DEVI AHILYA VISHWAVIDYALAYA, INDORE

SCHOOL OF STATISTICS

SESSION 2021-2025

Impact of Attendance on Academic Performance

A Project report submitted for partial fulfillment of the Degree of Bachelor of Science

GUIDED BY:

Dr. RASHMI AWAD

ASST. PROFESSOR

SCHOOL OF STATISTICS

SUBMITTED BY:

VANSHIKA DHAKAD

ST4A2108

B.Sc. (Hons)-ASA

Forward

I feel great pleasure to forward the project entitled "**Impact of Attendance on Academic Performance**" submitted by **Vanshika Dhakad** (5^{th} semester) of School of Statistics, DAVV, Indore for the degree of Bachelor of Science (Hons) Applied Statistics and Analytics.

Dr. V.B. GUPTA
HEAD
SCHOOL OF STATISTICS
DAVV, INDORE (M.P.)

CERTIFICATE

This is to certify that Vanshika Dhakad, a student of DAVV University, has successfully completed the project titled "Impact of Attendance on Academic Performance" under the supervision of Dr. Rashmi Awad. This project was conducted as a part of B.sc (Hons) degree during the academic year 2021-2025.

This certificate is awarded in recognition of the successful completion of the project and the demonstration of scholarly achievement.

Date: 23-12-2023

Dr. Rashmi Awad
Asst. Professor
School Of Statistics
DAVV, Indore

DECLARATION

I, Vanshika Dhakad, a student enrolled in B.SC (Hons) at DAVV University, solemnly declare that the project titled "Impact of Attendance on Academic Performance" is my original work. All aspects of this project, including the research design, data collection, analysis, and interpretation, have been carried out by me under the guidance of Dr. Rashmi Awad.

I further declare that:

The project has not been submitted for any other degree or academic qualification. Any external sources used in this project have been duly acknowledged through proper citation.

The data and findings presented in this project are authentic, and any discrepancies or inaccuracies are unintentional. I acknowledge that any act of plagiarism or academic dishonesty in this project is a serious violation of ethical standards, and I am fully aware of the consequences that may arise as a result of such misconduct.

I understand the importance of academic integrity and take full responsibility for the content and execution of this project.

Date: 23-12-2023

Name: Vanshika Dhakad

ROLL no: ST4A2108

ACKNOWLEDGEMENT

First and foremost, I express my deepest gratitude to Dr. Rashmi Awad for their unwavering guidance, insightful feedback, and constant support throughout the entire research process. Their expertise and commitment to academic excellence have been instrumental in shaping the direction and quality of this project.

I would like to thank the faculty and staff of the School Of Statistics at for providing an environment conducive to research and learning. Their dedication to fostering intellectual curiosity has greatly enriched my academic experience. Special thanks to my classmates and friends who provided invaluable assistance and encouragement.

I also express my gratitude to the participants of this study, whose willingness to share their attendance records and academic data contributed significantly to the depth and reliability of the findings.

This endeavor has been a culmination of collective efforts, support, and encouragement from various individuals, without whom this project would not have been possible.

Vanshika Dhakad

B.Sc (Hons)- Applied Statistics & Analytics

Devi Ahilya Vishwavidyalaya

CONTENT

Abstract:

The relationship between student attendance and academic performance has long been a subject of scholarly investigation within the educational landscape. The overarching aim is to discern whether a statistically significant difference exists in mean grades between students exhibiting high and low attendance. This study seeks to rigorously examine the impact of attendance on grades, employing a quantitative approach centered around a t-test analysis.

The study draws upon attendance records and corresponding grades of a diverse sample of students, utilizing predetermined criteria to categorize them into high and low attendance groups. A two-sample t-test is employed to compare the mean grades of these groups, and statistical significance is determined based on a predetermined alpha level, often set at 0.05.

The significance of this study lies in its potential to inform educational policies and practices. Understanding the statistical relationship between attendance and academic achievement can contribute to the development of targeted interventions, aiding educational institutions in enhancing student success and addressing potential barriers to learning.

Preliminary findings indicate that attendance may indeed play a role in determining academic success. The statistical analysis provides insights into whether students with higher attendance records consistently achieve higher grades, shedding light on the potential causal link between attendance and academic performance.

In conclusion, this study contributes to the ongoing discourse on effective educational strategies by providing empirical evidence on the impact of attendance on grades. The findings have implications for educators, administrators, and policymakers, offering insights that can inform decisions aimed at optimizing student outcomes within the academic environment. Further research may explore nuanced factors influencing attendance and the effectiveness of interventions designed to improve overall student engagement and success.

Introduction:

The relationship between attendance and academic performance has been a topic of interest and debate in educational research for many years. As students progress through their academic journeys, the question of whether regular attendance in classes directly correlates with higher grades remains a significant concern.

This project aims to explore and analyze the impact of attendance on grades, shedding light on the potential implications for both students and educators.

The significance of this study lies in its potential to inform educational policies and practices. Understanding the connection between attendance and academic achievement can help institutions develop strategies to enhance student success, identify students at risk of falling behind, and design interventions to improve overall educational outcomes.

The question of whether attendance plays a pivotal role in determining academic outcomes has practical implications for both educators and students. Understanding this relationship can guide educational institutions in developing targeted strategies to enhance student learning experiences and improve overall academic performance.

Objective:

1. Is there any relationship between class attendance and students' performance?

2. To what extent does the influence of classroom attendance affect the students' performance?

3. Are the final examination marks for students with high class attendance better than compared to those with low class attendance?

Research Problem:

Despite the widely acknowledged importance of student attendance in academic success, there remains a need for a comprehensive analysis to determine the specific impact of attendance on students' grades. While numerous studies suggest a positive correlation between attendance and academic performance, a detailed investigation using statistical methods, such as a T-Test, is required to ascertain the extent of this relationship and its implications for educational institutions.

Research Methodology:

Data collection:

We have taken the primary data for this project report. A structured questionnaire has been designed which contains some questions which are required to obtain the results. The questionnaire is prepared in google forms and all the respondents approached via social media platforms. The respondents are the students who are from different education qualification such as high school, under graduate, post graduate, PhD. Every efforts were made to ensure that the sample is the true representative of the population.

Data analysis:

Data analysis using graphs is a powerful way to visually represent and interpret information.

Tool used: Power Bi

Power BI is a powerful tool for data visualization, allowing users to create compelling and interactive visualizations that help in analyzing and understanding data. Power BI provides a wide range of visualization options, including bar charts, line charts, pie charts, tables, matrices, maps, and more. Users can choose from various visualization types to best represent their data.

In addition to the built-in visualizations, Power BI supports custom visuals created by the community and developers. This allows users to extend the range of visualizations available in Power BI. Power BI enables cross-filtering and highlighting interactions between visualizations. Selecting data points in one visualization can automatically filter or highlight related data in other visualizations on the same page.

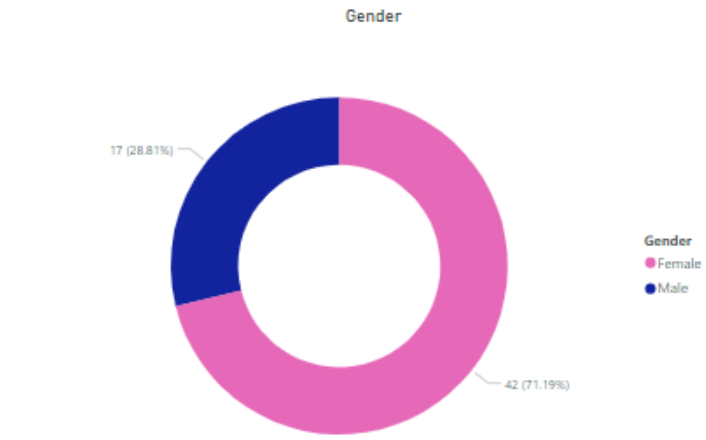
Slicers and filters allow users to interactively control what data is displayed in their visualizations. Users can filter data based on specific criteria, time ranges, or categories.

Power BI integrates seamlessly with Excel, enabling users to import data from Excel and use familiar Excel features for further analysis. Excel charts and tables can also be embedded directly into Power BI reports.

Whether you are a business analyst, data scientist, or decision-maker, Power BI's robust data visualization capabilities make it a valuable tool for transforming raw data into meaningful insights that drive informed decision-making.

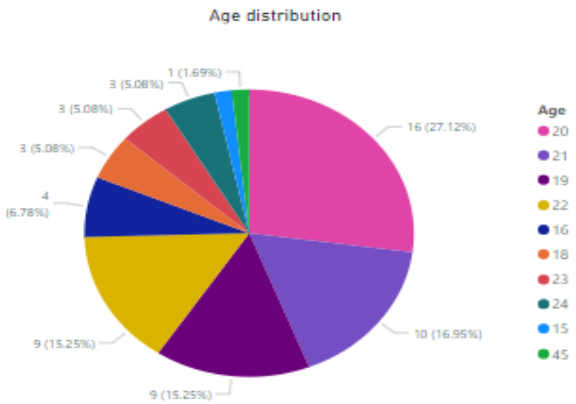
When investigating the impact of attendance on grades in a project or study, it's important to consider various variables to ensure a comprehensive analysis:

Gender: Each slice corresponds to a specific gender category, and the size of each slice is proportional to the percentage or count of individuals belonging to that gender. There are total 59 respondents in which 17 are male while 42 are female. Also entire pie represents 100%, with each slice representing its share of that total, 28.81% male and 71.19% female.

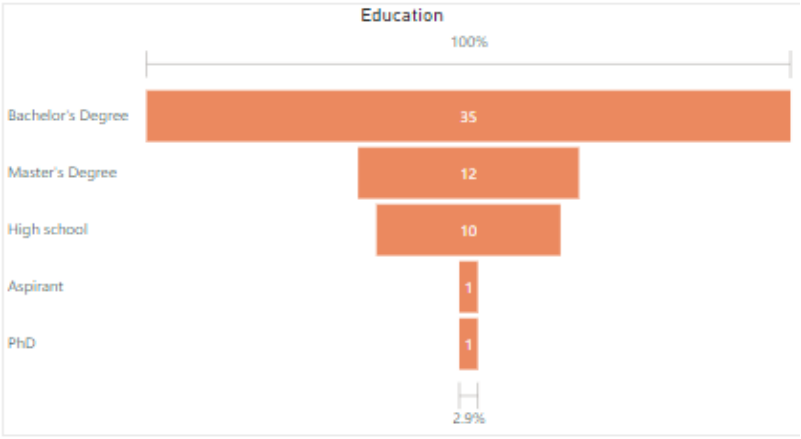


Age: The size of each slice is proportional to the percentage or count of individuals in the dataset belonging to that specific age group. Each age group is assigned a distinct color to visually differentiate between them. The larger the slice, the greater the percentage or count of individuals belonging to that age group, here 27.12% respondents belongs to age of 20 year old which is

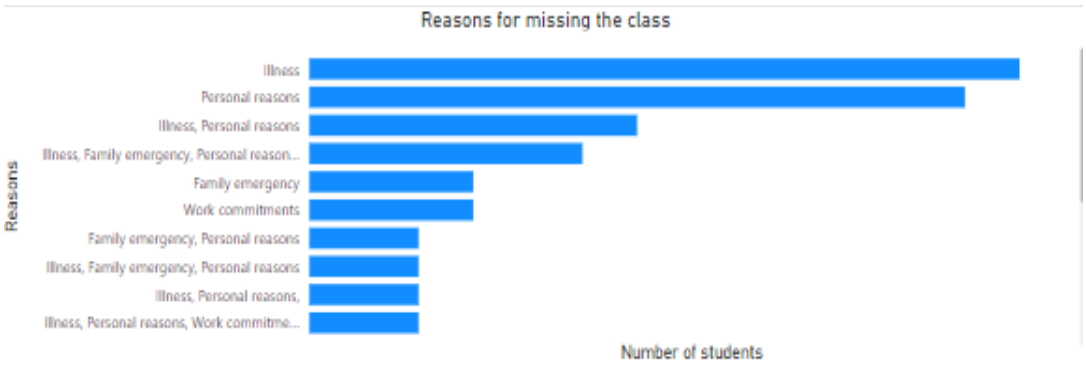
largest in proportion,16.95% are of age 21,15.25% are of age 19 and age 22,6.78% are of age16 and similarly we can see other.



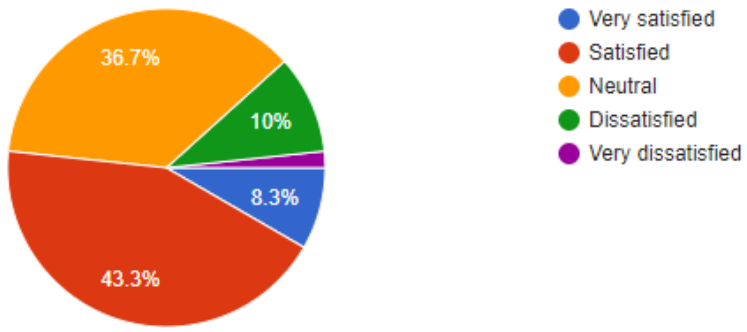
Education Qualification: Each bar in the graph represents a different education qualification, such as "High School Diploma," "Bachelor's Degree," "Master's Degree," and "Doctorate," among others. The vertical axis (y-axis) typically represents the different education qualifications, while the horizontal axis (x-axis) represents the percentage or count of individuals. Longer bars represent a higher percentage or count of individuals with that particular education qualifications .A total of 35 respondents are pursuing bachelor's degree,12-master's degree,10-high school, and a single respondent from PhD and a aspirant.



Reasons for missing classes: Each bar in the graph represents a specific reason for missing class. Common reasons might include "Illness," "Family Emergency," "Work Commitments," "Transportation Issues," "Personal Reasons," etc. The x-axis typically represents the frequency or percentage of students, while y-axis represents the different reasons for missing class. Taller bars represent a higher frequency or percentage of students citing that particular reason for missing class i.e. for illness.



Satisfaction levels: Here we represent the satisfaction level of respondents with their grades. Each slice in the pie chart represents a different satisfaction level. Common categories might include "Very Satisfied," "Satisfied," "Neutral," "Dissatisfied," and "Very Dissatisfied." The 8.3% of the respondents are very satisfied, 43.3% are satisfied, 36.7% are neutral and so on.

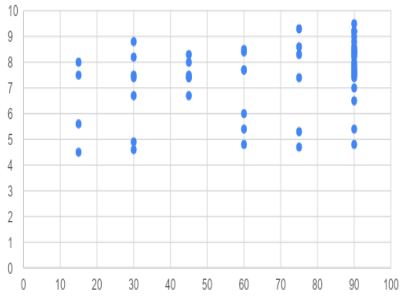


Statistical Analysis:

Correlation: Correlation is then being analyzed to determine whether there is a significant relationship between class attendance and final examination marks. Pearson correlation is then being used since both attendance and final examination marks are quantitative variables.

Pearson correlation

r=0.3521



Non-linear Relationship in the Scatter Plot: A non-linear relationship in the scatter plot implies that the association between attendance and grades doesn't follow a straight line. It could be a curve or some other pattern.

Correlation Coefficient (0.3521): The correlation coefficient of 0.3521 indicates a positive but weak linear relationship between attendance and grades.

Interpretation:

The positive correlation coefficient suggests that, on average, higher attendance tends to be associated with higher grades, even though the relationship is not very strong.
The non-linear pattern in the scatter plot implies that the relationship is not adequately described by a straight line.

Considerations:

Weak Linear Association: While there is a weak linear correlation, it's important to note that non-linear relationships may not be accurately represented by the correlation coefficient.
Other Influencing Factors: The weak correlation suggests that other factors, not captured by attendance alone, may also influence grades.

Causation: Correlation does not imply causation. Even with a correlation, it does not mean that higher attendance causes higher grades.

Conclusion:

The analysis of attendance and grades reveals a positive but weak linear relationship with a correlation coefficient of 0.3521. The scatter plot indicates a non-linear pattern, suggesting that the association between attendance and grades is not adequately captured by a simple straight line.

While there is evidence of a tendency for higher attendance to be associated with higher grades, the weak correlation implies that attendance alone is not a strong predictor of academic success. Other factors, not accounted for in this analysis, may play a significant role in determining grades.

Further exploration, potentially employing more sophisticated statistical techniques to capture non-linear relationships, is recommended to gain a more comprehensive understanding of the dynamics between attendance and academic performance. It is crucial to exercise caution in drawing causal conclusions and to consider the multifaceted nature of factors influencing student achievement.

Statistical test:

In order to analyze the impact of attendance on grades, a t-test can be employed to assess whether there is a statistically significant difference in the mean grades between students with high attendance and those with low attendance.

An introduction to t-test:

The t-test is a statistical method used to compare the means of two groups and determine if the observed differences are likely due to random chance or if they are statistically significant.

It is a parametric test commonly used in hypothesis testing when the data is approximately normally distributed and the variances are assumed to be equal.

A T-test studies a set of data gathered from two similar or different groups to determine the probability of the difference in the result than what is usually obtained. The accuracy of the test depends on various factors, including the distribution patterns used and the variants influencing the collected samples. Depending on the parameters, the test is conducted, and a T-value is obtained as the statistical inference of the probability of the usual resultant being driven by chance.

For example, if one wishes to figure out if the mean of the length of petals of a flower belonging to two different species is the same, a T-test can be done. The user can select petals randomly from two other species of that flower and come to a standard conclusion.

The final **T-test interpretation** could be obtained in either of the two ways:

- A null hypothesis signifies that the difference between the means is zero and where both the means are shown as equal.
- An alternate hypothesis implies the difference between the means is different from zero. This hypothesis rejects the null hypothesis, indicating that the data set is quite accurate and not by chance.

This T-test, however, is only valid and should be done when the mean or average of only two categories or groups needs to be compared.

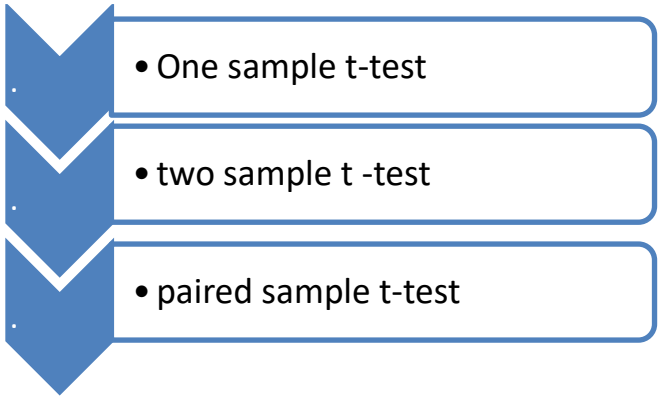
Assumptions:

The test runs on a set of assumptions, which are as follows:

- The measurement scale used for such hypothesis testing follows a set of continuous or ordinal patterns. The accounted parameters and variants influencing the samples and surrounding the groups are based on the standard consideration.
- The tests are completely based on random sampling. As no individuality is maintained in the samples, the reliability is often questioned.
- When the data is plotted with respect to the T-test distribution, it should follow a normal distribution and bring about a bell-curved graph.
- The variance should be such that the standard deviations of the samples are almost equal.

Types:

Some of the widely used **T-test types** are as follows:



#1 – One-Sample T-Test

While performing this test, the mean or average of one group is compared against the set average, which is either the theoretical value or means of the population. For example, a teacher wishes to figure out the average weight of the students of class 5 and compare the same against a set value of more than 45 kgs. The teacher first randomly selects a group of students and records individual weights to achieve this. Next, she finds out the mean weight for that group and checks if it meets the standard set value of 45+. The formula used to obtain one-sample t-test results is:

$$t = \frac{m - \mu}{s/\sqrt{n}}$$

Where,

- T = t-statistic
- m = mean of the group
- = theoretical mean value of the population
- s = standard deviation of the group
- n = sample size

#2 – Independent Two-Sample T-Test

This is the test conducted when samples from two different groups, species, or populations are studied and compared. It is also known as an independent T-test. For example, if a teacher wants to compare the height of male students and female students in class 5, she would use the independent two-sample test.

The **T-test formula** used to calculate this is:

$$t = \frac{m_A - m_B}{\frac{\sqrt{s^2}}{\sqrt{n_A}} + \frac{\sqrt{s^2}}{\sqrt{n_B}}}$$

Where,

- mA – mB = means of samples from two different groups or populations
- nA – nB = respective sample sizes
- s2 = standard deviation or common variance of two samples

#3 – Paired Sample T-Test

This hypothesis testing is conducted when two groups belong to the same population or group. The groups are studied either at two different times or under two varied conditions. The formula used to obtain the t-value is:

$$t = \frac{m}{s/\sqrt{n}}$$

Where,

- T = t-statistic
- m = mean of the group
- = theoretical mean value of the population
- s = standard deviation of the group
- n = sample size

Independent t-test: We divide the data collected into two groups –one of students achieved grades with high attendance and the other of students achieved grades low attendance. To compare the final grades obtained between students with class attendance high and class attendance low, independent t-test was carried out since two groups are considered. Independent sample t-test was carried out to compare the student's results between the two-attendance groups.

1. Formulate Hypotheses:

- **Null Hypothesis (H0):** There is no significant difference in grades between students with high attendance and low attendance.
  - $H_0 : \mu_1 = \mu_2$
- **Alternative Hypothesis (H1):** There is a significant difference in grades between students with high attendance and low attendance.
  - $H_1 : \mu_1 \neq \mu_2$ (two-tailed test)

OR

  - $H_1 : \mu_1 > \mu_2$ (one-tailed test, if you expect higher attendance to lead to higher grades)

OR

  - $H_1 : \mu_1 < \mu_2$ (one-tailed test, if you expect higher attendance to lead to lower grades)

2. Assumption: The independent t-test assumes that the populations from which the samples are drawn are normally distributed and have equal variance.

3. Perform t-test:

	High Attendence	*Low Attendence*
Mean	7.962068966	7.006666667
Variance	1.121009852	1.941333333
Observations	29	30
Pooled Variance	1.538367413	
Hypothesized Mean Differer	0	
df	57	
t Stat	2.957943448	
P(T<=t) one-tail	0.002251014	
t Critical one-tail	1.672028888	
P(T<=t) two-tail	0.004502027	
t Critical two-tail	2.002465459	

4. Significance level: The most common significance level is 0.05 (5%). This means that if the p-value obtained from a statistical test is less than or equal to 0.05, the null hypothesis is rejected at the 5% significance level.

5. Decision: The p-value is less than 0.05 at 5% we reject the null hypothesis i.e. there is no significance difference in grades between students with high attendance and low attendance. We find that there is a significance difference in grades between students with high attendance and low attendance.

Discussion:

Although a correlation was observed in this study (r = 0.3521; suggesting weak relationship between attendance and final examination marks), it is not a cause and effect relationship. The original data shows that there exist students with good attendance and still scored low marks in their final exam. No causality can be inferred from the weak correlation between attendance and final examination marks. It is likely that some students are not presence in class in view of that they can learn the material better by spending the same time studying in other approach such as online learning.

Conclusion:

In this study, we investigated the relationship between attendance and academic performance, specifically focusing on its impact on students' grades. The data was collected from respondents and the analysis aimed to explore whether there is a statistically significant association between attendance and grades. Students with higher attendance records tended to perform better academically, as evidenced by a positive correlation between attendance percentages and overall grades. This result suggests that consistent attendance is associated with improved academic outcomes. A rigorous statistical analysis, including t-test, revealed that there is a statistically significant relationship between attendance and grades.

Recommendations:

1. **Promote Awareness:**

- Increase awareness among students about the positive correlation between attendance and academic performance. Develop targeted campaigns to emphasize the importance of regular attendance in achieving higher grades.

2. **Implement Attendance Monitoring Systems:**

- Consider implementing an attendance monitoring system that provides real-time tracking of student attendance. This can help identify at-risk students early on and facilitate timely interventions.

3. **Early Intervention Strategies:**

- Develop and implement early intervention strategies for students with consistently low attendance. This could include personalized counseling, academic support programs, or mentoring initiatives to address underlying issues affecting attendance.

4. **Engagement Initiatives:**

- Design engagement initiatives to enhance the overall learning experience and encourage active participation in classes. Incorporate interactive teaching methods, group activities, and technologies to make classes more engaging and relevant.

5. **Flexible Learning Options:**

- Explore the feasibility of offering flexible learning options, such as online classes or recorded lectures, to accommodate students with unique scheduling challenges. This can contribute to improved attendance rates.

6. **Faculty Training:**

- Provide training and resources to faculty members on effective teaching methods that promote student engagement. Foster a positive and inclusive classroom environment that motivates students to attend regularly.

7. **Incentive Programs:**

- Consider implementing incentive programs to reward and recognize students with excellent attendance records. Positive reinforcement, such as certificates, small rewards, or acknowledgment ceremonies, can motivate students to prioritize attendance.

8. **Continuous Monitoring and Evaluation:**

- Establish a system for continuous monitoring and evaluation of attendance-related initiatives. Regularly assess the impact of implemented strategies on attendance rates and academic performance, and make adjustments as needed.

9. **Collaboration with Support Services:**

- Foster collaboration between academic departments and support services, such as counseling and student affairs. This interdisciplinary approach can address both academic and non-academic factors influencing attendance.

10. **Further Research:**

- Encourage further research to explore additional factors that may contribute to attendance patterns and academic performance. Investigate the effectiveness of specific interventions and gather feedback from students to inform continuous improvement.

References:

- Anikeeff, M. (1954). The Relationship between Class Absences and College Grades. Journal of Educational Psychology, 45, 244-249.
- Fundamentals of Applied Statistics (S. C. Gupta v. K. Kapoor).
- An Introduction to Statistical Learning with R by Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshiran.
- The Impact of Class Attendance on Student Performance Author: Credé, M. Journal: Journal of College Student Development..
- Class Attendance and Exam Performance: A Randomized Experiment" Author: Bettinger, E. P., Fox, L., Loeb, S., & Taylor, E. S.Journal: The Economic Journal.