

SmartInternz

<https://www.thesmartbridge.com/>

in association with

Tableau

<https://www.tableau.com/>

DataViz Challenge 2021

<https://smartinternz.com/tableau-dataviz-challenge-2021>

Workshop Date: 14-April-2021 to 18-April-2021

Challenge Date: 19-April-2021 to 30-April-2021

Project:

Data Visualization Using Tableau

By

Chakola Paul Johny

Email: paul.chakola@vidyaacademy.ac.in

Telephone: +91 949 675 4571

Vidya Academy of Science and Technology

AICTE Approved & Affiliated to the APJAK Technological University

Thalakkottukara PO, Thrissur - 680501, Kerala, India

Website: www.vidyaacademy.ac.in

Telephone: +91 (4885) 287751 / 287752

Contents

1. Background.	3
2. Introduction.	3
a) What is data visualization?.	3
b) What is Tableau?.	3
c) Overview..	3
d) Project	4
e) Purpose.	4
3. Literature Survey.	4
a) Existing problem..	4
b) Proposed solution.	4
4. Theoretical Analysis.	5
a) Tableau Sheets.	5
b) Tableau Dashboard / Story.	5
5. Experimental Investigation.	5
6. Flowchart	5
7. Result	6
8. Advantages and Disadvantages.	6
9. Applications.	6
10. Conclusion.	10
11. Future Scope.	10
12. Bibliography.	10

1. Background

This project is being done as part of the DataViz challenge 2021 organized by SmartInternz (Smart Bridge) in collaboration with Tableau. The hand done sessions were held online between 14-Apr-2021 and 18-Apr-2021. The project challenge for the participants was open between 19-Apr-2021 and 30-Apr-2021. For more details, please refer to:

<https://smartinternz.com/tableau-dataviz-challenge-2021>

2. Introduction

a. What is data visualization?

Data visualization is the graphical representation of information and data. By using visual elements like charts, graphs, and maps, data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data. In the world of Big Data, data visualization tools and technologies are essential to analyze massive amounts of information and make data-driven decisions. Our eyes are drawn to colors and patterns. We can quickly identify red from blue, square from circle. Our culture is visual, including everything from art and advertisements to TV and movies.

Data visualization is another form of visual art that grabs our interest and keeps our eyes on the message. When we see a chart, we quickly see trends and outliers. If we can see something, we internalize it quickly. It's storytelling with a purpose.

<https://www.tableau.com/learn/articles/data-visualization>

b. What is Tableau?

Tableau is a visual analytics platform transforming the way we use data to solve problems –

empowering people and organizations to make the most of their data. Tableau helps organization to be data driven.

<https://www.tableau.com/why-tableau/what-is-tableau>

C. Overview

The challenge entails developing and publishing a data visualization dashboard or story using a dataset from among the themes published on the site using data sourced from data.gov.in. Viz.

<https://smartinternz.com/tableau-dataviz-challenge-2021>

d. Project

The project entails visualizing the data provided at <http://data.gov.in> under the following head:

Released Under	: National Data Sharing and Accessibility Policy (NDSAP)
Contributor	: Ministry of Jal Shakti Department of Drinking Water and Sanitation
Granularity	: Annual
File Size	: 232.2 MB
NRDWP	: National Rural Drinking Water Programme.
Source	: Data taken from NRDWP website (http://indiawater.gov.in)
Data content	: The data refers to the list of habitations, its population in different caste category (SC, ST and GENERAL) and status of availability of potable drinking water (Covered or Partially covered) all over India.

Note : The data on the NRDWP website has been updated by respective State Governments & Ministry of Drinking Water and Sanitation, Government of India.

e. Purpose

The purpose of the project is to interpret the data and present it in a visual manner to convey the essence of the collated data to the users at a glance.

Tools to be used : Tableau Desktop

Display : Tableau Public

Project documents : Github

Data source :

https://data.gov.in/catalog/basic-habitation-information?filters%5Bfield_catalog_reference%5D=86138&format=json&offset=0&limit=6&sort%5Bcreated%5D=desc

Deliverables : Tableau dashboard, project document, a 4-minute video

The project data interpretation to be done visually using Tableau sheets which may be clubbed into a Tableau dashboard and published at Tableau Public website.

3. Literature Survey

a. Existing problem

Tools like Excel and other spreadsheets are often used for visual depiction of data. But these are basically spreadsheets and thus, there is need for a visualization tool. The questions to be answered include – does the tool support features for data discovery, data automation, data visualization and which tool best fits my organization's purpose.

b. Proposed solution

The proposed solution is to build a visual model using Tableau. Tableau is a visual analytics engine that makes it easier to create interactive visual analytics in the form of dashboards. These dashboards make it easier for non-technical analysts and end users to convert data into understandable, interactive graphics.

4. Theoretical Analysis

a. Tableau Sheets

Tableau uses a workbook and sheet file structure, much like Microsoft Excel. A workbook contains sheets. A sheet can be a worksheet, a dashboard, or a story. A worksheet contains a single view along with shelves, cards, legends, and the Data and Analytics panes in its side bar.

b. Tableau Dashboard / Story

A dashboard is a consolidated display of many worksheets and related information in a single place. It is used to compare and monitor a variety of data simultaneously. The different data views are displayed all at once.

Tableau story is a sequence of visualizations that work together to convey information. You can create stories to tell a data narrative, provide context, demonstrate how decisions relate to outcomes, or to simply make a compelling case. ... Each individual sheet in a story is called a story point.

5. Experimental Investigation

The data provided as part of the basic habitation scheme is the coverage of potable water among the various states at the habitation level. More specifically the data is collated at state, district, village, to habitat level for the categories SC, ST, and General.

Notes:

- Data is as of 01-Apr-2021
- State of Telangana was formed after this date so the data cannot be presently mapped for this and other states and union territories formed after this date.
- Over ten lakh records

- Use Tableau Desktop for creating sheets and dashboards
- Use Tableau Public to publish final dashboard

6. Flowchart

1. Prepare the basic Tableau sheet
 - a. Country level view of state wise population using a Map
 - b. Textual representation of data
 - c. Bar chart view of states with most pending population for coverage
 - d. View of state wise / district wise status with filter
 - e. Country level summary map with color coding.
2. Prepare the Tableau dashboard
3. Upload the dashboard on Tableau Public
4. Additionally, a story could be created; another way to communicate.

7. Result

The end users experience is provided at Tableau Public website. Please visit:
The solution URL:

<https://public.tableau.com/profile/paul.chakola#!/vizhome/CPJ-DataViz-Challenge-WaterAndSanitation/WaterandSanitationDashboard>

8. Advantages and Disadvantages

Advantages:

- Tableau is very institutive for rapid visualization work.
- With Tableau you can create powerful dashboards.
- With Tableau you can tell a story of the problem at hand and provide solution vistas.
- Tableau is fast at processing data.
- Tableau has features to publish the data and provide role-based access

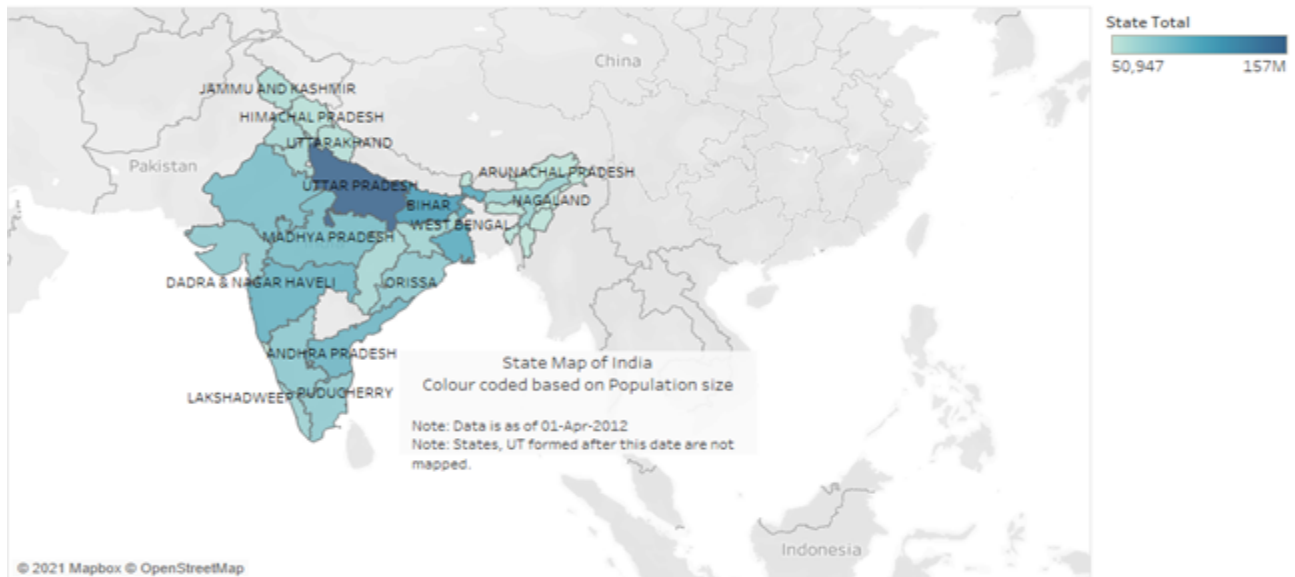
Disadvantages:

- Prior knowledge of how to use the tools.
- An understanding of what are dimensions and measures.

9. Applications

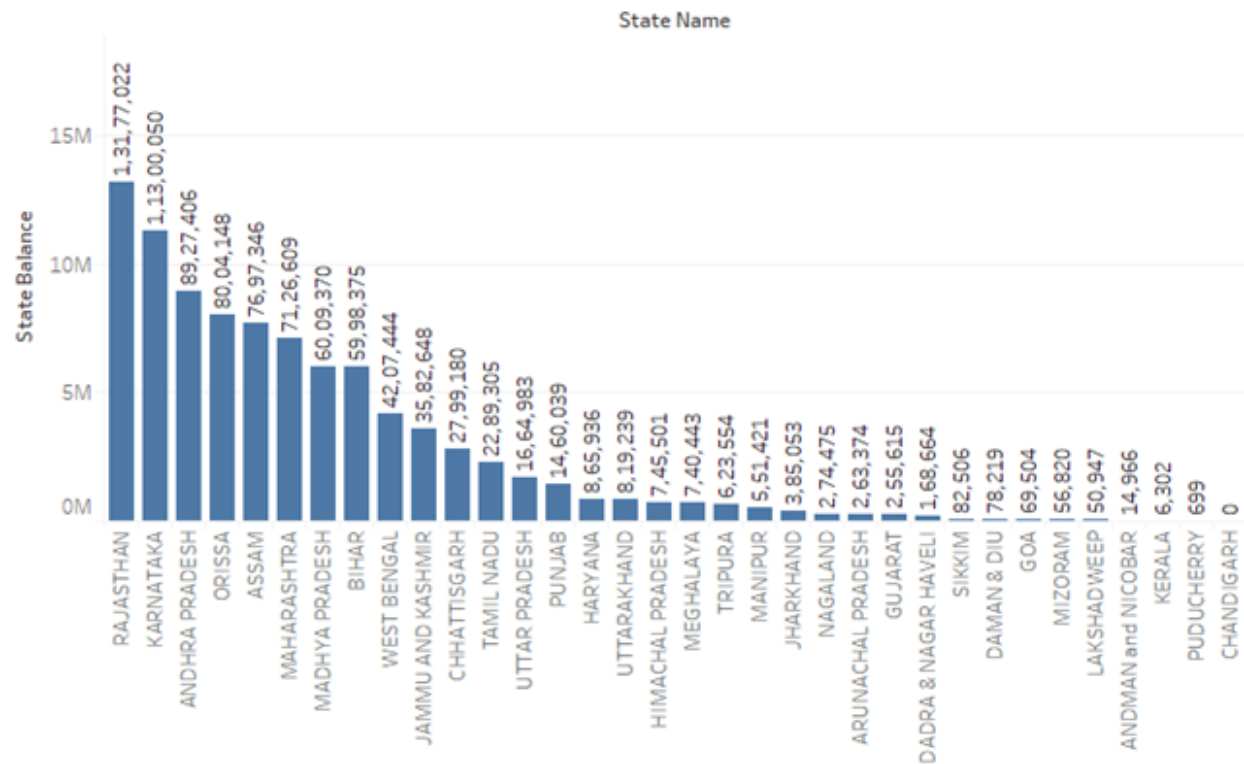
The concept can be made use of for building a data centric organization where views can be provided to users at various levels within the organization. Snap shot view of the sheets:

State Population Density



Map based on Longitude (generated) and Latitude (generated). Color shows sum of State Total. The marks are labeled by State Name. Details are shown for State Name.

State wise Balance Population to be Covered



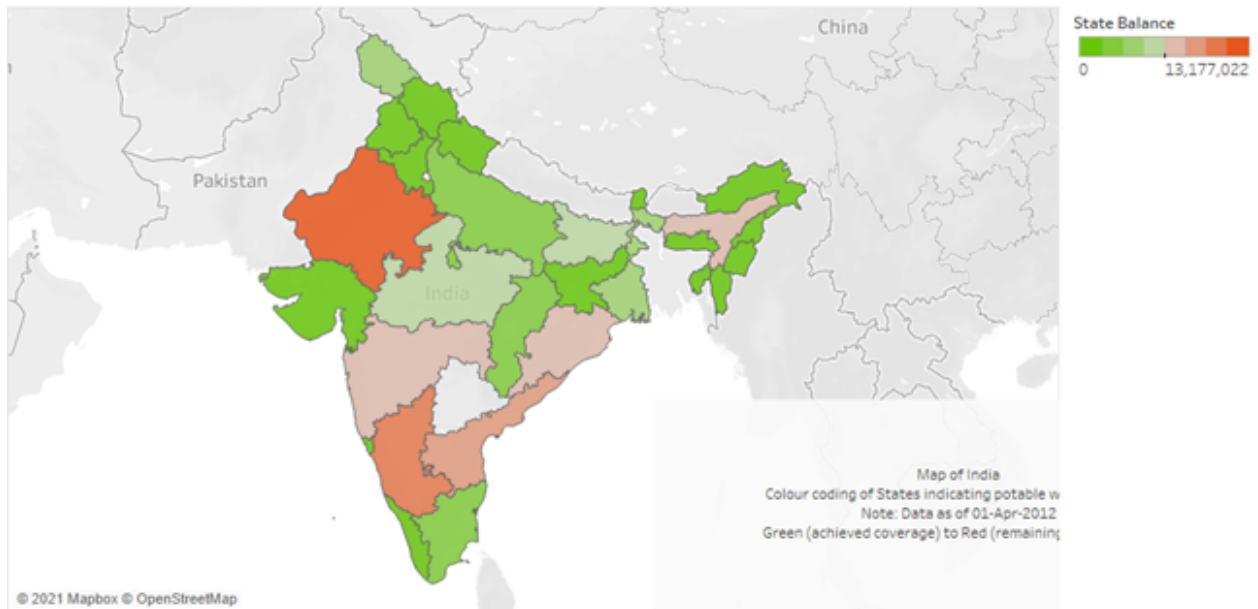
Sum of State Balance for each State Name. The marks are labeled by sum of State Balance.

State wise Water Distribution Coverage

State Name	State Total	State Covered	State Balance	State Covered Percentage
ANDHRA PRADESH	61,277,079	52,349,673	8,927,406	83
ANDMAN and NICOBAR	241,964	226,998	14,966	88
ARUNACHAL PRADESH	975,640	712,266	263,374	65
ASSAM	26,411,324	18,713,978	7,697,346	70
BIHAR	90,415,338	84,416,963	5,998,375	95
CHANDIGARH	81,397	81,397	0	100
CHHATTISGARH	18,350,304	15,551,124	2,799,180	85
DADRA & NAGAR HAVELI	168,664	0	168,664	0
DAMAN & DIU	78,219	0	78,219	0
GOA	754,931	685,427	69,504	92
GUJARAT	36,071,891	35,816,276	255,615	99
HARYANA	17,503,346	16,637,410	865,936	96
HIMACHAL PRADESH	6,228,454	5,482,953	745,501	86
JAMMU AND KASHMIR	10,089,910	6,507,262	3,582,648	59
JHARKHAND	22,965,457	22,580,404	385,053	97
KARNATAKA	38,351,278	27,051,228	11,300,050	70
KERALA	25,471,476	25,465,174	6,302	100
LAKSHADWEEP	50,947	0	50,947	0
MADHYA PRADESH	52,727,433	46,718,063	6,009,370	89
MAHARASHTRA	65,076,018	57,949,409	7,126,609	94
MANIPUR	2,478,377	1,926,956	551,421	73
MEGHALAYA	2,321,775	1,581,332	740,443	71
MIZORAM	522,543	465,723	56,820	96
NAGALAND	1,758,184	1,483,709	274,475	82
ORISSA	34,753,683	26,749,535	8,004,148	80
PUDUCHERRY	358,415	357,716	699	100
PUNJAB	18,056,763	16,596,724	1,460,039	87
RAJASTHAN	51,995,417	38,818,395	13,177,022	68
SIKKIM	540,848	458,342	82,506	84
TAMIL NADU	34,971,435	32,682,130	2,289,305	94
TRIPURA	2,812,940	2,189,386	623,554	72
UTTAR PRADESH	157,042,060	155,377,077	1,664,983	99
UTTARAKHAND	7,061,390	6,242,151	819,239	83
WEST BENGAL	75,088,307	70,880,863	4,207,444	96
Grand Total	863,053,207	772,756,044	90,297,163	88

State Total, State Covered, State Balance and State Covered Percentage broken down by State Name.

States needing Attention



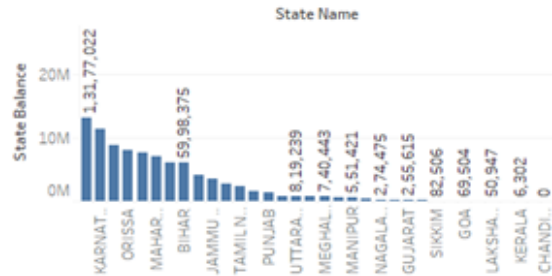
Map based on Longitude (generated) and Latitude (generated). Color shows sum of State Balance. Details are shown for State Name.

Water and Sanitation Dashboard

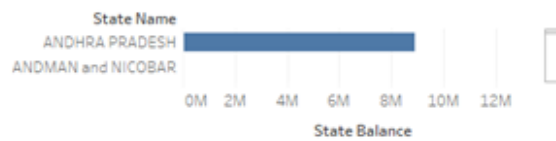
State wise Water Distribution Coverage

State Name	State Total	State Covered	State Balance
ANDHRA PR.	61,277,079	52,349,673	8,927,406
ANDAMAN & NICOBAR	241,964	226,998	14,966
ARUNACHAL PRADESH	975,640	712,266	263,374
ASSAM	26,411,324	18,713,978	7,697,346
BIHAR	90,415,338	84,416,963	5,998,375
CHANDIGARH	81,397	81,397	0
CHHATTISGARH	18,350,304	15,551,124	2,799,180
DADRA & NAGHROV	168,664	0	168,664

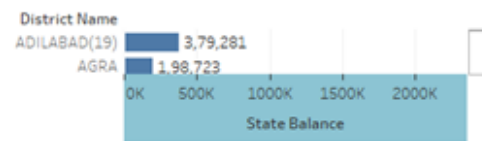
State wise Balance Population to be Covered



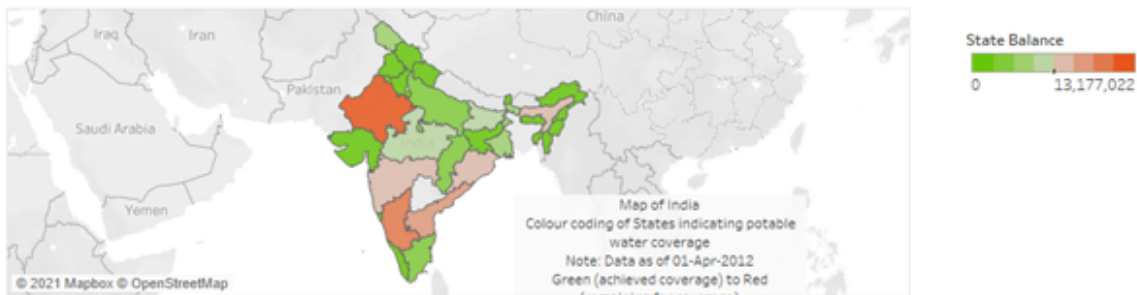
State with Balance Count



District with Balance Count



States needing Attention



10. Conclusion

Data visualization can be leveraged to provide a crisp overview of the problem at hand. The desktop component help in rapid creation of the necessary views and setup a dashboard. Once the authoring is accomplished and dashboard and sheets could be uploaded on the Tableau Public server for the consumption of general public. Tableau is powerful way to make a shift towards data centric decision making for better customer experience.

11. Future Scope

The model could be developed further by incorporating additional views to the dashboard:

1. Can be made to dynamically take data in an online manner
2. More filters could be added to give custom views
3. Sheets could be linked to provide dynamic views
4. Layout could be enhanced

12. Bibliography

<https://www.thesmartbridge.com/>

<https://www.tableau.com/>

<https://smartinternz.com/tableau-dataviz-challenge-2021>

13. Acknowledgements

Sincere thanks to:

- Smart Bridge and SmartInternz
- Tableau