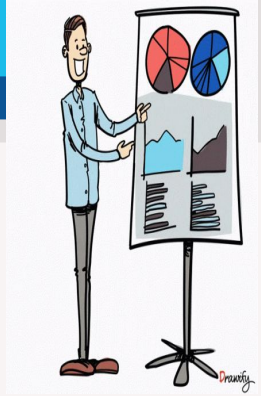


Google Slide link: Presentation Slide Show

Analysis of Key Factors Influencing Student Performance

A Data-Driven Approach to Understanding GPA Outcomes

By. Roberto and Nicholas

August 15,2024

Introduction Overview

Purpose of the Presentation:

The purpose of this presentation is to analyze key factors that influence student academic performance, particularly focusing on GPA. By leveraging data-driven insights, we aim to identify the impact of various elements such as study habits, parental involvement, absenteeism, gender, and ethnicity on student outcomes. The ultimate goal is to provide actionable recommendations that can enhance educational strategies and improve student success.

Agenda

1. **Objectives:**
   - **To understand how study habits, parental involvement, and gender influence GPA.**
   - **To evaluate specific hypotheses related to study time, absences, tutoring, and GPA.**
   - **To assess the relationship between parental education level and parental involvement.**
   - **To explore gender and ethnicity differences in GPA distribution.**
   - **To provide data-driven insights and recommendations for improving student performance.**
2. **Methodology:**
   - **Overview of the data collection process and tools used.**
   - **Description of the analytical methods employed, including correlation analysis, t-tests, and visualizations.**
3. **Analysis:**
   - **Detailed examination of the data to uncover significant trends and relationships.**
   - **Interpretation of key correlations between variables like study time, absences, and GPA.**
4. **Findings:**
   - **Presentation of key results, highlighting the most impactful factors on student performance.**
   - **Discussion of any surprising or unexpected patterns in the data.**

Objective

- Understand the impact of study habits, parental involvement, and gender on GPA.
- Test specific hypotheses related to study time, absences, tutoring, and GPA.
- Explore gender and ethnicity differences in GPA distribution.
- Provide data-driven insights and recommendations.

Methodology Pt.1

1. Data Collection and Preparation:

- **Dataset:** The data used in this analysis was collected from a dataset containing various attributes related to student performance, such as GPA, study time per week, absences, tutoring hours, parental involvement, gender, and ethnicity.
- **Preprocessing:**
  - The data was cleaned to handle any missing values, inconsistencies, or outliers that could skew the analysis.
  - Relevant variables were identified and selected for analysis. For instance, continuous variables like `StudyTimeWeekly`, `Absences`, and `GPA` were directly analyzed, while categorical variables like `Gender` and `Ethnicity` were appropriately encoded.

2. Correlation Analysis:

- **Purpose:** To measure the strength and direction of the linear relationship between pairs of variables, particularly in understanding how factors like study time, absences, and parental involvement relate to GPA.
- **Method:** Pearson's correlation coefficient was calculated for each pair of variables. This statistical measure ranges from -1 to 1, where:
  - **+1** indicates a perfect positive linear relationship.
  - **-1** indicates a perfect negative linear relationship.
  - **0** indicates no linear relationship.
- **Visualization:** A correlation heatmap was generated to visually represent these relationships, making it easier to identify strong correlations at a glance.

Methodology Pt.2

3. Hypothesis Testing:

- **Purpose:** To statistically evaluate specific hypotheses regarding the relationship between variables.
- **T-Tests:**
  - **Gender Differences in GPA:** Independent t-tests were conducted to compare the means of GPA between male and female students. This test helps determine if there is a statistically significant difference in GPA between genders.
- **Correlation Tests:**
  - Tests were conducted to assess the significance of correlations between study time, absences, tutoring, and GPA. For example, the correlation between absences and GPA was tested to confirm whether the observed negative correlation is statistically significant.

4. Visualizations:

- **Scatter Plots and Boxplots:** These were used to visualize relationships between variables. For instance:
  - A scatter plot was used to show the relationship between study time and GPA, highlighting the trend of higher study time being associated with higher GPA.
  - Boxplots were employed to compare the GPA distribution across different genders and ethnic groups, providing a clear visual of central tendencies and variability.
- **Correlation Heatmap:** As mentioned, this was a key tool for summarizing the correlations between multiple variables simultaneously.

Methodology Pt.3

Interpretation and Insights:

- **Data-Driven Insights:** The results from the correlation analysis, hypothesis testing, and visualizations were interpreted to identify key factors that have the most significant impact on student performance.
- **Actionable Recommendations:** Based on these insights, recommendations were formulated for educators, parents, and policymakers to enhance academic outcomes through targeted interventions.

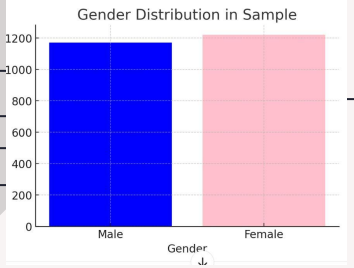
Data Overview

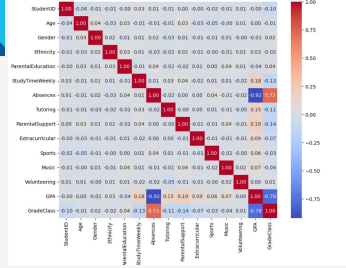
Dataset Description:

- The dataset analyzed in this study includes detailed information on student demographics, academic performance, and behavioral factors. The dataset comprises records from a sample of 2392 students, with each record containing key variables such as:
  - **GPA:** The primary measure of academic performance.
  - **Study Time per Week:** The number of hours students spend studying outside of class each week.
  - **Absences:** The total number of school days missed by each student.
  - **Tutoring Hours:** The amount of time students receive additional academic support.
  - **Parental Support:** A measure of the involvement and support provided by parents or guardians.
  - **Gender and Ethnicity:** Demographic variables that may influence academic outcomes

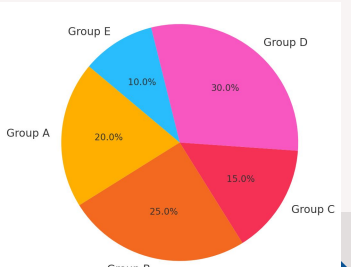
Data Preprocessing:

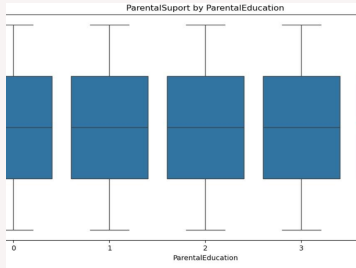
- Before analysis, the dataset was carefully prepared to ensure accuracy and reliability:
  - **Handling Missing Data:** Any missing values in critical variables were addressed either by imputation (for numerical variables) or by excluding records with insufficient data.
  - **Encoding Categorical Variables:** Categorical variables like gender and ethnicity were encoded to facilitate analysis. For example, gender was encoded as 0 for males and 1 for females.
  - **Outlier Detection:** Outliers in variables such as GPA and study time were identified and examined to determine if they should be excluded from the analysis or if they represent important variations in the data.



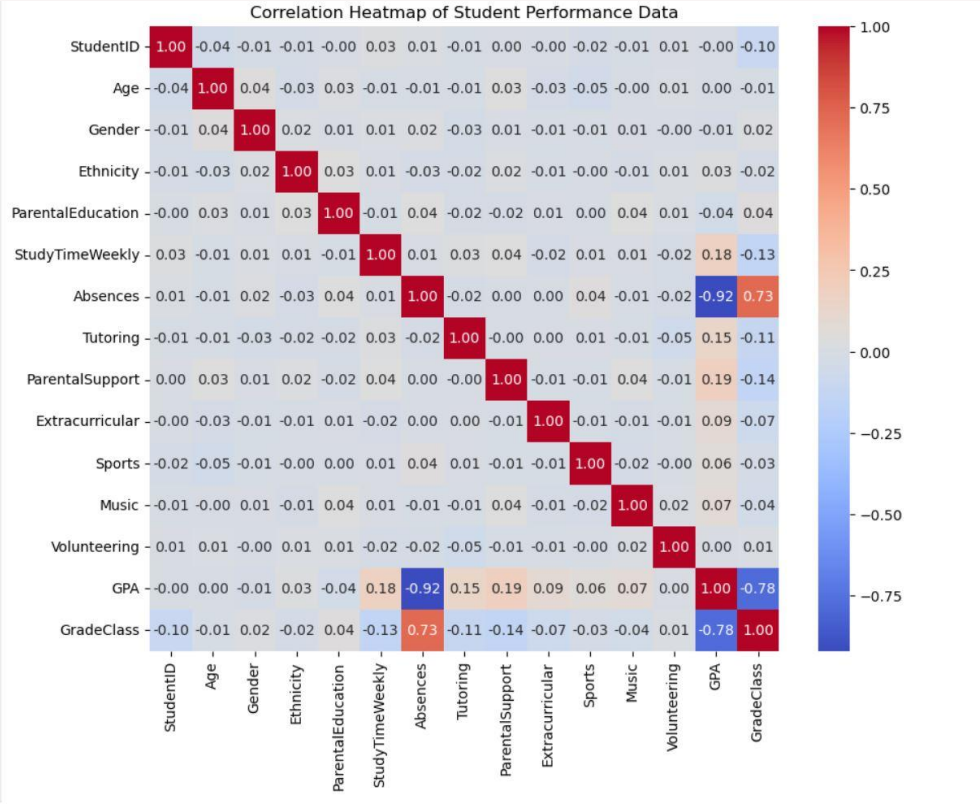


Variable	Description	Mean	Median	Std. Dev.
GPA	Grade Point Average	1.91	1.92	0.92
Study Time (hours)	Weekly study hours	9.77	9.00	5.65
Absences	Number of school days missed	14.54	14.00	8.47
Tutoring (hours)	Weekly tutoring hours	0.30	0.00	0.46
Parental Support	Level of parental involvement (0-4 scale)	2.12	2.00	1.12

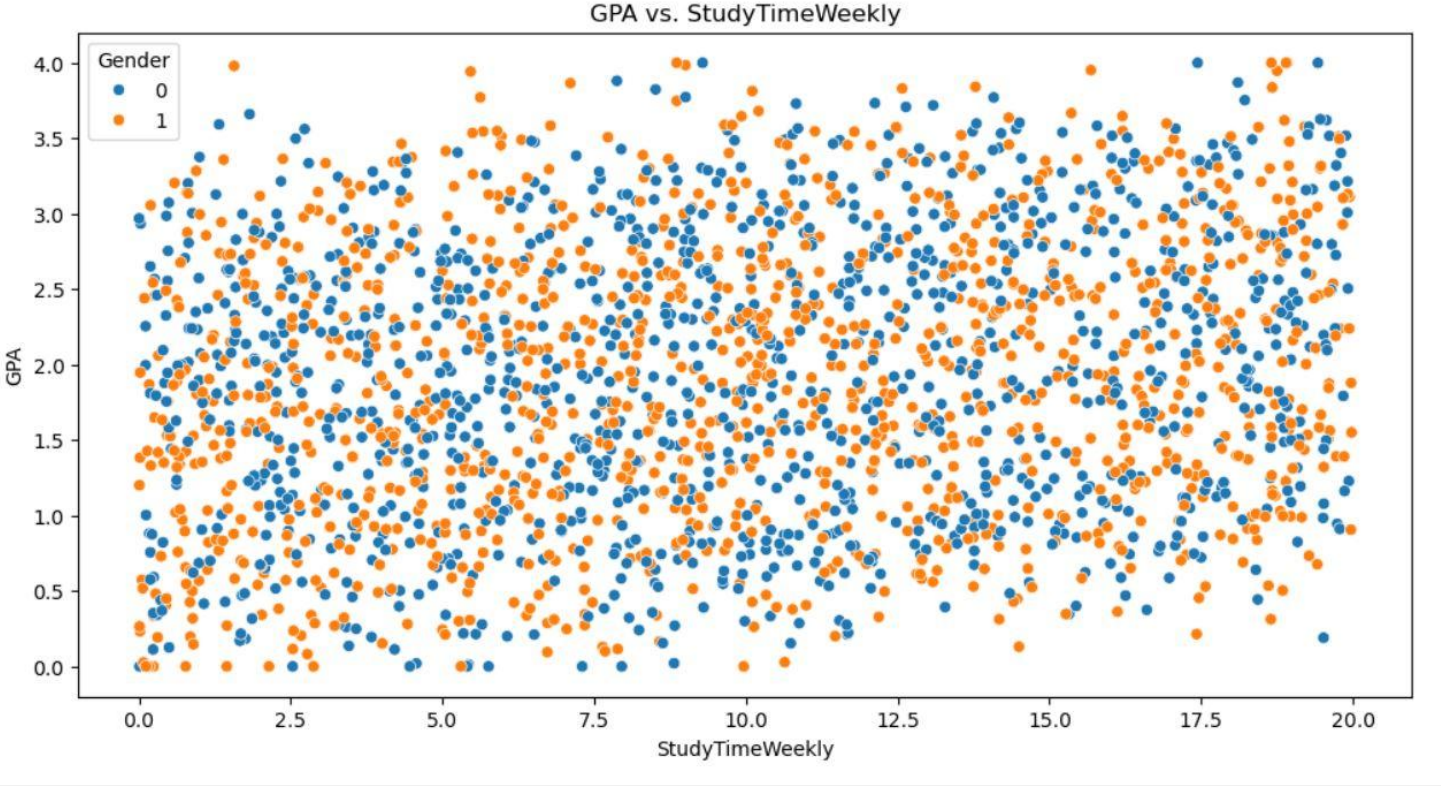




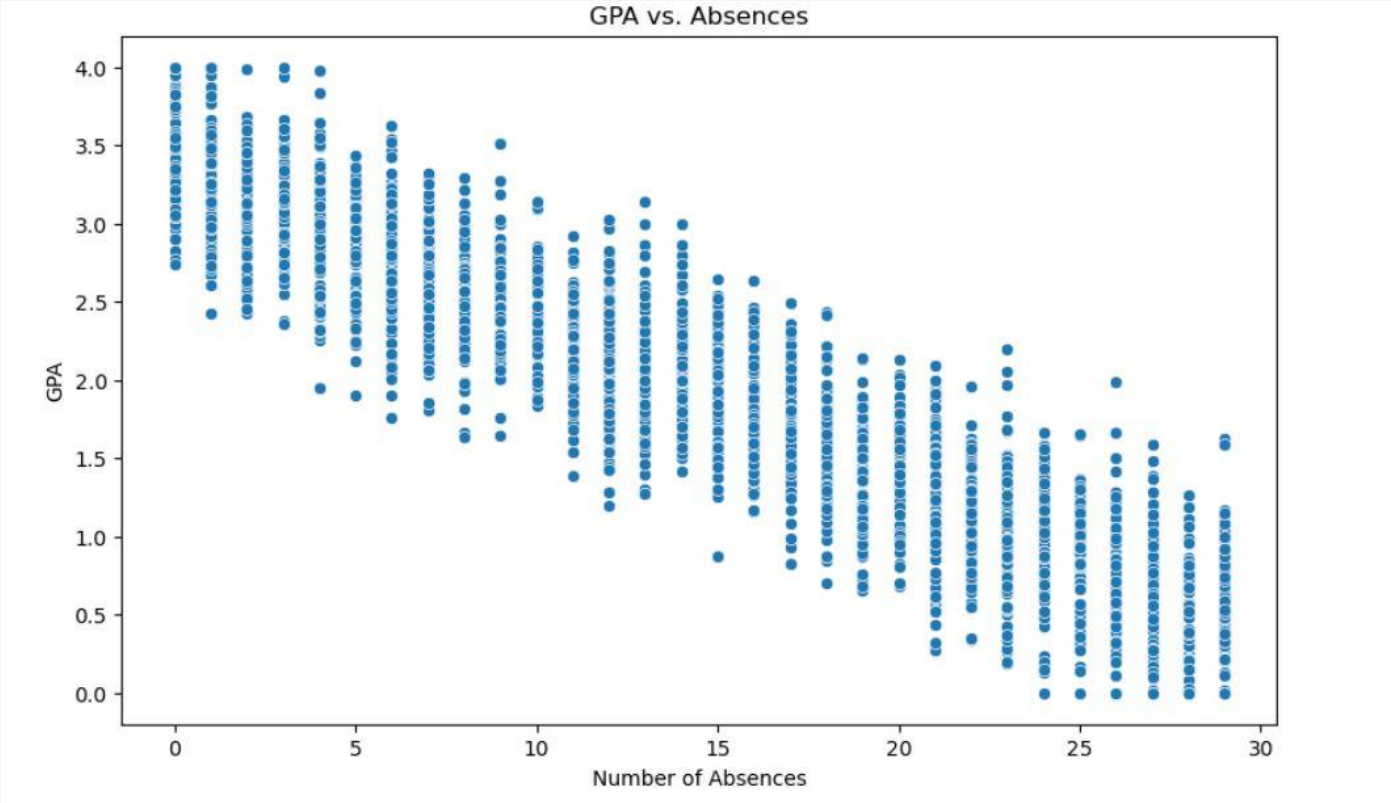
Correlation Heatmap



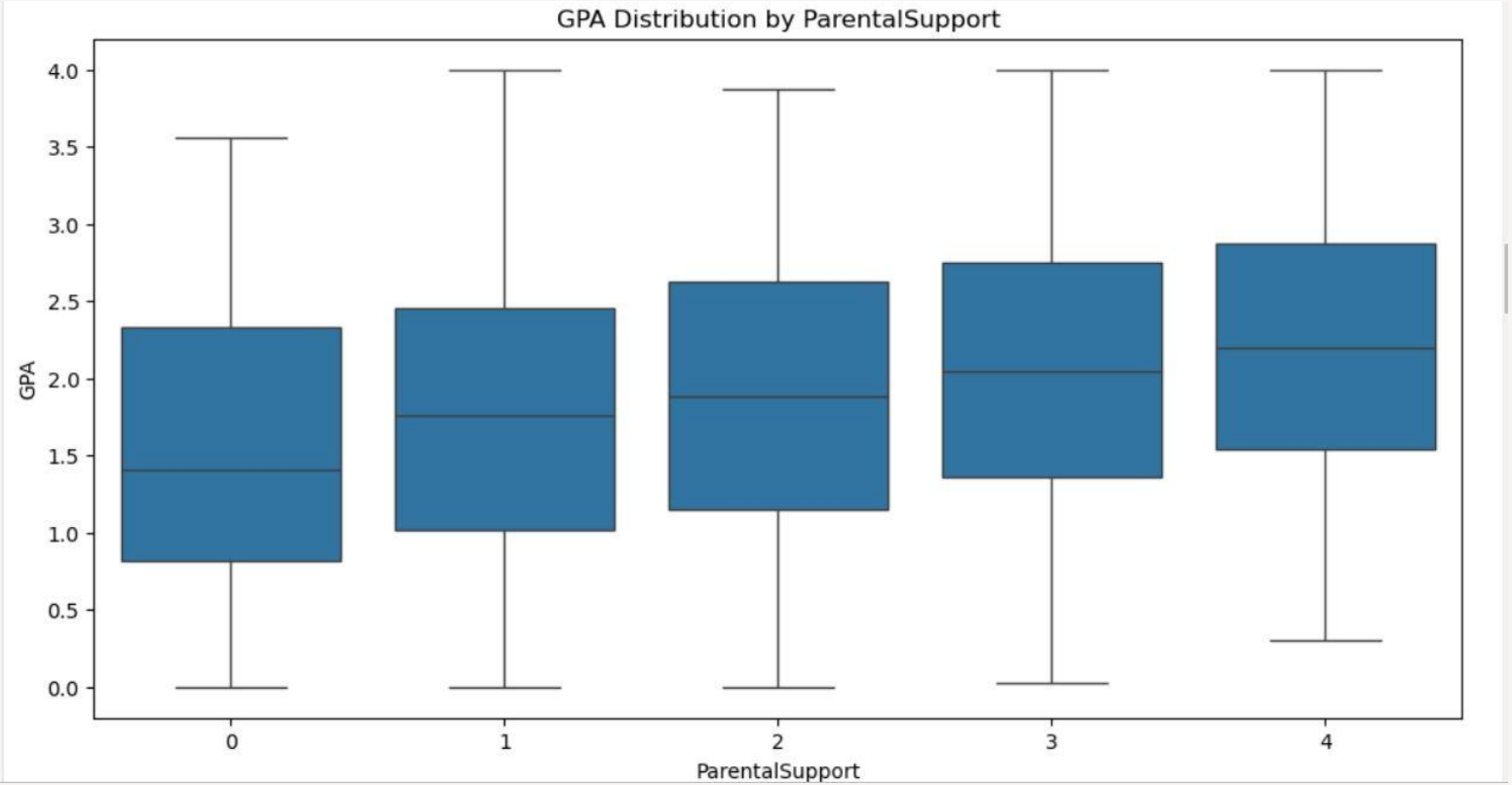
Study Time Weekly vs Gpa



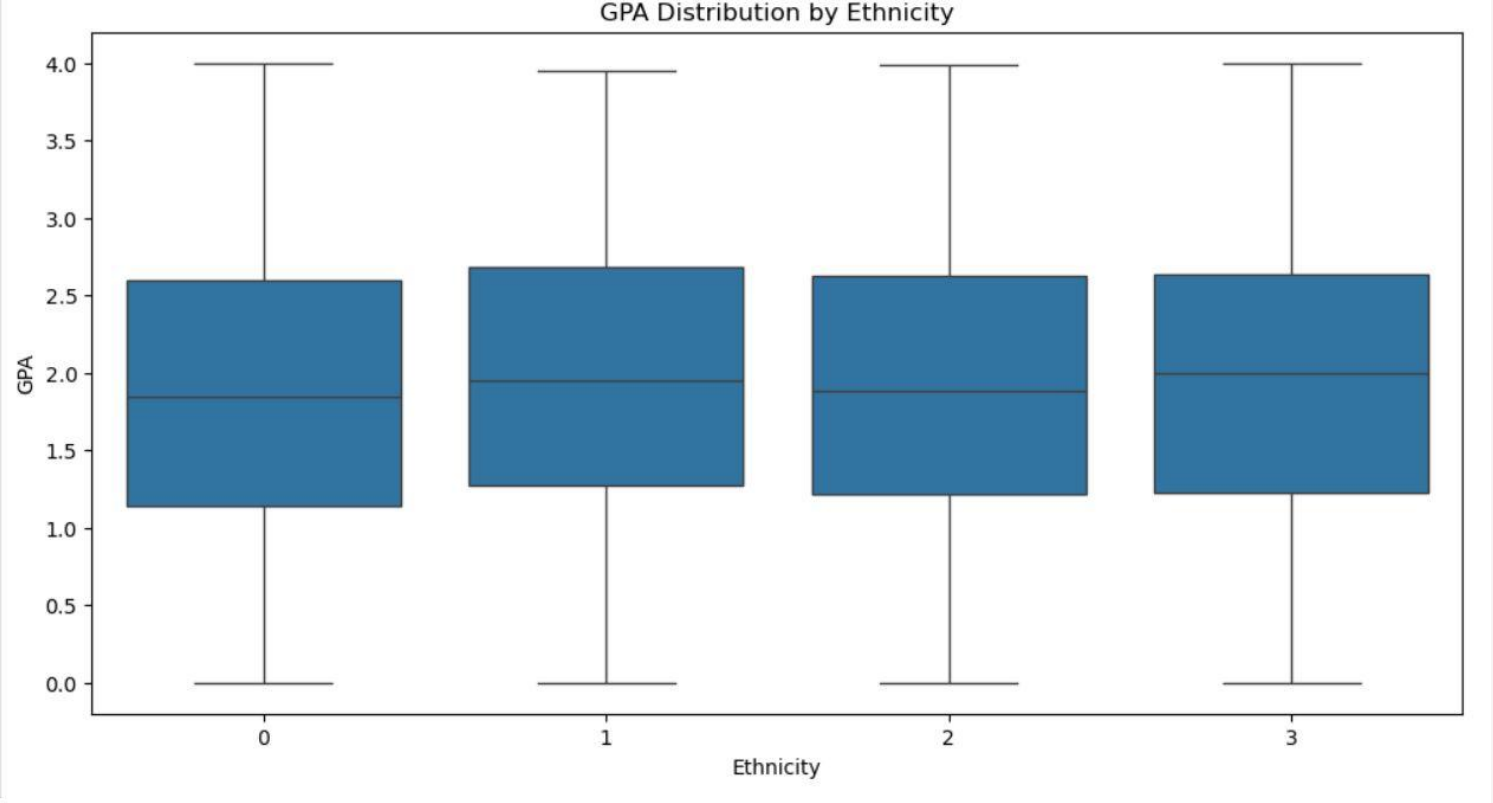
Absences and GPA



Parental Support and GPA



GPA and Ethnicity



What did we Find ?

Findings

Parental Involvement:

- **Key Result:** Parental support showed a positive correlation with student GPA. Students who reported higher levels of parental involvement and support tended to have better academic performance.
- **Significance:** This finding highlights the importance of engaging parents in their children's education. Schools might consider developing programs that foster greater parental involvement, especially for students who may be struggling academically.

Findings

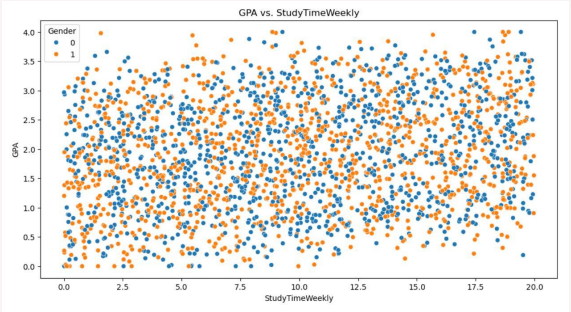
Impact of Study Habits on GPA:

- Key Result: The analysis revealed a positive correlation between study time and GPA. Students who spent more time studying each week generally had higher GPAs. This finding underscores the importance of cultivating effective study habits among students to improve academic outcomes.
- Significance: Encouraging students to allocate sufficient time for studying and providing them with tools and strategies to study more effectively could significantly enhance their academic performance.

Discussion of Surprising or Unexpected Patterns

Lack of Strong Correlation Between Study Time and GPA:

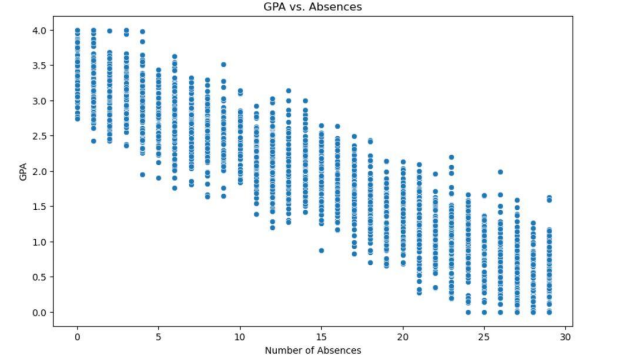
- **Unexpected Finding:** While study time is positively correlated with GPA, the scatter plot analysis showed a wide variation in GPA at different levels of study time. This indicates that simply increasing study time does not guarantee a higher GPA, suggesting that the quality of study time and other factors like student engagement, comprehension, and study strategies are also critical.



Discussion of Surprising or Unexpected Patterns

Absenteeism Strong Impact on GPA:

- Unexpected Finding: The strength of the negative correlation between absenteeism and GPA was particularly notable. This reinforces the idea that even small increases in absenteeism can have a disproportionately large impact on academic performance.



Discussion of Surprising or Unexpected Patterns

Minimal Gender Differences in GPA:

- **Unexpected Finding:** The minimal differences in GPA between genders were somewhat unexpected, given that many educational studies find gender-related performance differences in specific subjects or areas.