HyperionDev

Capstone Project – Data Analysis

Task

Visit our website

Introduction

This capstone project will consolidate what you have learnt thus far and allow you to put your newly acquired knowledge concerning exploratory analysis to the test.

Developer portfolio

Developers who have the edge are those who find ways to apply their newfound skills from the get-go. As you may know, a **developer portfolio** (a collection of online creations that you have made) allows you to demonstrate your skills rather than just telling people about them. It's a way of bringing your CV to life and introducing yourself to the world. As you learn more skills and put these into practice, each project that you complete will become more efficient and eye-catching.

These capstone projects give you the means to create projects for your very own developer portfolio, allowing you to walk away from this course with not only a certificate but, more importantly, with a head start to your tech career!

The task at hand

In this capstone project, you will be given the opportunity to showcase all the skills that you have been developing throughout this bootcamp. You will be tasked with analysing a dataset, and you will create an exploratory data analysis (EDA) report that contains several data visualisations that will clearly communicate your findings.

Take note

A key focus of this project will be ensuring that your code is correct, well-formatted, and readable. In this regard, make sure that you do the following before submitting your work:

1. Make sure that you have identified and removed all syntax, runtime, and logical errors from your code.

2. Make sure that your code is readable. To ensure this, add comments to your code, use descriptive variable names, and make good use of whitespace and indentation. Use the **PEP 8 style guide** to see how classes and methods should be named and how your program should be formatted.

3. Make sure that your code is modular. Create functions to perform specific units of work.

4. How you choose to write code to create the solution to the specified problem is up to you. However, make sure that you write your code as efficiently as possible.

5. Use defensive coding to make provisions for errors that may occur using exception-handling techniques.

6. Make sure that all output that your program provides to the user is easy to read and understand. Labelling all data that you output (whether in text files or to the screen) is essential to make the data your program produces more user-friendly.

   Output 1: Raw

```
admin, Register Users with taskManager.py, Use taskManager.py to add the usernam
es and passwords for all team members that will be using this program., 10 Oct 2
019, 20 Oct 2019, No
admin, Assign initial tasks, Use taskManager.py to assign each team member with
appropriate tasks, 10 Oct 2019, 25 Oct 2019, No
```

   Output 2: User-friendly

```
----------------------------------------------------------

Task:                   Assign initial tasks
Assigned to:            admin
Date assigned:          10 Oct 2019
Due date:               25 Oct 2019
Task Complete?          No
Task description:
 Use taskManager.py to assign each team member with appropriate tasks
----------------------------------------------------------
```

7. Label data visualisations with appropriate headings, axis labels, and legends.

Practical task

Create a copy of the **Capstone_Project_template.ipynb** file and save it as **automobiles.ipynb**.

- Follow the instructions to clean, sanitise, and explore the **automobile.txt** dataset.
- Compile an EDA report on the automobile dataset exported to a PDF named **EDA Report – Automobile Dataset** explaining your visualisations, investigations, and findings.

To do this, you will need to:

- load the dataframe into the notebook,
- clean the data,
- remove duplicate rows,
- discard entries with any missing values,
- manipulate certain columns to the correct data type, and
- answer the questions about the data provided in the template notebook.

Important: Be sure to upload all files required for the task submission inside your task folder and then click "Request review" on your dashboard.

Share your thoughts

Please take some time to complete this short feedback **form** to help us ensure we provide you with the best possible learning experience.