

Assignment-4

Data Analytics and Visualization (CS/IT312)

3 February, 2022

Instructions

1. Create “Assignment4_Branch_YourId.py” file. *YourId* will be the student ID and *Branch* will be “CS” or “IT”. For example, if student ID is 202018001 and Branch is CS, then file name will be “Assignment4_CS_202018001.py”.
 2. A student has to write the code for following tasks.
 3. Upload your “Assignment4_Branch_YourId.py” file into the Google form shared on Google class room. Fill the Google form.
 4. Please be present during lab session. The **deadline** of the submission is **11:59 pm, 11 February, 2022**.
-

Get help from:

Book: chapter 10: Fundamentals of Data Visualization

Data Source: <https://data.gov.in/>

Data:

- Data contains production of milk and egg from 2007 to 2012 in India (*Egg_Production_2007_2012.csv* and *Milk_Production_2007_2012.csv*).

Tasks

1. Merge two data into a data frame such that the new data frame has multi-level columns (like years under milk/eggs). Try to change the column names if required.
2. Present the production of milk in Gujarat, Kerala, Andhra Pradesh, Uttar Pradesh and Panjab on 2007-2008 as a Pie chart. The pie chart should consist of proportion in percentage and labels for each piece.
3. Plot five pie charts of egg production in Gujarat, Kerala, Andhra Pradesh, Uttar Pradesh and Panjab given five years range. Each pie chart should represent the egg production in above five states for a given year.
4. Plot Stacked Area Chart that represents state wise (i.e. states are in x axis) production of eggs (y-axis). There would be five stacked colors and each color represents the production of egg in a state. The egg production for each state should be normalized over the year range.