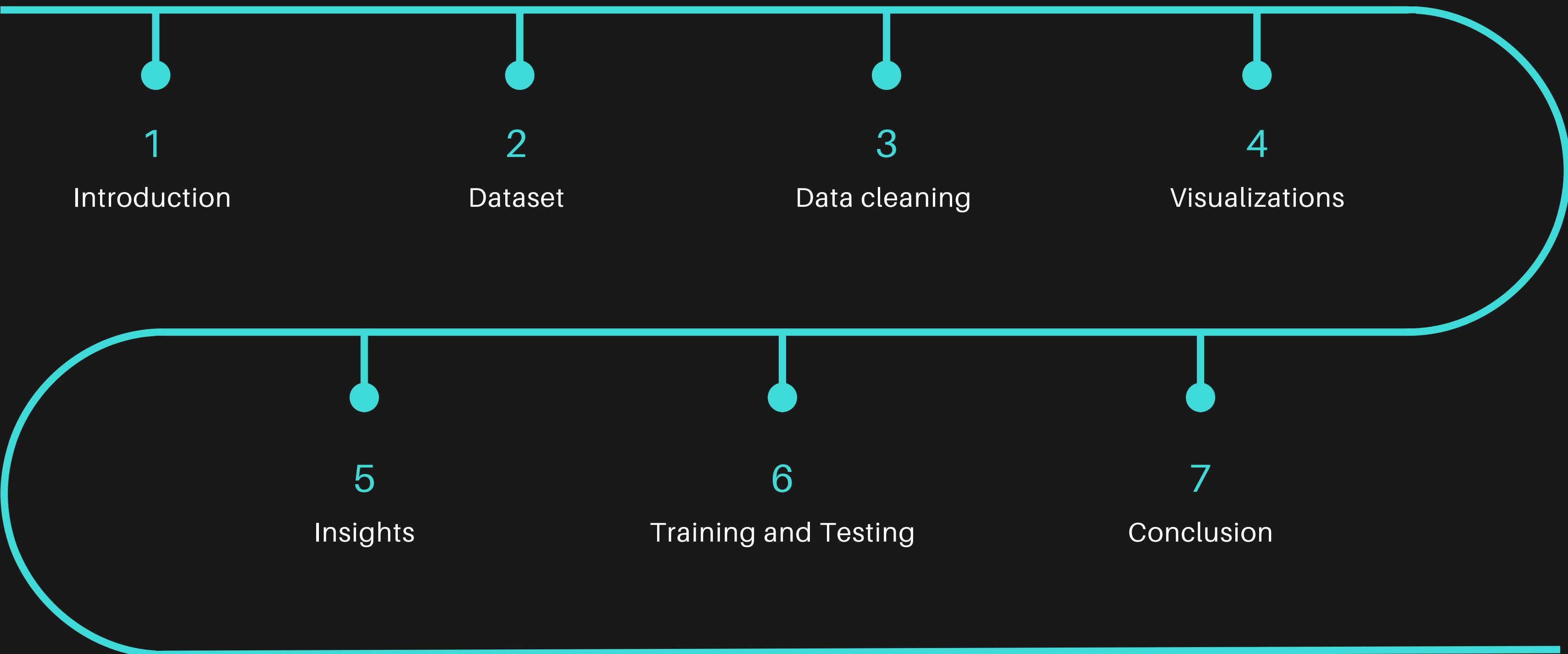




Analysis on Work Force in India from 2019-2020

ANALYZED AND PRESENTED BY
KOLADE ADEDIJI

Contents





Dataset

The dataset used for this analysis is sourced from kaggle data catalogue, accessible at the following link :

<https://www.kaggle.com/datasets/gokulrajkumar/unemployment-in-india>



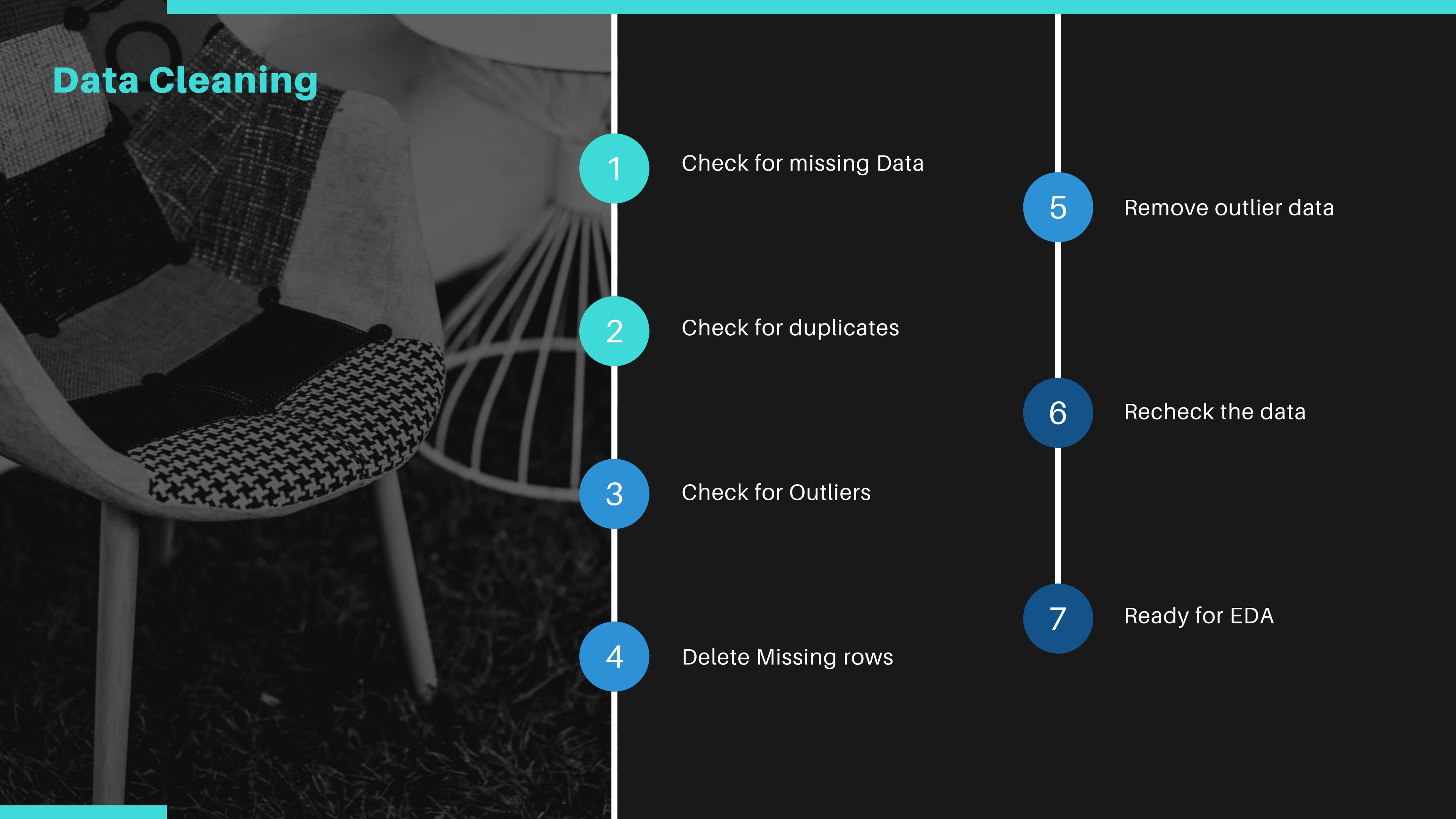


Dataset

The dataset contains 7 column namely Region, Date, Frequency, Estimated Unemployment rate, Estimated employed, Estimated labour participation rate. The dataset also contains 740 rows, contains 28 rows of missing data. The dataset contains 28 unique regions.

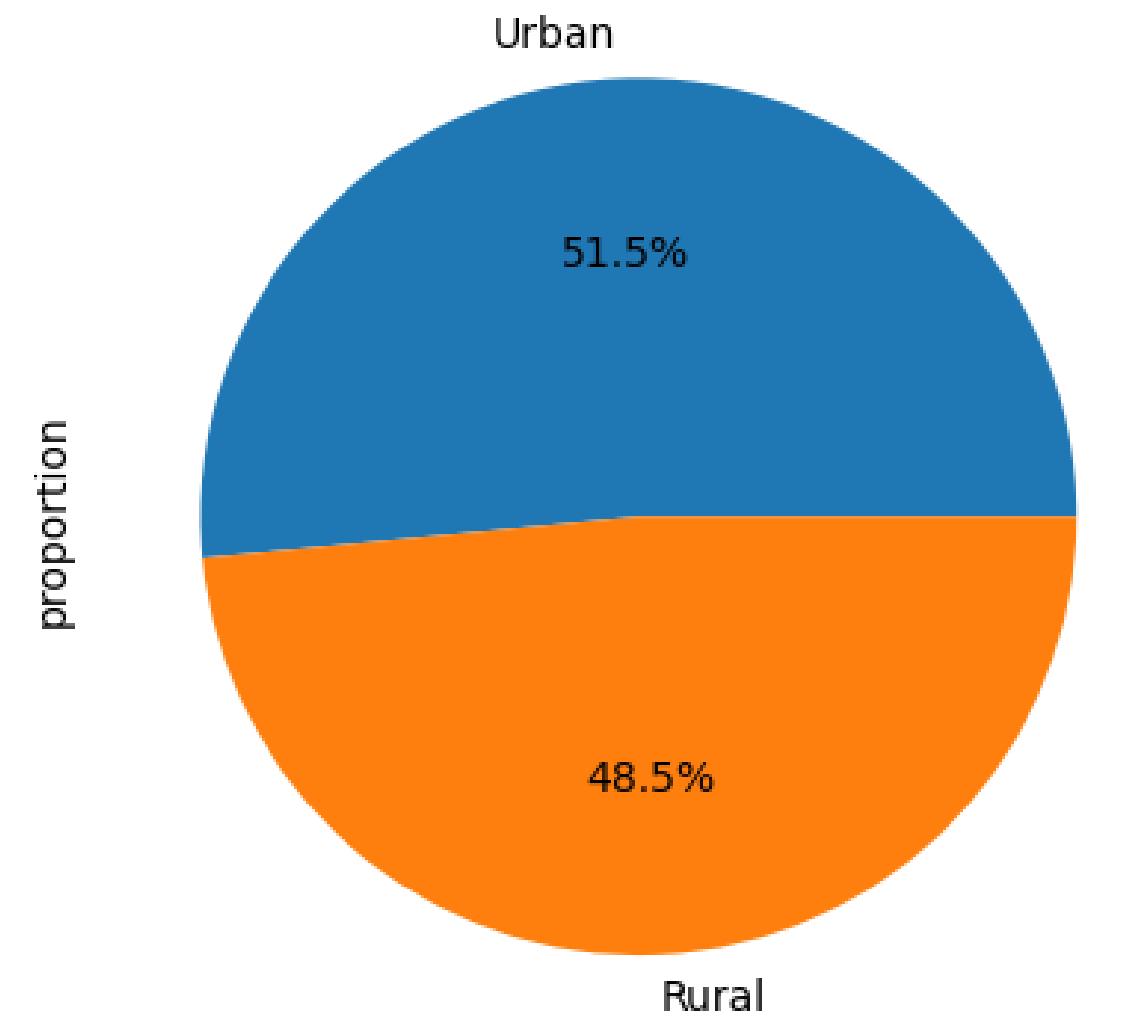


Data Cleaning

- 
- 1 Check for missing Data
 - 2 Check for duplicates
 - 3 Check for Outliers
 - 4 Delete Missing rows
 - 5 Remove outlier data
 - 6 Recheck the data
 - 7 Ready for EDA

Visualizations and Insights





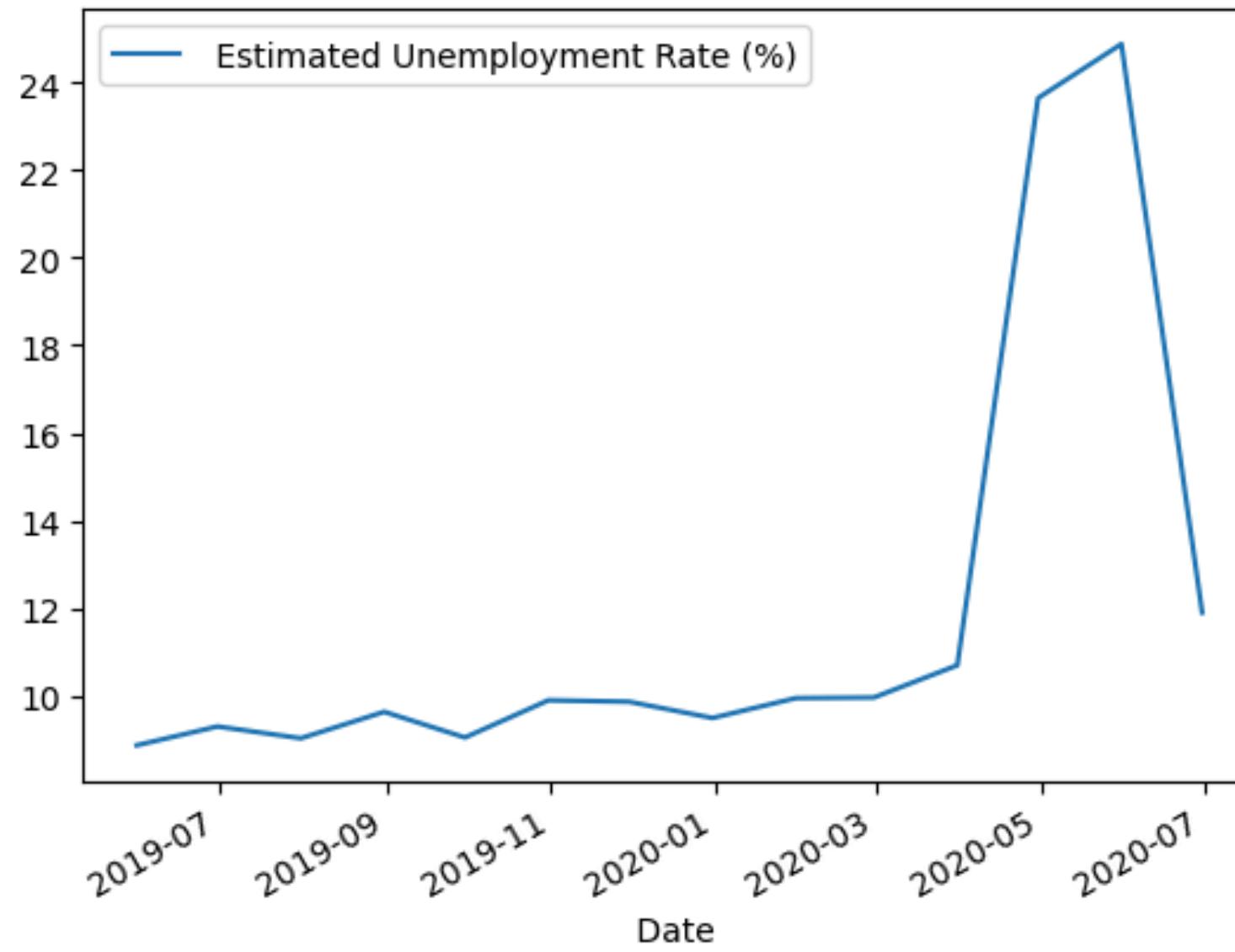
There are 51.55% urban cities in the dataset and 48.5% rural cities



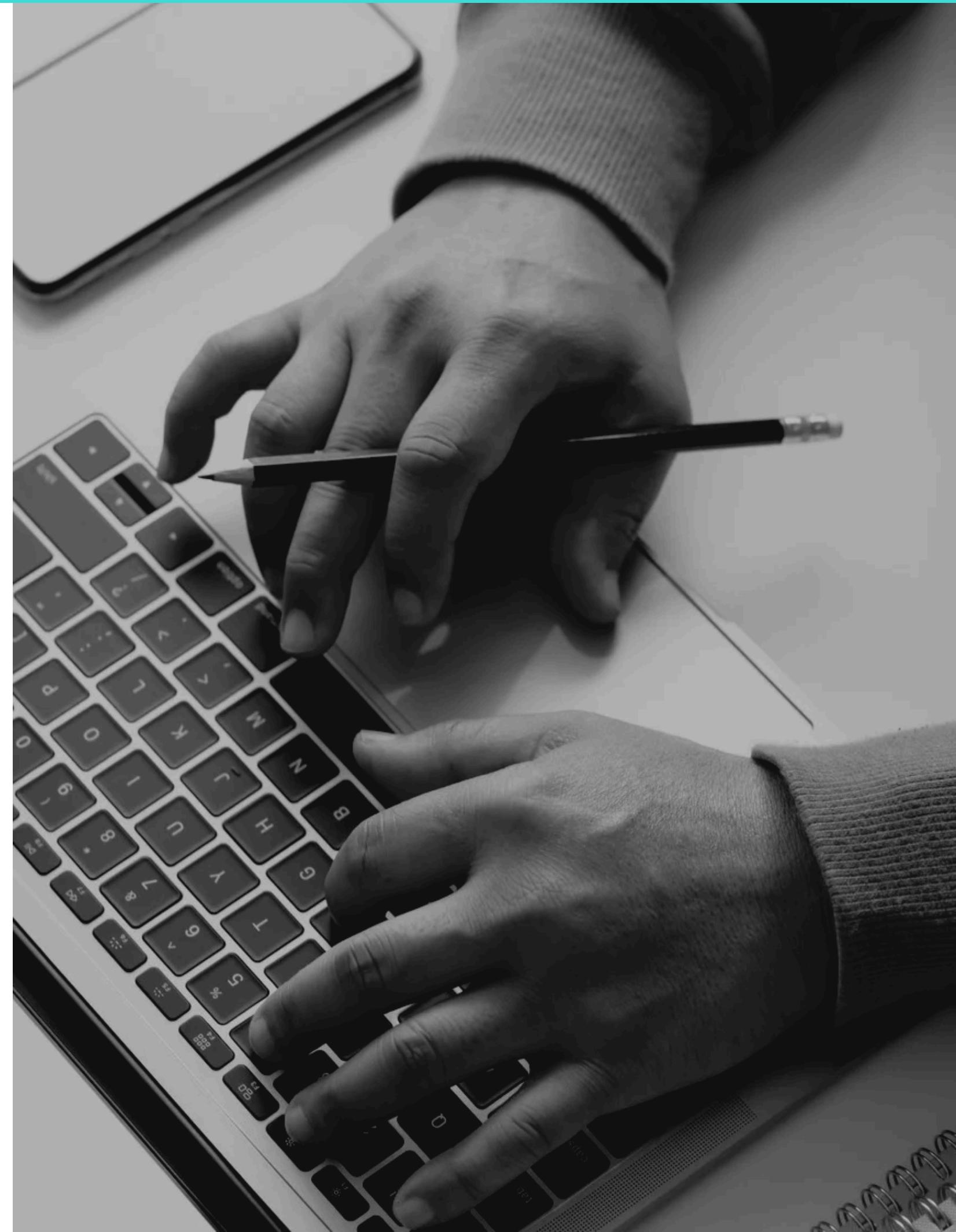


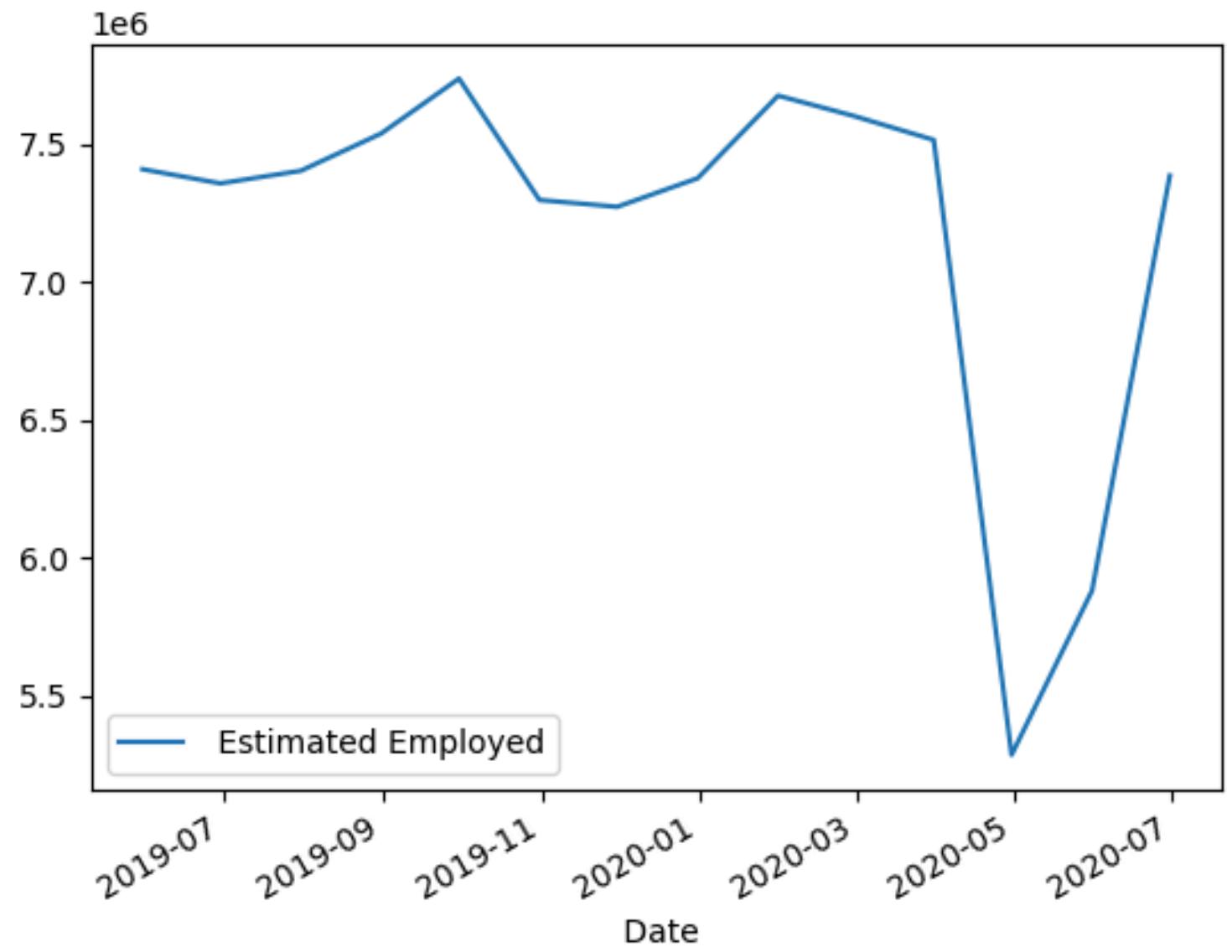
It shows that Estimated Unemployment rate increases, the Estimated Employed reduces. It also shows that an increase in people employed causes a little increase in Estimated Labour Participation rate





It shows that Estimated Unemployment rate increases over the year with the peak recorded in June 2020 which falls back to the original range back in July .





It shows that Estimated Employed has been up and down over the year with a peak fall recorded in May 2020 which also falls back to the original range back in July .



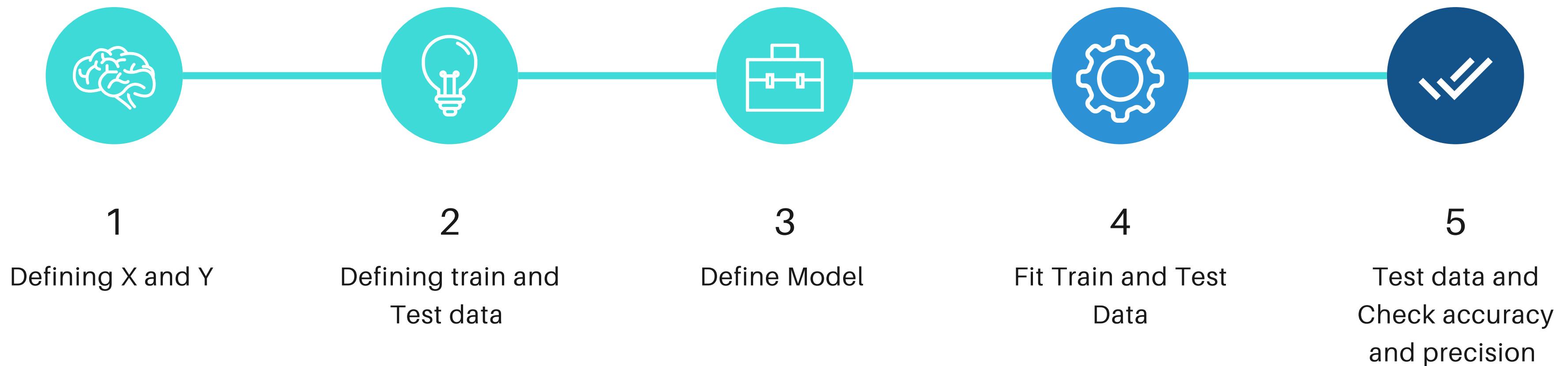


Insights

- The whole dataset contains 51% Urban region while there are 48.5% of rural regions
- Chandigarh does not have a Urban region
- There has been zero to little increase in number of people Employed In rural areas of Assam, Himachal Pradesh, Uttar Pradesh, Telangana and West Bengal and little to none in Urban areas of Delhi
- There has been fall in estimated number of people employed in almost all cities in May which in turn causes an increase in Unemployment rate in May
- The great unemployment rate fall in May is most visible in Andhra Pradesh and some rural parts of Assam, Bihar and Gujarat
- Region with highest unemployment rate is Tripura
- Region with lowest unemployment rate is Sikkim
- Region with highest people employed is Uttar Pradesh
- Region with lowest people employed is Sikkim



Model Training





Model Characteristics

TRAINING MSE

101.7016

TESTING MSE

139.0751

TRAINING R² SCORE

0.0521

TESTING R² SCORE

0.03939

Conclusion

The Model has an accuracy of about 75%

To check more about the analysis done, check
<https://github.com/Koladeadediji/Unemployment-Prices-In-India/>

