

IMPACT OF MENSTRUATION ON CONSUMER BEHAVIOUR

Krish Aggarwal¹, Krish Shah², Khushi Kataria³, Kavish Belmaan⁴, Kairavi Mehta⁵, Tejaswini Angre⁶

¹ Student, BBA, NMIMS, Maharashtra, India

² Student, BBA, NMIMS, Maharashtra, India

³ Student, BBA, NMIMS, Maharashtra, India

⁴ Student, BBA, NMIMS, Maharashtra, India

⁵ Student, BBA, NMIMS, Maharashtra, India

⁶ Professor, Statistics, NMIMS, Maharashtra, India

Menstruation is an important factor of human existence, and it is important to study it rigorously. This analysis is about consumer behaviour and factors related to the menstrual cycle. There are unnecessary myths created and passed on concerning the menstrual cycle, and it becomes difficult for females to talk about the natural processes. With this as a driving force, the report was made to know about consumer behaviour in the markets, majorly how income and age affect the behaviour was part of the study.

To get insights about the topic a questionnaire was prepared, and 151 responses were analysed. From the hypothesis considered the dependent variables were income and age, in the hypothesis where income is the dependent variable the independent variables were the frequency of buying sanitary napkins, expenditure on sanitary napkins, and expenditure on cravings.

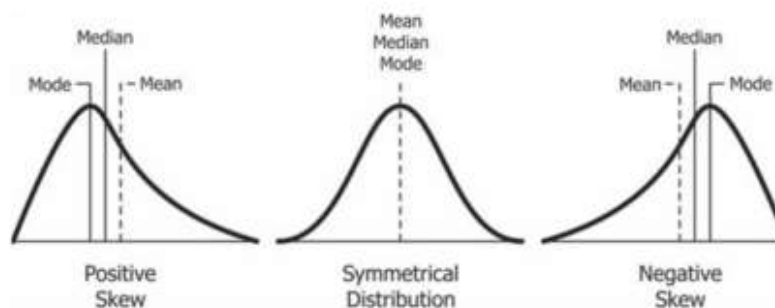
Keyword - Menstrual cycle, Consumer behaviour, Periods, Cravings, correlation, skewness

1. Abstract

This research paper has explored the biological phenomenon of menstruation and its impacts on economic factors like consumer behaviour. The research has shown how age an important variable affects menstruation, cravings, intake of painkillers along with the income of the household. Overall key insights were devolved on account of these variables.

2. Skewness

Skewness is a measure of the asymmetry of a distribution. A distribution is asymmetrical when its left and right side are not mirror images. A distribution can have right (or positive), left (or negative), or zero skewness. A right-skewed distribution is longer on the right side of its peak, and a left-skewed distribution is longer on the left side of its peak.



Variable	Skewness	Coefficient of Variation
Age	0.66002 - moderately positive skewness	46.02%
Income	0.54415 - moderately positive skewness	13.37%
Importance of Sanitary pads	-2.90494 – Highly negative skewness	17.37%
Importance of Quality of Sanitary Pads	-1.15510 - Highly negative skewness	48.28%
Frequency of buying Sanitary Pads	-0.40220 – Fairly symmetrical	49.22%
Expenditure On Sanitary Napkins	-0.12613 – Fairly symmetrical	42.10%
Painkillers	0.33465 – Fairly symmetrical	29.76%

3. Coefficient of Determination

R-squared is a statistical measure that represents the proportion of the variance for a dependent variable that is explained by an independent variable or variables in a regression model.

SUMMARY OUTPUT	
Regression Statistics	
Multiple R	0.750315224
R Square	0.562972936
Adjusted R Square	0.554054016
Standard Error	22259.40659
Observations	151

Variable Assigned		Coefficient	P-value
	Intercept	7371.0600	0.22947
A	How frequently do you buy sanitary pads?	149.7939	0.73468
B	Expenditure on Sanitary Napkins	47.9124	1.763E08
C	Expenditure on Cravings	64.4528	3.717E14

Income (Y) – Independent variable

$$A = 7371.06007674876 + 149.793884877111 Y$$

$$B = 7371.06007674876 + 47.9123705392423 Y$$

$$C = 7371.06007674876 + 64.4527771180795 Y$$

		Coefficient	P-value
Variable Assigned	Intercept	47.0371	3.81784E-47
D	Do you take painkiller s?	3.5465	3.303E-09
E	Do you have cravings?	-6.1047	3.301E-31

Age (Z) – Independent variable

$$D = 47.0370739066722 + 3.54645058770427 Z$$

$$E = 47.0370739066722 + -6.10467174918023 Z$$

4. Hypothesis Testing

It is a method of statistical inference which tests an assumption regarding a population parameter. A statistical hypothesis test is used to decide whether the data at hand sufficiently supports a particular hypothesis about the population.

Chi-squared test of independence (X2 TOI), in statistics, is a method of testing hypotheses to determine whether two variables are related in the population. It compares observed results with expected results to determine the dependency of the two variables

Application of hypothesis testing in this project

The **central limit theorem (CLT)** states that the distribution of sample mean approximates a normal distribution as the sample size gets larger, regardless of the population's distribution.

- Sample sizes equal to or greater than 30 are often considered sufficient for the CLT to hold true. (Sample size in this report is 150)
- A chi-square with many degrees of freedom (N-1) is approximately equal to the standard normal variable, as the central limit theorem states. (N is the number of observations in the sample)

A **confidence interval**, in statistics, refers to the probability that a population parameter will fall between a set of values for a certain proportion of times.

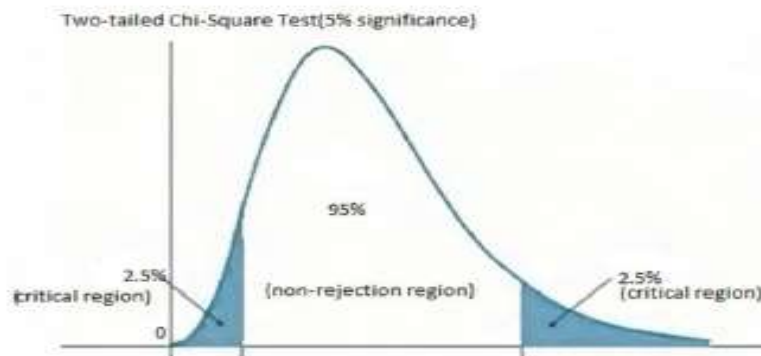
Confidence interval = 95%

Hence, 5% would be the 'α' (Type I error)

This means that we assume there is a 5% chance that the conclusions made using the hypothesis testing would be wrong.

Null Hypothesis is	True	False
	Rejected	Type I error (α)
Not Rejected	Correctly accepting the null hypothesis (1- α)	Type II error (β)

Type I & type II errors



- How frequently one buys sanitary pads
 H0: How frequently one buys sanitary pads is independent of the income
 H1: How frequently one buys sanitary pads is dependent on the income
 As the p-value is greater than 0.05, we accept the null hypothesis H0.

- Expenditure on Sanitary Napkins
 H0: Expenditure on Sanitary Napkins is independent of the income
 H1: Expenditure on Sanitary Napkins is dependent on the income
 As the p-value is less than 0.05, we reject the null hypothesis and accept alternate hypothesis H1.

- Expenditure on Craving
 H0: Expenditure on cravings is independent of the income
 H1: Expenditure on cravings is dependent on the income
 As the p-value is less than 0.05, we reject the null hypothesis and accept alternate hypothesis H1.

- Painkillers
 H0: How frequently one has painkillers is independent of age.
 H1: How frequently one has painkillers is dependent on age.
 As the p-value is less than 0.05, we reject the null hypothesis and accept alternate hypothesis H1.

- Cravings
 H0: How frequently one has cravings is independent of age.
 H1: How frequently one has cravings is dependent on age.
 As the p-value is less than 0.05, we reject the null hypothesis and accept alternate hypothesis H1.

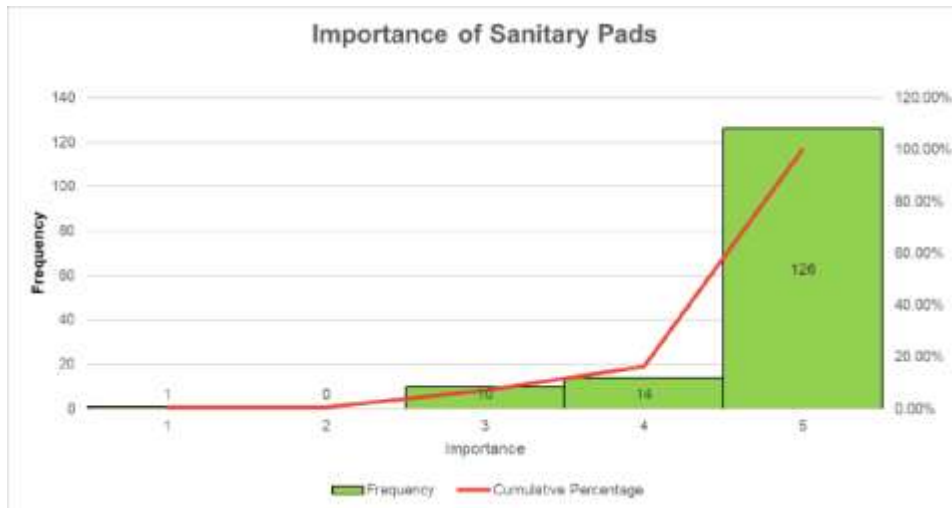
5. Brand Data

The brand data was obtained by assigning values to each brand and using vlookup to assign said values to each brand to convert it into quantitative data and then correlation was performed.

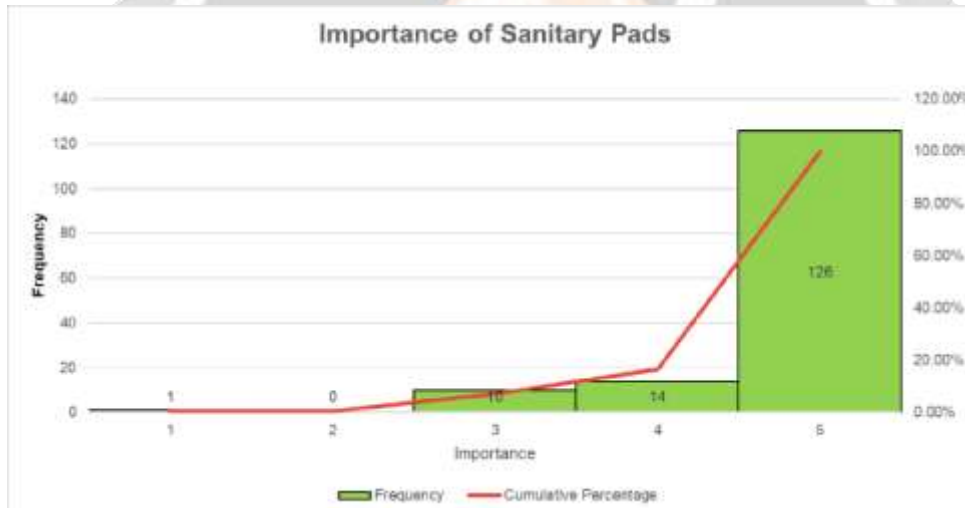
The correlation between brand and the importance of the quality of sanitary pads is equal to 0.1228. This is a low positive correlation depicting that the change in brand and importance given to the quality of sanitary pads don't affect each other too much.

6. Histogram

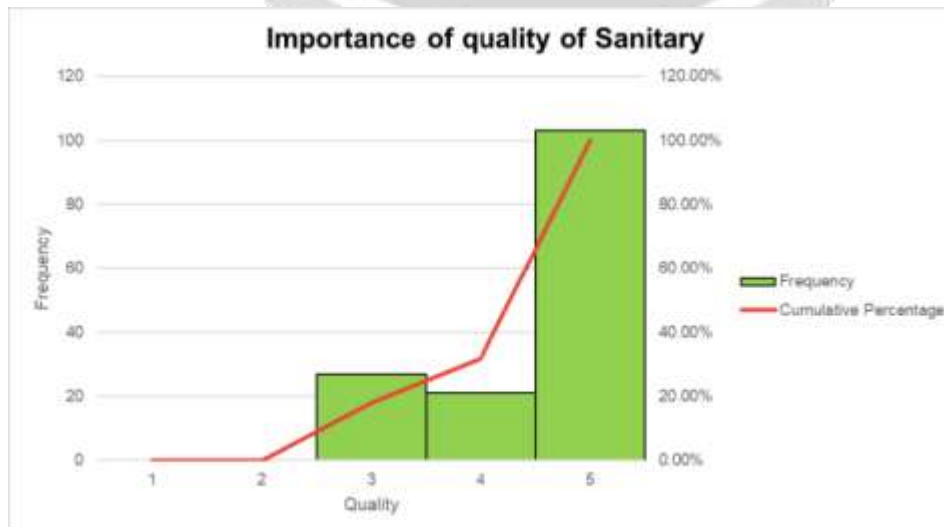
6.1.



6.2.



6.3.



7. Correlation

- The correlation between income and the importance of using sanitary pads is equal to 0.1504. This is a low positive correlation depicting that an increase in income does not affect much of an increase in the importance of the use of sanitary pads.
- The correlation between income and the importance of quality sanitary pads is equal to 0.4974. This is a moderate positive correlation depicting an increase in income corresponding to increased importance given to the quality of sanitary pads.
- The correlation between income and frequency of buying sanitary pads is equal to 0.0570. This is a low positive correlation as the value is very close to 0 depicting that the frequency of buying sanitary pads and income is independent of each other and are not correlated.
- The correlation between income and expenditure on sanitary napkins is equal to 0.5944. This is a high positive correlation depicting that as income increases the expenditure on sanitary napkins also increases.
- The correlation between income and expenditure on cravings is equal to 0.6753. This is a high positive correlation depicting that as income increases, the expenditure on cravings also increases.
- The correlation between age and how frequently one takes painkillers (pertaining to menstruation) is equal to 0.5991. This is a high positive correlation depicting that females of more age tend to take painkillers more frequently.
- The correlation between age and how frequently one has cravings is equal to -0.8213. This is a high negative correlation depicting that females of more age tend to have cravings less frequently.

8. Findings

- 8.1. There is a moderate positive correlation between age and intake of painkillers signifying that with age painkiller intake increases. Menstrual pain gets worse during menopause and pain tolerance decreases with age. Moreover, it depends on person to person regardless of age. Hence, a moderate positive correlation.
- 8.2. High negative correlation between age and cravings showcases that as age increases cravings decrease. An increase in age causes a slowdown of metabolism and decreases appetite. Hence, a negative correlation.
- 8.3. Low positive correlation between income and brand showcases how the sanitary napkin market is close to perfect competition. The prices of different brands are almost the same. Regardless of income, consumers have a preference. Hence, a low positive correlation.
- 8.4. Moderate positive correlation between income and spending on cravings shows that the higher income bracket has the luxury to afford irrational cravings. As income increases, affordability increases too. Hence, a positive correlation.

9. CONCLUSIONS

This study helped us understand the various variables that come into play in the menstrual cycle. It shows how individuals make their choices, some being collectively similar while others completely different. It helps understand the intricacies of consumer behaviour which can be capitalized upon by the producer if they're aware of it.

10. Limitations

1. Data points - 151 - Collection of data was a limitation because population was limited to females.
2. Due to limited data the coefficient of variation increases, to overcome that limitation collection from the population again becomes a problem.
3. To work with ranges, one has to take into account the midpoints which might not be the most accurate representation of data for every individual. Also, in case of options like '150,000+' choosing mid-point according to the class interval could lead to a wrong conclusion.
4. Survey couldn't reach the lower income bracket people. That limited the variety of data which was analysed in this project. That could have provided a basis for comparison with respect to the independent variable 'income'.

11. Recommendations

1. Responses can sometimes be biased, which means that the responses provided a respondent does not reflect his true actions or thoughts. The response is affected by some or the other factor. Like often respondent are hesitant to share their actual income so they might chose a different income bracket in the survey. However, biases are not always intentional; it is just human nature as human being are not able to always think and act rationally (the idea of bounded rationality)
2. Just for the sake of finishing the survey people might randomly click the options which affects the analysis as whole, because the survey is designed in such a way that the variables are interrelated.
3. Providing numerous smaller ranges would have given better insights of the consumer behaviour. Because having smaller intervals for each question would have segregated each costumer more precisely but it was not done because a longer survey (with more option to chose from) often stops the respondent from filling up the form as the questionnaire looks much longer and time-taking.
4. If the survey could have been sent to females with different economic backgrounds, that could have provided a more accurate result with respect to the level of awareness amongst the women because as of 2020, 65% of the Indian population lives in rural areas. So, the results would then depict a more accurate picture of the society.

12. ACKNOWLEDGEMENT

This paper and the research behind it would not have been possible without the exceptional support of our professor, Tejaswini Angre. Her knowledge and exacting attention to all the details have been an inspiration and helps us keep our work on track and draft of this paper. Under her guidance, we were able to apply the concepts taught in our course and do a research based on it. The generosity and expertise of our professor has improved this study in innumerable ways and improved certain errors, bringing us with this final document for our research work.

13. REFERENCES

- a. <https://www.google.com/url?sa=i&url=https%3A%2F%2Fanalystprep.com%2Fcfca-level-1-exam%2Fquantitative-methods%2Fchi-square-test-example%2Fattachment%2Fchi-square-distribution-two-tailedtest%2F&psig=AOvVaw2aaVFmEOzbnC9g4rrsyCXp&ust=1670856282947000&source=images&cd=vfe&ved=0CBEQjhXqFwoTCOii8-m8fsCFQAAAAAdAAAAABAD>
- b. <https://www.investopedia.com/terms/h/hypothesistesting.asp>
- c. <https://pubmed.ncbi.nlm.nih.gov/9221991/>
- d. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7357764/>
- e. <https://www.ncbi.nlm.nih.gov/books/NBK279324/>