STATS 101B - Final Project Report: Effect of Different Music Genres on Short-Term Muscle Performance

Introduction

The goal of this research project is to address the research question: Does listening to certain musical genres before working out affect short-term muscle performance? We aim to investigate whether listening to specific musical genres shortly before exercise influences short-term muscle performance, which is measured by the number of arm curls completed in 30 seconds using a 2kg weight. The treatment factor is musical genre with five levels: classical, country, dance, heavy metal, and no music. To control for individual differences in physical ability, particularly due to age, a randomized complete block design (RCBD) was implemented using five age-based blocks: 18-24, 25-31, 32-38, 39-45, and 46-52.

To ensure a representative and unbiased sample, participants were selected through a multi-stage randomization process. First, we used a random number generator (1–3) to randomly choose one of the three islands: Ironboard, Providence, or Bonne Santé. After selecting an island, another random number generator was used to choose a city within that island: 1–6 for Ironboard, 1–9 for Providence, and 1–12 for Bonne Santé. Once a city was selected, we accessed the birth records from the hall and identified individuals within the targeted age group. For each age block (18–24, 25–31, 32–38, 39–45, 46–52), we used a random number generator to select among the eligible individuals listed by birth date. For example, when identifying individuals in the 18–24 age group, we counted all individuals born between 27/341 and 28/348, and then randomly selected one from that list. If the selected individual was female or declined to participate, the process was repeated until a suitable male participant consented to the study. This process continued until we obtained a total of 175 male participants, with 35 individuals evenly distributed across the five age blocks. We restricted the study to male participants aged 18 to 52 to reduce variability in muscle strength that may arise due to sex and age-related physiological differences. By narrowing the demographic range, we aimed to more precisely isolate the effect of musical genre on short-term muscle performance.

Understanding how music may influence immediate physical performance has important implications for personal fitness and athletic performance. Music is already commonly used in workout environments, but its specific impact on muscle performance is underexplored. This study aims to quantify the short-term effects of musical genre on arm muscle performance and assess whether certain genres can enhance physical performance more effectively than others.

The population model for this study is expressed as

$$y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij} \text{ for } i = 1, \ldots, 5 \text{ and } j = 1, \ldots, 5$$

where y_{ij} is the number of arm curls completed by the j^{th} participant in the i^{th} treatment group, μ is the overall mean, τ_i represents the effect of the i^{th} musical genre, β_j accounts for the effect of the j^{th} age block, and ε_{ij} is the random error term.

To determine whether musical genre has a statistically significant effect on short-term muscle performance, we tested the null hypothesis that all treatment effects are equal:

$$H_0 : \tau_1 = \tau_2 = \tau_3 = \tau_4 = \tau_5 = 0$$

against the alternative hypothesis that at least one treatment has a nonzero effect:

$$H_1 : \tau_i \neq 0 \text{ for at least one } i$$

Design of Experiment

The experimental design that we chose to answer our investigative question is the Randomized Complete Block Design because there was only one treatment factor that we were interested in and one nuisance factor that we believed was crucial to our data. At first, we wanted to use the Latin Square Design or the Graeco-Latin Square Design for our experiment and include more blocking factors, but this prevented us from including replication in our experiment. By attaining the mean square error, we were able to find an effect size in order to run the pwr.anova.test.

```
                  Df Sum Sq Mean Sq F value Pr(>F)
factor(treatment)  4   37.6    9.40   0.618 0.6522
factor(block)      4  201.2   50.30   3.307 0.0195 *
Residuals         41  623.7   15.21
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

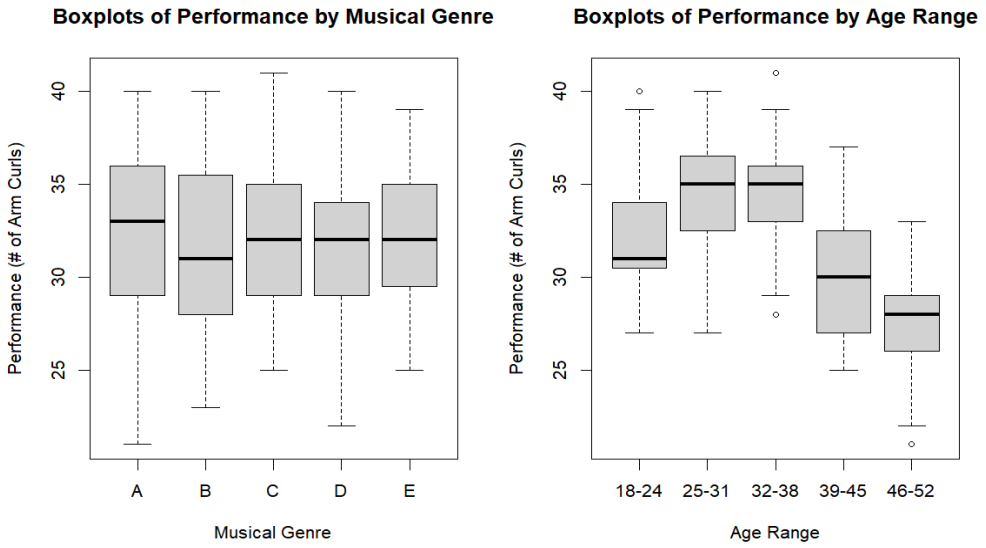
As for the effect size, we used a value of d = 3 because we wished to detect a maximum difference in mean response arm curls of 3 reps with n = 35 replicates at the 0.05 significance level and a value of k = 5 for the number of treatment groups. The effect size was calculated by dividing d=3 by the square root of the MSE which is shown above.

```
     Balanced one-way analysis of variance power calculation

              k = 5
              n = 35
              f = 0.7692308
      sig.level = 0.05
          power = 1

NOTE: n is number in each group
```

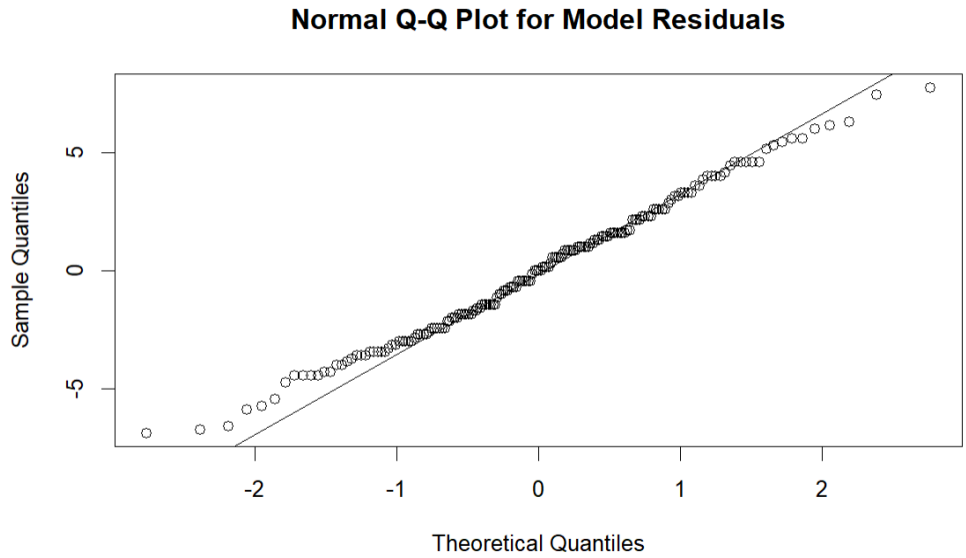
To check whether our experiment was well designed and able to detect the true treatment effects, we ran a pwr.anova.test with n = 35 replicates to find the power of the test (the probability of correctly rejecting the null when it is false). Since the test showed the power of the test was equal to 1, this confirmed that our experiment was well-designed.

Results and Interpretation

```
              Df Sum Sq Mean Sq F value Pr(>F)
genre          4    8.1     2.0   0.209  0.933
age_range      4 1374.7   343.7  35.333 <2e-16 ***
Residuals    166 1614.6     9.7
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

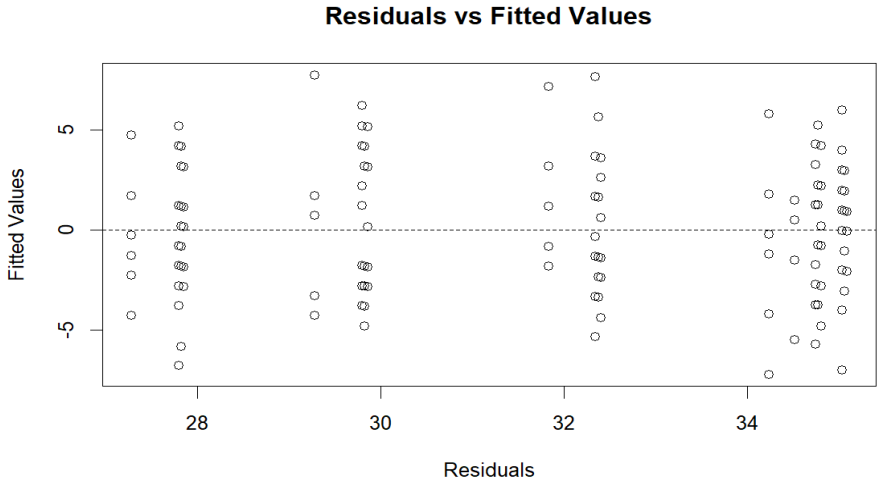


The ANOVA table for the basic model represented by the formula Performance ~ Musical Genre + Age Range provides a p-value of 0.933 > α = 0.05, which means it does not have a statistically significant effect on performance. This suggests that there is no statistically significant evidence that the type of music listened to before exercising affects the number of arm curls performed. However, this does not necessarily mean that music has no effect – it is possible that the sample size or variability in individual response masked a potential effect. The boxplots further support these findings: performance appears fairly consistent across music genres, while clear differences are observed across age groups. Overall, we cannot confidently conclude that musical genre has no effect without further investigation or a larger sample size.



To evaluate whether the residuals from our ANOVA model follow a normal distribution, we analyzed the Normal Q-Q plot. This diagnostic plot compares the ordered residuals to the expected quantiles from a standard normal distribution. In the plot, most of the residuals lie closely along the 45-degree reference line, particularly in the central region, indicating that the residuals are approximately normally distributed. Although there are minor deviations at the

tails, these are not severe and do not suggest strong departures from normality. This is important because ANOVA assumes that residuals are normally distributed, and violations of this assumption can affect the accuracy of p-values and confidence intervals. Given the overall linear pattern in the Q-Q plot, we conclude that the normality assumption is sufficiently satisfied, and the ANOVA results can be considered valid within this context.



To evaluate whether the assumptions of constant variance (homoscedasticity) and model fit are satisfied, we examined the Residuals vs Fitted Values plot. In this plot, the residuals are scattered roughly symmetrically around the horizontal line at zero across the entire range of fitted values. This pattern indicates that the variance of the residuals remains relatively constant and does not increase or decrease systematically with the fitted values. There is no clear funnel shape or curvature, which suggests that the model does not suffer from heteroscedasticity or non-linearity. The residuals also appear randomly dispersed rather than forming any visible pattern, supporting the assumption that the model adequately captures the linear relationship between the response and predictors. While a few residuals fall slightly farther from the center, they do not represent extreme outliers or systematic deviations. Therefore, we are able to conclude that the assumption of homoscedasticity is reasonably met and that the linear model is appropriate for the data.

t-test

To further investigate the influence of musical genre and age range on arm curl performance, we examined the output of the linear regression using the formula perf ~ genre + age_range. In this model, genre A was treated as the reference level for the genre variable, and the 18 - 24 age group served as the reference for age range. The intercept, estimated at approximately 32.34, represents the expected number of arm curls performed by individuals in this baseline group.

```
Call:
lm(formula = perf ~ genre + age_range, data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-7.2286 -2.0571 -0.0286  1.7429  7.7143

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)     3.234e+01  7.073e-01  45.729  < 2e-16 ***
genreB         -5.143e-01  7.455e-01  -0.690 0.491265
genreC          1.030e-14  7.455e-01   0.000 1.000000
genreD          2.857e-02  7.455e-01   0.038 0.969476
genreE          5.714e-02  7.455e-01   0.077 0.938996
age_range25-31  2.400e+00  7.455e-01   3.219 0.001546 **
age_range32-38  2.686e+00  7.455e-01   3.602 0.000416 ***
age_range39-45 -2.543e+00  7.455e-01  -3.411 0.000813 ***
age_range46-52 -4.543e+00  7.455e-01  -6.093 7.46e-09 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.119 on 166 degrees of freedom
Multiple R-squared:  0.4613,    Adjusted R-squared:  0.4354
```

Looking at the music genre coefficients, none of the levels (genre B ~ genre E) showed statistically significant differences compared to genre A. All associated p-values were greater than 0.49 (specifically, 0.4913 for genre B, 1.0000 for genre C, 0.9695 for genre D, and 0.9390 for genre E). This indicates there is no evidence that listening to a particular genre of music, compared to genre A, significantly affects short-term arm curl performance. This result aligns with our earlier ANOVA analysis and suggests that, under the conditions of this study, musical genre is not a major factor in immediate muscular output.

Blocking by age was a necessary and thoughtful design choice. Physical ability is known to vary by age, and blocking helped reduce variability and isolate the treatment effects of musical genre. Without this control, individual differences in strength could have confounded the results.

Finally, the model's Multiple R-squared value of 0.4613 and Adjusted R-squared of 0.4354 suggest that approximately 43.5% of the variation in performance is explained by the combination of musical genre and age range. This reflects a moderately good model fit, but it also highlights that age can possibly be a more influential factor than music, though not consistently significant across all analyses, in explaining short-term upper-body exercise performance within this project.

Tukey Test

From Tukey's pairwise comparison, we were able to clearly see that the blocking factor had an effect, which was expected. This finding further supports our decision to include the blocking factor of age in the experimental design, as it helped account for variability and improve the precision of pairwise comparisons. The problem was that we could not find any treatment pairwise comparisons that were statistically significant, which was the purpose of the experiment.

```
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = perf ~ genre + age_range, data = data)

$genre
             diff       lwr      upr     p adj
B-A -5.142857e-01 -2.570443 1.541872 0.9584575
C-A  1.065814e-14 -2.056158 2.056158 1.0000000
D-A  2.857143e-02 -2.027586 2.084729 0.9999995
E-A  5.714286e-02 -1.999015 2.113301 0.9999921
C-B  5.142857e-01 -1.541872 2.570443 0.9584575
D-B  5.428571e-01 -1.513301 2.599015 0.9496966
E-B  5.714286e-01 -1.484729 2.627586 0.9398340
D-C  2.857143e-02 -2.027586 2.084729 0.9999995
E-C  5.714286e-02 -1.999015 2.113301 0.9999921
E-D  2.857143e-02 -2.027586 2.084729 0.9999995

$age_range
                  diff        lwr         upr     p adj
25-31-18-24  2.4000000  0.3438423  4.45615767 0.0132427
32-38-18-24  2.6857143  0.6295566  4.74187195 0.0037676
39-45-18-24 -2.5428571 -4.5990148 -0.48669948 0.0071810
46-52-18-24 -4.5428571 -6.5990148 -2.48669948 0.0000001
32-38-25-31  0.2857143 -1.7704434  2.34187195 0.9953873
39-45-25-31 -4.9428571 -6.9990148 -2.88669948 0.0000000
46-52-25-31 -6.9428571 -8.9990148 -4.88669948 0.0000000
39-45-32-38 -5.2285714 -7.2847291 -3.17241376 0.0000000
46-52-32-38 -7.2285714 -9.2847291 -5.17241376 0.0000000
46-52-39-45 -2.0000000 -4.0561577  0.05615767 0.0608547
```

However, even though the results and p-values show that the mean comparisons are not statistically significant, something interesting that we did find was that all the pairwise comparisons that involved dance music (treatment C) yielded relatively lower p-values compared to the other treatment pairs, most of which had p-values close to 1.00. This was interesting because it logically made sense that the mean for dance music should be different compared to all the other music genres, which generally featured lower tempos and more calm moods. This observation aligns with our initial hypothesis. Although the results did not meet the threshold for statistical significance, the pattern observed may suggest a potential trend worth further investigation. These preliminary findings motivate us to continue this line of research, as experimenting with a larger sample size may yield more definitive results and potentially confirm the differences we originally anticipated.

Discussion

Through this project, our goal was to gain insights into whether listening to certain musical genres before working out the arms affects muscle performance within a short time. We first started off gathering our data by sampling 175 male individuals from the three islands and then running an ANOVA table to which our results suggests that there isn't a statistically significant effect on performance. The ANOVA table showed that the residual sum of squares is still very large and accounts for over half the variability in the data. Additionally, we created boxplots to verify our results from the ANOVA table, and the results do correlate to our findings: performance seems fairly consistent across music genres, and the age groups seem to capture more variance in the performance data (according to the sum of squares value in the ANOVA table earlier). To further evaluate our results from the ANOVA table, we created two plots: Residuals vs Fitted Values and Q-Q plot to test the normality assumption of our data. From these two plots, there were no major concerns that violate the homoscedasticity, as our residuals look to have constant variance and no apparent patterns. Because our plots looked normally distributed, we could conclude that our ANOVA results were consistent with the normality assumption. We then ran a t-test to check whether there is a significant difference between the different music genres, and our results indicate that there was no evidence that shows listening to a particular music genre would significantly affect short-term arm curl performance. Likewise, the results from our Tukey's test also show there isn't any significant difference between different music genres. Based on our experiment, we can say that we don't have enough evidence to suggest that music does have an effect on short-term muscle performance, but we aren't fully concluding that music doesn't affect muscle performance. Future studies could therefore test factors in addition to age range to capture more of that variability.

Real World Applicability

From our results, we do not have enough evidence to suggest that music affects short-term muscle performance, since our results indicate no statistically significant effect of different music on performance. As for listening to different types of music while working out, much literature and studies show a significant effect between the two. In "Lose yourself in the music: music's effect on strength training," Bryce Lewis summarized findings from three different studies to conclude that music does enhance physical performance for individuals working out. Again, our study doesn't provide certain evidence that different music types affect short-term muscle performance, but analyzing existing articles that do support this indicates that we should not definitively rule out such an effect.

Limitations

The main limitation of our design is that individual physical conditions can vary from day to day, which may influence performance. For example, if a participant is feeling unwell or is recovering from illness, their results may not accurately reflect their typical muscle ability. Additionally, while we used age as a blocking factor to control for variability, this may not fully account for other influential factors such as underlying medical conditions, occupational demands, or regular physical activity. Participants with physically demanding jobs or those who exercise regularly may naturally perform better than others in the same age group, potentially reducing the effectiveness of our age-based blocks in isolating treatment effects.