**Factors influencing the focus of Second-Year IT Students at LSPU SCC and Their Impact on Study Habits.**

**Members:**

* Salvador, Jobert D.
* Ramiro, Ivan Geonyl L.
* Oracion, John Eric E.
* Tolentino, Sasha Angela P.
* Villarosa, Mark Anthony B.

**Section:** BSIT-2C

**Date:** 01/04/2024

**1. Introduction**

* The goal of our project is to determine the factors that are affecting the focus of the Second-Year IT Students at LSPU SCC and how it affects the way the students study.
* The Data that we’re visualizing are hours of the student that are being poured into either their hobby or studying. The data came from an online survey that we made.

**What questions are you hoping to answer with this visualization?**

* The questions that we are hoping to be answered in this visualization is, how big is the impact of the hobbies that the students are doing, to their focus in studying.

**2. Data**

**Number of rows and columns:**

- Our dataset consists of 36 rows and 7 columns.

**Data types of each variable:**

**The survey captures various data types:**

- "How many people live in your home?" - Integer

- "The average income of the family as a whole" - Categorical

-"How many hours a day do you spend studying?" - Categorical

-“What aspects of studying do you think have the biggest impact on your ability to focus? (What pastimes do you enjoy?)” – Categorical

- “In a week, how much homework do you typically have?” - Categorical

**Missing values and how they were handled:**

- No missing values were observed in our dataset. All survey responses were complete.

**Outliers and how they were addressed:**

- As the survey primarily deals with categorical data, outliers are not applicable in this context.

**Data transformations performed:**

- Income ranges were transformed into categorical variables for better analysis. For instance, "10k - 20k" became a categorical variable representing the income bracket.

**3. Visualization Techniques**

**List the types of visualizations used:**

* **Bar Chart:** Utilized for categorical data, such as "What aspects of studying do you think have the biggest impact on your ability to focus? (What pastimes do you enjoy?)" and "In a week, how much homework do you typically have?"
* **Pie Chart or Stacked Bar Chart:** Employed to illustrate the distribution of income ranges.

**Justify your choice of visualization for each type of data:**

* **Bar Chart:** Enables clear comparison between different categories, suitable for categorical data.
* **Pie Chart or Stacked Bar Chart:** Effectively represents proportions or comparison between different categories, providing a visual overview.

**Explain how the visual elements encode the data:**

* **Color:** Used to differentiate between categories or highlight specific trends within the visualizations.
* **Size:** Adjusted to represent the magnitude of variables, such as larger bars indicating more hours of study.

**Mention any additional libraries or packages used for visualization:**

* **Matplotlib and Seaborn:** Employed for creating static, informative, and aesthetically pleasing visualizations in Python.

**4. Implementation in Google Collab**

**Provide snippets of the relevant Python code used to create the visualizations:**

* import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

* survey\_data = pd.read\_csv("Survey2.csv")

plt.figure(figsize=(10, 6))

sns.countplot(x='In a week, how much homework do you typically have?', data=survey\_data, order=['1-2', '3-4', '5-6'], palette='muted')

plt.title('Number of Pending Homework')

plt.xlabel('Homework Count')

plt.ylabel('Count')

plt.show()

* plt.figure(figsize=(14, 8))

sns.countplot(x='What aspects of studying do you think have the biggest impact on your ability to focus? (What pastimes do you enjoy?)', hue='How frequent?', data=survey\_data, palette='viridis')

plt.title('Impact of Studying Aspects on Ability to Focus by Frequency')

plt.xlabel('Studying Aspects')

plt.ylabel('Count')

plt.xticks(rotation=45, ha='right')

plt.legend(title='Studying Frequency')

plt.show()

* plt. figure(figsize=(10, 6))

sns.countplot(x='How many hours a day do you spend studying?', data=survey\_data, order=['1-2hrs', '2-3hrs', '4-5hrs'], palette='pastel')

plt.title('Distribution of Study Hours per Day')

plt.xlabel('Study Hours')

plt.ylabel('Count')

plt.show()

* plt.figure(figsize=(10, 6))

sns.countplot(x='How many people live in your home?', data=survey\_data)

plt.title('Factors Influencing Focus: Number of People Living in Homes')

plt.xlabel('Number of People')

plt.ylabel('Count')

plt.show()

* plt.figure(figsize=(8, 8))

survey\_data['The average income of the family as a whole'].value\_counts().plot.pie(autopct='%1.1f%%', startangle=90)

plt.title('Impact of Income on Study Habits: Distribution of Income Ranges')

plt.ylabel('') # Remove the default label

plt.show()

**Data loading and cleaning:**

* Load survey2 data into a Pandas DataFrame (survey\_data) using pd.read\_csv("Survey2.csv").

**Visualization construction:**

* We use Seaborn and Matplotlib functions to create visualizations based on survey questions.

**Customization and styling:**

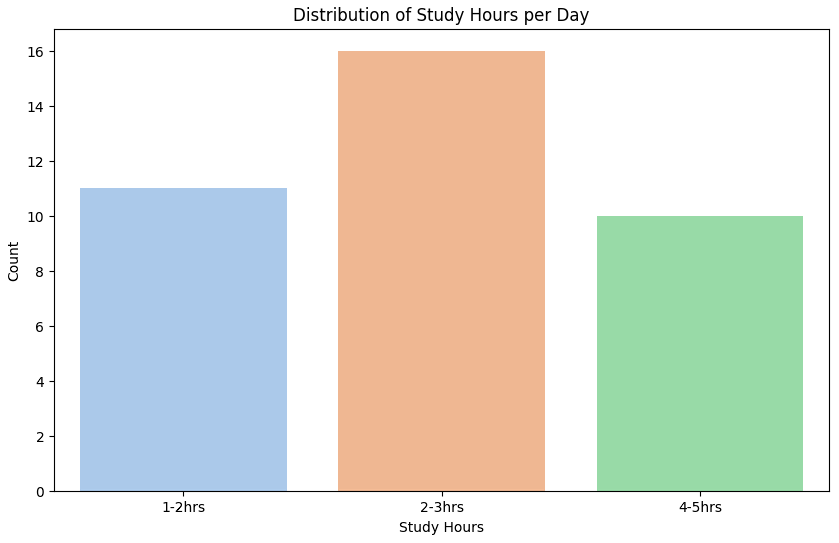
* We didn’t customize the data.

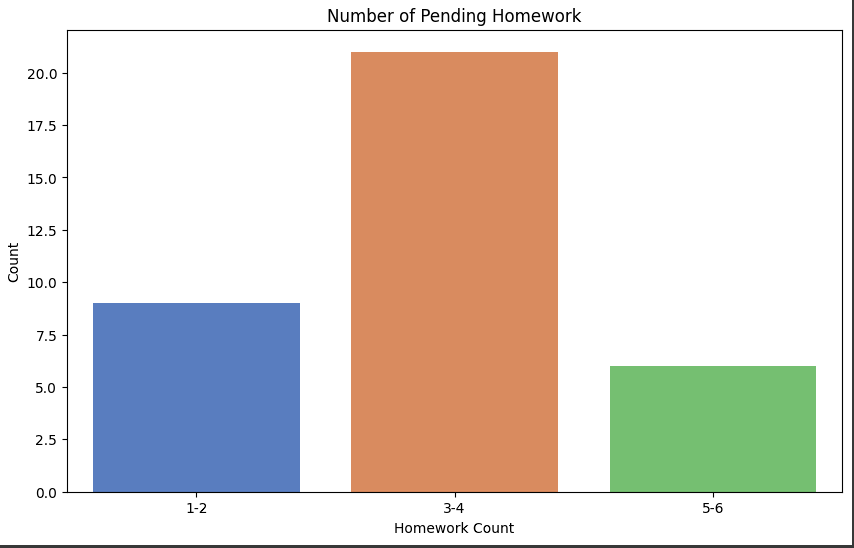
**Challenges:**

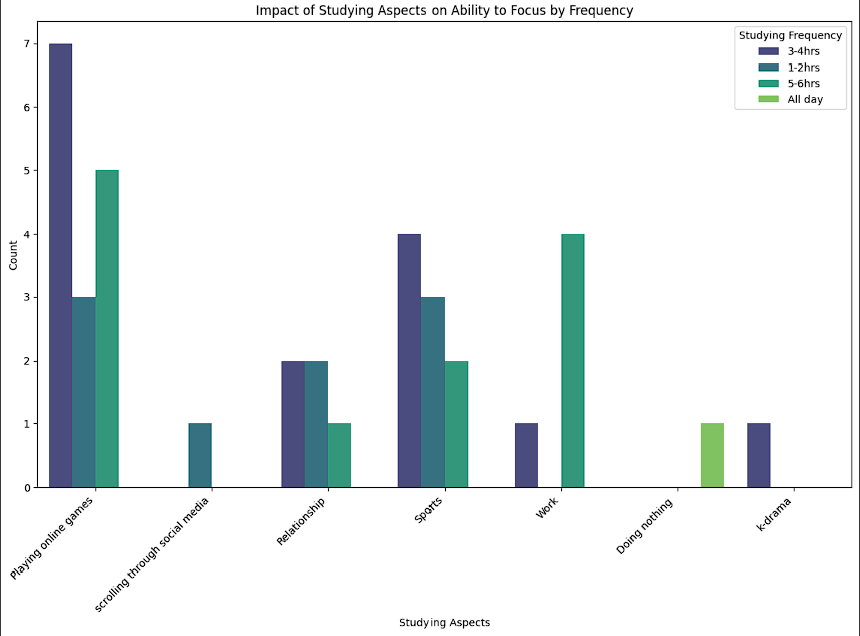
* Ensure the file path is correct.
* Choosing appropriate visualization types.

**Solutions:**

* Double-check file path and file name to ensure correct data loading.
* Thoroughly assess each question's nature to select suitable visualization types.

**5. Results and Interpretation**

****

****

* As you can see from our graphics, the leading hobby is playing online games then the most homework counts are 3-4 and also the leading study hours is 2-3 hours.
* We can see that the hobbies such as playing mobile games has a pretty big impact in the pending home works and the study hours because the homework count's leading one is the 3-4 and the study hours the lowest is the longest study hours also.
* The visualizations answer the initial project questions by presenting that the online gaming is a big factor that affects the focus of the students in studying because instead of studying they spend their time playing online games.
* While our visualizations provide valuable insights, it's crucial to acknowledge potential limitations. The data relies on self-reported responses from second-year IT students at LSPU SCC, introducing the possibility of response bias. Participants may understate or overstate certain aspects, impacting the accuracy of the findings and also the study is based on a specific time frame and may not capture long-term trends or changes in students' habits and preferences. External factors such as academic workload variations or shifts in popular hobbies over time could influence the results.

**6. Conclusion**

Our analysis of the survey data sheds light on the key factors influencing the focus of second-year IT students at LSPU SCC and their impact on study habits. The following key findings and takeaways emerge from our visualizations:

**Leading Hobbies:** The leading hobby among the surveyed students is playing online games.

**Homework Counts:** The majority of students report having 3-4 hours of homework per week.

**Study Hours:** The leading study hours per day fall within the 2-3 hours range.

**Impact of Playing Mobile Games:** Our analysis suggests a notable impact of playing mobile games on both pending homework counts and study hours. Students reporting this hobby tend to have a higher count in the 3-4 hours homework range, and their study hours are concentrated in the lower brackets, with the lowest count observed in the longest study hours category.

Our visualizations provide valuable insights into the dynamics of student life, particularly highlighting the potential impact of online gaming on academic responsibilities. It is essential to acknowledge certain limitations in our study, including the reliance on self-reported responses, which introduces the possibility of response bias. Participants may unintentionally misrepresent certain aspects, impacting the accuracy of our findings.

Additionally, our study is based on a specific time frame, and the evolving nature of student habits and preferences may not be fully captured. External factors, such as variations in academic workloads or changes in popular hobbies, could influence the observed results over time.

**Future Improvements and Extensions:**

To enhance the robustness of our findings and address potential limitations, future iterations of this project could consider:

**Longitudinal Studies:** Conducting longitudinal studies to track changes in study habits and hobbies over an extended period, providing a more comprehensive understanding of trends.

**Objective Measures:** Supplementing self-reported data with objective measures, such as academic performance metrics, to validate and enhance the reliability of our findings.

**Expanded Demographic Analysis:** Expanding the demographic scope to include a more diverse sample, considering factors such as gender, socioeconomic status, or academic majors, to uncover potential variations in study habits.

**Qualitative Insights:** Incorporating qualitative research methods, such as interviews or focus groups, to gain deeper insights into the motivations and experiences of students with different hobbies.

In conclusion, while our current analysis offers valuable insights into the interplay between hobbies and academic focus among second-year IT students, continued research and refinement of methodologies will contribute to a more comprehensive understanding of these dynamics.

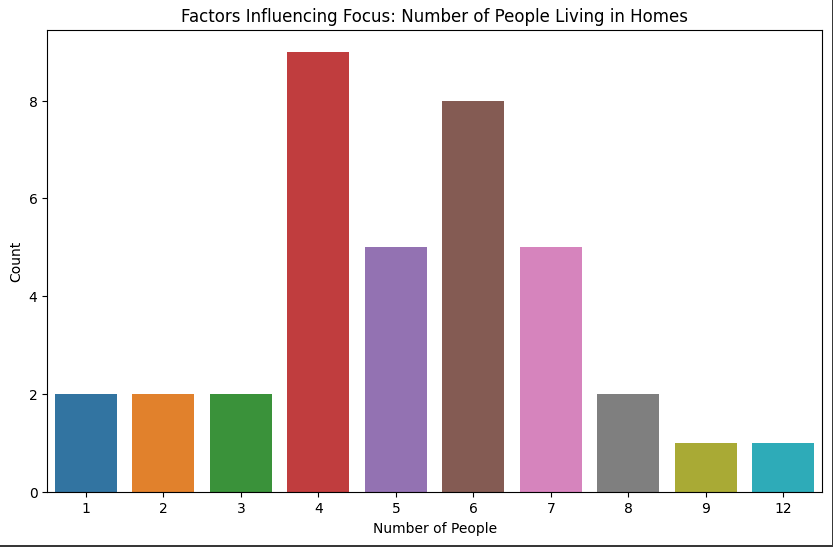
**7. Appendix**

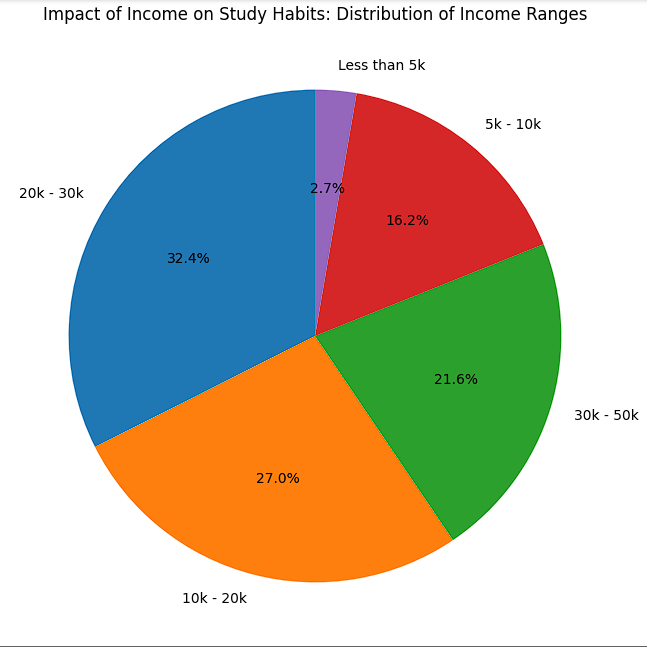
**Data Dictionary:**

**Below is a summary of the key variables used in our analysis:**

* **How Many People Live In Your Home:** Number of people living in the respondent's home.
* **Income:** Average income of the family as a whole.
* **Studying Aspects Impact:** Aspects of studying that have the biggest impact on the ability to focus (response to the question on pastimes enjoyed).
* **Pending Homework Per Week:** Number of hours of homework typically done in a week.
* **Study Hours per Day:** Number of hours spent studying per day.

**Additional Visualizations and Analysis:**





* The additional visualization that we presented here are questions also from our survey that may affect the focus of the students aside from the hobbies that we gave.