Investigating Noise Levels in Campus Study Locations

Humaira Nafisa Ahmed, Paige Elizabeth Cooper, Kai Shuen Neo, R M Shahbab

Abstract

Noise levels in different study locations affects student experiences. In relation to study locations on the University of Arizona campus we would like to address the following: 1) Which study location is the most comfortable in terms of noise and distraction?, 2) How does the expected noise of each location differ during the morning and afternoon?, and 3) Does the expected noise for each location change based on day of the week (beginning, middle, end)? To address these key questions, we will perform a Full Factorial experimental design, fit an ANOVA model, create interaction plots and perform formal tests with Tukey's test for additivity, check model assumptions with model diagnostics and compute Tukey-adjusted confidence intervals to make recommendations. We subsequently came to the finding that the Common Room is the quietest of all study locations, while the Student Union is the loudest. We also concluded the morning is generally quieter than the afternoon across all locations.

Introduction

Noise affects study preferences. Some students prefer a quiet, calm environment, while others appreciate background noise. The environment can significantly impact a student's ability to concentrate. This project evaluates noise levels in popular campus study locations to help students choose optimal study environments. Other important factors that could impact how noisy different study locations are the time of day and the day of the week. We have added these factors into our model to explore how noise can affect these locations during the morning and afternoon (AM/PM), and the beginning, middle, and end of the week (Mon/Wed/Fri).

The study examines the noise levels in various popular study locations on campus. The motivation behind this study is to assist students in selecting the most suitable study spots based on actual data that can enhance their study experience. Noise levels can fluctuate depending on the time of day and the day of the week, yet their effects on study locations have not been documented for different areas on campus. The experiment is conducted to answer the following research questions.

- Which study location is the most comfortable in terms of noise and distraction?
- How does the expected noise of each location differ during the morning and afternoon?
- Does the expected noise for each location change based on day of the week (beginning, middle, and end)?

Two different measurement methods were used to measure the noise. The first method involves objective ratings, using a mobile app to measure the noise level in average decibel readings. The second method is subjective rating, where each observer rates the noise level on a scale from 0 to 10. We have used Score as a response variable, combining both the subjective ratings and the objective ratings. To normalize the objective and subjective ratings on the same scale, we use the following formulas for any observation x:

$$\text{Score}_{\text{subjective}} = \frac{x}{\text{Subjective Rating}_{\text{max}}}$$

$$\text{Score}_{\text{objective}} = \frac{x}{\text{Objective Rating}_{\text{max}}}$$

We then average their normalized scores to obtain the final composite score, assigning a weight of 0.5 to both the subjective and objective ratings:

$$\text{Composite Score} = w_{\text{subjective}} \cdot \text{Score}_{\text{subjective}} + w_{\text{objective}} \cdot \text{Score}_{\text{objective}}$$

A full factorial ANOVA was conducted to evaluate the effects of different factors (e.g., location, time of day, and day of the week) on noise levels. Residual plots, normality checks, and homoscedasticity tests were performed to validate the assumptions of ANOVA. Tukey-adjusted pairwise comparisons was done to investigate specific group difference.

Experimental design

We have used a full factorial design with 3 factors, one with a = 3 levels (Location), one with b = 3 levels (Day of Week), and one with c = 2 levels (Time of Day), repeated within each of d = 4 blocks (Observer), each with n = 2 replicates. The total number of observations is a×b×c×d×n=144. The ordering of the locations visited each day within each observer was randomized to protect against bias based on the order in which each location was surveyed.

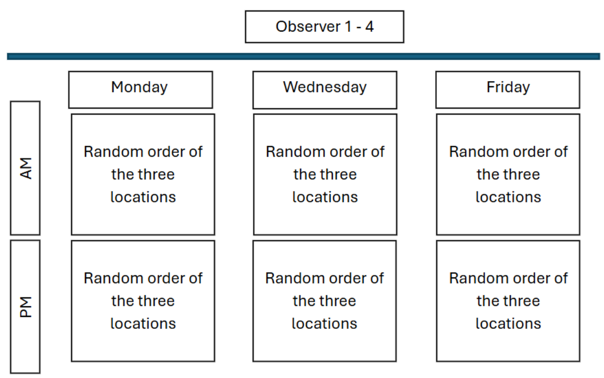

Figure 1: Proposed Experimental Design

The diagram above illustrates how the design is structured for each observer. This design was replicated twice for every observer, spanning two weeks, and repeated for each of the four members in our group, resulting in a total of eight replicates. This design is appropriate because it enables us to estimate all the main effects and explore potential interactions. The full factorial design allows us to easily formulate and test numerous hypotheses using ANOVA. This approach have assisted us in addressing the questions outlined previously and in creating study recommendations based on our results. By blocking runs for each observer, the design mitigates any bias related to personal opinion or differences in mobile device sensitivity, allowing the results to be generalized to the student population at the university.

Statistical Analyses

To assess the best location and time for study in the campus, we performed a statistical test with mixed-effect model where the composite score representing both noise and distraction level was the response variable. The model also included three main effects - Day, Time, and Location, and their interactions. Additionally, to account for the variability in collected response, we considered the observers as random effect. We also performed two replications when collecting the data. The factors and their levels can be summarized as follows.

Factors	**No. of levels**	**Specific levels**
Time of the day	2	AM; PM
Days of the week	3	Monday; Wednesday; Friday
Study locations	3	Student Union; Main Library; Common Room
Observers	4	Paige; RM; Humaira; Kai

Table 1: Factors and levels used in the noise and distraction score study.

Statistical Model

The statistical model used for the proposed design is as follows.

$$Y_{ijklm} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + \delta_l + \epsilon_{ijklm}$$

where:

- Y_{ijklm} is the observed response for the m-th replicate in the combination of i-th Location, j-th Day of Week, k-th Time of Day, and l-th Observer.
- μ is the overall mean of the response.
- α_i is the fixed effect of Location ($i = 1, 2, 3$).
- β_j is the fixed effect of Day of Week ($j = 1, 2, 3$).
- γ_k is the fixed effect of Time of Day ($k = 1, 2$).
- $(\alpha\beta)_{ij}$, $(\alpha\gamma)_{ik}$, $(\beta\gamma)_{jk}$, and $(\alpha\beta\gamma)_{ijk}$ represent the interaction effects.
- δ_l is the random effect of the Observer (Block) ($l = 1, 2, 3, 4$), assumed to be normally distributed as $\delta_l \sim N(0, \sigma_\delta^2)$.
- ϵ_{ijklm} is the random error, assumed to be normally distributed as $\epsilon_{ijklm} \sim N(0, \sigma_\epsilon^2)$.

The following constraint is imposed to ensure identifiability of parameters:

$$\alpha_1 = \beta_1 = \gamma_1 = (\alpha\beta)_{11} = (\alpha\gamma)_{11} = (\beta\gamma)_{11} = (\alpha\beta\gamma)_{111} = 0$$

where:

- $\alpha_1 = 0$: Baseline effect for Location (level 1)
- $\beta_1 = 0$: Baseline effect for Day of Week (level 1)
- $\gamma_1 = 0$: Baseline effect for Time of Day (level 1)
- $(\alpha\beta)_{11} = 0$: Interaction between baseline Location and Day
- $(\alpha\gamma)_{11} = 0$: Interaction between baseline Location and Time

- $(\beta\gamma)_{11} = 0$: Interaction between baseline Day and Time
- $(\alpha\beta\gamma)_{111} = 0$: Three-way interaction between all baseline levels

Under these constraints, the remaining parameters represent:

α_i $(i > 1)$: Difference in response between Location i and Location 1 (baseline), averaged over other factors.

β_j $(j > 1)$: Difference between Day j and Day 1 (baseline), averaged over other factors.

γ_k $(k > 1)$: Difference between Time k and Time 1 (baseline), averaged over other factors.

$(\alpha\beta)_{ij}$: Additional effect of combining Location i and Day j beyond their individual effects.

$(\alpha\gamma)_{ik}$: Additional effect of combining Location i and Time k beyond their individual effects.

$(\beta\gamma)_{jk}$: Additional effect of combining Day j and Time k beyond their individual effects.

$(\alpha\beta\gamma)_{ijk}$: Three-way interaction: How the Location×Day interaction varies by Time k.

The model is appropriate because the full factorial design considers the main effects- location, day, and time- and their interactions, such as whether the location effect varies by day or time. This approach helps identify which combinations of day, time, and location yield the highest or lowest levels of comfort or noise. The interactions can reveal if certain days or times amplify the effect of location on the response variable. Additionally, in this study, including random effects for observers is essential as it accounts for their variability. Different observers may have distinct sensitivities to noise and might interpret the survey scales differently. By modeling these variations as random effects, each observer can have their own unique baseline, which enhances the accuracy of the fixed effect estimate.

Model Assumptions

The assumptions we made about the model and how they were evaluated using appropriate diagnostic techniques can be listed as follows.

- Normality of Residuals: The residuals from the model are approximately normally distributed. This assumption was assessed using the normal plot of the residuals shown in Appendix Figure 2. The points followed a nearly straight line with only slight deviations at the tails, therefore, suggesting the validity of the normality assumption.
- Homoscedasticity (constant variance): The residuals have constant variance across all levels of the predictors. This assumption was examined using the residuals vs. fitted values plot shown in Appendix Figure 3. Additional boxplots of residuals seen in peripheral Figures A4, A5, and A6, that help identify the Student Union as a location with slightly higher variance and a potential source of heteroscedasticity. In our case, the overall analysis suggested no conspicious violation of the constant variance assumptions.
- Additivity: The effects of the predictors combine additively unless an interaction is explicitly included. This assumption was evaluated through interaction plot in Appendix Figure 1 and using Tukey's Test for Nonadditivity. We found that interaction effect was significant and unavoidable for the model.

Results

Initially we fit the full factorial model with all the factors listed in Table 1, except for the Observers because they were considered as the random effects. From the full factorial model we observed that the three-way interactions Day × Time × Location has no significant effect on the score. To remove this higher order interaction effect we simplified our model with a second-order model including only the main effects and two-way interactions. Analyzing this model revealed that one main effect 'Day' and its associated interactions with other main effects are also not statistically significant. Therefore, we decided to remove this effect as well as the associated interactions and we further simplified our model with only two main effects 'Time' and 'Location' and their interactions. Please note that random effect was considered for all the models. The analysis result for the final simplified model can be summarized as follows.

Effect	**Sum Sq**	**Mean Sq**	**Num DF**	**Den DF**	**F value**	**Pr(>F)**
Time	13.20	13.201	1	135	24.1785	2.51×10^{-6}
Location	384.19	192.096	2	135	351.8344	$< 2.2 \times 10^{-16}$
Time:Location	3.93	1.966	2	135	3.6016	0.02993

Table 2: Type III Analysis of Variance Table on the final model using Satterthwaite's method.

The results shown in Table 2 suggests that all the main effects, 'Time', 'Location' and their interaction 'Time × Location' are statistically significant due to having P-values lower than the significance level 0.05. We confirmed the presence of this interaction through the interaction plot in Appendix Figure 1 and Tukey's Test for Non-additivity. Both methods suggested that the interaction cannot be assumed to be zero and should be included in the model. Additional insights from Figure A1 include that the Common Room consistently has the lowest score, both in AM and PM indicating it as the best location to study, and the Student Union has the highest distraction levels any time of the day, therefore, is the worst place to study, specially at PM.

Contrast	**Estimate**	**p-value**
Common Room – Main Library	-1.79	$< .0001$
Common Room – Student Union	-3.99	$< .0001$
Main Library – Student Union	-2.20	$< .0001$
AM – PM	-0.606	$< .0001$
AM Common Room – PM Common Room	-0.428	0.3432
AM Common Room – AM Main Library	-1.846	$< .0001$
AM Common Room – PM Main Library	-2.165	$< .0001$
AM Common Room – AM Student Union	-3.674	$< .0001$
AM Common Room – PM Student Union	-4.742	$< .0001$
PM Common Room – AM Main Library	-1.417	$< .0001$
PM Common Room – PM Main Library	-1.737	$< .0001$
PM Common Room – AM Student Union	-3.245	$< .0001$
PM Common Room – PM Student Union	-4.314	$< .0001$
AM Main Library – PM Main Library	-0.320	0.6660
AM Main Library – AM Student Union	-1.828	$< .0001$
AM Main Library – PM Student Union	-2.897	$< .0001$
PM Main Library – AM Student Union	-1.508	$< .0001$
PM Main Library – PM Student Union	-2.577	$< .0001$
AM Student Union – PM Student Union	-1.069	$< .0001$

Table 3: Tukey-adjusted pairwise comparisons across Location, Time, and Time × Location combinations.

To further investigate specific group differences and support our interpretation of the interaction effects made earlier, we conducted Tukey-adjusted pairwise comparisons. These comparisons help identify which combinations of study time and location differ significantly in terms of their estimated noise and distraction scores. The results are summarized in Table 3.

From Table 3, it can be observed that Common Room is significantly quieter (lower distraction) than the others while Student Union being the loudest. AM time of the day is usually quieter and therefore, a better time to go for studying in the campus. The AM $\times$ Common Room is the quietest interaction combination. However, the difference in AM $\times$ Common Room and PM $\times$ Common Room is not significant. Therefore, based on our analysis, **the Common Room is recommended as the most suitable location for studying on campus, regardless of the day of the week or time of day**.

Conclusion

Through this statistical experiment, we have identified the campus locations and times for effective studying by comparing noise levels across the Common Room, the Student Union and the Main Library at the University of Arizona, from which we determined the optimal study conditions based on time and location. Key findings are that Common Room in the AM is the quietest, which the Student Union is the loudest. AMs are also generally quieter than the PMs across all locations.

Some limitations of the study are that external factors such as the use of microphones or loudspeakers in different events (especially at the Student Union) could influence the objective rating. In addition, there is inconsistency in the noise level app installed on different smartphones. Data collection might also differ slightly based on specific dates and times.

Possible future studies include increasing the sample size and including more time slots. Additionally, we could also incorporate a qualitative survey for subjective ratings. We could control external factors by accounting for campus events and include noise level meters (decibel meters) instead of the mobile app.

References

Montgomery, Douglas C. (2013). Design and Analysis of Experiments (8th Edition). John Wiley & Sons.