

# Brazilian Forest Fires Data Analysis



In Brazil, there is a lot of open nature and forest. Official organizations try their best to protect the rainforests and natural habitats within the country, but still deal with multiple forest fires each year.

For this project, you will be given a dataset on forest fires in Brazil called **amazon.csv**.

## Requirements

1. Import all required libraries
2. Read the csv file and if you experience 'unicodedecodeerror' then give second parameter in read\_csv method as 'encoding=iso-8859-1'
3. Check data type of each column
  - a. Find what is the type of 'date' column is
  - b. If it is of 'object' type, we need to convert the type into datetime
  - c. For conversion you need to make changes again in read\_csv method, put third parameter into it for parsing dates (you need to figure out by yourself regarding third parameter)
  - d. Please check after making changes what is the data type of date column is (it should be datetime)
4. Display top 5 rows of the dataset.

5. Check the last 5 rows.
6. Find the shape of our dataset (number of rows and number of columns).
  - a. Kindly print number of rows and columns separately
7. Getting information about our dataset like total number rows, total number of columns, datatypes of each column and memory requirement.
8. Check for duplicate data and drop them.
  - a. You can use duplicated () method
  - b. Please mention how many duplicate values are removed
9. Check null values in the dataset.
  - a. Please add a comment if you find any null value
10. Get overall statistics about the data frame.
  - a. Here you will find that you are getting only numeric columns
  - b. Kindly show categorical columns as well, if you find any warning then kindly add extra parameters in method.
11. Rename month names to english.
  - a. You will some columns in different language, you need to change them into english
12. Total number of fires registered.
13. In which month was the maximum number of forest fires reported?
  - a. You need to create proper stats and showing them with the help of bar plot
  - b. Mention in your comment what did you observe
14. In which year was the maximum number of forest fires reported?
  - a. Please visualize your output with barplot
15. Which state has the maximum number of forest fires reported?
  - a. Visualization should be there
  - b. If find that barplot is overlapping, then please fix it using rotation
16. Find total number of fires were reported in amazonas (state of brazil)
  - a. Add in comment the number of reports
17. Display number of fires were reported in amazon (year-wise)
  - a. Kindly visualize it using barplot after having groupby

- b. Please take care of plot overlapping (it should be neat and clean)
- 18. Display number of fires were reported in amazon (day-wise)
  - a. You need to visualize the result using weekdays name
  - b. Don't forget to mention your observation in comment
- 19. Find the total number of fires reported in 2015 and visualize data based on each 'month'.
  - a. Showing output with beautiful barplot
  - b. Mention your observation in comment
- 20. Find the average number of fires reported from highest to lowest (state-wise).
  - a. Visualization should be there with proper observation
- 21. To find the state names where fires were reported in 'dec' month.

Bonus: no limit for adding extra analysis.

## Due Date

You have **1 week** to complete this project, after which you will present your project.