Python Programming Certification Course

Certification Project

edureka!

Domain – 911 Calls

focus – data analysis and visualization

overview

Congratulations!!!! for making it so far. This is the capstone project for Python Programming Course and here we will use all the concepts learned so far in this course.

Business challenge/requirement

For this capstone project we will be analyzing 911 call data from Kaggle

This data is from Montgomery Country in the Pennsylvania State of USA.

911 is the most important social security feature of USA. It is the no., which citizens can call in case of any emergencies such as crime, medical, traffic, fire etc.

As a data analyst you have to analyze and visualize the data and answer the questions in section **Approach to Solve**

Key issues

Data should be analyzed accurately

considerations

NONE

data volume & Description

- Approx 260K records – file 911.csv

Fields in Data are:

- lat : String variable, Latitude
- lng: String variable, Longitude
- desc: String variable, Description of the Emergency Call
- zip: String variable, Zipcode
- title: String variable, Title
- timeStamp: String variable, YYYY-MM-DD HH:MM:SS
- twp: String variable, Township
- addr: String variable, Address
- e: String variable, Dummy variable (always 1)

additional information

- NA

business benefits

Better utilization of resources based on the density of 911 calls.

Approach to Solve

You have to use fundamentals covered till Module 8 and answer following 8 questions

- Compute -- What are the top 10 Zipcodes for 911 & Question 1: Are Zipcodes 19446 and 19090 presents?
- Compute -- What are the top 4 townships (twp) for 911 calls & Question 2: Which of the following township are not present? -- LOWER POTTSGROVE, NORRISTOWN, HORSHAM, ABINGTON
- Compute -- Create new features & Question 3: What is the most common Reason for a 911 call based on Reason Column? Which comes second
- Compute -- Plot barchart using matplot for 911 calls by Reason & Question 4: How can you plot the bars horizontally?
- Do data manipulation & Question 5: Which day got maximum calls for EMS and how many?
- Compute -- Create a countplot of the Day of Week column with the hue based of the Reason column & Question 6: On which day traffic calls were lowest?
- Compute -- Create a countplot month wise -- Question 7: Which month saw highest calls for fire?
- Compute -- Create Web Map for Traffic Calls & Question 8: Why some areas seem to have lower or almost zero traffic calls? Hint: Zoom the map

Answers:

1. Yes
2. LOWER POTTSGROVE, HORSHAM
3. EMS -- 133234, Traffic – 93400
4. Change plt.bar to plt.barh
5. Friday -- 19938
6. As expected, -- Sunday -- don't you prefer remaining inside :-)
7. June -- Barbeque and Holidays time?
8. These are areas like Airport, State Park etc.