7COM1079 – Team Research and Development Project

Final Report Title: *Investigating Differences in Online Activity Across Student Performance Levels in the xAPI-Edu-Data Dataset*

Dataset Number: *183*

Prepared By:

- [Ajay vadde -24094194]
- [Siddu battu -23092967]
- [Uday Kiran jalapati - 24098972]
- [Sadhik Mohammad - 24065240]
- [Akshay kumar Reddy gadampally - 24096317]

University of Hertfordshire

Hatfield, 2025

Table of Contents

1. Introduction

1.1 Problem Statement and Research Motivation

Learning analytics data is becoming more popular in educational institutions to gain an insight into student use of digital learning environments. Raised hands, resource visits, announcement views and discussion participation are some of the measurable behavioural indicators that could be associated with academic results of students. Determining whether there is a difference in engagement between higher and lower levels of performance is necessary in the design of early-warning systems and interventions (Tora et al., 2025). Previous studies have always documented that pattern of online activity correlates with student success, but few studies have made elaborate statistical comparisons with the xAPI-Edu-Data benchmark dataset. This paper hence tests the question of whether there is any significant difference in distribution of online activity within performance classes.

1.2 The Dataset

The data set, which is the xAPI-Edu-Data, has 480 records of students and 17 variables, such as demographic, engagement behaviour, and performance classes (Low, Medium, High). The variables of numeric activities that were analysed were raised hands, VisITedResources, Announcements View, and Discussion. R was used to clean the data by removing text fields and changing the right variables into numeric or factor. The analysis of descriptive statistics indicates that there are distinct differences in the mean level of engagement with the different performance classes and this warrants further hypothesis testing.

1.3 Research Question

RQ: *Do students' online activity measures (raised hands, visited resources, announcements viewed, and discussion participation) differ significantly across the three performance classes (Low, Medium, High) in the xAPI-Edu-Data dataset?*

The question aims at the identification of the fact that the engagement behaviours have a systematic correlation with level of student performance.

1.4 Hypotheses

To test the research question, the hypotheses that were tested under each variable of activities were: Null Hypothesis (H 0) The distribution of the online activity variable is not significantly different between the classes of Low, Medium, and High performance (Althaqafi, Saleem and

AL-Ghamdi, 2025). H1: There exists a significant difference in distribution of online activity variable across the three performance classes. Since initial diagnostics revealed the breach of normality and homogeneity, the hypotheses were tested with the Kruskal Wallis non-parametric test, and the Dunn post-hoc tests with Bonferroni correction.

2. Background Research

2.1 Relevant literature Review

The interest of students in the academic outcomes is a long-established predictor of academic performance, and various studies have shown how behavioural interaction data may be used to facilitate early-warning and intervention systems. Some of the most common measures used in research based on online learning environments include access to resources, engagement in discussions, and engagement with an instructor as key predictors of student performance. Research on the analysis of datasets similar to xAPI-Edu-Data indicates that high performing students tend to have more regular patterns of activities and a higher level of interaction intensity compared to low performing students. The variables of engagement like VisITedResources, raisedhands, etc., tend to be among the most effective predictors in machine learning models that assess the result of students (Alkan et al., 2025).

A number of studies have employed a behavioural log within learning management system to establish the relationship between level of activities and overall final achievement categories. These reports indicate statistically significant differences in activity by performance level, which supports the idea that student activity does not happen randomly, but it is significantly correlated with academic level. It is also observed by the literature in the past that non-normal distributions exist in behavioural data, hence the application of non-parametric techniques to compare data. Results in the area encourage hypothesis-based statistical analysis of behavioural indicators especially in structured educational data as is available through xAPI-Edu-Data.

2.2 Why the Research Question Is interesting

Though previous literature is always consistent in associating engagement behaviour with academic performance, most of the studies are concerned with predictive modelling as compared to statistical comparison of performance groups. As a result, a limited analysis is conducted to study the existence of specific engagement variables that vary significantly in Low, Medium, and High-performance classes in the xAPI-Edu-Data benchmark dataset. This gap would give a better understanding of the behavioural patterns that define the levels of performance and would be useful in making informed decisions based on data to the educators. The results of the current research can be used to explain the importance of engagement behaviours that demonstrate the highest levels of variation across performance groups, and, therefore, their use in strategies and monitoring students and creating specific academic support.

3. Visualisation

3.1 Main Plot and Explanation

Boxplots were created to compare the performance variable of each engagement variable (raisedhands, VisITedResources, AnnouncementsView, Discussion) to the three performance classes. Boxplots are suitable since they indicate variation in medians, forms of distribution and outliers between groups. These generated R-visualisations are good initial evidence as to whether there is consistent variation in patterns of activities by class.

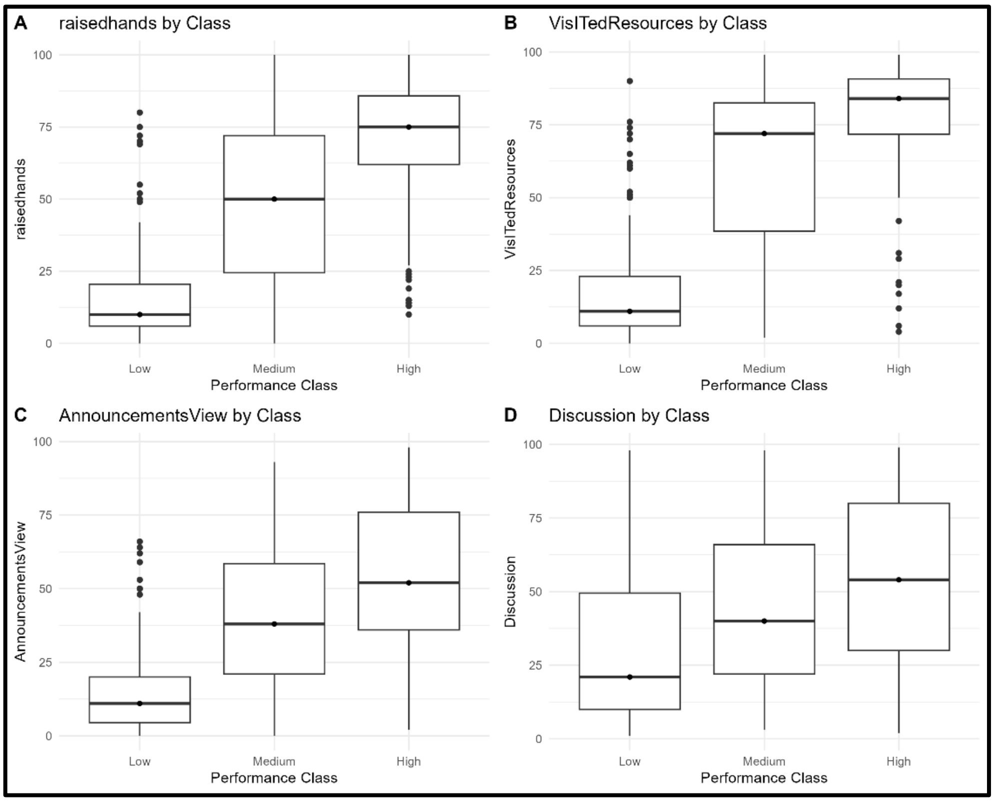

Figure 1: Boxplots of Activity Variables by Performance Class

3.2 Additional Information Relating to Understanding the Data

Where further details are necessary in order to comprehend the data, they are presented here (roughly 50 words). Additional visualisation was made to facilitate interpretation of distribution shapes and categorical patterns (Geeksforgeeks, 2024). The non-parametric tests are motivated by a non-parametric histogram of raisedhands. Gender grouped bar chart by class gives an insight on demographic composition that may affect engagement.

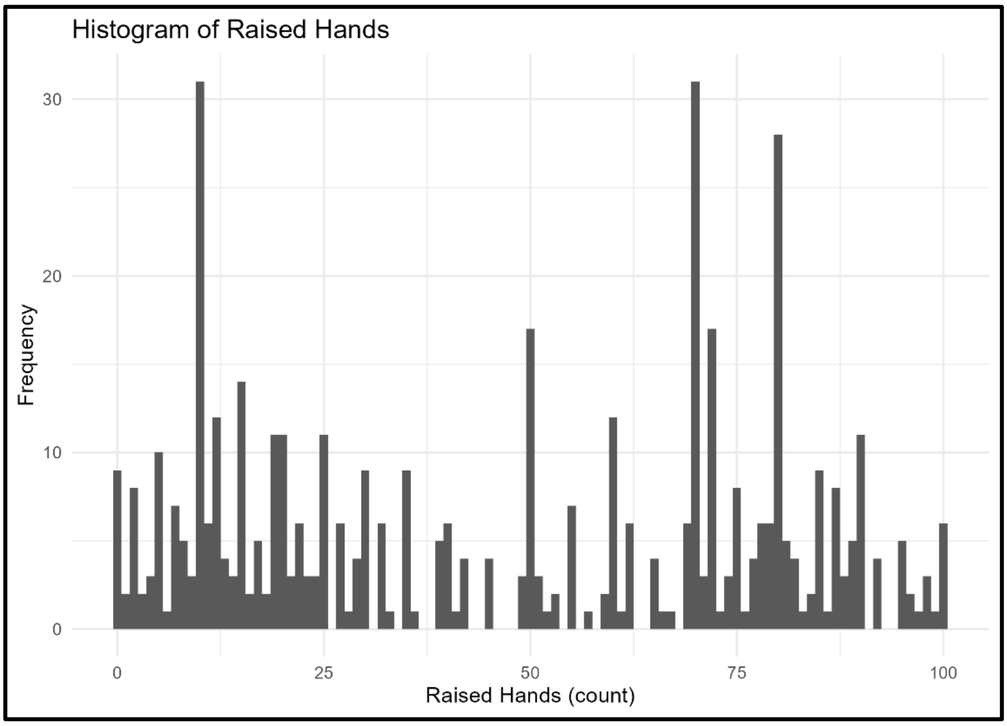

Figure 2: Histogram of Raised Hands

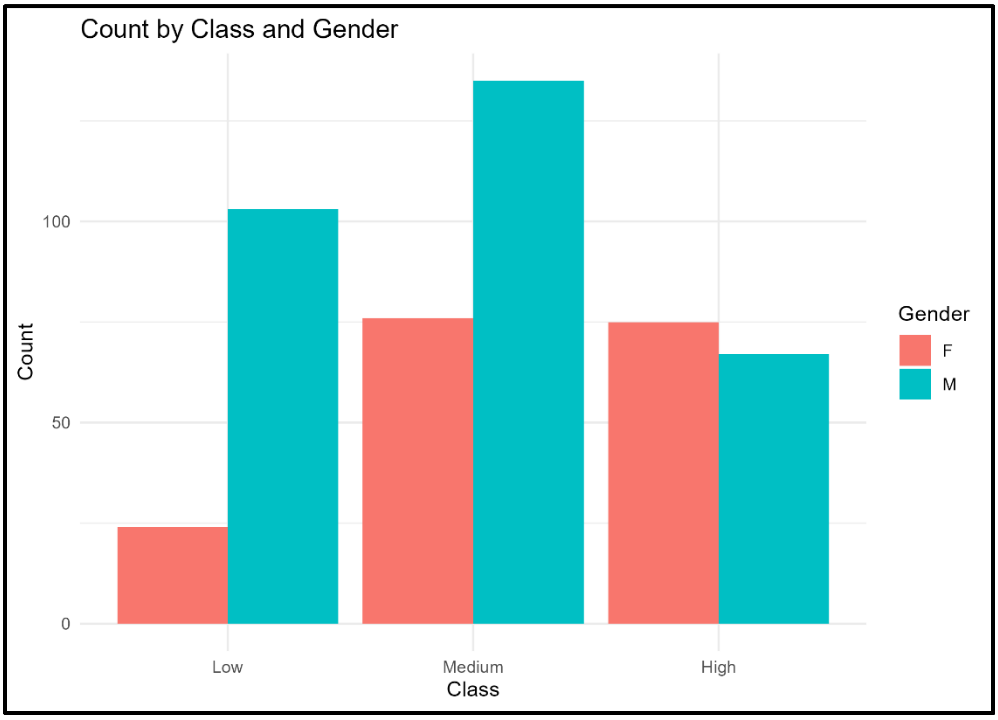

Figure 3: Bar Plot of Gender by Class

3.3 Helpful Information to Data Understanding

In all four boxplots, the engagement rises with the Low classes to the Medium and the Higher classes, implying high behavioural segregation between classes. The histogram contains a multi-modal distribution, with a high level of variability and the gender-class bar chart indicates the disproportionate gender distribution within the categories provided as the supplementary information to the Chi-square test.

4. Analysis

4.1 Statistical Test to be used to test the Hypotheses

Initial diagnostics has demonstrated that all the four variables of engagement were not normative (Shapiro. Wilk p-values less than 0.001 in all classes) and three variables were not homogeneous of variances (Levene test p less than 0.001). Thus, Kruskal-Wallis test was chosen as a suitable non-parametric alternative of one-way ANOVA. In every variable of activity, Kruskal-Wallis was used to compare the distribution of the variable amongst Low, Medium and High-performance classes (Geeksforgeeks, 2020). Dunn has used post-hoc tests with Bonferonni correction when there is significant difference to determine the specific pair-wise differences.

4.2 Decision on hypothesis based on p-value

The differences in all four variables of activities were highly significant in the performance classes.

- **raisedhands:** KW 2 = 207.82, p < 2.2e-16; all pairwise comparisons significant (e.g., Low High p. Adjust = 5.13e-46).
- **VisITedResources:** KW 2 = 222.10, p = less than 2.2e-16; all pairwise p. Adjust = less than 1e-7.
- **AnnouncementsView:** KW 2 =154.85, p<2.2e-16; all pairwise comparisons significant.
- **Discussion:** KW 2 = 49.648, p = 1.656e-11; all pairwise differentials significant on adjustment.

The p-values of all engagement variables are rejected because all the tests yield p-values much less than 0.05. There are high differences in the engagement behaviour with Low, Medium, and High-performance classes.

5. Evaluation – Group's Experience at 7COM1079

5.1 What Went Well

The project was a pleasant experience, and the team was able to share duties in the areas of data cleaning, visualisation, statistical testing, and documentation. R enhanced the accuracy of the analytical results of the report and made the outputs reproducible. Teamwork was also enhanced by checking in frequently and checking in on files. The workflow also enabled team members to confirm each other on their work, particularly with the statistical tests, which made the final results more credible and the entire quality of the report more effective.

5.2 Points for Improvement

Early agreement on the research question and analytical focus is one of the areas of improvement because initial exploration generated scope expansion before narrowing on the final hypothesis (Casula, Rangarajan and Shields, 2021). The staff also might do better in terms of coding style and documentation to help new members of the team to follow the script. More time would have been spent on polishing visuals and verifying assumption diagnostics. Projects in future must have set milestones to prevent the last-minute workload congestion.

5.3 Group's Time Management

Once the group had good time management because of the establishment of clear task ownership. Even though the initial work was slow, the next level of coordination significantly enhanced the efficiency of work. Meetings were made more formal with quicker decision making. In general, the group achieved all the in-house deadlines and finished statistical analysis and report writing within the period of time.

5.4 Project's Overall Judgement

The module requirements were achieved with the project as a result of generating clean and reproducible R analysis and interpretable results. The results evidently indicate differences in behaviour between performance classes as well as good group cooperation. The major report is coherent, analytically sound, and meaningful insights have been given which are in line with the research question.

5.5 Comment on GitHub Log Output

The GitHub log output is showing a clear and structured development process with meaningful commits that indicate the progression of the project. The commit history reveals that the repository was updated in logical steps, thus benefiting from the aspects of transparency and reproducibility. The following are three major commits that greatly depict this development:

- **"Data Cleaning and Factor Conversion" –** The intention of this commit was to create a reliable analytical dataset by taking care of missing values, changing the category variable to a factor, and maintaining structural uniformity. It laid the groundwork for all the following analysis to be conducted.
- **"Visualisation Updates" –** The addition of the crucial R-made visual outputs such as boxplots, histograms, and bar charts were the main point of this commit. The visualisations were very helpful in detecting the patterns of student engagement according to the performance classes and were the basis of the choice of appropriate statistical tests.
- **"Statistical Testing Implementation"** – The implementation of the Kruskal–Wallis tests, Dunn's post-hoc comparisons, and the chi-square test for gender differences was the main point of this commit. It marks the project's analytical core, connecting the hypotheses to formal statistical proof.

All in all, the GitHub actions mirror the principles of a collaborative workflow, such as the use of version control, documentation, and gradual improvement. The updates are conducted in a logical order and provide a traceable route from raw data to final results, thus reinforcing open, reproducible, and professional research practice.

6. Conclusions

6.1 Results Explained

The statistical test revealed significant and obvious difference in engagement behaviours among the three classes of performance. The Kruskal-Walli's test outputs of all four activity variables raisedhands, VisITedResources, AnnouncementsView, and Discussion showed a highly significant result ($p < 0.001$), and Dunn post-hoc results indicated that each of the classes was significantly different. Also, descriptive statistics revealed a steady upward trend between Low and Medium and High-performance groups on all engagement measures that also supported the strength of the observed relationships.

6.2 Interpretation of the Results

The results show that students with better achievements have been recorded to exhibit high degrees of online activities compared to their underperformers. This indicates that academic performance is closely correlated with the exposure to learning materials and involvement in activities in the classroom. Such behavioural patterns can include better study habits, greater motivation or greater engagement in the learning process. The findings indicate that the tracking of engagement indicators can be a valuable component of institutional early-warning systems that may help to predict students that are vulnerable to academic failure.

6.3 Limitations and Future Work

This research only uses observation as its data and thus cannot be able to draw a causal conclusion. The comparative analysis did not concentrate on additional demographic or contextual variables, which can affect results. The multivariate modelling, effect-size analysis, cross-dataset validation, and fairness assessments should be introduced into future works to provide fair interpretation of engagement-based interventions.

7. Reference

Alkan, B.B., Serafettin Kuzucuk, Alkan, N. and Sinan, A. (2025). Using machine learning to predict student outcomes for early intervention and formative assessment. *Scientific Reports*, [online] 15(1). doi:https://doi.org/10.1038/s41598-025-23409-w.

Althaqafi, T., Saleem, F. and AL-Ghamdi, A.A.-M. (2025). Enhancing student performance prediction: the role of class imbalance handling in machine learning models. *Discover Computing*, 28(1). doi:https://doi.org/10.1007/s10791-025-09576-4.

Casula, M., Rangarajan, N. and Shields, P. (2021). The Potential of Working Hypotheses for Deductive Exploratory Research. *Quality & Quantity*, 55(1), pp.1703–1725. doi:https://doi.org/10.1007/s11135-020-01072-9.

GeeksforGeeks (2020). *KruskalWallis test in R Programming*. [online] GeeksforGeeks. Available at: https://www.geeksforgeeks.org/r-language/kruskal-wallis-test-in-r-programming/.

GeeksforGeeks (2024). *Charts and Graphs for Data Visualization*. [online] GeeksforGeeks. Available at: https://www.geeksforgeeks.org/data-visualization/charts-and-graphs-for-data-visualization/.

Tora, T.T., Andarge, L.T., Abebe, A.T., Utallo, A.U., Meshesha, D.T. and Wubie, A.F. (2025). An expert-based assessment of early warning systems effectiveness in South Ethiopia Regional State. *Discover Sustainability*, 6(1). doi:https://doi.org/10.1007/s43621-025-00943-1.

8. Appendix A — R Code Used for Analysis and Visualisation

The full R script used in this project is included below.

A.1 Full R Script

```
# ---- 0. Setup ----

packages <- c("tidyverse", "ggplot2", "car", "nnet", "broom",
"rstatix", "ggpubr")

installed <- rownames(installed.packages())

for(p in packages){

  if(!(p %in% installed)) install.packages(p, repos =
"https://cloud.r-project.org")

}


library(tidyverse)

library(ggplot2)

library(car)

library(nnet)

library(broom)

library(rstatix)

library(ggpubr)


# ---- 1. Load data ----

data_path <- "D:/HEN Harrier/Material/MZ52064/xAPI-Edu-
Data.csv"

df <- read.csv(data_path, stringsAsFactors = FALSE)



# Quick structure check
```

```
str(df)
summary(df)

# ---- 2. Basic cleaning & factor conversions ----
# Trim whitespace from all character columns
df <- df %>%
  mutate(across(where(is.character), ~ trimws(.)))

# Convert variables used in analysis to appropriate types
df <- df %>%
  mutate(
    gender = factor(gender),
    StageID = factor(StageID),
    GradeID = factor(GradeID),
    SectionID = factor(SectionID),
    Topic = factor(Topic),
    Semester = factor(Semester),
    Relation = factor(Relation)
  )

# Fix Class: map "L" -> "Low", "M" -> "Medium", "H" -> "High"
safely
df$Class <- toupper(df$Class)
df$Class <- ifelse(df$Class == "L", "Low",
                   ifelse(df$Class == "M", "Medium",
                          ifelse(df$Class == "H", "High", NA)))
df$Class <- factor(df$Class, levels = c("Low","Medium","High"))
```

```r
# Numeric variables for activity
activity_vars <- c("raisedhands","VisITedResources","AnnouncementsView","Discussion")
df[activity_vars] <- lapply(df[activity_vars], function(x) as.numeric(as.character(x)))


# ---- 3. Descriptive statistics ----
cat("Total rows:", nrow(df), "\n")
cat("Class distribution:\n")
print(table(df$Class))


desc_by_class <- df %>%
  group_by(Class) %>%
  summarize(
    n = n(),
    mean_raisedhands = mean(raisedhands, na.rm = TRUE),
    sd_raisedhands = sd(raisedhands, na.rm = TRUE),
    mean_resources = mean(VisITedResources, na.rm = TRUE),
    sd_resources = sd(VisITedResources, na.rm = TRUE),
    mean_announcements = mean(AnnouncementsView, na.rm = TRUE),
    sd_announcements = sd(AnnouncementsView, na.rm = TRUE),
    mean_discussion = mean(Discussion, na.rm = TRUE),
    sd_discussion = sd(Discussion, na.rm = TRUE)
  )
```

```
cat("\n=== Descriptive statistics by Class ===\n")
print(desc_by_class)
write.csv(desc_by_class, "descriptive_by_class.csv", row.names
= FALSE)
cat("Descriptive statistics saved as
'descriptive_by_class.csv'\n")

# ---- 4. Visualisations ----
p_list <- lapply(activity_vars, function(var){
  ggplot(df, aes_string(x = "Class", y = var)) +
    geom_boxplot() +
    stat_summary(fun = median, geom = "point", shape = 20, size
= 2, color = "black") +
    labs(title = paste(var, "by Class"),
         x = "Performance Class",
         y = var) +
    theme_minimal()
})

main_plot <- ggarrange(plotlist = p_list, ncol = 2, nrow = 2,
labels = "AUTO")
print(main_plot)   # Display in RStudio Plots pane
ggsave("boxplots_activity_by_class.png", main_plot, width = 10,
height = 8)

hist_plot <- ggplot(df, aes(x = raisedhands)) +
  geom_histogram(binwidth = 1, closed = "left") +
  labs(title = "Histogram of Raised Hands",
```

```
       x = "Raised Hands (count)",
       y = "Frequency") +
  theme_minimal()
print(hist_plot)   # Display in Plots pane
ggsave("hist_raisedhands.png", hist_plot, width = 7, height =
5)


tab_gender_class <- table(df$gender, df$Class)
cat("\n=== Gender vs Class table ===\n")
print(as.data.frame(tab_gender_class))
write.csv(as.data.frame(tab_gender_class),
"gender_class_table.csv", row.names = FALSE)
cat("Gender-Class table saved as 'gender_class_table.csv'\n")


bar_gender_class <- as.data.frame(tab_gender_class) %>%
  ggplot(aes(x = Var2, y = Freq, fill = Var1)) +
  geom_bar(stat = "identity", position = "dodge") +
  labs(title = "Count by Class and Gender", x = "Class", y =
"Count", fill = "Gender") +
  theme_minimal()
print(bar_gender_class)   # Display in Plots pane
ggsave("bar_gender_class.png", bar_gender_class, width = 7,
height = 5)


# ---- 5. Statistical testing ----
test_results <- list()
```

```r
for(var in activity_vars){
  cat("\n\n=== Analysis for:", var, "===\n")

  if(length(levels(df$Class)[!is.na(levels(df$Class))]) < 2){
    cat("Skipping", var, "- Class factor has less than 2 levels.\n")
    next
  }

  shapiro_by_class <- df %>%
    group_by(Class) %>%
    summarize(
      n = sum(!is.na(.data[[var]])),
      shapiro_p = ifelse(n >= 3, shapiro.test(.data[[var]])$p.value, NA_real_)
    )
  print(shapiro_by_class)

  lev <- leveneTest(as.formula(paste(var, "~ Class")), data = df)
  print(lev)

  shapiro_ok <- all(na.omit(shapiro_by_class$shapiro_p) > 0.05)
  levene_p <- lev$`Pr(>F)`[1]

  if(!is.na(shapiro_ok) && shapiro_ok && levene_p > 0.05){
    cat("Using one-way ANOVA for", var, "\n")
```

```
    aov_mod <- aov(as.formula(paste(var, "~ Class")), data = df)
    print(summary(aov_mod))
    tuk <- TukeyHSD(aov_mod)
    print(tuk)
    test_results[[var]] <- list(method = "ANOVA", model = aov_mod, posthoc = tuk)
  } else {
    cat("Using Kruskal-Wallis for", var, "\n")
    kw <- kruskal.test(as.formula(paste(var, "~ Class")), data = df)
    print(kw)
    dunn <- dunn_test(df, as.formula(paste(var, "~ Class")), p.adjust.method = "bonferroni")
    print(dunn)
    test_results[[var]] <- list(method = "Kruskal-Wallis", test = kw, posthoc = dunn)
  }
}

# ---- 6. Chi-square test ----
cat("\n\n=== Chi-square test: Gender vs Class ===\n")
if(length(levels(df$Class)[!is.na(levels(df$Class))]) < 2){
  cat("Cannot perform Chi-square: Class factor has less than 2 levels.\n")
} else {
  chi <- chisq.test(tab_gender_class)
  print(chi)
  if(any(chi$expected < 5)){
```

```
    cat("Some expected counts < 5; running Fisher's Exact Test:\n")
    fisher_res <- fisher.test(tab_gender_class, simulate.p.value = TRUE)
    print(fisher_res)
  }
}

# ---- 7. Multinomial logistic regression ----
mod_df <- df %>%
  select(Class, all_of(activity_vars), gender) %>%
  drop_na()

if(length(levels(mod_df$Class)[!is.na(levels(mod_df$Class))]) >= 2){
  multinom_mod <- multinom(Class ~ raisedhands + VisITedResources + AnnouncementsView + Discussion + gender,
                           data = mod_df, trace = FALSE)

  cat("\n\n=== Multinomial logistic regression tidy results ===\n")
  tidy_multinom <- broom::tidy(multinom_mod, exponentiate = FALSE, conf.int = FALSE)
  print(tidy_multinom)
  write.csv(tidy_multinom, "multinom_model_results.csv", row.names = FALSE)
  cat("Multinomial regression results saved as 'multinom_model_results.csv'\n")
```

```
  # Odds ratios
  or <- exp(coef(multinom_mod))
  cat("\nOdds ratios:\n")
  print(or)

} else {
  cat("Cannot fit multinomial regression: Class factor has less than 2 levels.\n")
}

# ---- 8. Summary / Quick Interpretations ----
cat("\n\n=== Summary for report ===\n")
for(var in activity_vars){
  res <- test_results[[var]]
  if(!is.null(res)){
    if(res$method == "ANOVA"){
      p_anova <- summary(res$model)[[1]][["Pr(>F)"]][1]
      cat(sprintf("%s: ANOVA p = %.4g\n", var, p_anova))
    } else {
      p_kw <- res$test$p.value
      cat(sprintf("%s: Kruskal-Wallis p = %.4g\n", var, p_kw))
    }
  }
}
if(exists("chi")) cat("\nChi-square Gender vs Class p =", chi$p.value, "\n")
```

```
cat("\nAll plots displayed in Plots pane. Descriptive tables and
model outputs saved as CSV.\n")
```

A.2 Generated Files from R

The following files were produced by the script and must be attached as separate files or inserted into the Appendix:

- **descriptive_by_class.csv**

Class	n	mean_rais	sd_raisedh	mean_resc	sd_resourc	mean_ann	sd_annour	mean_disc	sd_discussion
Low	127	16.88976	17.20937	18.32283	19.187	15.5748	15.31444	30.83465	25.71016
Medium	211	48.93839	26.89364	60.63507	28.22975	40.96209	24.08692	43.79147	26.12031
High	142	70.28873	22.54344	78.74648	19.35647	53.38028	25.05504	53.66197	27.19587

- **gender_class_table.csv**

Var1	Var2	Freq
F	Low	24
M	Low	103
F	Medium	76
M	Medium	135
F	High	75
M	High	67

- **multinom_model_results.csv**

y.level	term	estimate	std. error	statistic	p. Value
Medium	(Intercept)	-2.045899	0.4837503	-4.2292462	2.34E-05
Medium	raisedhands	0.028762	0.008626	3.33432971	0.000855052
Medium	VisITedResources	0.033099	0.0067268	4.92042282	8.64E-07
Medium	AnnouncementsView	0.029932	0.0098828	3.02868534	0.002456204
Medium	Discussion	0.006815	0.0065231	1.04474398	0.296141314
Medium	gender	-0.718883	0.3637104	-1.9765257	0.048095271
High	(Intercept)	-5.506629	0.7113733	-7.7408431	9.88E-15
High	raisedhands	0.051962	0.0099642	5.21490199	1.84E-07
High	VisITedResources	0.055626	0.0088614	6.27733849	3.44E-10
High	AnnouncementsView	0.030122	0.0112008	2.68928304	0.007160568
High	Discussion	0.015561	0.0077746	2.00146866	0.045341908
High	genderM	-1.426689	0.4221847	-3.3793016	0.000726702

B. Appendix B — Git Log Output

```
Changes to be committed:
  (use "git rm --cached <file>..." to unstage)
        new file:   .gitignore
        new file:   A53_Code.R
        new file:   A53_Report.docx
        new file:   bar_gender_class.png
        new file:   boxplots_activity_by_class.png
        new file:   descriptive_by_class.csv
        new file:   gender_class_table.csv
        new file:   hist_raisedhands.png
        new file:   multinom_model_results.csv


C:\Users\Dell\Downloads\A53>git commit -m "Initial commit: Add code, report, and analysis outputs"
Author identity unknown

*** Please tell me who you are.

Run

  git config --global user.email "you@example.com"
  git config --global user.name "Your Name"

to set your account's default identity.
Omit --global to set the identity only in this repository.

fatal: unable to auto-detect email address (got 'Dell@ajay.(none)')

C:\Users\Dell\Downloads\A53>https://github.com/ajayvadde1507/A53-team-research-and-development-.git
'https:' is not recognized as an internal or external command,
operable program or batch file.

C:\Users\Dell\Downloads\A53>git commit -m "first commit"
Author identity unknown
```

```
  git config --global user.name "Your Name"

to set your account's default identity.
Omit --global to set the identity only in this repository.

fatal: unable to auto-detect email address (got 'Dell@ajay.(none)')

C:\Users\Dell\Downloads\A53> git config --global user.email "av24abz@herts.ac.uk"

C:\Users\Dell\Downloads\A53> git config --global user.name "ajayvadde1507"

C:\Users\Dell\Downloads\A53>git commit -m "first commit"
[master (root-commit) ec61ac6] first commit
 9 files changed, 251 insertions(+)
 create mode 100644 .gitignore
 create mode 100644 A53_Code.R
 create mode 100644 A53_Report.docx
 create mode 100644 bar_gender_class.png
 create mode 100644 boxplots_activity_by_class.png
 create mode 100644 descriptive_by_class.csv
 create mode 100644 gender_class_table.csv
 create mode 100644 hist_raisedhands.png
 create mode 100644 multinom_model_results.csv

C:\Users\Dell\Downloads\A53>git branch -M main

C:\Users\Dell\Downloads\A53>git remote add origin https://github.com/ajayvadde1507/A53-team-research-and-development-.git

C:\Users\Dell\Downloads\A53>git remote -v
origin  https://github.com/ajayvadde1507/A53-team-research-and-development-.git (fetch)
origin  https://github.com/ajayvadde1507/A53-team-research-and-development-.git (push)
```

```
C:\Users\Dell\Downloads\A53>git push -u origin main
info: please complete authentication in your browser...
Enumerating objects: 11, done.
Counting objects: 100% (11/11), done.
Delta compression using up to 8 threads
Compressing objects: 100% (10/10), done.
Writing objects: 100% (11/11), 325.25 KiB | 14.78 MiB/s, done.
Total 11 (delta 0), reused 0 (delta 0), pack-reused 0 (from 0)
To https://github.com/ajayvadde1507/A53-team-research-and-development-.git
 * [new branch]      main -> main
branch 'main' set up to track 'origin/main'.

C:\Users\Dell\Downloads\A53>
```