

# README

For project of CS595 – Data Visualization Lab at IIT Guwahati

By, Ode Deepkumar Rameshbhai (204161011).

---

## 1. How to reproduce graphics (with given datasets)?

- i. Navigate to 'Code' folder. Open 'main.Rmd'. Do not run yet.
- ii. Check the first code chunk – setup code chunk/cell. This code chunk does two things: (a) Install required packages and (b) Download & unzip shape file for India's map. You might want to comment packages which are already installed in your environment.

You need to run this code chunk/cell only once. Run this code cell. If there are troubles downloading packages, please refer below to "Troubleshooting".

Note that, after running this code cell, some 26 files related to Indian map will be unzipped in this same folder. After running setup code, it is advisable to comment it.

- iii. You can run all rest of code cells now. Graphics will be visible in .Rmd itself. In addition, a new directory called 'images' will be created in which 17 out of 18 visualizations will be saved automatically. The remaining treemap which is visible in .Rmd file, can't be saved using ggsave() and hence it's screenshot was used in the report.

## 2. Where to download datasets? How to use downloaded datasets in this code?

The list of sources along with links is given below. All these files must be renamed to names in column 4 and should be placed in same folder as 'main.Rmd'.

Sr. No.	Source	Name of dataset	Rename dataset to	Link
1.	Census of India, 2011	C-08 Educational Level By Age And Sex For Population Age 7 And Above (Total, SC/ST) (India & States/UTs-District Level)	ed_level_above_7.xlsx	Click on "India – C-08" from <a href="#">link</a>
2.	Census of India, 2011	B-1 Main workers, Marginal workers, Non-workers and those marginal workers, non-workers seeking/available for work classified by age and sex(Total, SC/ST)	workers.xls	Click on "India – B-1" from <a href="#">link</a>
3.	Open Government Data Portal	Crime against Women during 2013	recovered_crime1.xls	Download as xls from <a href="#">link</a> . If you're prompted to

				recover file, click "Yes".
4	Ministry of Statistics and Programme Implementation	State Domestic Product and other aggregates, 2011-2012 series	gdp.xls	Download from <a href="#">link</a> .
5	Open Government Data Portal	State-wise Population, Decadal Population Growth rate and Population Density - 2011	pop_growth.csv	Download csv file from <a href="#">link</a> .
6	Census of India, 2011	B-17 Female Workers Classified By Industrial Category And Marital Status - 2011(India/State/UTs)	f_workers.xlsx	Download from <a href="#">link</a> . Meaning of various categories can be read from this <a href="#">link</a> .
7	Union Budget, Wikipedia	Literacy as seen in 2001 census.	lit_rate_national.xlsx	Original source (pdf) is <a href="#">down</a> . Wikipedia cites same source and table is copied from <a href="#">there</a> .

### 3. Troubleshooting

The code was tested on Windows – 10, 64 bit machine with R version 4.0.3.

If you're having trouble installing packages, close RStudio without saving image of workspace. Now open RStudio again but from its installation place (not by opening any .Rmd file). The main thing is to not load any project. Now copy paste all `install.packages()` line in the console itself.

If you're running some other OS than Windows, there might be problems with sf package. For any package problems, you can always download binaries(.tar or .gz file) from source (CRAN) manually. Link for doing so for sf package is [this](#). Check this [link](#) to see how to use these binary files.

If you have any more queries, feel free to contact me on +91 91067 72052 (calls) or +91 98241 11074 (WhatsApp) or [d.ode@iitg.ac.in](mailto:d.ode@iitg.ac.in).

Thank you!

Deep Ode

M. Tech in Data Science (204161011)