AN ANALYSIS REPORT BASED ON THE DATA FROM AITECH STUDENT SURVEY.

INTRODUCTION.

This report explores the key aspects of student satisfaction at AiTech. Leveraging a comprehensive survey conducted among students, this analysis provides insights into various critical areas. This analysis focused on five key questions:

1. What does the data look like overall? Are there any outliers or interesting patterns in the distribution of key variables? (**Use Descriptive Statistics**)
2. (a) Does the overall quality of the learning environment differ between Program A and Program B? (**Use Independent Samples T-Test**)
   (b) Does the quality of instruction vary depending on how long students have been enrolled? (**Use One-Way ANOVA**)
3. Is there a connection between how students heard about the program and the program they enrolled in? (**Use Chi-Square Test**)
4. To what extent is the quality of instruction related to the overall perceived quality of the learning environment? (**Use Correlation Analysis**)
5. How are students distributed across different programs in terms of their enrollment duration? (**Use Crosstabulation**).

The analysis used information gathered by the AITECH Data team, which included 42 students enrolled in different AITECH programs. The data included variables on:

- Program Duration
- About Us
- Instruction Clarity on Instructors
- Course Content Relevance
- Availability of Learning Materials
- Classroom Discussions
- Staff Support
- Learning Management System
- Learning Environment
- Teaching Methods
- Program Enrolled

STEPS.

I started by importing the CSV file, after which I updated the original names of each column and assigned labels. Additionally, I verified that the measures matched the data types.

To get rid of duplicates, I went back to the data view and recoded the categorical variables. However, I made the mistake of changing the unique ID to a string, which prevented it from displaying. I had to go through the entire process again to get the desired results. After that, I

formatted my working data in a new table, ungrouped the unnecessary data, and got my new table ready for analysis.

FINDINGS.

DESCRIPTIVE STATISTICS

Descriptive Statistics

	N	Range	Minimum	Maximum	Sum	Mean		Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Quality of instruction clarity knowledge engagement of instructors	42	4	1	5	85	2.02	.138	.897	.804	1.230	.365	2.249	.717
Relevance and usefulness of course content	42	4	1	5	86	2.05	.163	1.058	1.120	1.328	.365	1.580	.717
Availability and comprehensiveness of learning materials	42	4	1	5	99	2.36	.163	1.055	1.113	.658	.365	.237	.717
Effectiveness of classroom discussions and activities	42	4	1	5	78	1.86	.143	.926	.857	1.458	.365	2.667	.717
Accessibility and helpfulness of support staff	42	3	1	4	82	1.95	.136	.882	.778	.767	.365	.101	.717
Usability and effectiveness of the Learning Management System	42	4	1	5	100	2.38	.156	1.011	1.022	.639	.365	-.115	.717
Overall quality of the learning environment including physical	42	3	1	4	86	2.05	.118	.764	.583	.609	.365	.560	.717
Teaching methods employed by instructors in your program	42	4	1	5	84	2.00	.149	.963	.927	1.033	.365	1.172	.717
Valid N (listwise)	42												

I started with Descriptive analysis, trying to get an overview of the data distribution. I noticed the mean fell between 1.86 & 2.38, indicating a positive perception of the learning environment, and then most skewness values were close to 0 suggesting a symmetrical distribution of responses. I also noticed some kurtosis values were slightly above 3, meaning the distributions might have more chances of having outliers with scores concentrated around the mean. The standard deviation suggested a high level of agreement among participants in their answers.

INDEPENDENT SAMPLE T-TEST

The mean comparison using the independent sample t-test was to determine if there was a difference in the learning environment between students enrolled in Cyber Security and Data Analysis.

Double-click to activate

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Overall quality of the learning environment including physical	Equal variances assumed	.462	.503	.322	27	.750	.117	.362	-.626	.859
	Equal variances not assumed			.290	5.297	.783	.117	.402	-.899	1.132

Null hypothesis: There is no difference between the learning environment and the students enrolled in the Cyber Security and Data Analysis program.
Alternative Hypothesis: There is a difference between the learning environment and the students enrolled in the Cyber Security and Data Analysis program.
Significance Level: 0.05
P-Value: 0.750

Initially, the output result compared the overall quality of the learning environment among students enrolled in Cyber Security and Data Analysis.
Levene's Test for equality of variances shows;
F: 0.462
Significance level: 0.503

Since the significance level here is greater than 0.05, we fail to reject the assumption of equal variances between the two programs, therefore we can interpret the results based on equal variances assumed.
t = 0.322
df = 27
Sig. (2-tailed) = 0.750

The P-value (0.750) is much greater than the significance level (0.05), therefore we don't have enough evidence to reject the null hypothesis, hence there's no significant difference or relationship between the learning environment and students enrolled in the Cyber Security and Data Analysis program.

In non-technical terms, the learning environment does not influence the students enrolled in Cyber Security and Data Analysis programs.

ONE-WAY ANOVA TEST

The mean comparison using the One-way ANOVA was to determine the difference in the quality of instruction among students categorized by their enrollment duration.

Null hypothesis: There is no difference in the quality of instruction among students categorized by their enrollment duration.
Alternative Hypothesis: There is a difference in the quality of instruction among students categorized by their enrollment duration.
Significance Level: 0.05
P-Value: 0.963

The null hypothesis cannot be rejected since the P-value (0.963) is higher than the significance level (0.05), therefore, there is no difference in the quality of instruction among students categorized by their enrollment duration. This data suggests that students' views in general on the quality of instruction seem to be consistent over enrollment periods.

CHI-SQUARE TEST

The chi-square test is supposed to determine the association between how students heard about AITECH programs and the program they enrolled in

Chi-Square Test

Frequencies

How did you hear about AiTech

	Observed N	Expected N	Residual
A friend	3	4.7	-1.7
AiTech Staff	2	4.7	-2.7
Facebook	4	4.7	-.7
Friend	1	4.7	-3.7
Google	3	4.7	-1.7
Instagram	21	4.7	16.3
Past Student	4	4.7	-.7
Present Student	3	4.7	-1.7
Relative	1	4.7	-3.7
Total	42		

What program are you currently enrolled in

	Observed N	Expected N	Residual
Cyber Security	5	10.5	-5.5
Data Analysis and Science	24	10.5	13.5
DevOps	2	10.5	-8.5
Software Engineering	11	10.5	.5
Total	42		

Test Statistics

	How did you hear about AiTech	What program are you currently enrolled in
Chi-Square	66.429[a]	27.143[b]
df	8	3
Asymp. Sig.	<.001	<.001

a. 9 cells (100.0%) have expected frequencies less than 5. The minimum expected cell frequency is 4.7.

b. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 10.5.

Null hypothesis: There is no significant association between how students heard about AITECH programs and the program they enrolled in
Alternative Hypothesis: There is a significant association between how students heard about AITECH programs and the program they enrolled in
Significance Level: 0.05
P-Value: 0.001

From the Output, the chi-square of both categories (66.429 & 27.143) are relatively high values meaning they suggest a strong relationship between the two variables.
The residual column also highlights the difference between the expected and observed values. The large positive residual in the Instagram column indicates that probably many more students than expected heard about the program through Instagram.

So, since the asymptotic distribution is (0.001) for both tests, we reject the null hypothesis in both cases. This confirms a significant association between how students heard about the program and the program they enrolled in.

CORRELATION ANALYSIS

The correlation analysis is supposed to examine the relationship between equality of instruction with overall quality of the learning environment.

Null hypothesis: There is no significant relationship between equality of instruction with overall quality of the learning environment
Alternative Hypothesis: There is a significant relationship between equality of instruction with overall quality of the learning environment
Significance Level: 0.05
P-Value: 0.001

Since P-Value (0.001) is lower than the significance level (0.05), we reject the null hypothesis. This means we have evidence that there's a significant relationship between these two variables.
The correlation coefficient (0.568) being positive indicates a positive linear relationship on the quality of instructions and the overall quality of the learning environment.

CROSSTABULATION ANALYSIS

The crosstabulation analysis explores the distribution of students' duration of enrollment across different programs.

Null hypothesis: There is no significant relationship between program duration and program enrolled.
Alternative Hypothesis: There is a significant relationship between program duration and program enrolled.
Significance Level: 0.05
P-Value: 0.010

The chi-square test shows the significant P-value (0.010) and it allows us to reject the null hypothesis, meaning there's evidence that there's a significant relationship between program duration and program enrolled.
In non-technical terms, the program duration influences the program students get enrolled in.