Gaming Behavior Analysis during the COVID-2019 Pandemic: A Comprehensive Study of Digital Gaming Habits in 2020.

Author Name: Anowar Hussain

Students ID: 151934177

Email: anowar.hussain@tuni.fi

Abstract:

This study investigates changes in digital gaming habits during the COVID-19 pandemic in 2020. A cross-sectional survey collected data from 794 participants. The research explores the influence of demographic factors and perceived pandemic impacts on gaming behavior and imposed restriction correlate with gamin behavior. Descriptive and inferential statistics were applied using R and two hypotheses are tested: 1) Increased leisure time due to lockdowns led to a rise in digital gaming activities. 2) Gaming behavior varied based on age, gender, and the perceived impact of the pandemic on individual lifestyles. Statistical analyses were performed using R, including t-tests, ANOVA tests and chi-square tests and correlation plot and scatter plot.

Table of Contents

1. **Introduction:**

The COVID-19 pandemic profoundly influenced leisure activities, leading to increased interest in digital gaming. This Study explores the dynamics of digital gamic habits during the pandemic, aiming to understand the underlying patterns and the impact of demographic factors and perceived pandemic effects on gaming behavior.

2. **Aims:**

• Research Questions:

1. How did the frequency and types of digital game usage change during the COVID-19 pandemic in 2020?

2. What demographic factors and perceived impacts of the pandemic influenced gaming behaviour, and how were these behaviours correlated with the imposed restrictions and social isolation measures?

• Hypotheses:

1. Increased leisure time due to lockdowns led to a rise in digital gaming activities.

2. Gaming behaviour varied based on age, gender, and the perceived impact of the pandemic on individual lifestyles.

3. **Material and Methods:**

Sample Description: The data set comprises 793 observations surveyed during COVID-19 lockdowns. The sample includes individuals of various ages, gender, restrictions, and diverse gaming habits.

Variables and Measurement scales: The analysis focused on variables such as age(q39) gender(q40), imposed restrictions (q27 to q33), survey responses (survey) and weekly time spent gaming (q42).

Data Collection Methods: Data was collected via a cross-sectional survey from the FINNISH SOCIAL AND DATA ARCHIVED (FSD3547 Playing Video Games during the COVID-19 Pandemic: Survey 2020)

Statistical Analysis Methods: Descriptive statistics, t-tests, Anova, chi-square tests, regression analysis and correlation analysis and scattered plot were performed.

R Version and Packages Used: In this project to analyze the data I used R version 4.3.1 (2023-06-16 ucrt) Dplyr, ggplot2, car packages were utilized for data manipulation, data visualization and regression analysis.

4. Result

Summary of Statistics: This result was the summary of data which was used to analyze the model.

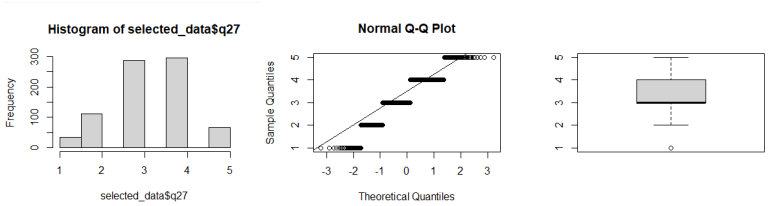
```
      q27             q28             q29             q30             q31             q32       
 Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000  
 1st Qu.:3.000   1st Qu.:3.000   1st Qu.:3.000   1st Qu.:3.000   1st Qu.:2.000   1st Qu.:3.000  
 Median :3.000   Median :3.000   Median :3.000   Median :3.000   Median :3.000   Median :3.000  
 Mean   :3.314   Mean   :3.248   Mean   :3.253   Mean   :3.095   Mean   :2.927   Mean   :3.098  
 3rd Qu.:4.000   3rd Qu.:4.000   3rd Qu.:4.000   3rd Qu.:4.000   3rd Qu.:4.000   3rd Qu.:4.000  
 Max.   :5.000   Max.   :5.000   Max.   :5.000   Max.   :5.000   Max.   :5.000   Max.   :5.000  

      q33             q39             q40             q42                survey     
 Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :    0.00   Min.   :1.000  
 1st Qu.:3.000   1st Qu.:3.000   1st Qu.:1.000   1st Qu.:    5.00   1st Qu.:4.000  
 Median :3.000   Median :4.000   Median :1.000   Median :   15.00   Median :5.000  
 Mean   :3.259   Mean   :3.845   Mean   :1.478   Mean   :   86.32   Mean   :4.832  
 3rd Qu.:4.000   3rd Qu.:5.000   3rd Qu.:2.000   3rd Qu.:   25.00   3rd Qu.:6.000  
 Max.   :5.000   Max.   :9.000   Max.   :4.000   Max.   :44654.00   Max.   :7.000  
```

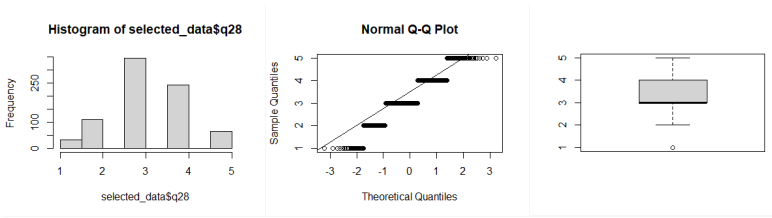
Demographic Characteristics (frequency table for Categorical Variables):

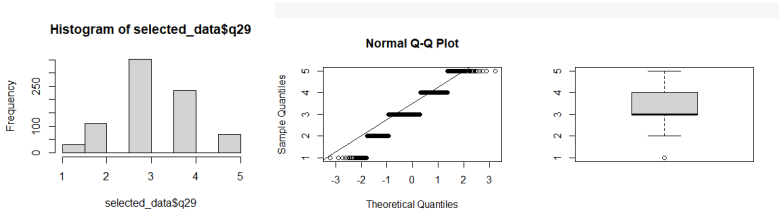
```
Frequency Table for Restriction (q27):

  1   2   3   4   5 
 34 111 286 296  66 
Frequency Table for Restriction (q28):

  1   2   3   4   5 
 32 109 346 242  64 
Frequency Table for Restriction (q29):

  1   2   3   4   5 
 30 109 352 234  68 
Frequency Table for Restriction (q30):

  1   2   3   4   5 
 46 126 378 193  50 
Frequency Table for Restriction (q31

  1   2   3   4   5 
 74 193 282 205  39 
Frequency Table for Restriction (q32)

  1   2   3   4   5 
 52 119 356 231  35 
Frequency Table for Restriction (q33):

  1   2   3   4   5 
 34 126 296 275  62 
Frequency Table for Age Group (q39):

  1   2   3   4   5   6   7   8   9 
  6  88 272 189 129  40  18  10   9 
Frequency Table for Gender (q40):

  1   2   4 
458 313  22 
Frequency Table for Survey Response:

  1   2   3   4   5   6   7 
 96  28  28  79 250 154 158 
```

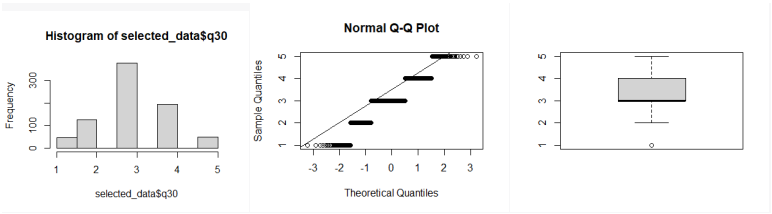
Distribution Plots and Shapiro-Wilk Tests:

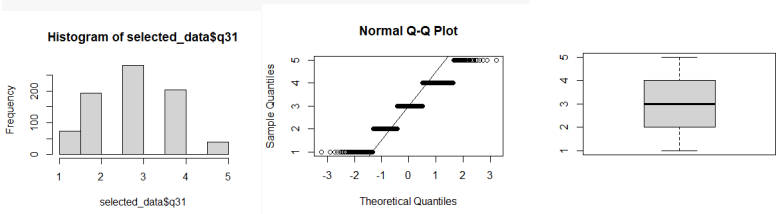
These plots (histogram, qqline, boxplot) are for all the variables to check the normality.

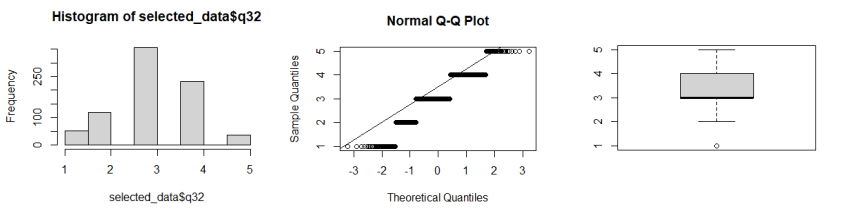


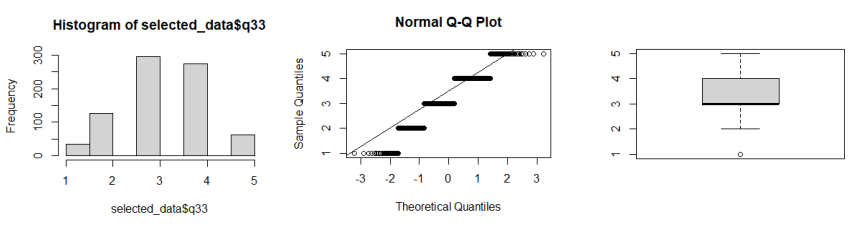


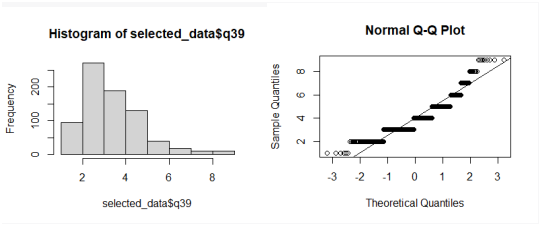


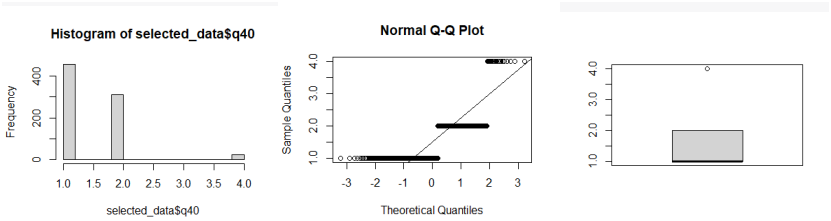


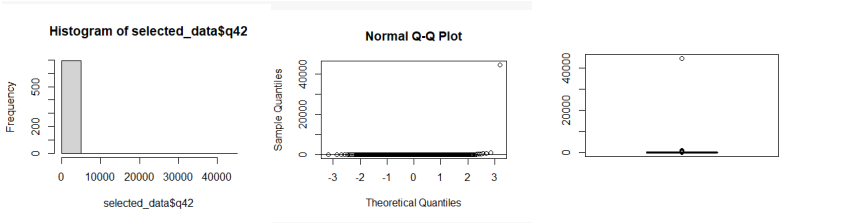


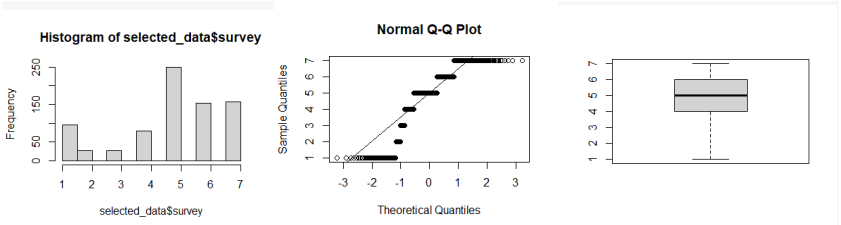












The result of Shapiro-Wilk normality test for all variables:

```
data:  selected_data$q27
W = 0.89271, p-value < 2.2e-16
data:  selected_data$q28
W = 0.89355, p-value < 2.2e-16
data:  selected_data$q29
W = 0.89353, p-value < 2.2e-16
data:  selected_data$q31
W = 0.91, p-value < 2.2e-16
data:  selected_data$q32
W = 0.88568, p-value < 2.2e-16
data:  selected_data$q33
W = 0.89804, p-value < 2.2e-16
data:  selected_data$q39
W = 0.88973, p-value < 2.2e-16
data:  selected_data$q40
W = 0.64724, p-value < 2.2e-16
data:  selected_data$q42
W = 0.020123, p-value < 2.2e-16
data:  selected_data$survey
W = 0.85595, p-value < 2.2e-16
```

Inferential Statistics:

Chi-Square test of Demographic Variables:

Table-1: Chi-Square Test Result for Demographic Variables with Gaming Behavior

Demographic Variable	P-Value
Age	0.0001535
Gender	0.0601419
Restriction_1	0.0000120
Restriction_2	0.0000000
Restriction_3	0.1564096
Restriction_4	0.0015557
Restriction_5	0.0024070
Restriction_6	0.1977177
Restriction_7	0.0001084
Survey	0.0267995

Regression Analysis for Gaming Behavior Correlation:

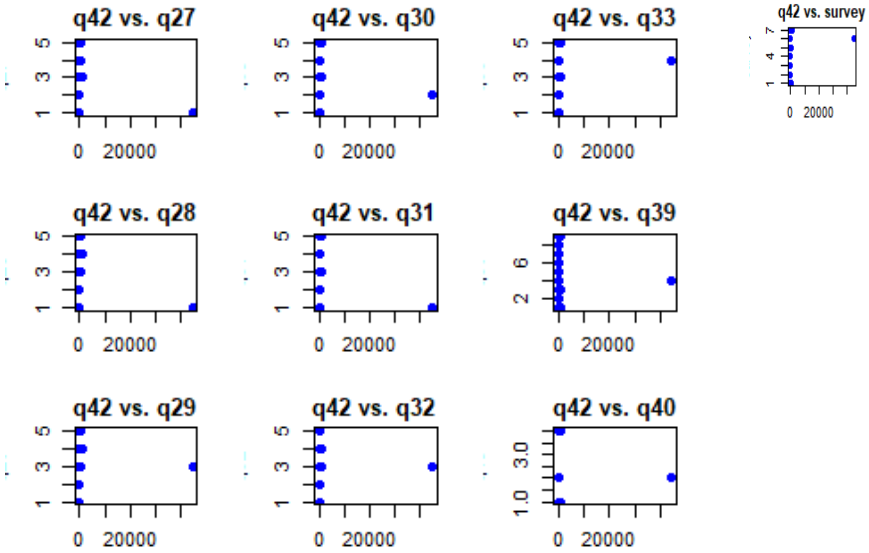
Table-2: Regression Analysis Results for Gaming Behavior

	Variable	Coefficient	SE	T-Value	P-Value
(Intercept)	Intercept	273.44359	518.83338	0.5270355	0.5983447
q27	Restriction_1	-204.17113	86.26574	-2.3667696	0.0182296
q28	Restriction_2	-171.95242	91.70763	-1.8750068	0.0612298
q29	Restriction_3	116.98820	88.65357	1.3196107	0.1874191
q30	Restriction_4	-27.43788	81.47013	-0.3367845	0.7363856
q31	Restriction_5	-114.74591	69.84913	-1.6427679	0.1009037
q32	Restriction_6	56.62206	86.86182	0.6518636	0.5147144
q33	Restriction_7	154.11043	82.45940	1.8689249	0.0620726
q39	Age	18.21673	49.14496	0.3706734	0.7109988
q40	Gender	140.82077	114.76405	1.2270460	0.2202395
survey	Survey	29.92712	37.35382	0.8011796	0.4233138

Correlation to the variable time spent weekly(q42) with other variables:

```
          q27       q28       q29        q30        q31       q32       q33        q39
q42 0.1480054 0.1674922 0.2134695 0.06816177 0.05962924 0.1068246 0.0550501 0.01845988
             q40     survey
q42 -0.006207299 -0.1873153
```

The scatter plots for the variable time spent weekly (q42) with others variables:



Discussion:

Interpretation of the result:

Summary statistics revealed variations in gaming habits across different age groups, genders, imposed restrictions and survey. Frequency tables provided insights into the distribution of gaming time among categorical variables. Distribution plots and Shapiro-Wilk tests assessed the normality of continuous variables.

From the Distribution plots (Histograms, QQline, Boxplots) observed the clear patterns of all the continuous categorical variables and from the Shapiro-Wilk tests result observed that the p-value of all the variables is less the 0.05 so all the variables are not normally distributed.

The frequency table interprets the distribution of characteristics of the sample and frequencies according with their categorical values.

Data Analysis of the result:

Chi-Square test (Table-1):

Age(p=0.0001535) has a significant association with weekly gaming time spent(q42) and younger age groups might spend more time on gaming compared to older age groups during the lockdowns.

Gender(p=0.0601419) is not statistically significant association with weekly gaming time, and it differs between genders in the survey population.

Restrictions_1, Restrictions_2, Restrictions_4 and Restrictions_7 have significant associations and other restrictions dose not have any significant association with gaming behavior (q42) that means some restrictions might lead to increased activities.

Survey (p=0.0267995) has a significant association with weekly gaming time spent(q42).

Regression Analysis (Table-2):

Restriction_1, Restriction_2, Restriction_5 have negative significant coefficients, indicating their influence leading to reducing the weekly gaming time.

Restriction_3, Restriction_6, Restriction_7 have positive significant coefficients, indicating their influence leading to increase the weekly gaming time.

Restriction_4, age, gender, and survey are not statistically significant though they have positive coefficient, because their p-values are greater than 0.05. Suggesting no influence on gaming time.

Correlation Matrix:

The variables q27, q28, q29, q30, q31, q32 have strong positive correlation suggesting that if they increase the weekly gaming time tends to increase. Especially q29 is the highest positive correlated.

The variables q33, q39(age) have weaker positive correlation with the weekly gaming time.

The variables gender(q40) and survey have negative correlation with the weekly gaming time(q42). It suggests that if these variables increase, the weekly gaming time tends to decrease.

Conclusions:

This study provides valuable insights into gaming behavior during lockdowns. While certain restrictions age played significant roles, the influence of gender and survey responses highlights further exploration. The finding results emphasize the importance of considering diverse factors in understanding leisure activities during lockdown periods.

6. Appendices

Appendix 1: List of Variables

- q39: Age Group (1-5)
- q40: Gender (1, 2, 4)
- q27 to q33: Imposed Restrictions (1-5)
- survey: Survey Responses (1-7)
- q42: Weekly Time Spent Gaming

Appendix 2: R Code Used for Analysis

```
daF3547 <- read.csv2("~/Biostatistics/daF3547.csv", header=TRUE)
names(daF3547)
class(daF3547)
head(daF3547)  # Display the first few rows of the data
str(daF3547)   # Display the structure of the data frame
View(daF3547)
summary(daF3547)

library(dplyr)  # for data manipulation
library(ggplot2)  # for data visualization
library(car)  # for regression analysis

# Subset the data for selected variables
selected_data <- daF3547 %>% dplyr::select(q27, q28, q29, q30, q31, q32, q33, q39, q40,
q42, survey)
# Summary statistics for selected variables
summary(selected_data)

# Frequency table for demographic variables
print("Frequency Table for Restriction (q27):")
print(table(daF3547$q27))
print("Frequency Table for Restriction (q28):")
print(table(daF3547$q28))
print("Frequency Table for Restriction (q29):")
print(table(daF3547$q29))
print("Frequency Table for Restriction (q30):")
print(table(daF3547$q30))
print("Frequency Table for Restriction (q31):")
print(table(daF3547$q31))
print("Frequency Table for Restriction (q32):")
print(table(daF3547$q32))
print("Frequency Table for Restriction (q33):")
print(table(daF3547$q33))
print("Frequency Table for Age Group (q39):")
print(table(daF3547$q39))
print("Frequency Table for Gender (q40):")
print(table(daF3547$q40))
print("Frequency Table for Survey Response:")
print(table(daF3547$survey))

#Distribution Plots and Shapiro-Wilk Tests:
```

```
hist(selected_data$q27)
qqnorm(selected_data$q27)
qqline(selected_data$q27)
boxplot(selected_data$q27)
shapiro.test(selected_data$q27)

hist(selected_data$q28)
qqnorm(selected_data$q28)
qqline(selected_data$q28)
boxplot(selected_data$q28)
shapiro.test(selected_data$q28)

hist(selected_data$q29)
qqnorm(selected_data$q29)
qqline(selected_data$q29)
boxplot(selected_data$q29)
shapiro.test(selected_data$q29)

hist(selected_data$q30)
qqnorm(selected_data$q30)
qqline(selected_data$q30)
boxplot(selected_data$q30)
shapiro.test(selected_data$q30)

hist(selected_data$q31)
qqnorm(selected_data$q31)
qqline(selected_data$q31)
boxplot(selected_data$q31)
shapiro.test(selected_data$q31)

hist(selected_data$q32)
qqnorm(selected_data$q32)
qqline(selected_data$q32)
boxplot(selected_data$q32)
shapiro.test(selected_data$q32)

hist(selected_data$q33)
qqnorm(selected_data$q33)
qqline(selected_data$q33)
boxplot(selected_data$q33)
shapiro.test(selected_data$q33)

hist(selected_data$q39)
qqnorm(selected_data$q39)
qqline(selected_data$q39)
boxplot(selected_data$q39)
shapiro.test(selected_data$q39)

hist(selected_data$q40)
qqnorm(selected_data$q40)
```

```
qqline(selected_data$q40)
boxplot(selected_data$q40)
shapiro.test(selected_data$q40)

hist(selected_data$q42)
qqnorm(selected_data$q42)
qqline(selected_data$q42)
boxplot(selected_data$q42)
shapiro.test(selected_data$q42)

hist(selected_data$survey)
qqnorm(selected_data$survey)
qqline(selected_data$survey)
boxplot(selected_data$survey)
shapiro.test(selected_data$survey)


#Data Analysis:
# Perform chi-square test for categorical variables (age, gender, restrictions, survey)
chi_square_age <- chisq.test(selected_data$q39, selected_data$q42)
chi_square_gender <- chisq.test(selected_data$q40, selected_data$q42)
chi_square_restriction_1 <- chisq.test(selected_data$q27, selected_data$q42)
chi_square_restriction_2 <- chisq.test(selected_data$q28, selected_data$q42)
chi_square_restriction_3 <- chisq.test(selected_data$q29, selected_data$q42)
chi_square_restriction_4 <- chisq.test(selected_data$q30, selected_data$q42)
chi_square_restriction_5 <- chisq.test(selected_data$q31, selected_data$q42)
chi_square_restriction_6 <- chisq.test(selected_data$q32, selected_data$q42)
chi_square_restriction_7 <- chisq.test(selected_data$q33, selected_data$q42)
chi_square_survey <- chisq.test(selected_data$survey, selected_data$q42)
chi_square_results <- data.frame(
  Variable = c("Age", "Gender", "Restriction_1", "Restriction_2", "Restriction_3",
               "Restriction_4", "Restriction_5", "Restriction_6", "Restriction_7", "Survey"),
  P_Value = c(chi_square_age$p.value, chi_square_gender$p.value,
              chi_square_restriction_1$p.value, chi_square_restriction_2$p.value,
              chi_square_restriction_3$p.value, chi_square_restriction_4$p.value,
              chi_square_restriction_5$p.value, chi_square_restriction_6$p.value,
              chi_square_restriction_7$p.value, chi_square_survey$p.value)
)
library(knitr)
# Print the Chi-Square test results as Table-1
kable(chi_square_results,
      col.names = c("Demographic Variable", "P-Value"),
      caption = "Table-1: Chi-Square Test Result for Demographic Variables with Gaming
Behavior",
      align = "c")

# Perform regression analysis
regression_model <- lm(q42 ~ q27 + q28 + q29 + q30 + q31 + q32 + q33 + q39 + q40 +
survey, data = selected_data)
# Store regression summary results in a data frame
```

```
regression_results <- summary(regression_model)$coefficients
# Extract coefficients, standard errors, t-values, and p-values
coefficients <- regression_results[, 1]
standard_errors <- regression_results[, 2]
t_values <- regression_results[, 3]
p_values <- regression_results[, 4]

# Create a data frame for regression analysis results
regression_table <- data.frame(
  Variable = c("Intercept","Restriction_1", "Restriction_2", "Restriction_3", "Restriction_4",
               "Restriction_5", "Restriction_6", "Restriction_7",
               "Age", "Gender", "Survey"),
  Coefficient = coefficients,
  SE = standard_errors,
  T_Value = t_values,
  P_Value = p_values
)
# Print the regression analysis results as Table-2
kable(regression_table,
      col.names = c("Variable", "Coefficient", "SE", "T-Value", "P-Value"),
      caption = "Table-2: Regression Analysis Results for Gaming Behavior",
      align = "c")


# Calculate correlation matrix between q42 and all other variables
correlation_matrix <- cor(selected_data[, c('q42')], selected_data[, c('q27', 'q28', 'q29', 'q30',
'q31', 'q32', 'q33', 'q39', 'q40', 'survey')], use="complete.obs", method = "spearman")
rownames(correlation_matrix)[1] <- "q42"
print(correlation_matrix)


# Create scatter plots for q42 against all other numeric variables

numeric_columns <- sapply(selected_data, is.numeric) & !(names(selected_data) %in%
c("q42"))

# Get the names of numeric columns (excluding q42)
numeric_column_names <- names(selected_data)[numeric_columns]

# Create scatter plots for q42 against all other numeric variables
par(mfrow=c(3, 3))  # Set the layout for the plots (adjust rows and columns as needed)
for (col in numeric_column_names) {
  plot(selected_data$q42, selected_data[[col]], main=paste("q42 vs.", col),
       xlab="q42", ylab=col, pch=16, col="blue")
}
```