In 2022, there were high expectations for the growth of the 365 company and increased student engagement based on the introduction of new website platform features. Some of these features included an XP system that enabled students to track their progress, level up, and earn rewards by completing various learning objectives. The platform also offered in-app coins that could be exchanged for special awards, a leaderboard where students could compete for top positions in different divisions, earning weekly rewards and advancing up the ladder, and streaks to motivate students to maintain consistent learning habits.

**Task 1**

Your task is to provide insights into the relative engagement levels in Q4 2021 and Q4 2022.

* **Paid-plan Students**

|  |  |  |
| --- | --- | --- |
| paid | min\_wat\_21 | min\_wat\_22 |
| mean | 33.79695842 | 273.0205033 |
| median | 26.33 | 40.28 |
| standard deviation | 28.213543 | 854.5752369 |
| skewness | 0.625157717 | 7.065595612 |
| kurtosis | -0.849282209 | 58.4849581 |

* **Mean**: Among students who watched between 1 and 100 minutes in 2021, the average minutes watched by paid-plan students increased significantly from Q4 2021 to Q4 2022, from approximately 33.80 minutes to about 273.02 minutes. This suggests a substantial increase in engagement among this group of initially low-engagement-paid-plan students.
* **Median**: The median minutes these low-engagement-paid-plan students watched increased from Q4 2021 to Q4 2022, from 26.33 minutes to 40.28 minutes. While this increase is not as dramatic as the increase in the mean, it indicates that the typical student in this group (i.e., the student in the middle of the distribution) also increased their engagement. This suggests that the increase in engagement was more widespread among paid-plan students and not solely driven by a few outliers.
* **Standard Deviation**: The standard deviation for these low-engagement-paid-plan students increased substantially from 28.21 minutes in Q4 2021 to 854.58 minutes in Q4 2022. This indicates a much larger variability in the minutes watched by these students in Q4 2022 compared to Q4 2021. This could be due to a broader range of engagement levels among the students in Q4 2022, with some students watching very little content and others watching a lot of content.
* **Free-Plan Students**

|  |  |  |
| --- | --- | --- |
| free | min\_wat\_21 | min\_wat\_22 |
| mean | 25.39167554 | 117.6374576 |
| median | 14.17 | 11.83 |
| standard deviation | 26.2315369 | 468.9346649 |
| skewness | 1.16773938 | 15.06167594 |
| kurtosis | 0.360033015 | 315.7647882 |

* **Mean:** Among students who watched between 1 and 100 minutes in 2021, the average minutes watched by free-plan students increased from about 25.39 minutes in Q4 2021 to about 117.64 minutes in Q4 2022. This suggests that overall engagement among these initially low-engagement-free-plan students increased during this period. But the extent of this increase is less than what was observed for similar low-engagement-paid-plan students, suggesting that while these free-plan students are watching more content, they're still not as engaged as the equivalent group of paid-plan students.
* **Median:** Interestingly, the median minutes watched by these low-engagement-free-plan students decreased from Q4 2021 to Q4 2022, from 14.17 minutes to 11.83 minutes. This indicates that engagement decreased for the typical student in this group (i.e., the student in the middle of the distribution). The increase in the mean might be driven by a small number of free-plan students who significantly increased their engagement in Q4 2022, while the majority did not increase their engagement or even reduced it.
* **Standard Deviation:** The standard deviation for the low-engagement-free-plan students increased from 26.23 minutes in Q4 2021 to 468.93 minutes in Q4 2022. This indicates a more significant variability in the minutes watched by these students in Q4 2022 compared to Q4 2021. The behavior of these students then became more diverse in Q4 2022, with some watching a lot of content and others watching very little.

**Task 2**

Calculate the skewness and kurtosis of students who watched content in Q4 2021 and Q4 2022.

For **paid-plan students**, the skewness increased from 0.63 in Q4 2021 to 7.07 in Q4 2022.

For paid-plan students, the kurtosis increased from -0.85 in Q4 2021 to 58.48 in Q4 2022.

|  |  |  |
| --- | --- | --- |
| skewness | 0.625157717 | 7.065595612 |
| kurtosis | -0.849282209 | 58.4849581 |

The skewness for **free-plan students** increased from 1.17 in Q4 2021 to 15.06 in Q4 2022, indicating positive skewness.

The kurtosis increased from free-plan students—from 0.36 in Q4 2021 to 315.76 in Q4 2022.

|  |  |  |
| --- | --- | --- |
| skewness | 1.16773938 | 15.06167594 |
| kurtosis | 0.360033015 | 315.7647882 |

the increasing skewness and kurtosis for both groups from Q4 2021 to Q4 2022 suggest a growing number of students watching significantly more content than the majority. This is especially true for free-plan students with a higher skewness and kurtosis in Q4 2022 than paid-plan students.

**Task 3**

Determine your sample size

Calculate the sample’s mean value.

Determine the sample’s standard deviation

Estimate the significance level

Obtain the z-score from the standard normal distribution table for probability (p) = 1 - α/2.

Calculate the standard error.

Calculate the margin of error.

Calculate the confidence interval by adding and subtracting the margin of error from the mean value.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | **Paid-Plan Students** | | **Free-plan Students** | |
|  |  | *min\_wat\_21* | *min\_wat\_22* | *min\_wat\_21* | *min\_wat\_22* |
| 1 | sample size | 3433 | 5104 | 32171 | 120658 |
| 2 | mean | 332.502508 | 368.3547139 | 133.9333129 | 69.14765544 |
| 3 | standard deviation | 485.8634701 | 596.4051599 | 367.2624454 | 255.6234424 |
| 4 | significance level | 0.05 | 0.05 | 0.05 | 0.05 |
| 5 | z-score | 1.96 | 1.96 | 1.96 | 1.96 |
| 6 | standard error | 8.292344684 | 8.348069563 | 2.047595853 | 0.735906464 |
| 7 | margin of error | 16.25299558 | 16.36221634 | 4.013287872 | 1.44237667 |
| 8 | Confidence Intervals  high | 348.7555036 | 384.7169303 | 137.9466008 | 70.59003211 |
| 9 | Confidence Intervals  low | 316.2495124 | 351.9924976 | 129.920025 | 67.70527877 |
| 10 | Confidence level | 95% |  | Confidence level | 95% |

Students with a paid-plan subscription watch substantially more than those without. The confidence interval for the average minutes watched in Q4 2022 was 61.71 to 70.59 minutes for free-plan students and 351.99 to 384.72 minutes for paid-plan students. We then can be 95% confident that paid-plan students watched significantly more minutes than free-plan students in Q4 2022. This aligns with the expectation that paid-plan students who have invested in the platform tend to be more engaged than free-plan users.

**Task 4**

You want to reach a data-driven customer engagement decision on whether the platform’s new features contribute to the increase of minutes watched on the platform for both free-plan and paying students i.e., the rise in student engagement in their study process.

* For free-plan students, perform a two-sample t-test assuming unequal variances.
* For paying students, conduct a two-sample t-test assuming unequal variances.

|  |  |  |
| --- | --- | --- |
| F-Test Two-Sample for Variances For Paid Students | | |
|  |  |  |
|  | *Variable 1* | *Variable 2* |
| Mean | 332.502508 | 368.3547139 |
| Variance | 236063.3116 | 355699.1148 |
| Observations | 3433 | 5104 |
| df | 3432 | 5103 |
| F | 0.663660104 |  |
| P(F<=f) one-tail | 0 |  |
| F Critical one-tail | 0.949796198 |  |

|  |  |  |
| --- | --- | --- |
| F-Test Two-Sample for Variances Free Students | |  |
|  |  |  |
|  | *Variable 1* | *Variable 2* |
| Mean | 133.9333129 | 69.14765544 |
| Variance | 134881.7038 | 65343.34428 |
| Observations | 32171 | 120658 |
| df | 32170 | 120657 |
| F | 2.06419958 |  |
| P(F<=f) one-tail | 0 |  |
| F Critical one-tail | 1.014667161 |  |

The p-value indicates the probability of obtaining the observed f-value if the null hypothesis (equal variances) were true. The sample variances are not identical since the p-value in both cases is 0.

**Paid-Plan Students**

|  |  |  |
| --- | --- | --- |
| Paid Students | minutes\_watched\_21 | minutes\_watched\_22 |
| Mean | 332.502508 | 368.3547139 |
| Standard Deviation | 485.8634701 | 596.4051599 |
| Sample Size | 3433 | 5104 |
|  |  |  |
| T-Score | -3.05 |  |
| Degree of Freedom (df) | 8229 |  |
| Critical Value | 1.65 |  |

If −3.35≤−1.65, then reject (Null Hypothesis)

**Reject** because the calculated **t-statistic** is lower than the critical value.

|  |  |  |
| --- | --- | --- |
| t-Test: Two-Sample Assuming Unequal Variances | | |
|  |  |  |
|  | minutes\_watched\_21 | minutes\_watched\_22 |
| Mean | 332.502508 | 368.3547139 |
| Variance | 236063.3116 | 355699.1148 |
| Observations | 3433 | 5104 |
| Hypothesized Mean Difference | 0 |  |
| df | 8229 |  |
| t Stat | -3.046942872 |  |
| P(T<=t) one-tail | 0.001159572 |  |
| t Critical one-tail | 1.645038819 |  |
| P(T<=t) two-tail | 0.002319144 |  |
| t Critical two-tail | 1.960252308 |  |

If p−value ≤ 0.05, Reject (Null Hypothesis)

**Reject**because the p-value is lower than the specified significance level α (0.05).

With a t-statistic of -3.05 (less than the critical value of -1.645), you would reject the null hypothesis because the negative t-statistic indicates that (the mean minutes watched by students in Q4 2021) is significantly smaller than (the mean minutes watched by students in Q4 2022). This is contrary to the null, so we reject it. Of course, rejecting the null hypothesis does not confirm the alternative hypothesis; it suggests that the data provide enough evidence against the null hypothesis.

**Free-Plan Students**

|  |  |  |
| --- | --- | --- |
| Free Students | minutes\_watched\_21 | minutes\_watched\_22 |
| Mean | 133.9333129 | 69.14765544 |
| Standard Deviation | 367.2624454 | 255.6234424 |
| Sample Size | 32171 | 120658 |
|  |  |  |
| T-Score | 29.78 |  |
| Degree of Freedom (df) | 40836 |  |
| Critical Value | 1.64 |  |

If 29.78≤−1.65, then reject (Null Hypothesis)

**Fail to Reject** because the calculated **t-statistic** is higher than the critical value.

|  |  |  |
| --- | --- | --- |
| t-Test: Two-Sample Assuming Unequal Variances | |  |
|  |  |  |
|  | Variable 1 | Variable 2 |
| Mean | 133.9333129 | 69.14765544 |
| Variance | 134881.7038 | 65343.34428 |
| Observations | 32171 | 120658 |
| Hypothesized Mean Difference | 0 |  |
| Df | 40836 |  |
| t Stat | 29.77523819 |  |
| P(T<=t) one-tail | 4.7441E-193 |  |
| t Critical one-tail | 1.644890942 |  |
| P(T<=t) two-tail | 9.4881E-193 |  |
| t Critical two-tail | 1.960022079 |  |

For free-plan students: With a t-statistic of 29.78 (greater than the critical value of -1.645), you would fail to reject the null hypothesis. This means there’s not enough evidence to conclude that μ1μ1 is smaller than μ2μ2. So, the data supports the null hypothesis that μ1μ1 is larger than or equal to μ2μ2.

These results align with previous findings from the confidence intervals and further underscore the difference in engagement patterns between paid- and free-plan students.

Regarding the second part of the question, a Type I error (false positive) occurs when you reject the null hypothesis, which is true. In our case, this would mean concluding that engagement in 2022 is higher when it's not. The probability of making this error is the level of significance, α. Since you (the researcher) choose the significance level of the hypothesis test, the responsibility for making this error lies solely on you.

Note that the significance level is closely related to the confidence level, representing our degree of certainty in the estimated results. It’s equal to (1 − α). For example, a significance level of 5% for a hypothesis test means a 5% probability of rejecting a true null hypothesis, corresponding to a 95% confidence level.

A Type II error (false negative) occurs when you fail to reject the null hypothesis, but it’s false. In our case, this would mean that the engagement in 2022 is not higher than it is.

The cost to the company of each type of error would depend on the implications of incorrectly concluding that engagement has increased—potentially leading to over-investment in certain features or complacency about needing to improve features—versus incorrectly concluding that engagement has not increased—potentially missing out on recognizing successful features or identifying areas that need improvement.

**Task 5**

Your last task is determining whether the average number of minutes watched in the US is similar to that in India.

|  |  |  |
| --- | --- | --- |
| F-Test Two-Sample for Variances | |  |
|  |  |  |
|  | *minutes\_watched\_22\_US* | *minutes\_watched\_22\_IN* |
| Mean | 73.07053569 | 78.42208628 |
| Variance | 95208.64187 | 101975.5527 |
| Observations | 6459 | 21210 |
| Df | 6458 | 21209 |
| F | 0.933641833 |  |
| P(F<=f) one-tail | 0.000347535 |  |
| F Critical one-tail | 0.967314359 |  |

The p-value indicates the probability of obtaining the observed f-value if the null hypothesis (equal variances) were true. The sample variances are not identical since the p-value is lower than 0. We must perform a left-tailed t-test assuming unequal variances:

|  |  |  |
| --- | --- | --- |
|  | *minutes\_watched\_22\_US* | *minutes\_watched\_22\_IN* |
| Mean | 73.07053569 | 78.42208628 |
| Standard Error | 3.839330887 | 2.192692341 |
| Median | 9.47 | 3.57 |
| Mode | 0.05 | 0.05 |
| Standard Deviation | 308.5589763 | 319.3361124 |
| Sample Variance | 95208.64187 | 101975.5527 |
| Kurtosis | 599.5710496 | 443.6430945 |
| Skewness | 19.73855899 | 16.10708335 |
| Range | 11449.63 | 13188.37 |
| Minimum | 0.05 | 0.05 |
| Maximum | 11449.68 | 13188.42 |
| Sum | 471962.59 | 1663332.45 |
| Count | 6459 | 21210 |
| T-Score | -1.21 |  |
| Degree of Freedom (df) | 11001 |  |
| Critical Value | 1.65 |  |

If −1.21≤−1.65 , then reject H0

**Fail to** **Reject** because the calculated **t-statistic** is higher than the critical value.

|  |  |  |
| --- | --- | --- |
| t-Test: Two-Sample Assuming Unequal Variances | |  |
|  |  |  |
|  | *minutes\_watched\_22\_US* | *minutes\_watched\_22\_IN* |
| Mean | 73.07053569 | 78.42208628 |
| Variance | 95208.64187 | 101975.5527 |
| Observations | 6459 | 21210 |
| Hypothesized Mean Difference | 0 |  |
| Df | 11001 |  |
| t Stat | -1.210387573 |  |
| P(T<=t) one-tail | 0.113078106 |  |
| t Critical one-tail | 1.644992151 |  |
| P(T<=t) two-tail | 0.226156213 |  |
| t Critical two-tail | 1.960179649 |  |

Decision Rule: If p−value ≤ 0.05, Reject (Null Hypothesis)

**Fail to Reject** because the p-value is higher than the specified significance level α (0.05).

If the hypothesis that US students watch more or an equal amount of content as Indian students is rejected, this suggests that US students watch less content on average than students in India.

This could have the following implications.

**Market Differences:**These details might indicate that the platform is more engaging or relevant to students in India than e US students. Understanding the reasons behind this could be valuable. Are the platform’s specific features, content, or aspects particularly appealing to Indian students? Such questions need to be addressed further but are beyond the scope of this analysis.

**Growth Opportunities:** If US engagement is lower, this could represent a growth opportunity. The 365 company might seek ways to increase engagement among US students, involving marketing efforts, adding more content relevant to US students, or other strategies.

**Resource Allocation**: This information could be helpful when deciding where to allocate resources. For example, if Indian students are more engaged, investing in more content and features targeted toward this audience might make sense.