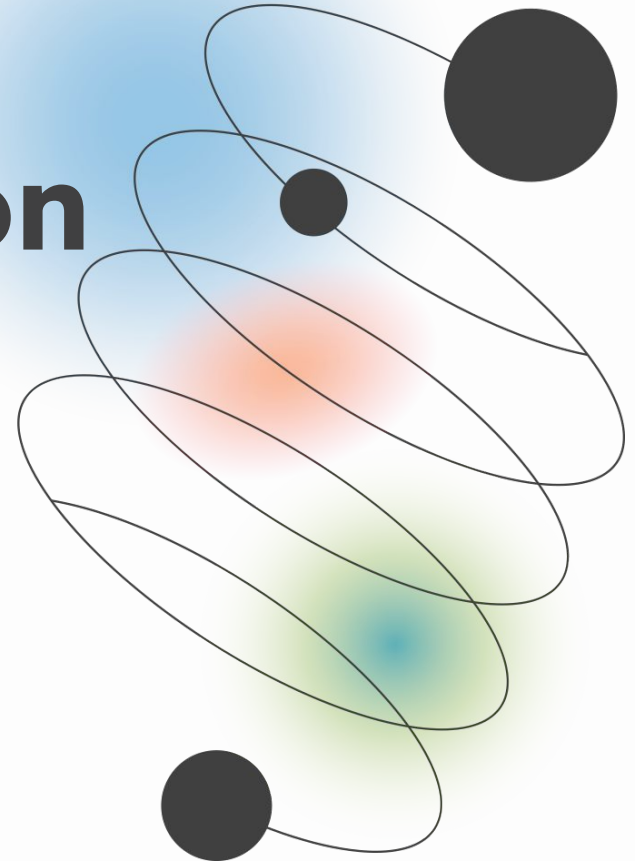# DA332 : Data Visualization

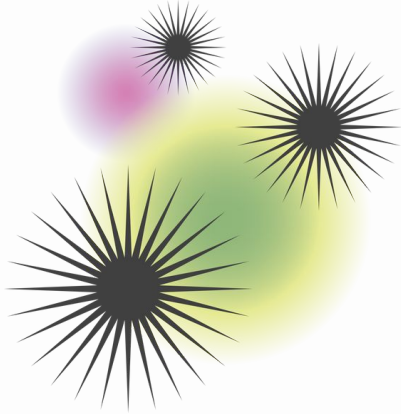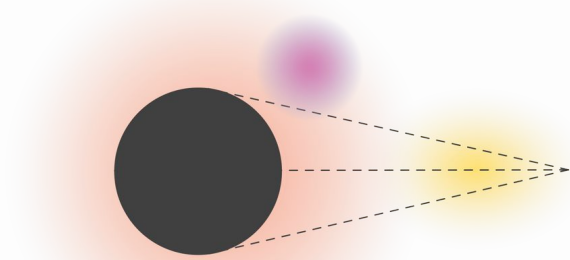**Topic** : <u>Education - A Brief Analysis Of Enrollment Statistics</u>

Presented By :

Palthiya Laanith Chouhan

210150014

Motivation :

This project is fueled by the drive to analyze the complex patterns of student enrollment through insightful data visualization. By delving into class-wise and district-wise enrollment data spanning multiple years, it aims to uncover meaningful insights that inform educational policy and practice.

Datasets

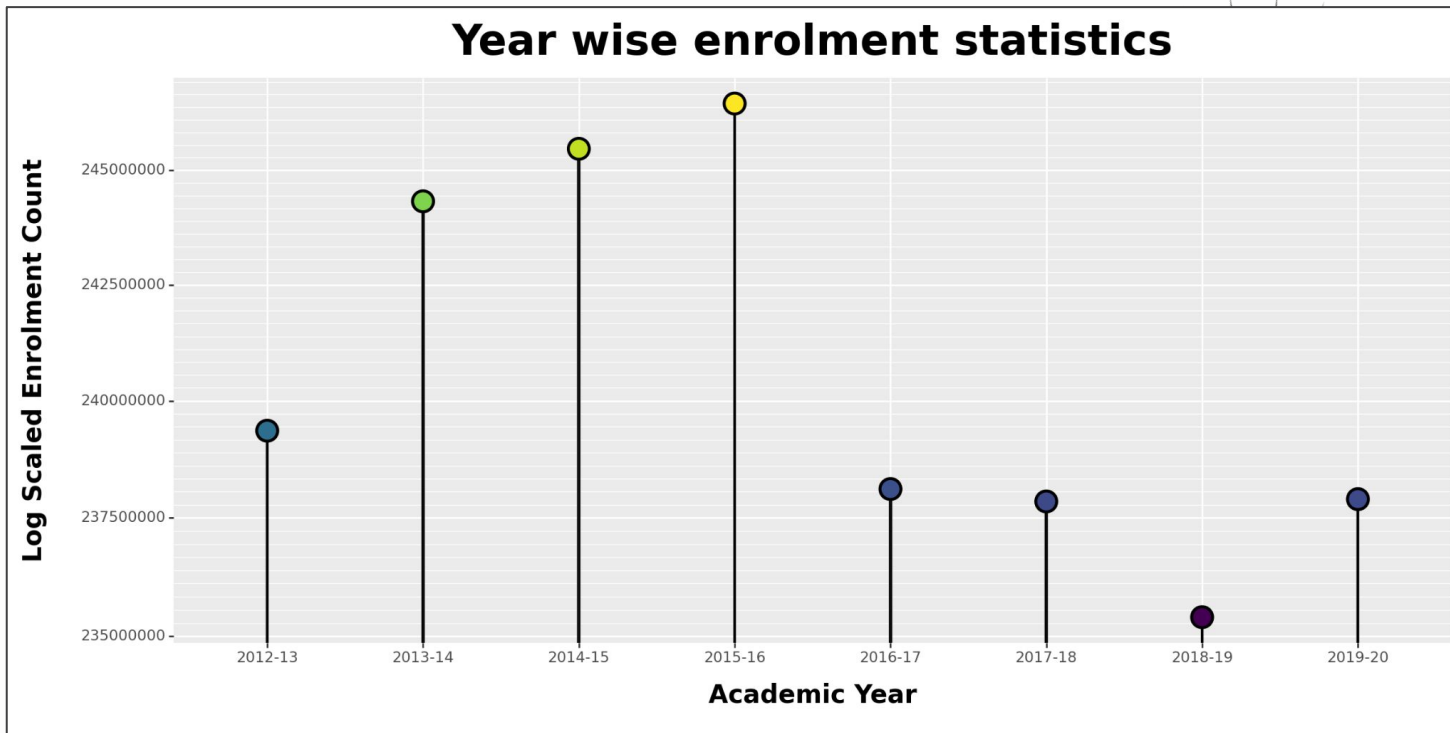The dataset is found on data.gov.in.
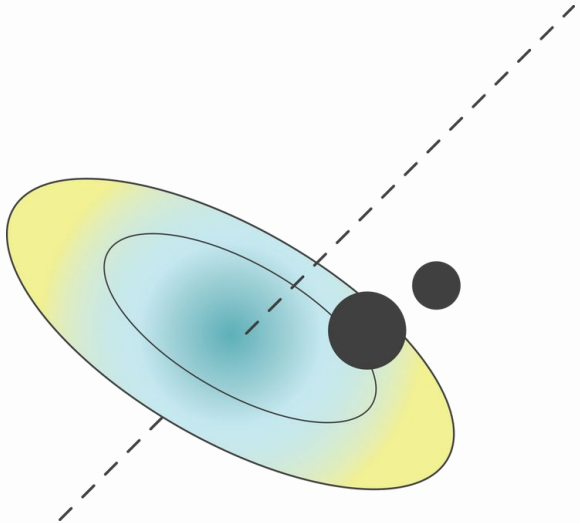
A set of 8 CSV (Comma Separated Value) files that show district wise and state wise, class and gender-wise enrolments from the year 2012-2020

A short demographic overview :

- 24 columns of names class_<class_no>_<gender>
- Each of those columns spanning over roughly 780 districts.

Year Wise Enrolment Statistics


Year wise enrolment statistics

- Utilizing a categorical chart type is ideal for analyzing categorical data, leading us to employ custom made Lollipop chart using Plotnine Library.
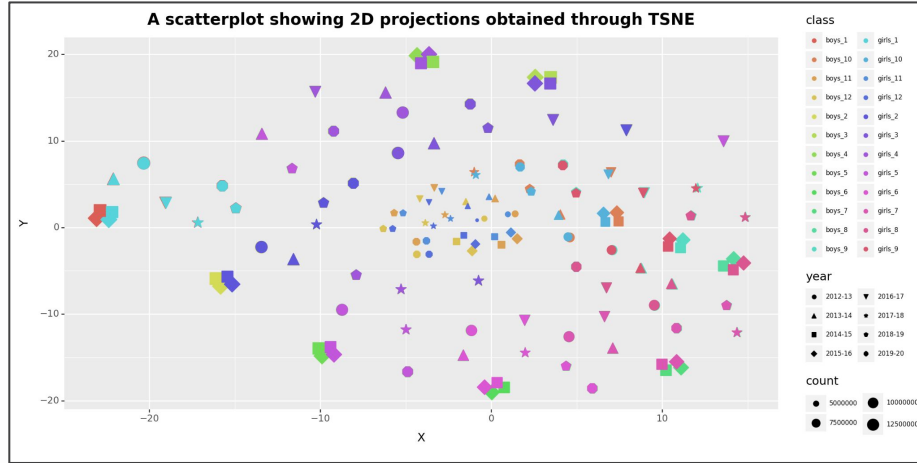
Observing the Distribution of Data in greater detail

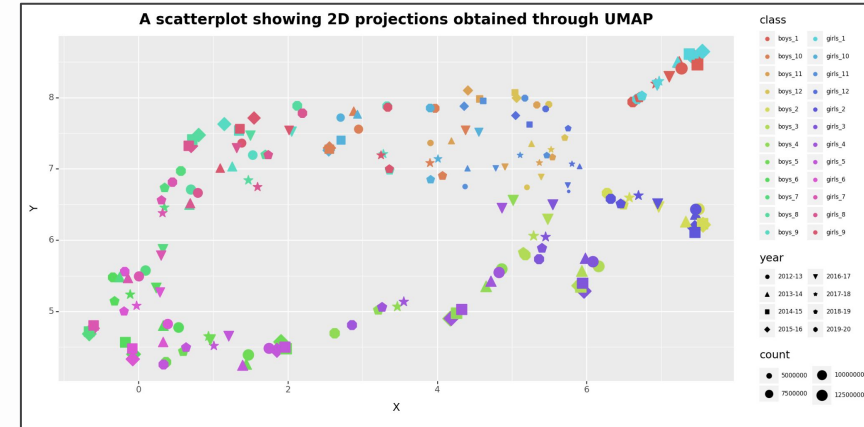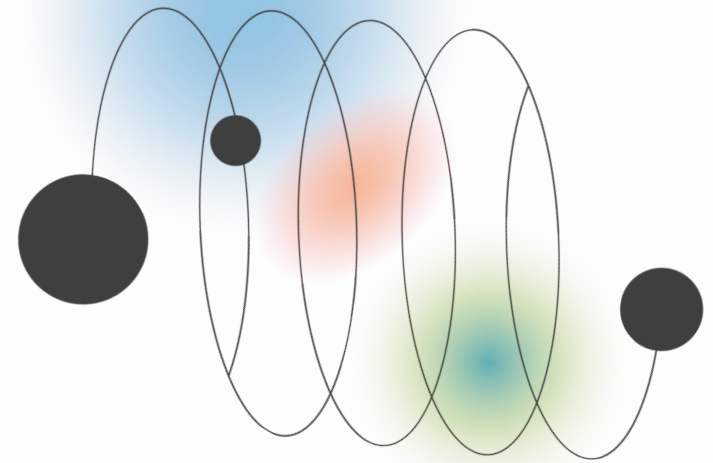A heatmap showing distribution of various classes across various years

2D Projections

A scatterplot showing 2D projections obtained through TSNE
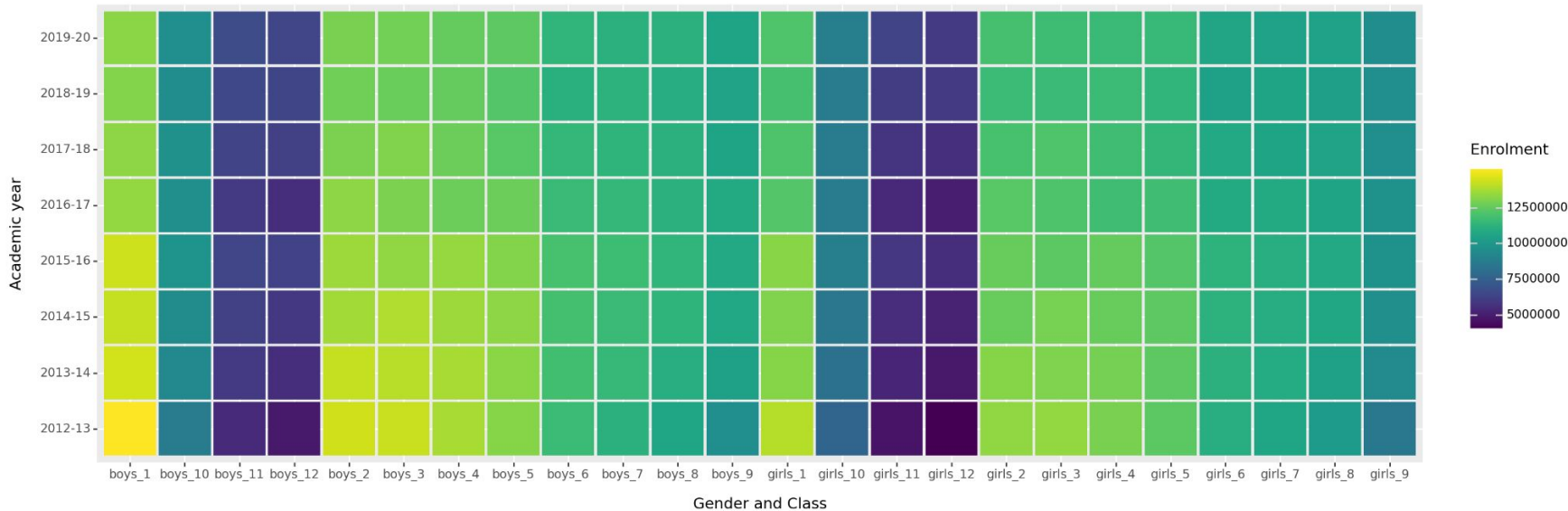
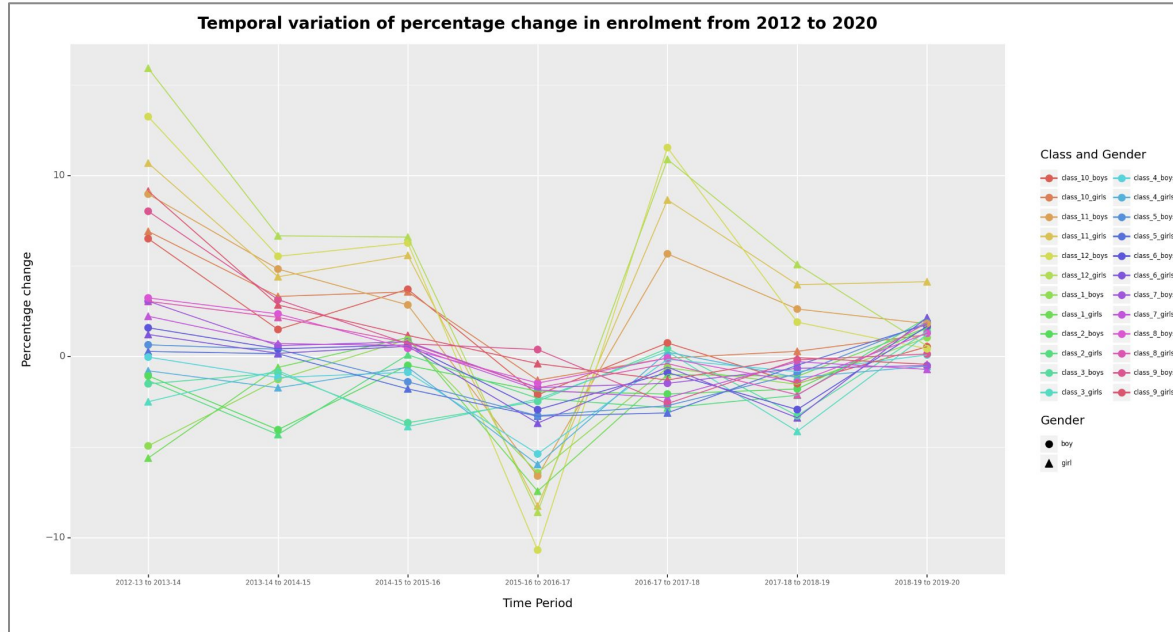Points of same shape and adjacent colors in legend are either overlapping or in proximity to each other.They are also of same size.
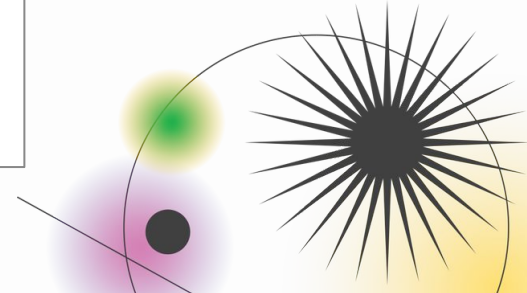This indicates that the enrolment distribution of boys and girls across the states and districts is quite similar, and this backs up the previous plot shown by the heatmap.

A scatterplot showing 2D projections obtained through UMAP

A Temporal Anomaly

Temporal variation of percentage change in enrolment from 2012 to 2020

Observing an intriguing anomaly in the plot, there's a noticeable decline in enrollment rates for several classes, including class-12 boys, class-12 girls, and class-11 girls. Further investigation revealed a correlation with the highest average annual dropout rate for secondary schools in India, which peaked at 17.06 percent in 2016. While this sheds light on the enrollment decline that year, the underlying causes for the surge in dropout rates remain a mystery.

A look into the Outliers

A unique usage of jitterplot to display outliers in class distribution in the academic year 2019-20
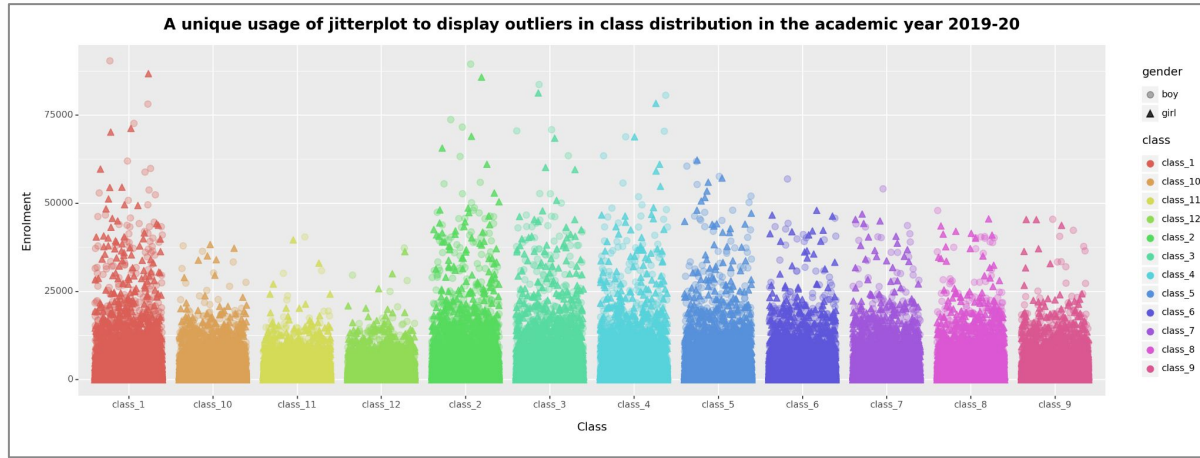
Analyzing district-wise data distribution poses challenges, especially with many districts having zero class enrollment, significantly impacting the dataset. Traditional methods like box plots and violin plots exhibit bias towards these districts, resulting in skewed distributions. To mitigate this bias, a jitter plot is utilized, assigning distinct shapes and opacity to distinguish between boys and girls.

Insights :
Insights reveal that maximum enrollment in classes 11th and 12th does not surpass approximately 37,500, with enrollment declining as students progress to higher grades.

# Percentage contribution of Each Gender

<u>Insights :</u>

The ordered stacked bar chart for 2019-20 enrollment across Indian states reveals a consistent gender disparity, with male enrollment consistently higher than female enrollment. Uttar Pradesh exhibits the highest gender gap at 52.31% male enrollment, while West Bengal shows the lowest gap at 50.86% female enrollment. Overall, Uttar Pradesh records the highest total enrollment at 45,362,059, while Andhra Pradesh reports the lowest at 8,136,933.
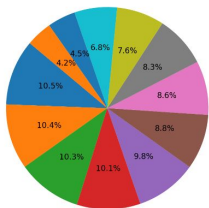

Ordered - stacked bar chart showing percentage of male and femake enrollment in 2019-20
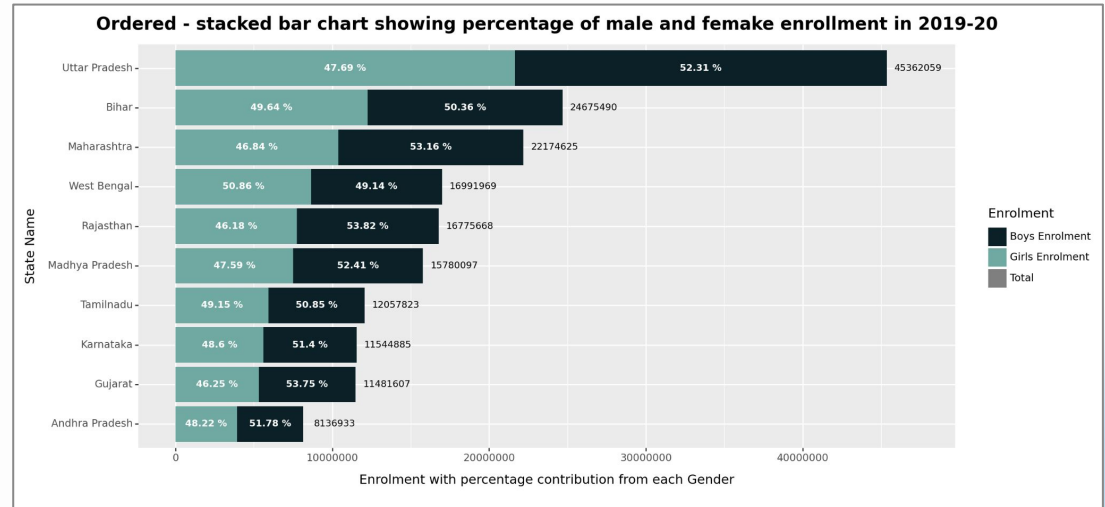
A view of Data through Pie Charts

Percentage distribution over the years among classes
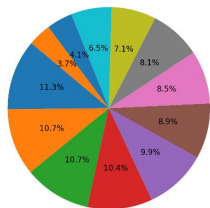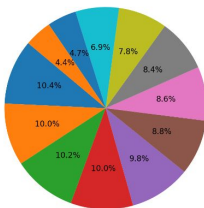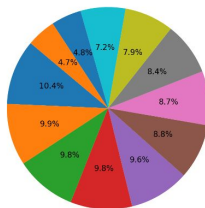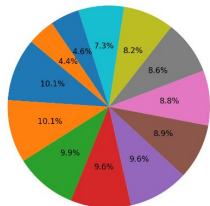
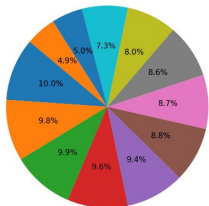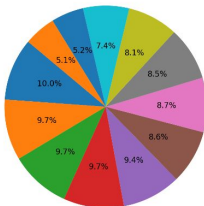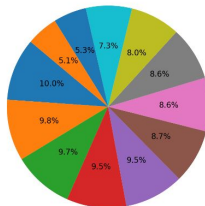- The pie charts depict percentage enrollment distribution across different classes in India from 2012-13 to 2019-20.
- Class 1 consistently shows the highest enrollment percentage, indicating a focus on primary education.
- Enrollment percentages gradually decline as students progress to higher grades, with classes 6 to 12 exhibiting lower percentages overall.
- Class 12 consistently records one of the lowest percentages, suggesting significant dropout rates at the higher secondary level.
- Despite minor fluctuations, distribution patterns remain relatively stable, highlighting persistent challenges in enrollment and dropout rates that require targeted interventions for improved educational outcomes.

Thank you for attending!