Test 1

Regression Statistics	
Multiple R	0.287419905
R Square	0.082610202
Adjusted R Square	0.040910665
Standard Error	0.287203202
Observations	24

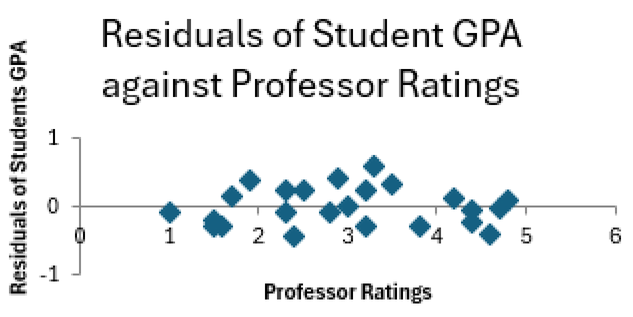
ANOVA

	df	SS	MS	F	Significance F
Regression	1	0.163410896	0.163410896	1.981082024	0.173249259
Residual	22	1.814684937	0.082485679		
Total	23	1.978095833			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	2.980072952	0.166611443	17.88636422	1.36E-14	2.634541968	3.325603937	2.634541968	3.325603937
X Variable 1	0.073681806	0.052349078	1.407509156	0.173249259	-0.034883536	0.182247149	-0.034883536	0.182247149

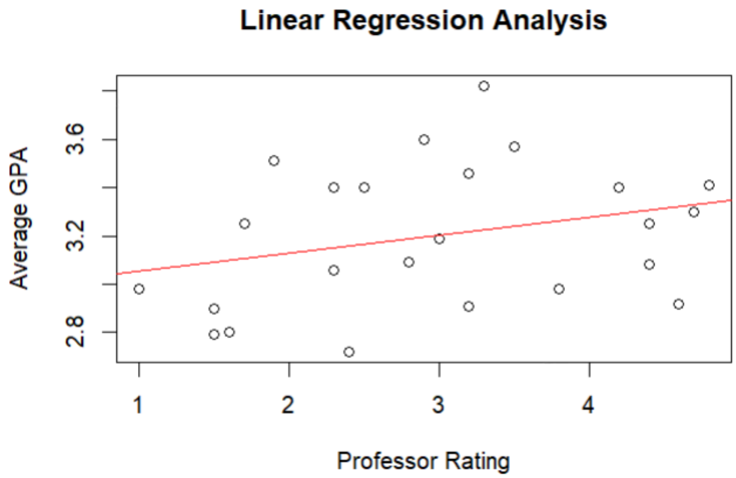
RESIDUAL OUTPUT

Observation	Predicted Y	Residuals
1	3.156909287	-0.436909287
2	3.333745622	0.076254378
3	3.223222913	0.596777087
4	3.105332023	0.144667977
5	3.3042729	-0.0542729
6	3.201118371	-0.011118371
7	3.090595662	-0.190595662
8	3.149541107	0.250458893
9	3.326377442	-0.026377442
10	3.237959274	0.332040726
11	3.149541107	-0.089541107
12	3.090595662	-0.300595662
13	3.289536538	0.110463462
14	3.164277468	0.235722532
15	3.215854732	0.244145268
16	3.120068384	0.389931616
17	3.215854732	-0.305854732
18	3.3042729	-0.2242729
19	3.19375019	0.40624981
20	3.053754759	-0.073754759
21	3.260063816	-0.280063816
22	3.319009261	-0.399009261
23	3.097963842	-0.297963842
24	3.18638201	-0.09638201



We expected that professors with higher rating would have higher GPAs as they are more liked by the students, however based on the data there is no obvious correlation between average GPA and professor rating. The purpose of this test was to see if the ratings a professor receives on ratemyprofessor.com would be any indication of the average GPA of their students. We figured that professors with high ratings would be more effective teachers and that would lead to higher average GPAs in the classroom.

We have also done a linear regression test in R. The results of that can be seen below



However our data shows that there is little correlation between the two variables of professor ratings and average GPAs. The correlation coefficient is 0.287, indicating a weak positive linear relationship between the independent variable and the dependent variable. The R-squared value is 0.082, meaning that the regression model only explains about 8.2% of the variance in the dependent variable. This implies that the independent variable you're considering average professor rating has a very limited effect on the dependent variable average class GPA 2023. Also, the F-statistic is low, and the corresponding p-value is greater than 0.05, indicating that the model is not statistically significant. This means there's a high chance that any relationship observed between the independent variable and the dependent variable could be due to random chance rather than an actual predictive relationship.

R Code for linear regression test

```
# Define the data
professor_ratings <- c(2.4, 4.8, 3.3, 1.7, 4.4, 3, 1.5, 2.3, 4.7, 3.5, 2.3, 1.5, 4.2, 2.5, 3.2, 1.9, 3.2, 4.4, 2.9, 1, 3.8, 4.6, 1.6, 2.8)
average_gpa <- c(2.72, 3.41, 3.82, 3.25, 3.25, 3.19, 2.9, 3.4, 3.3, 3.57, 3.06, 2.79, 3.4, 3.4, 3.46, 3.51, 2.91, 3.08, 3.6, 2.98, 2.98, 2.92, 2.8, 3.09)

# Create a data frame from the vectors
data <- data.frame(professor_ratings, average_gpa)

# Perform linear regression
model <- lm(average_gpa ~ professor_ratings, data = data)

# Print the summary of the regression model
summary(model)

# Plot the data and the regression line
```

```
plot(average_gpa ~ professor_ratings, data = data, main = "Linear Regression Analysis", xlab =
"Professor Rating", ylab = "Average GPA")
abline(model, col = "red")
```

Test 2
Step 1: Set up hypotheses:
Null hypothesis (H0): There is no difference in mean GPAs between the two groups.
Alternative hypothesis (H1): There is a difference in mean GPAs between the two groups.

Step 2: Choose a significance level (α):
Let's assume $\alpha = 0.05$ for this test.

Step 3: Calculate sample statistics:
Calculate the sample mean ($\bar{x}$) and sample standard deviation (s) for each group.
Determine the sample size (n) for each group.

Step 4: Perform the t-test:
Calculate the test statistic (t-value) using the formula:
\[t = \frac{{\bar{x}_1 - \bar{x}_2}}{{\sqrt{\frac{{s_1^2}}{{n_1}} +
\frac{{s_2^2}}{{n_2}}}}} \]
Determine the degrees of freedom (df) using the formula:
\[df = \frac{{\left(\frac{{s_1^2}}{{n_1}} +
\frac{{s_2^2}}{{n_2}}\right)^2}}{{\frac{{\left(\frac{{s_1^2}}{{n_1}}\right)^2}}{{n_1 - 1}}
+ \frac{{\left(\frac{{s_2^2}}{{n_2}}\right)^2}}{{n_2 - 1}}}} \]
Use the t-distribution or statistical software to find the p-value associated with the calculated t-value.

Step 5: Interpret the results:
If the p-value is less than α, reject the null hypothesis and conclude that there is a significant difference in mean GPAs between the two groups.
If the p-value is greater than or equal to α, fail to reject the null hypothesis and conclude that there is not enough evidence to support a difference in mean GPAs between the two groups.

Code: R-Code

```
&gt; # Import data for MEEM department professors
&gt; meem_data &lt;- data.frame(
+ professor = c(&quot;Mahesh Gupta&quot;, &quot;Ossama Abdelkhalik&quot;, &quot;Fernando
Ponta&quot;, &quot;Susanta Ghosh&quot;,
&quot;Ashok Ambardar&quot;,
+ &quot;Nasser alarje&quot;, &quot;Shiyan Hu&quot;, &quot;james davis&quot;, &quot;Rupak
Rajachar&quot;, &quot;megan frost&quot;,
+ &quot;Jeremy Goldman&quot;, &quot;Bruce Lee&quot;, &quot;Paul Nelson&quot;,
&quot;Dana Johnson&quot;, &quot;Latha
Poonamallee&quot;,
+ &quot;Jonathan Leinonen&quot;, &quot;David poplawski&quot;, &quot;Charles Wallace&quot;,
&quot;kim tracey&quot;, &quot;yakov
nekrich&quot;,
```

```
+ &quot;renfang jiang&quot;, &quot;elizabeth reed&quot;, &quot;juergen bierbrauer&quot;,
&quot;iosif pinelis&quot;),
+ gpa = c(2.72, 3.41, 3.82, 3.25, 3.25, 3.19, 2.9, 3.4, 3.3, 3.57,
+ 3.06, 2.79, 3.4, 3.4, 3.46, 3.51, 2.91, 3.08, 3.6, 2.98,
+ 2.98, 2.92, 2.8, 3.09)
+ )
&gt;
&gt; # Import data for professors from the unspecified department
&gt; unspecified_data &lt;- data.frame(
+ professor = c(&quot;Prof1&quot;, &quot;Prof2&quot;, &quot;Prof3&quot;, &quot;Prof4&quot;,
&quot;Prof5&quot;, &quot;Prof6&quot;, &quot;Prof7&quot;, &quot;Prof8&quot;,
&quot;Prof9&quot;,
&quot;Prof10&quot;),
+ gpa = c(3.1, 3.3, 3.5, 3.2, 3.4, 3.6, 3.0, 3.8, 3.2, 3.7)
+ )
&gt;
&gt; # Perform two-sample t-test
&gt; t_test_result &lt;- t.test(meem_data$gpa, unspecified_data$gpa)
&gt;
&gt; # Print the test result

&gt; print(t_test_result)

Welch Two Sample t-test

data: meem_data$gpa and unspecified_data$gpa
t = -1.7482, df = 18.573, p-value = 0.09693
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-0.39675248 0.03591915
sample estimates:
mean of x mean of y
3.199583 3.380000
```