**EXPLORATORY DATA ANALYSIS-14**

**Exploratory Data Analysis (EDA) on Global Missing Migrants Dataset**

**Introduction:** You are provided with a dataset containing information about missing migrants worldwide. Your task is to perform an exploratory data analysis using Python and pandas. This assignment will help you practice various data manipulation and visualization techniques.

[Dataset:](https://drive.google.com/file/d/1Viq2kLIAuS4QksqMrBN-Ywsrq0s-iyuC/view?usp=sharing)

* Dataset Name: Global Missing Migrants Dataset
* Columns:
  1. Incident Type
  2. Incident year
  3. Reported Month
  4. Region of Origin
  5. Region of Incident
  6. Country of Origin
  7. Number of Dead
  8. Minimum Estimated Number of Missing
  9. Total Number of Dead and Missing
  10. Number of Survivors
  11. Number of Females
  12. Number of Males
  13. Number of Children
  14. Cause of Death
  15. Migration route
  16. Location of death
  17. Information Source
  18. Coordinates
  19. UNSD Geographical Grouping

**Instructions:**

1. Load the dataset into a pandas DataFrame.
2. Perform data cleaning, handling missing values, and data type conversions if necessary.
3. Answer the following questions using pandas functions and visualizations where applicable.
4. Provide explanations and interpretations of your findings.
5. Comment your code for clarity.

**Questions:**

1. How many rows and columns are there in the dataset?
2. What are the data types of each column?
3. Are there any missing values in the dataset? If yes, handle them appropriately.
4. Calculate the summary statistics for numeric columns (e.g., Number of Dead, Number of Survivors).
5. What is the average total number of dead and missing?
6. Visualize the distribution of reported years.
7. In which year did the highest number of incidents occur?
8. List the unique incident types in the dataset.
9. For each incident type, calculate the average number of survivors.
10. Visualize the distribution of the number of dead.
11. What is the most common cause of death?
12. Calculate the total number of dead and missing for each region of origin.
13. Identify the top 5 regions of origin with the highest total number of dead and missing.
14. Create a new column 'month' derived from 'Reported Month' and analyze the monthly trends.
15. Calculate the total number of dead and missing for each month.
16. Visualize the trends over months using a line plot.
17. Calculate the correlation between 'Number of Dead' and 'Number of Survivors'.
18. Find the migration route with the highest number of incidents.
19. Calculate the total number of incidents for each location of death.
20. Visualize the distribution of incidents by location of death.
21. Identify the top 5 information sources with the highest number of incidents reported.
22. What is the average number of males per incident type?
23. Find the top 3 causes of death for incidents involving children.
24. Calculate the average total number of dead and missing for each reported year.
25. Identify the top 3 UNSD Geographical Groupings with the highest total number of dead and missing.