# **Homework Report**

|  |  |  |  |
| --- | --- | --- | --- |
| **Individual Contribution** | | | |
| **CWID** | **Name** | **Contribution** | **Percent Contribution** |
| A20563445 | Liyin Shi | One visualization and report writing | 33.3% |
| A20563456 | Haowen Zheng | One visualization and data gathering | 33.3% |
| A20563412 | Kaixiao Duan | Two visualizations | 33.3% |

1. **Data Description**

This dataset contains demographic information, gaming habits and academic performance of 112 students. The attributes are as follows:

**Sex**: The student's gender (0 for 'Female' and 1 for 'Male').

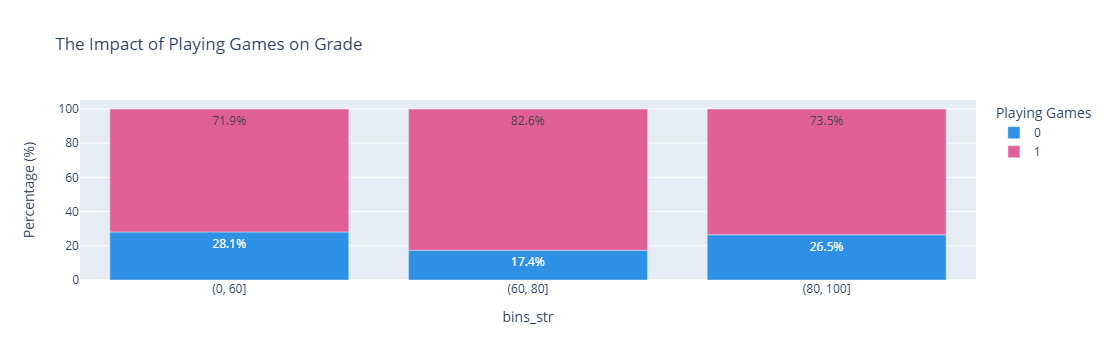
**Playing Years**: The number of years the student has been playing games.

**Playing Often**: A rating of how often the student plays games, likely on a scale from 1 to 5.

**Playing Hours**: The number of hours the student spends playing games per day.

**Playing Games**: The number of types of games the student plays (e.g., 0: not playing games, 1: playing games).

**Grade**: The student's academic grade.

1. **Data Visualizations & Details**
2. **The Impact of Playing Games on Grade**

**Visualization:**

This bar chart provides an explicit visual demonstration of how Playing Games has an impact on Grades by showing the percentage of students who play games across different grade ranges.

To visualize this part, we split and counted the data by different grade ranges and whether students play games, calculated the percentage of students whose Playing Games is 1 within each score range, and then used bar function of plotly to create the chart.

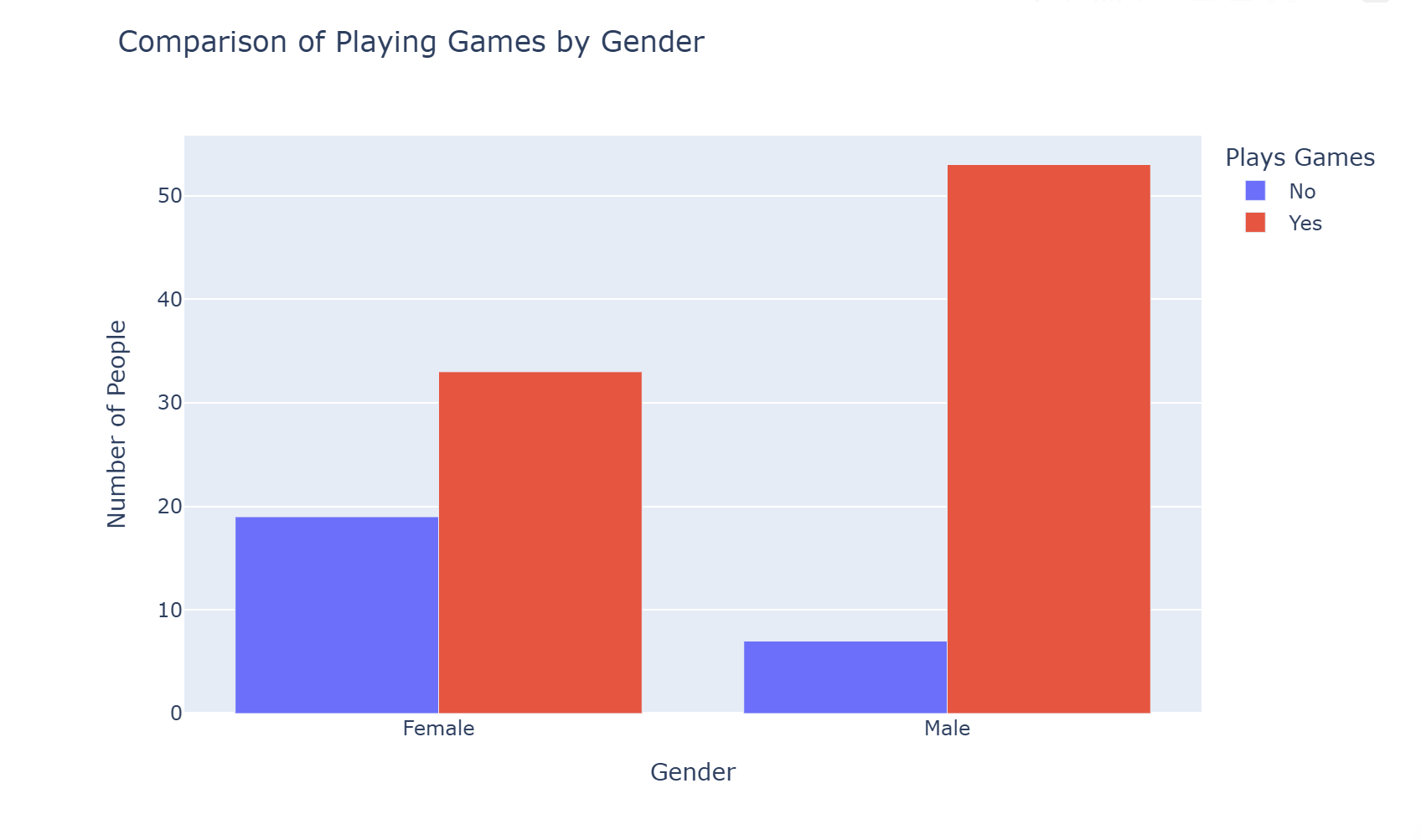
**Libraries used:**

plotly, pandas.

**Analysis:**

Before the data visualization, we made a guess that playing games might have a negative impact on grades. However, the chart shows that the proportion of students who play games is relatively low and similar in both the low and high score ranges, while the middle score range has a slightly higher proportion. Therefore, there doesn’t appear to be a direct correlation between playing games and grades.

1. **The Relationship between Sex and Playing Games**



### Visualization:

In this data visualization, we used a **Grouped Bar Chart** to explore the relationship between **Sex** and **Playing Games**.

First, we performed data preprocessing by reading the csv file of the dataset and selecting the two attributes: **Sex** and **Playing Games** for analysis. The original values of **Sex** were 0 (Female) and 1 (Male). To enhance readability, we mapped them to **"Female"** and **"Male"**, respectively.  
 Similarly, **Playing Games** was originally 0 (Not Playing) and 1 (Playing), which we mapped to **"No"** and **"Yes"** accordingly. We then counted the number of individuals in each gender category who either **play games** or **do not play games**.

For data visualization, we used **bar function of plotly** to create a grouped bar chart for **Gender vs. Playing Games**, where:

x="Sex": The X-axis represents gender (Male & Female).

y="Count": The Y-axis represents the number of people who play or do not play games.

color="Playing Games": Different colors distinguish between individuals who play games ("Yes") and those who do not ("No").

barmode="group": Displays different categories (Yes/No) side by side for easy comparison.

### Libraries Used:

During the data processing and visualization, we primarily used the following Python libraries:

**pandas**: For loading the CSV data and performing data cleaning and processing (such as mapping **Sex** and **Playing Games** values).

**plotly.express**: For creating an **interactive grouped bar chart**, making the data visualization more intuitive and visually appealing.

### Analysis:

From the grouped bar chart, we can clearly observe the distribution of **playing games vs. not playing games** among different genders. The proportion of **males (Male) playing games is higher**, with the majority of males choosing to play games ("Yes"). In contrast, the proportion of **females (Female) playing games is lower**, and compared to males, more females choose **not** to play games ("No").

Overall, males are more likely to play games than females, suggesting that gender may have some influence on the choice to play games.

### Conclusion:

The results indicate that **males are more inclined to play games than females**, while a significant proportion of females choose **not** to play games. This may reflect certain **cultural or societal trends**, such as the general perception that males tend to have a stronger interest in gaming, whereas females may prefer other forms of entertainment.

1. **The Impact of Playing Years and Playing Often on Grades**



**Visualization:**

1.X-axis: This axis categorizes years of play into three subgroups: low, medium and high.

2. Y-axis: This axis categorizes the frequency of play into three subgroups: short, medium, and long.

3. Colors of the heat map: The shades of colors represent the average scores of the different subgroups. The darker the color, the higher the score; the lighter the color, the lower the score.

The playing year (Playing Year) and playing frequency (Playing Often) are discretized into three levels: low, medium and high. Then group the game year and frequency combinations by the discretized boxes, calculate the average score of each group, and set the heat color to map the average score to generate a heat map.

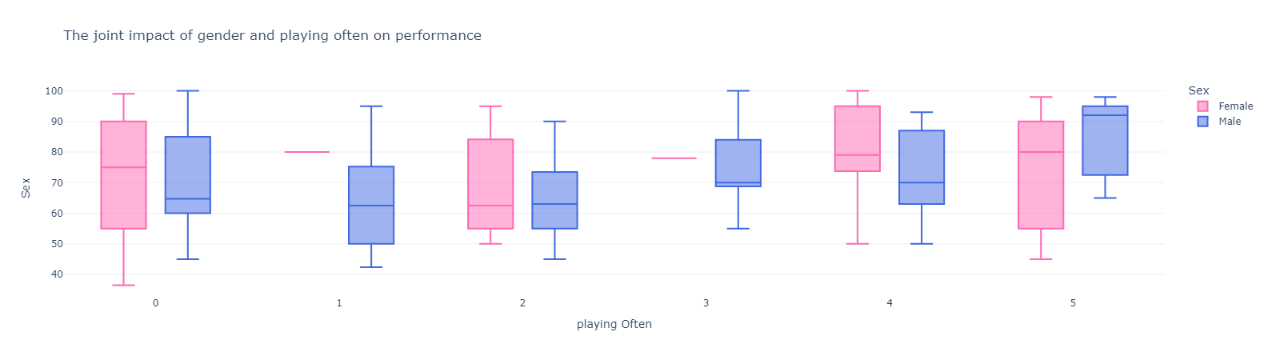
**Libraries Used:**

plotly, pandas, numpy.

**Analysis:**

As you can see from the graph: The combination of Short, High (Short, High) game and High Years (Short, High) performs better with a score of around 74. We can get that proper games can help improve the score.

1. **The Impact of Sex and Playing Often on Grades**



**Visualization:**

0 indicates female and 1 indicates male in the source dataset. Specify the playing Often as the x-axis and Grade as the y-axis, and then plot a box-and-line plot according to the gender grouping.

1. X-axis (game frequency): The X-axis is labeled with values from 0 to 5, indicating different game frequency levels.

2. Y-axis: Y-axis values range from 40 to 100 and represent grades.

3. Colors indicate gender: pink for women and blue for men.

**Libraries Used:**

plotly, pandas, numpy.

**Analysis:**

We can observe: Trends in performance change are not exactly the same for males and females at different frequencies of play. It is showed that there is a difference in the impact of games on the sexes.

1. **Results & Conclusions**

For the dataset we studied, we primarily analyzed the relationships between sex, gaming, and grades.

Through our visualizations, we found that there isn't a particularly noticeable correlation between students' gaming habits and their grades, which does not align with our initial hypothesis. However, the association between sex and gaming strongly conforms to the prevailing stereotypes.

Through our research we noticed that, firstly, for the group of students in our dataset, the experience of playing games did not have a negative impact on their academic performance; secondly, the experience of playing games may have a positive effect on the performance of some individuals; finally, among this group, the impact of games on them shows sex-based differences.