

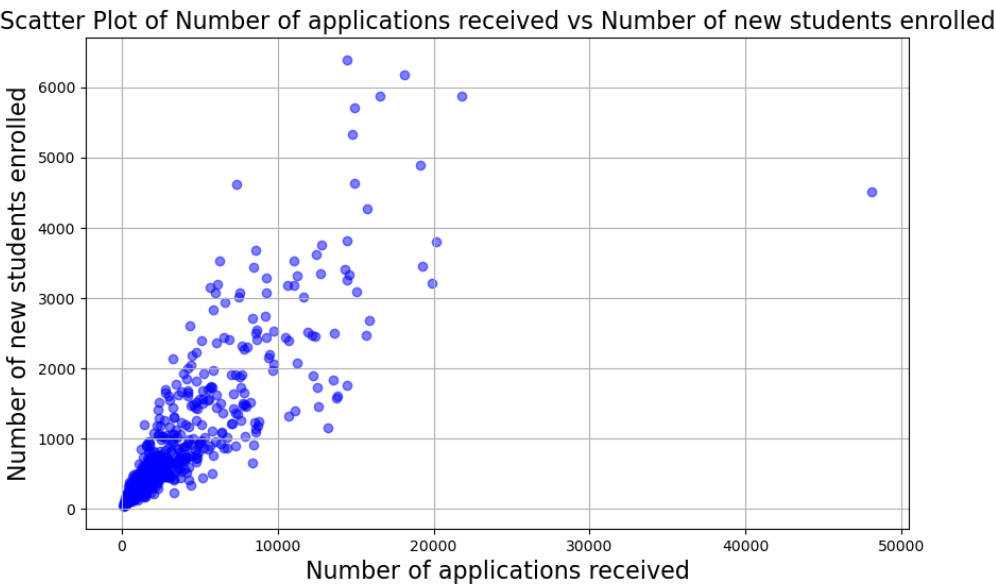
Name: Md Salahuddin Chowdhury, I'd: 23029856, Github: <https://github.com/sajal329/Data-Clustering-and-Fitting>

Dataset Description and Report Overview:

This report analyzes the "clustering_assignment_data.csv" (Tarek Muhammed, 2024) dataset, which contains information on various metrics and characteristics of universities. The dataset includes variables such as the number of applications received, number of applications accepted, new student enrollment figures, tuition fees, faculty qualifications, graduation rates, and indicators of whether the university is private or public. The report aims to explore the relationships between these variables, uncover patterns, and gain insights into factors influencing university admissions and academic performance.

Scatter Plot Analysis:

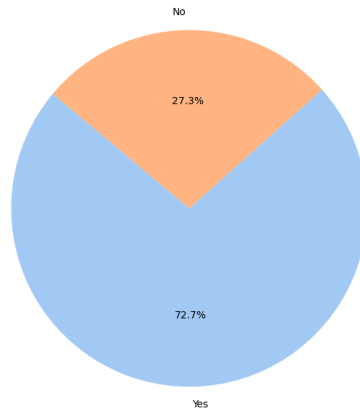
The scatter plot depicts a positive correlation between the number of applications received and the number of new students enrolled across universities. However, there is considerable variation in the data, suggesting the influence of other factors beyond just application numbers on enrollment figures. The plot shows a dense cluster of institutions with lower application volumes and smaller enrollment, while institutions with higher application numbers exhibit a wider range of enrollment outcomes.



Pie Chart Analysis:

The Pie Chart of Private vs Non-Private Universities presents the proportion of universities that are private versus non-private in the dataset. With blue representing private universities and orange representing non-private universities, the chart provides a clear visual representation of the distribution of university types. The predominance of private universities, comprising 72.7% of the dataset, underscores the importance of considering private institutions' characteristics in the analysis and decision-making processes. This analysis highlights the need to account for the diversity of university types when interpreting findings and drawing conclusions from the dataset.

Pie Chart of Private Universities

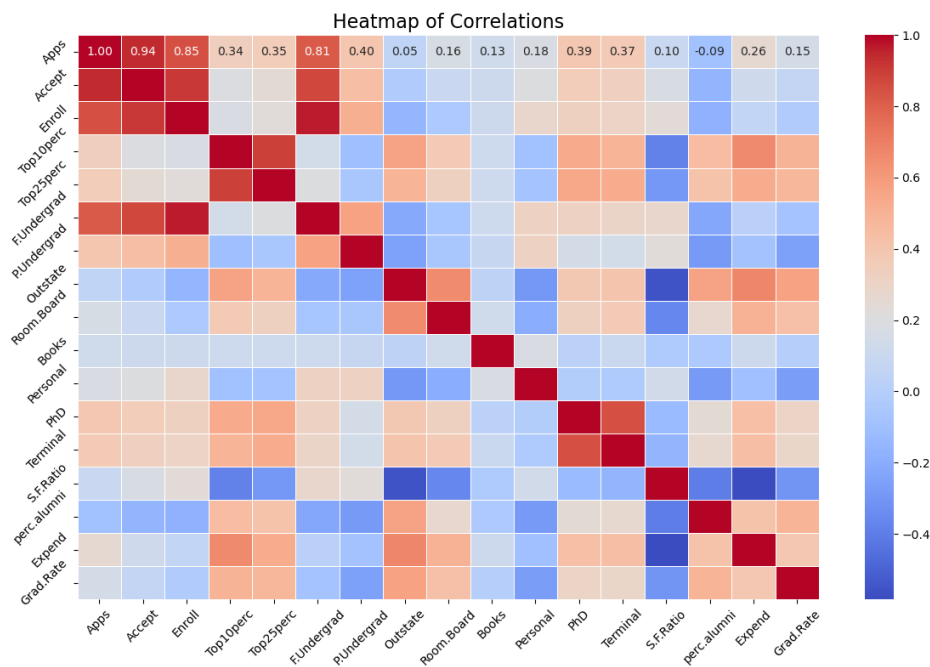


Correlation Analysis:

The correlation heatmap provides insights into the relationships between different variables in the dataset. A notable observation is the strong positive correlation between the number of applications received and the number of applications accepted, indicating that universities with higher application volumes tend to accept more applicants.

Additionally, there is a moderate negative correlation between the graduation rate and the percentage of alumni who donate, as well as the instructional expenditure per student. This suggests that universities with higher graduation rates may have lower alumni donation rates and potentially allocate fewer resources per student for instructional purposes.

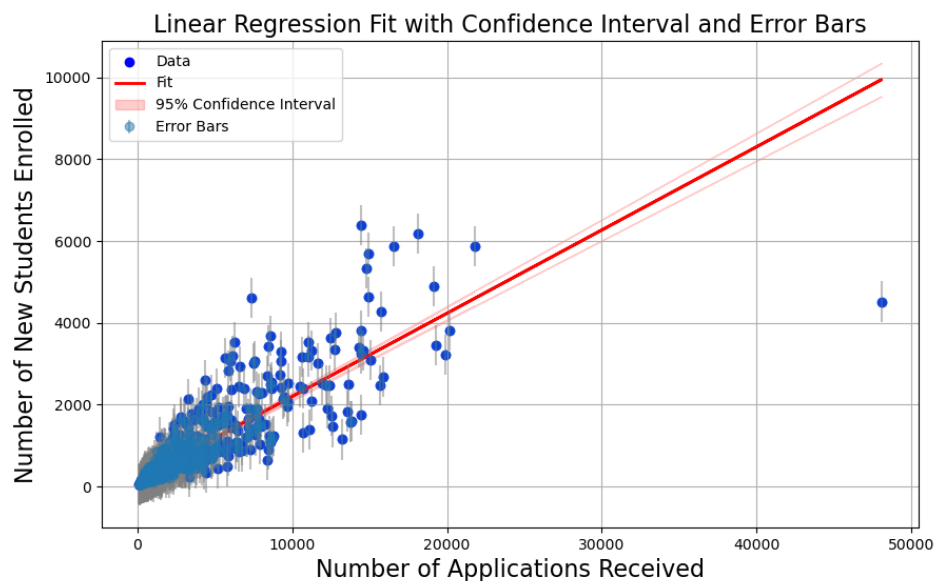
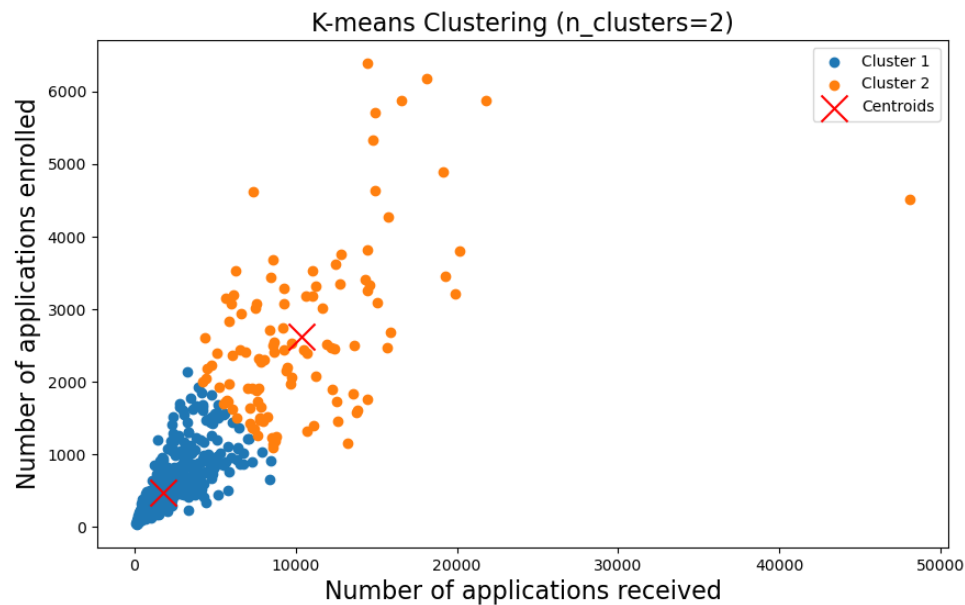
These correlations offer valuable insights into factors that may influence university admissions, academic performance, and resource allocation, providing a foundation for further analysis and decision-making in the higher education sector.



Clustering and Fitting Analysis:

The k-means clustering analysis with two clusters reveals one cluster comprising institutions with lower application volumes and correspondingly lower enrollment numbers, suggesting similar characteristics within this group. The second cluster consists of institutions with higher application volumes and a wider range of enrollment figures,

indicating greater variability in their ability to convert applications into enrollments. The linear regression fit shows a positive trend, where an increase in applications received corresponds to an increase in new students enrolled. However, the scattered data points around the fitted line suggest that additional factors beyond application numbers influence enrollment outcomes.



Conclusion:

This analysis provides valuable insights into university metrics and characteristics, guiding strategic decision-making in higher education. By examining relationships between variables, distributions, and potential groupings, stakeholders can gain a comprehensive understanding of factors influencing university admissions and academic performance. These insights contribute to informed decision-making and policy formulation in higher education, enabling data-driven approaches to address challenges and opportunities in the sector.

References

Tarek Muhammed. (2024). University Students Data. Kaggle. Available at:
<https://www.kaggle.com/tarekmuhammed/university-students-data>