4. Exploring and Visualizing NIRF Rankings 2024

Objective:
The goal of this experiment is to explore the NIRF 2024 rankings data through various visualizations to gain insights about the distribution of ranks, scores, and the performance of institutions across different states.

Dataset Description: NIRF_Overall_2024

The NIRF 2024 Overall Ranking Dataset provides a comprehensive ranking of higher education institutions in India based on a range of performance parameters. The dataset is part of the National Institutional Ranking Framework (NIRF), which is an initiative by the Ministry of Education, Government of India. The rankings are designed to evaluate institutions across various domains such as engineering, management, medical, law, and overall performance.

Key Features of the Dataset:

- **Institute ID:** A unique identifier for each institution listed in the ranking.
- **Name:** The name of the institution.
- **City:** The city in which the institution is located.
- **State:** The state where the institution is situated.
- **Score:** The overall score of the institution based on various performance indicators such as teaching, learning, and resources (TLR), research and professional practice (RP), graduation outcomes (GO), outreach and inclusivity (OI), and perception (PR).
- **Rank:** The overall rank of the institution based on the calculated score.

Task: Create a new column titled 'Zone / Zonal Region' and categorize each institute into one of the following six zones based on the state:

1. **Northern Zone** - Chandigarh, Delhi, Haryana, Himachal Pradesh, Jammu and Kashmir, Ladakh, Punjab, and Rajasthan
2. **North Eastern Zone** - Assam, Arunachal Pradesh, Manipur, Meghalaya, Mizoram, Nagaland, Sikkim and Tripura
3. **Central Zone-** Chhattisgarh, Madhya Pradesh, Uttarakhand and Uttar Pradesh
4. **Eastern Zone** - Bihar, Jharkhand, Odisha, and West Bengal
5. **Western Zone** - Dadra and Nagar Haveli and Daman and Diu, Goa, Gujarat, and Maharashtra
6. **Southern Zone** - Andhra Pradesh, Karnataka, Kerala, Puducherry, Tamil Nadu, and Telangana, Andaman and Nicobar

Visualization Questions:

1. Among the top 5 states, which state has the highest number of ranked institutions in the NIRF rankings?
2. Which states should the government focus on to improve their presence in the top 100 NIRF rankings? Identify the states that could benefit from increased investment in higher education.
3. Which zone has the highest representation in the NIRF rankings?
4. Create two visualizations of the same data to discuss the “Lie Factor” in this context and explain how the visualization could mislead viewers.
5. Visualize the distribution of scores among all institutions.

Submission Guidelines:

- You can choose the programming language that best suits your approach and skills.
- Please avoid using commercial software (e.g., MATLAB, Tableau) for your visualizations. We encourage the use of open-source tools like Python, R, or JavaScript libraries
- If you are working in python, submit the assignment as a .ipynb (Jupyter Notebook) file. **Include a .pdf export** of the notebook. (Click Here If you do not know how to export PDF from .ipynb notebook)
- For other languages, Export all the codes, comments and results into a **PDF file** and submit it
- Name your file as DS304_LabAssignmentX_StudentID.ipynb (where X is the lab number).
- Include comments in your code to explain your logic and approach.
- Use Markdown cells for detailed explanations, if necessary, especially for complex code blocks.
- Ensure that all code cells have been executed and the outputs are visible.
- Do not clear the outputs before submission.
- Submit the assignment via the designated Google Classroom link.
- Submit your work before the end of the day.
- Late submissions may incur penalties
- Avoid Plagiarism; ensure the work you submit is your own.