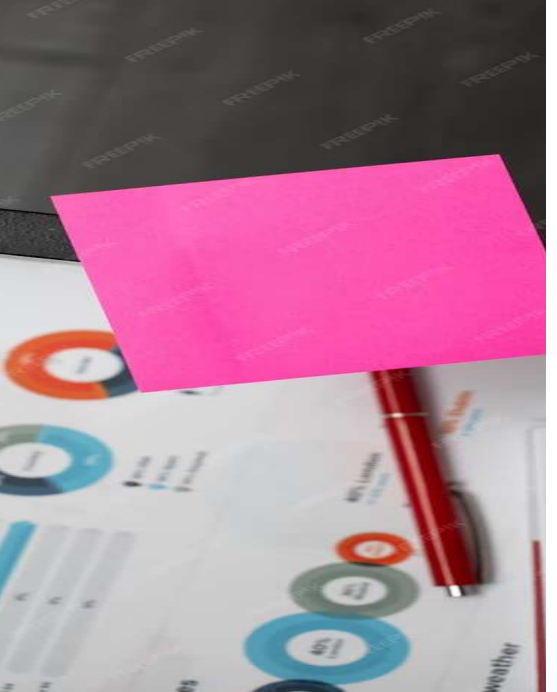
Unveiling the Patterns: An In-depth Analysis of Students' Exam Scores



Introduction

Welcome to the comprehensive dataset capturing the academic landscape of students at a (fictional) public school. Within these digital corridors of information, you will find a wealth of data encompassing not only the students' test scores but also a myriad of personal and socio-economic factors that weave together in intricate patterns, potentially influencing their academic performance. This collection serves as a window into the dynamic interplay between individual characteristics and external variables, shedding light on the nuanced relationships that shape educational outcomes. As we delve into this dataset, we embark on a journey to uncover the multifaceted dimensions of student achievement and the intricate web of factors that contribute to their scholastic journey.

Students Exam Scores Data Analysis

About Dataset

This dataset includes scores from three test scores of students at a (fictional) public school and a variety of personal and socio-economic factors that may have interaction effects upon them.

Data Dictionary (column description)

1. Gender: Gender of the student (male/female)
2. EthnicGroup: Ethnic group of the student (group A to E)

14. WritingScore: writing test score(0-100)

In [12]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

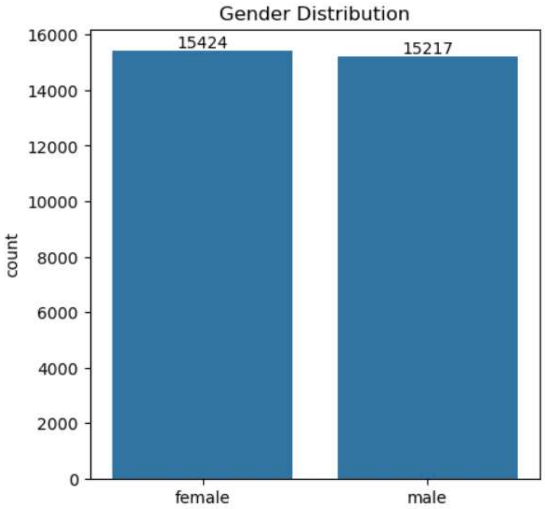
In [13]:
```
df = pd.read_csv("student_scores.csv")
```

In [14]:
```
print(df.head())
```

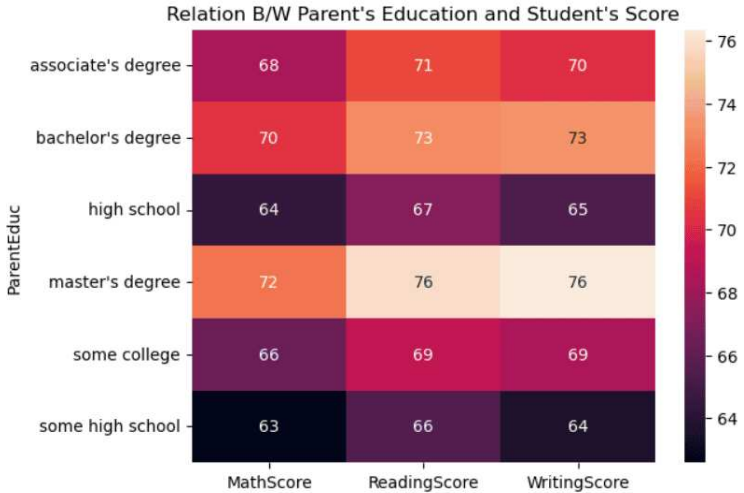
```
   Unnamed: 0  Gender EthnicGroup          ParentEduc     LunchType TestPrep  \
0           0  female         NaN   bachelor's degree      standard     none
1           1  female     group C        some college      standard      NaN
2           2  female     group B     master's degree      standard     none
3           3    male     group A  associate's degree  free/reduced     none
4           4    male     group C        some college      standard     none

  ParentMaritalStatus PracticeSport IsFirstChild  NrSiblings TransportMeans  \
0             married     regularly          yes         3.0     school_bus
1             married     sometimes          yes         0.0            NaN
2              single     sometimes          yes         4.0     school_bus
3             married         never           no         1.0            NaN
4             married     sometimes          yes         0.0     school_bus

  WklyStudyHours  MathScore  ReadingScore  WritingScore
0            < 5         71            71            74
1        5 - 10         69            90            88
2            < 5         87            93            91
3        5 - 10         45            56            42
4        5 - 10         76            78            75
```

Gender Distribution

In [36]:
```
plt.figure(figsize=(5,5))
ax = sns.countplot(data = df, x = "Gender" )
ax.bar_label(ax.containers[0])
plt.title("Gender Distribution")
plt.show()
```

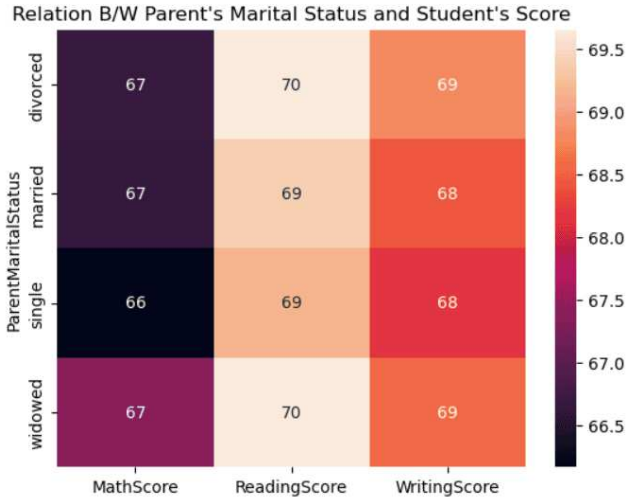


In [37]:
```
sns.heatmap(gb, annot = True)
plt.title("Relation B/W Parent's Education and Student's Score")
plt.show()
```



from the above chart we have concluded that the education of the parents have the good impact on their scores.

In [38]:
```
sns.heatmap(gb1, annot = True)
plt.title("Relation B/W Parent's Marital Status and Student's Score")
plt.show()
```

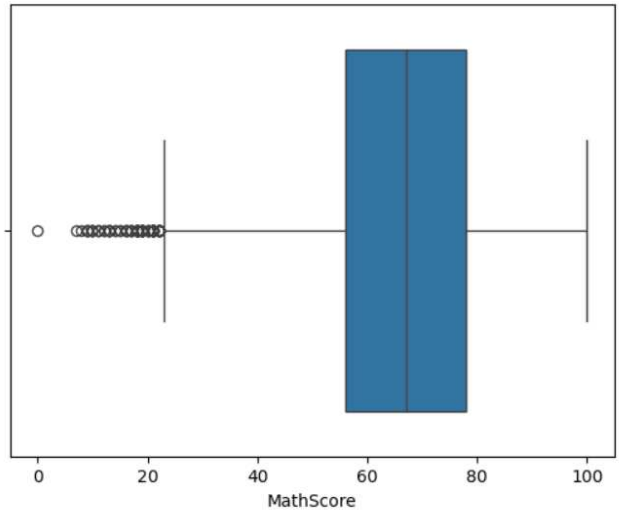


from the above chart we have concluded that there is no/negligible impact on the student scores due to their parent's marital status.

MathScore ReadingScore WritingScore

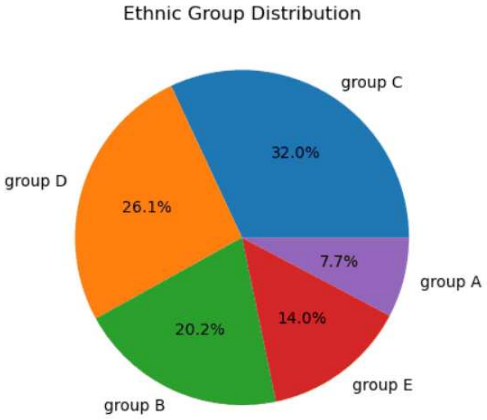
from the above chart we have concluded that there is no/negligible impact on the student scores due to their parent's marital status.

In [40]:
```
sns.boxplot(data = df, x = "MathScore")
plt.show()
```

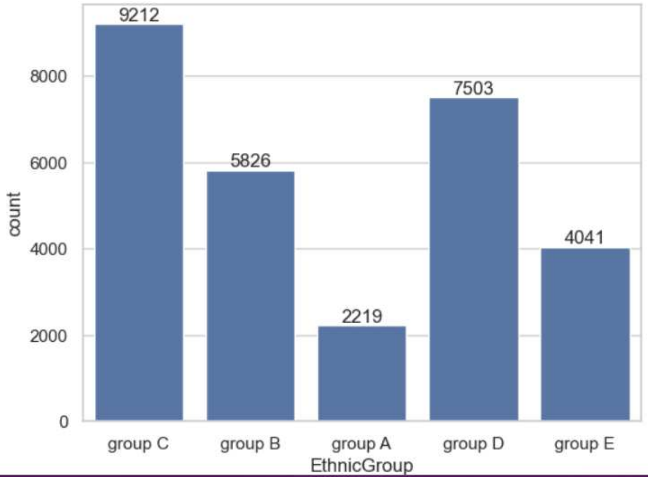


Distribution Of Ethnic Groups

In [49]:
```
# Assuming 'EthnicGroup' is a column in your DataFrame
ethnic_counts = df['EthnicGroup'].value_counts()

labels = ethnic_counts.index
values = ethnic_counts.values

plt.pie(values, labels=labels, autopct='%1.1f%%')
plt.title('Ethnic Group Distribution')
plt.show()
```



```
sns.countplot(data=df, x='EthnicGroup')

# Get the current Axes object
ax = plt.gca()

# Add labels to the bars
for container in ax.containers:
    ax.bar_label(container)

plt.show()
```



Data Analysis

- The analysis suggests a potential correlation between parents' education and students' academic performance.
- Gender distribution indicates a higher representation of females in the dataset.
- Parental marital status seems to have a minimal impact on students' scores.
- Box plots provide insights into the spread and central tendency of test scores.
- Ethnic group distribution is visualized to understand the composition of the dataset.

Conclusion

while factors like parental education showcase a strong association with academic achievement, other socio-economic factors within the dataset might necessitate further investigation to understand their impact comprehensively on students' test scores. The dataset offers valuable insights into potential areas for educational intervention and tailored support for students based on their backgrounds and circumstances.

Thanks!

Do you have any questions?
mk1423141@gmail.com
+917587805856
https://manishbelsare.github.io/Manish.Kumar.belsare.github.io/

www.linkedin.com/in/manishkumarbelsare