



# APPLE STORE REVIEW STATISTICAL ANALYSIS

**PRESENTED BY: SARANSH GARG**





# **PROBLEM STATEMENT**

**This project aims to analyze Apple Store product reviews to better understand user ratings and engagement. It involves examining the average ratings and overall trends, analyzing the range and patterns in purchase amounts, exploring the relationship between likes and ratings, and comparing user perceptions of Instagram and WhatsApp through data analysis. The goal is to identify key trends and provide actionable recommendations for app developers to enhance user satisfaction and engagement.**

# RATING ANALYSIS



**Mean:** 2.869

The average rating is just below 3 because many users gave very low ratings, like 1, which pulls the average down.

**Mode:** 1

The most common rating is 1, showing a lot of users gave poor ratings.

**Median:** 3

The middle rating is 3, meaning half the ratings are below 3 and half are above.

**Conclusion:**

The median is the best measure of the typical rating since it isn't affected by very low ratings and shows the middle user experience.

# VARIANCE AND STANDARD DEVIATION OF LIKES



**Variance:** 822.8547

Variance measures how much likes differ from the average, showing the overall spread in likes.

**Standard Deviation:** 28.6854

The standard deviation, calculated in the same units as the data, shows a high spread in likes, meaning user engagement varies a lot.

**Conclusion:**

A high standard deviation means users reacted very differently—some reviews got very few likes, while others received a lot.

# SPREAD OF PURCHASE AMOUNT



## Range: 19.97

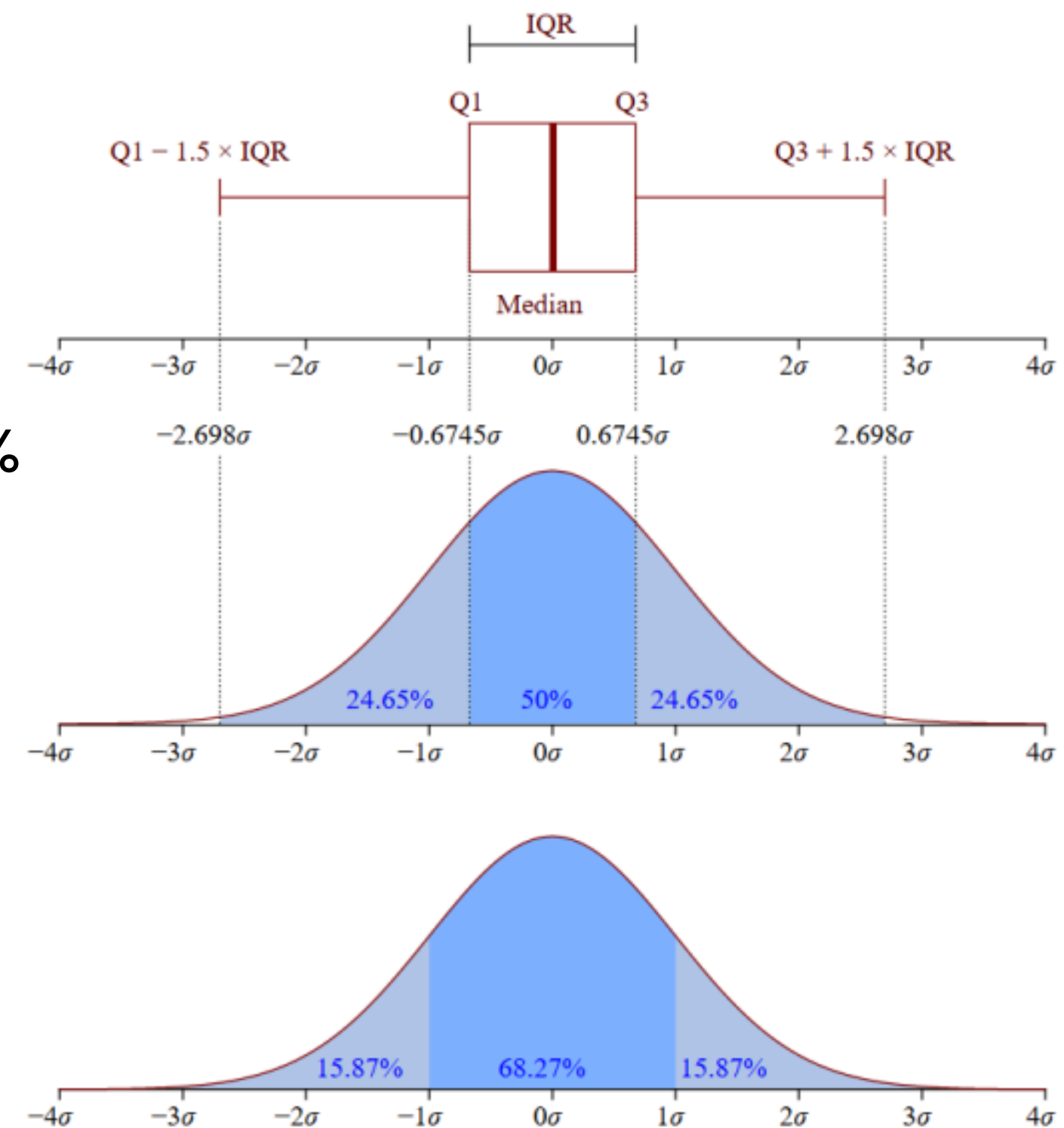
The range of 19.97 shows the difference between the largest and smallest purchase amounts, providing an overall view of variation.

## IQR: 10.1925

The IQR of 10.1925 represents the range of the middle 50% of purchase amounts, focusing on typical values and ignoring outliers.

## Conclusion:

The range gives a broad picture of how much purchase amounts vary, while the IQR highlights the spread of the most common values, helping to detect unusual patterns or outliers.





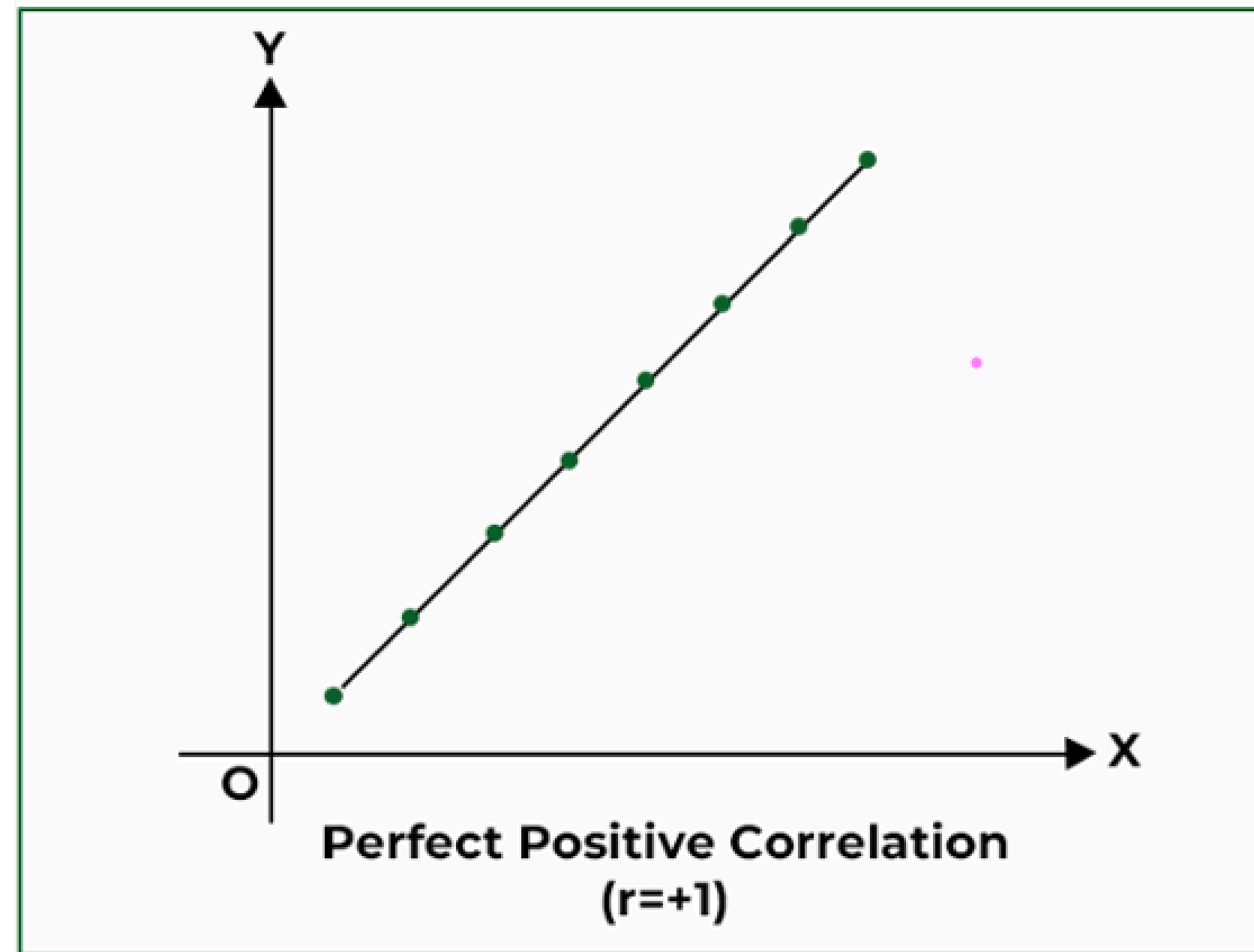
# CORRELATION BETWEEN LIKES AND RATING

**Correlation Coefficient:** 0.8425

A correlation of 0.8425 shows a strong positive relationship between likes and ratings, meaning that more likes usually come with higher ratings.

## **Conclusion:**

This strong connection suggests that user satisfaction (ratings) and engagement (likes) are closely linked.





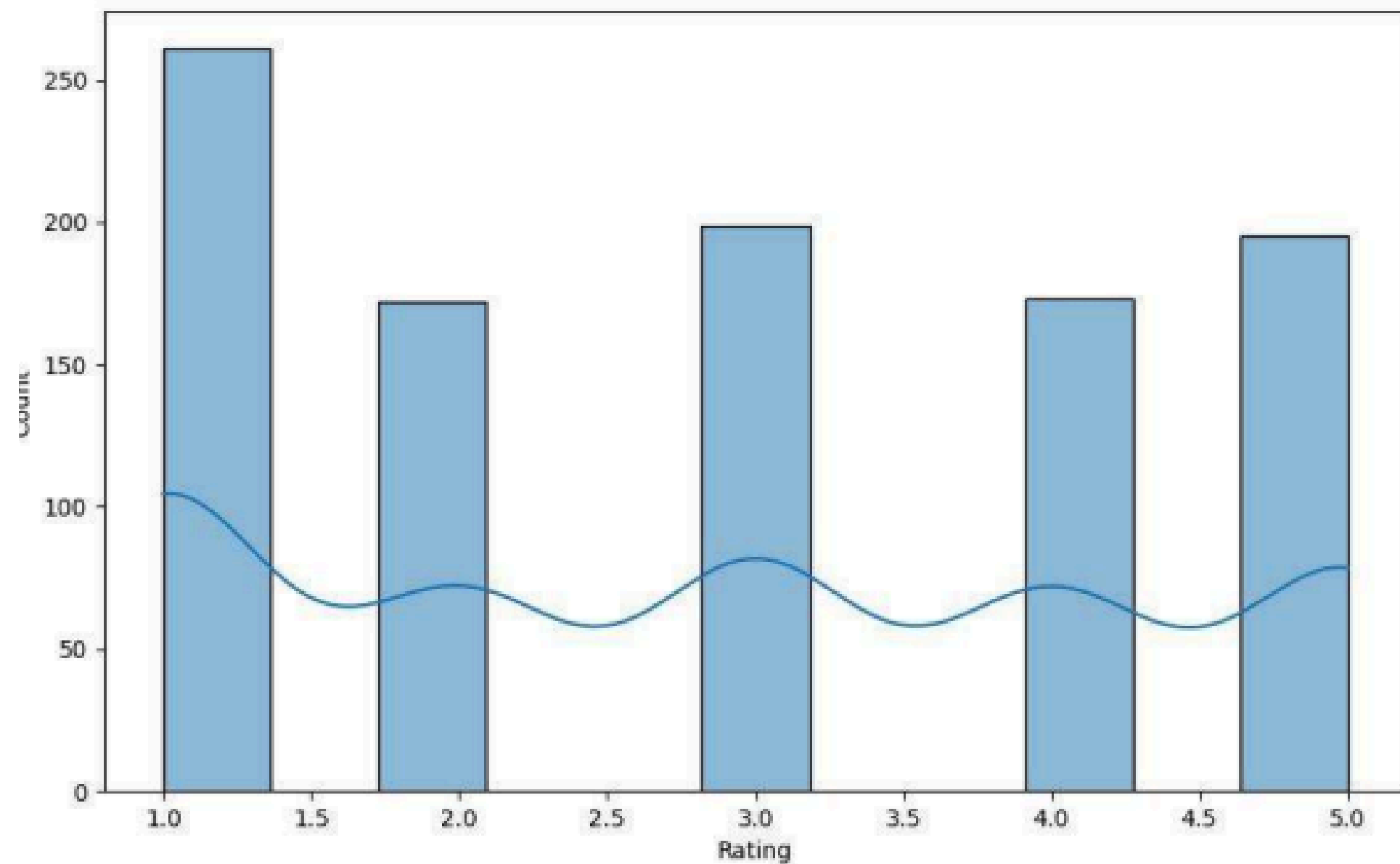
# DISTRIBUTION OF APP RATINGS

## Skewness (based on visual analysis):

If the distribution is skewed to the right (positively), it means most users gave low ratings, and a few high ratings are increasing the average. If it's skewed to the left (negatively), it means most users gave high ratings.

## Conclusion:

A right-skewed distribution could point to overall dissatisfaction, while a left-skewed distribution would suggest user satisfaction.





# HYPOTHESIS TEST FOR INSTAGRAM VS. WHATSAPP RATINGS



## Hypothesis Testing Framework:

### Null Hypothesis ( $H_0$ ):

The average rating for Instagram is equal to or less than the average rating for WhatsApp.

$$H_0: \mu_{\text{Instagram}} \leq \mu_{\text{WhatsApp}}$$

### Alternative Hypothesis ( $H_1$ ):

The average rating for Instagram is greater than the average rating for WhatsApp.

$$H_1: \mu_{\text{Instagram}} > \mu_{\text{WhatsApp}}$$

**T-statistic: -0.7967**

**P-value: 0.7868**

### Interpretation:

A high p-value means we fail to reject the null hypothesis.

### Conclusion:

There is no significant difference between the average ratings for Instagram and WhatsApp, indicating that users perceive both apps similarly.



# SAMPLING DISTRIBUTION AND CENTRAL LIMIT THEOREM



## Sampling Distribution:

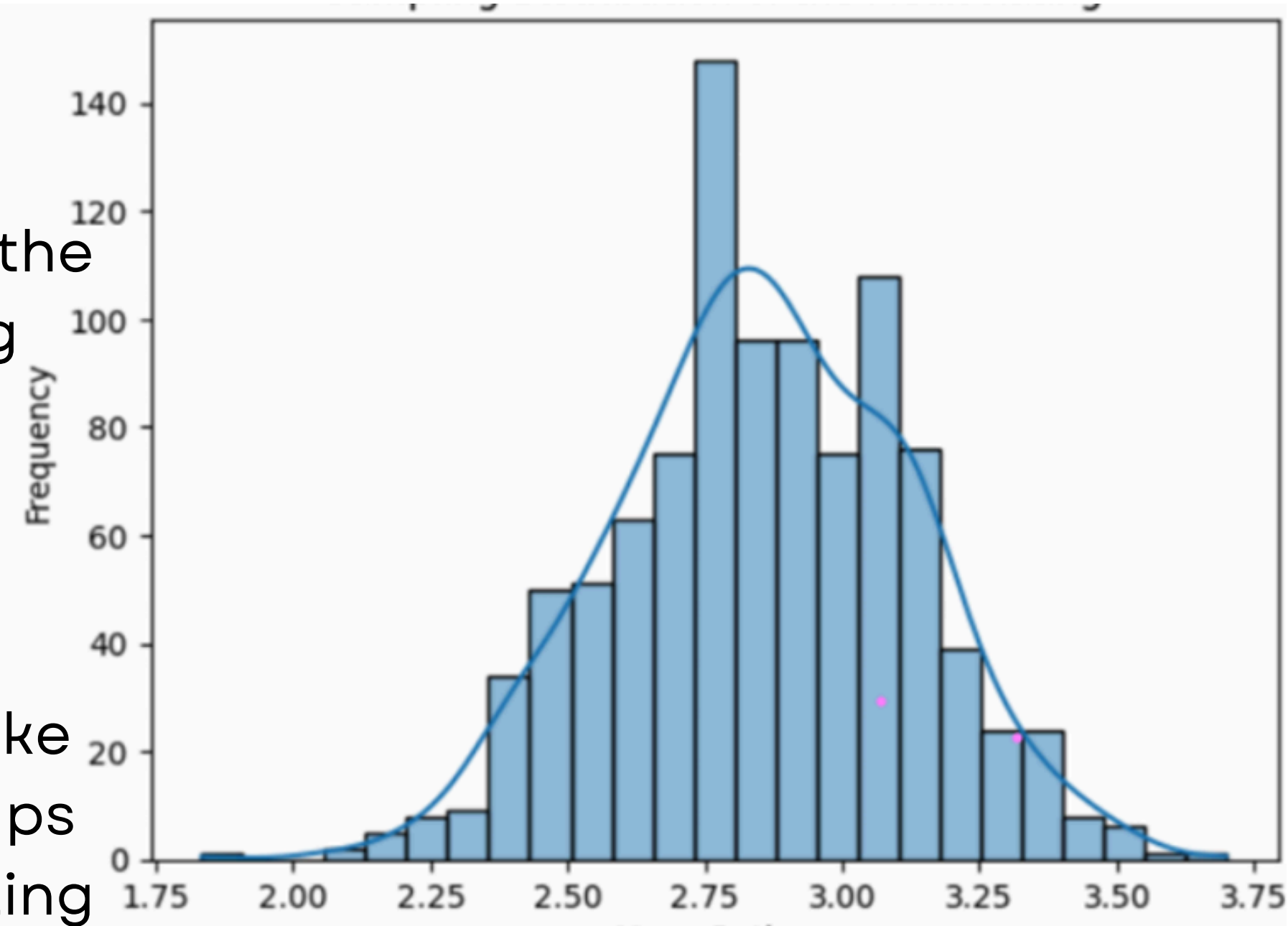
The histogram of sample means shows a roughly normal distribution.

## Analysis:

According to the Central Limit Theorem (CLT), if the sample size is large enough ( $n \geq 30$ ), the sampling distribution of the mean will be normally distributed. The histogram here shows a clear normal shape, confirming the CLT applies.

## Conclusion:

This means we can use normal distribution to make inferences about the population mean, which helps in conducting reliable hypothesis tests and creating confidence intervals.





# CONCLUSION:

- The strong positive correlation (0.8425) between likes and ratings indicates that reviews with higher ratings tend to get more likes.
- Hypothesis testing showed no significant difference between Instagram and WhatsApp ratings.
- The range and IQR of purchase amounts indicate moderate variation in spending habits.
- A significant variance in likes reflects different levels of user engagement with app reviews.
- The median rating of 3 shows mixed user satisfaction, with many low ratings (mode of 1).
- The Central Limit Theorem confirmed reliable inferences, aiding in improvements to app features and user satisfaction.



# LET'S CONNECT WITH ME!



<https://www.linkedin.com/in/saranshgarg11/>



<https://github.com/Saransh-garg>



[gargsaransh11@gmail.com](mailto:gargsaransh11@gmail.com)