



# Sardar Patel Institute of Technology

(Autonomous Institute Affiliated to University of Mumbai)  
Bhavan's Campus, Munshi Nagar, Andheri (West), Mumbai-400058-India

## EXPERIMENT 10

NAME:	Sarthak Gharat
ROLL NO:	2021700025
BATCH:	CSE(DS) - D

### Aim:

Design Big Data Dashboards using Tableau on the dataset - Women empowerment / Gender participation

- Basic - Bar chart, Pie chart, Histogram, Time line chart, Scatter plot, Bubble plot
- Advanced - Word chart, Box and whisker plot, Violin plot, Regression plot (linear and nonlinear), 3D chart, Jitter
- Write observations from each chart

### DATASET:

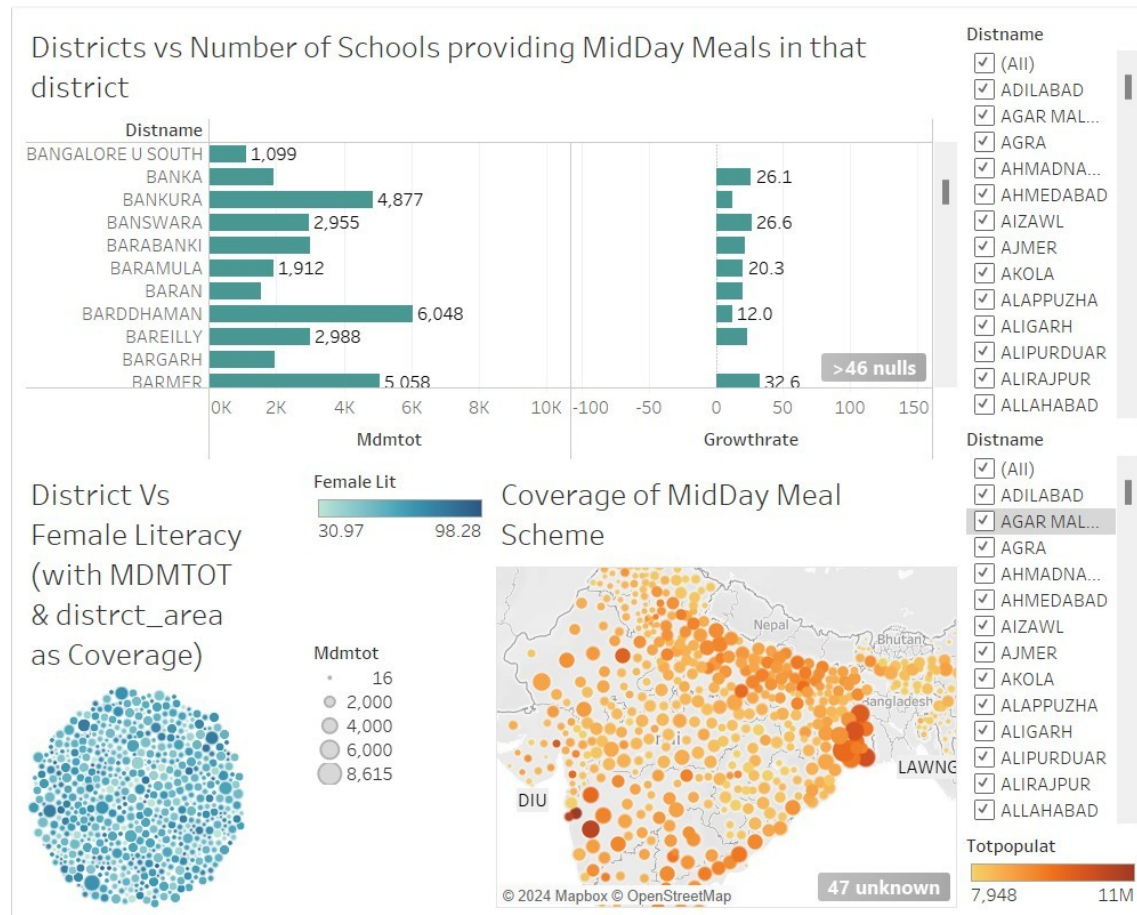
[https://www.kaggle.com/datasets/rajanand/education-in-india?select=2015\\_16\\_Statewise\\_Secondary.csv](https://www.kaggle.com/datasets/rajanand/education-in-india?select=2015_16_Statewise_Secondary.csv)

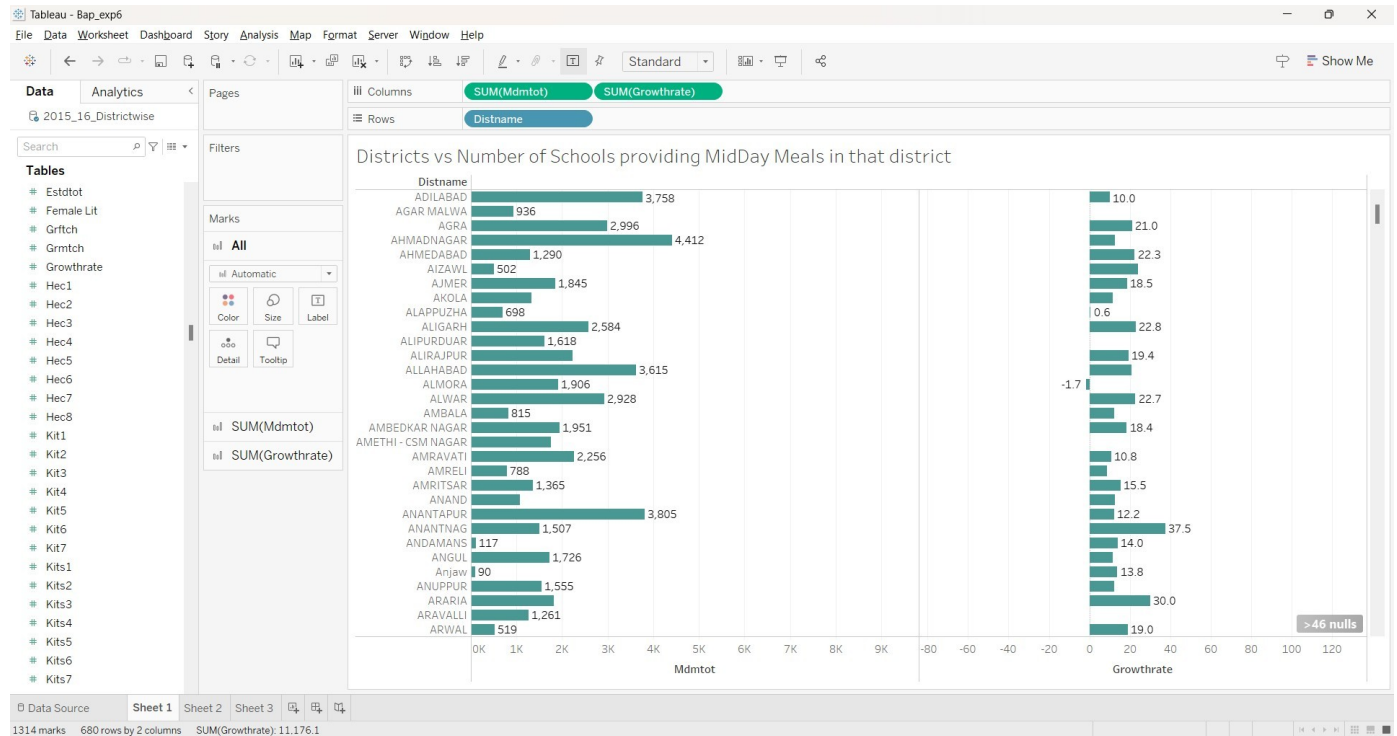
- **ac\_year:** Represents the academic year, with a single unique value of "2015-16" across all records, indicating the data is specific to that year.
- **statcd:** A unique code for each state or region, ranging from 1 to 36, likely used to uniquely identify each geographic entity.
- **statname:** The name of each state or region, with 36 unique names, showing that data is provided for each state in this dataset.
- **area\_sqkm:** The area in square kilometers for each region, with values ranging from 8,249 to 342,239 sq km. The average area is approximately 93,900 sq km, with a large variation, suggesting a mix of small and large states.
- **tot\_population:** Total population of each state, with figures ranging from 1.38k to 200k (likely in thousands). The average population is around 34.6k, again showing a wide variation.
- **urban\_population:** Urban population of each state, ranging from 8.36 to 77 (presumably in thousands or as a percentage). The mean urban population is 26.8, indicating some states have significantly higher urban populations.
- **growth\_rate:** Annual growth rate, with values between -0.47 and 55.5, indicating both positive and negative growth. The mean growth rate is 19%, with a high standard deviation, showing variability across states.
- **sexratio:** Sex ratio per 1,000 males, ranging from 618 to 1,084. The mean ratio is 929, with a standard deviation of 78.5, suggesting most states have a relatively balanced gender ratio, with a few outliers.
- **sc\_population:** Proportion of the Scheduled Caste (SC) population in each state, with values from 0 to 31.9.

The mean is 11.8, indicating SC populations vary considerably between states.

- **st\_population:** (Likely) Proportion of the Scheduled Tribe (ST) population, similar to the SC column. The specifics are not provided, but it would likely follow a similar pattern with variations by state.

### Dashboard: *Positive effects of implementing Mid Day Meal Scheme*

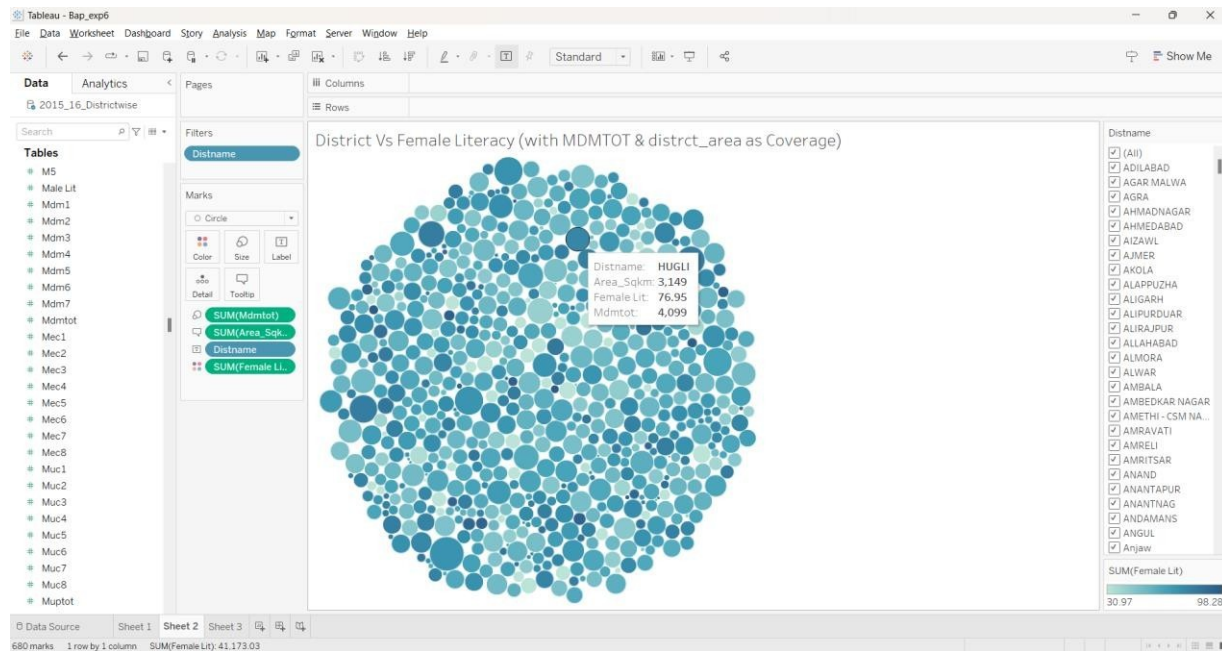




## 1. Horizontal Bar Graphs:

- District-wise comparison of mdmtot (midday meal scheme total) and growth rate.
- Horizontal bar graph shows mdmtot values for each district.
- Length of bars represents the growth rate, indicating rate of change in mdmtot over time.
- Helps identify districts with higher mdmtot and significant growth rates.

This Tableau plot displays a comparison of districts based on two key metrics: mdmtot (midday meal scheme total) and growth rate. Districts with a higher number of schools offering the midday meal scheme show an average growth rate of around 15%, which is significantly higher the districts with fewer schools providing the scheme. This indicates a positive correlation between the implementation of the midday meal scheme and the overall growth rate in education within these districts.



## 2. Packed Bubbles:

Packed bubbles are a data visualization method that represents hierarchical or grouped data using circles (bubbles) of varying sizes. Key points include:

- **Quantitative Representation:** Bubble size indicates values, allowing quick comparison.
- **Grouping and Clustering:** Bubbles are grouped to show relationships or categories.
- **Color and Additional Dimensions:** Colors within bubbles add more information.
- **Interactive Features:** Enables drill-down and exploration of specific data points.
- **Insightful Visuals:** Visually engaging, effective for hierarchies and comparisons.

This packed bubble plot includes variables like district name, female literacy percentage, area of dist (in sqkms) and total no. of schools offering mid day meal service in the district. The size is represented by total number of schools offering mid day meal service, the greater the number for a particular district, the greater is the size of the bubble for respective district. The color of this districts is influenced by the female literacy rate (darker the shade, higher is the literacy %age), We can generally see that bigger bubbles have darker shade colors, indicating a positive correlation between the implementation of the midday meal scheme and the female literacy rate in education within these districts

## 3. MAP Plot:

A map plot visually represents geographical data:

- **Geographical Representation:** Displays locations on a map based on their coordinates.
- **Color and Size Encoding:** Uses color and size to convey additional data attributes.
- **Interactive Features:** Allows zooming, panning, and tooltips for detailed exploration.

- **Layering and Overlays:** Includes boundaries, labels, and other overlays for context.
- **Data Analysis:** Enables spatial analysis and identification of geographical patterns.
- **Communication:** Effectively communicates geographical insights to a broad audience.

#### **Analysis:**

- Districts with a higher total population tend to have more schools implementing midday meal schemes. This correlation suggests that areas with larger populations may have a greater need for such programs to support the nutritional requirements of students.
- Areas with higher population densities and more midday meal schemes may experience positive outcomes in terms of improved attendance, academic performance, and overall health of students.
- The correlation between population, midday meal schemes, and district names can provide insights into socio-economic factors influencing the implementation and success of such programs. Districts with lower populations but higher mdmtot may indicate targeted efforts in specific communities or socio-economic backgrounds.