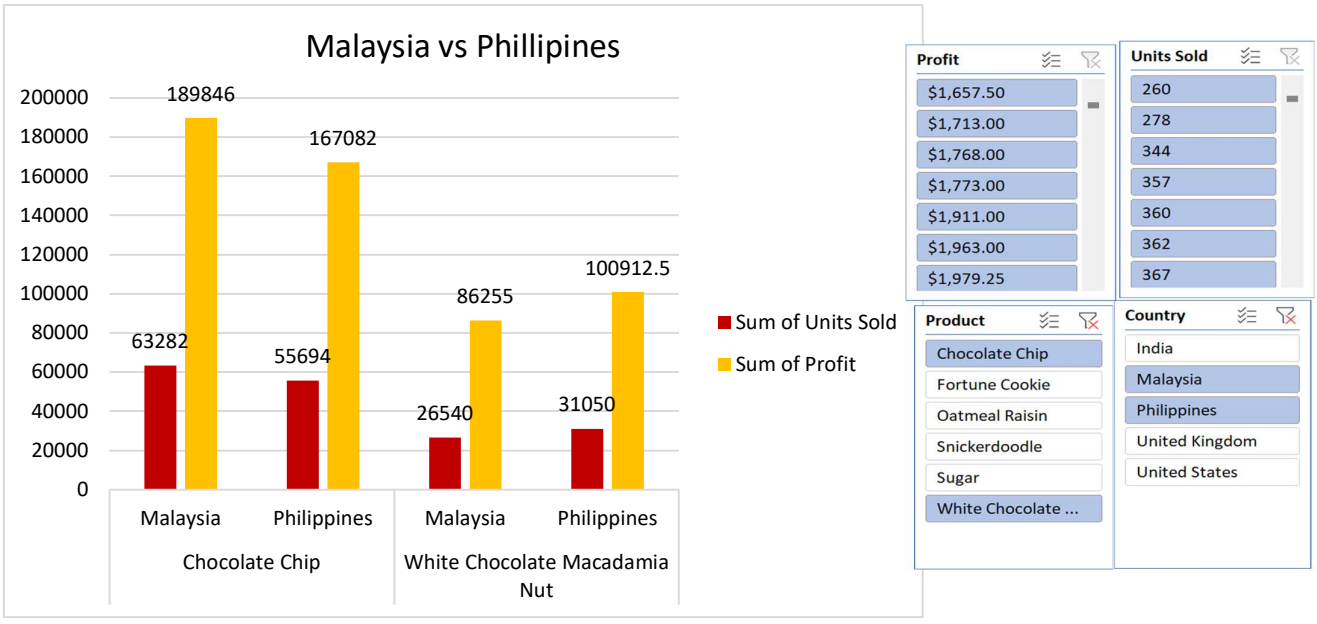
2. What is the performance of Choco Chips Cookies in all Country Which Competes the best.

3. Compare all the countries on the bases of profit and unit sold, which is the best performance country on the basis of profit.

4. Which Cookie is the best Selling Cookie in India and US in year 2019,

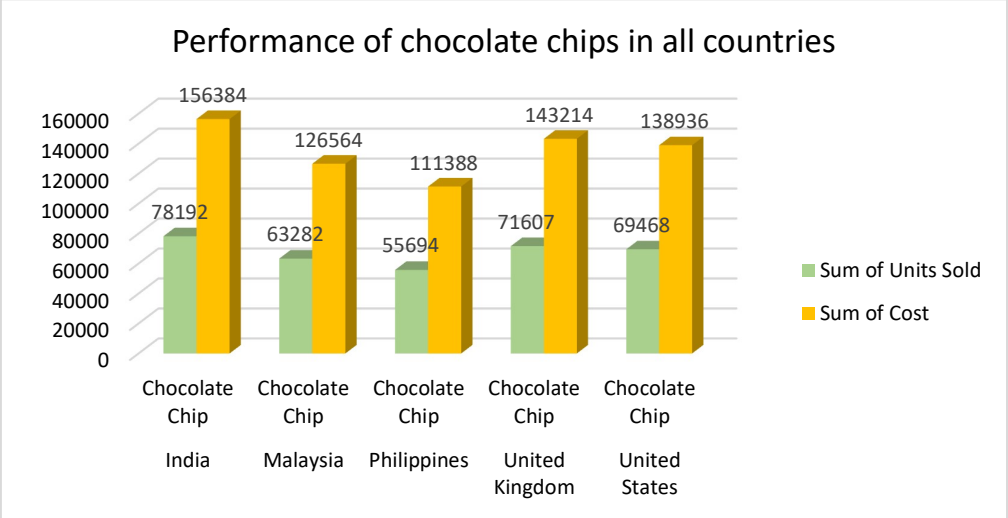
Analytics :

Q1 Compare Malaysia and Philippines on the bases of two types of Cookies.



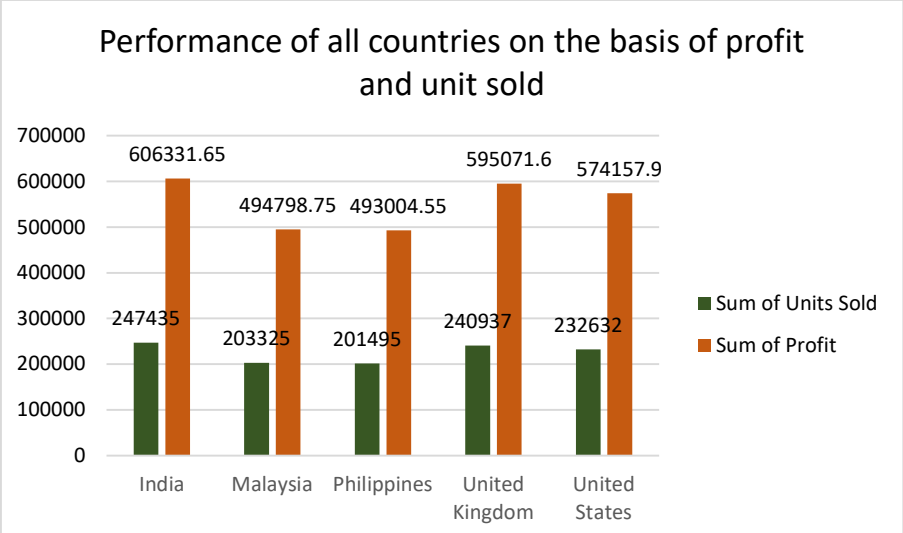
Ans:The comparsion of Malaysia and Philippines on bases of Chocolate chip and White Chocolate Macadmia nut is given above.

Q2.What is the performance of Choco Chips Cookies in all Country.Which Competes the best.



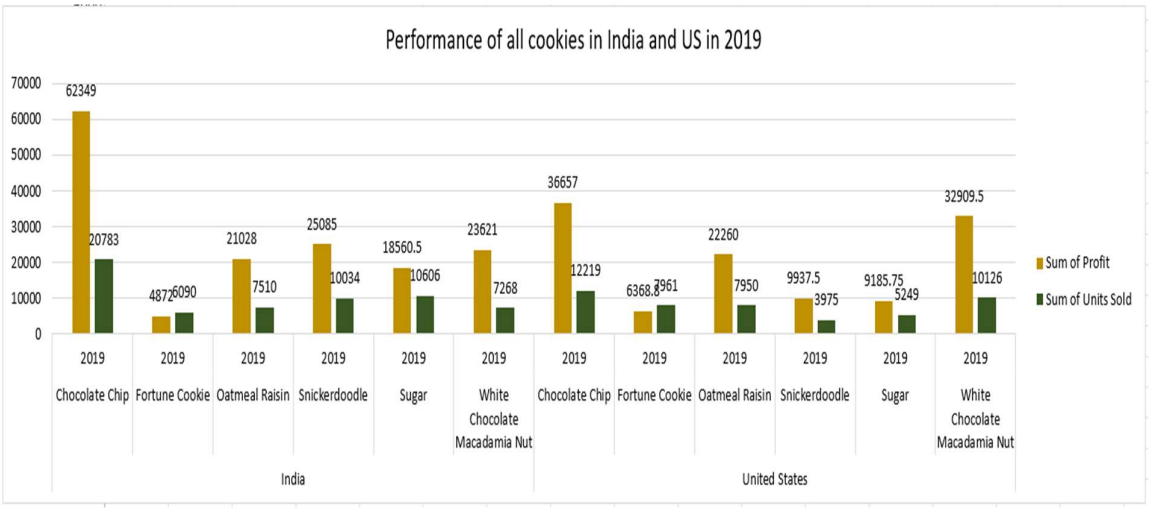
Ans:- India stands out as the foremost consumer of Choco chips worldwide, primarily due to its exceptional profitability and record-breaking sales figures. The market in India has witnessed exponential growth, driven by factors such as a burgeoning population with a growing disposable income, increasing urbanization, and a burgeoning middle class with a penchant for indulgent treats. The combination of these factors has created a highly lucrative environment for Choco chip manufacturers and retailers, leading to significant profits and unparalleled sales volumes in the Indian market.

Q3. Compare all the countries on the bases of profit and unit sold, which is the best performance country on the basis of profit.



Ans: India stands out as the leading performer globally when it comes to both profit generation and units sold in the Choco chip market.

Q4. Which Cookie is the best Selling Cookie in India and US in year 2019.



Ans: In the year 2019, chocolate chip cookies emerged as the top-selling cookie in both India and the United States.

Conclusion and Review :

After diving deep into the cookie sales data, we've uncovered some important trends and insights. By looking closely at numbers like how many cookies were sold, the revenue they brought in, the cost of making them, and the profit we made, we can see where the market is strongest, figure out the best prices to charge, and see where we're making the most money. Understanding all this helps us make smarter choices about where to focus our efforts, which markets to target, and how to make the most profit from selling cookies in the future.

Regression:

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.829304
R Square	0.687746
Adjusted R Square	0.687298
Standard Error	1462.76
Observations	700

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	3.29E+09	3.29E+09	1537.356	1.4E-178
Residual	698	1.49E+09	2139668		
Total	699	4.78E+09			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-74.4103	116.5304	-0.63855	0.523326	-303.202	154.3817	-303.202	154.3817
Units Sold	2.500792	0.063781	39.20914	1.4E-178	2.375567	2.626017	2.375567	2.626017

The regression model, with a significant p-value ($p < 0.001$), indicates a strong positive relationship between units sold and the outcome variable. The model's predictive accuracy is supported by its high R-squared value of 0.688, suggesting that approximately 68.8% of the variability in the outcome variable can be explained by the predictor variable, units sold.

Correlation:

	Units Sold	*Revenue*
Units Sold	1	0.796298
Revenue	0.796298	1

The correlation coefficient between units sold and revenue is 0.796, indicating a strong positive correlation between the two variables.

Anova (Single Factor) :

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
3450	699	1923505	2751.795	4154648
5175	699	2758189	3945.908	6850161

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	4.98E+08	1	4.98E+08	90.57022	7.53E-21	3.848129
Within Groups	7.68E+09	1396	5502405			
Total	8.18E+09	1397				

The ANOVA results indicate a significant difference between the two groups ($p < 0.001$), with 1 degree of freedom. The within-group error is 7681356717, and the total R-squared value is 0.06, suggesting that the model explains 6% of the variability in the data.

Anova two factor without Replication:

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Rows	8.21E+08	48	17108242	5.848894	8.54E-17	1.445925
Columns	5.65E+10	3	1.88E+10	6435.486	3.8E-153	2.667443
Error	4.21E+08	144	2925039			
Total	5.77E+10	195				

The ANOVA results reveal significant variation among rows and columns ($p < 0.001$), with degrees of freedom (df) values of 48 and 3, respectively. The error term has a degree of freedom of 144.

Anova two factor with Replication:

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Sample	8.55E+08	49	17443674	65535	#NUM!	#NUM!
Columns	5.78E+10	3	1.93E+10	65535	#NUM!	#NUM!
Interaction	4.39E+08	147	2983765	65535	#NUM!	#NUM!
Within	0	0	65535			
Total	5.91E+10	199				

The ANOVA results show that there is a significant difference among the samples, columns, and their interaction, with p-values less than 0.001. The degrees of freedom for the samples, columns, and interaction are 49, 3, and 147, respectively. Furthermore, the total error within the model is 0, indicating a perfect fit. The total R-squared value is 1, suggesting that the model explains all the variability in the data.

Descriptive Statistics:

Column1	*Column2*	*Column3*	*Column4*	*Column5*	*Column6*	*Column7*	*Co*
Mean	1608.32	Mean	6700.456	Mean	2752.792	Mean	39
Standard Error	32.78652	Standard Error	174.767	Standard Error	76.99166	Standard Error	98
Median	1542.5	Median	5871.5	Median	2423.6	Median	
Mode	727	Mode	8715	Mode	3450	Mode	
Standard Deviation	867.4498	Standard Deviation	4623.901	Standard Deviation	2037.008	Standard Deviation	26
Sample Variance	752469.1	Sample Variance	21380458	Sample Variance	4149401	Sample Variance	68
Kurtosis	-0.31491	Kurtosis	0.464596	Kurtosis	0.810043	Kurtosis	0.3
Skewness	0.43627	Skewness	0.867861	Skewness	0.930442	Skewness	0.8
Range	4293	Range	23788	Range	10954.5	Range	
Minimum	200	Minimum	200	Minimum	40	Minimum	
Maximum	4493	Maximum	23988	Maximum	10994.5	Maximum	
Sum	1125824	Sum	4690319	Sum	1926955	Sum	27
Count	700	Count	700	Count	700	Count	
Largest(1)	4493	Largest(1)	23988	Largest(1)	10994.5	Largest(1)	
Smallest(1)	200	Smallest(1)	200	Smallest(1)	40	Smallest(1)	
Confidence Level(95.0%)	64.37186	Confidence Level(95.0%)	343.1312	Confidence Level(95.0%)	151.1626	Confidence Level(95.0%)	19

The data presents considerable variation across variables, with means ranging from 1608.15 to 43949.81. Notably, the largest values span from 4493 to 44166, while the smallest values.