Coursework - Practical Report

The Problem:

In this coursework, you are provided with a dataset which includes information on the various aspects of a building location and construction. The dataset includes lots of information such as location of the building, it's age, foundation, etc. Using the data, one can help predict the level of damage to the building that can be caused by an earthquake.

Such datasets can help concerned government agencies design better governance plans, including but not limited to improvement to existing infrastructure, better disaster management plans, etc.

Impressed by your previous research report, you are now tasked to provide a practical implementation and analysis report of the dataset, addressing the various identified research questions. You are expected to report your observations, inferences and a set of recommendations as part of this report.

See Canvas for submission deadlines.

---

[1] Refer to the Canvas page here on where to register your group and how your final individual mark will be calculated based on the Individual Contribution Report.

Practical Project:

You are asked to perform an exploratory data analysis of the provided dataset in a Jupyter Notebook. Your practical project **must** address the research questions you mentioned in your Research Report. You should demonstrate use of:

- NumPy Arrays
- Pandas Dataframe
- EDA using Pandas functions
- Matplotlib for Visualization
- appropriate graphs for data presentation
- proper variable names, methods and comments

A minimum of the following analysis must be present in this report:

Initial Data Analysis:
1) Quality of data:
   a) Frequency counts
   b) Descriptive/Summary statistics (mean, median, standard deviation, etc.)
   c) Normality (frequency histograms)
2) Quality of measurements, if applicable
3) Data transformation: data merging (features and labels), transpose, data type change, data sorting, data deletion, etc.
4) Characteristics of dataset:
   a) Printing top and bottom 8 rows
   b) Basic plots
   c) Correlation analysis

Exploratory Data Analysis:
1. Finding null values and dealing with missing values with appropriate justification.
2. Correlation structures, distribution characteristics
3. Trends in data using appropriate visualization tools
4. Using appropriate plots, present the results of the research questions previously identified in the research report.

Note: For visualization, you **must** use Matplotlib for at least 5 of the plots. For the rest, you can use seaborn, if suitable.

The practical report is for you to demonstrate the understanding and use of at least the basic EDA presented in the lectures and workshops. You should use your learning and work from the Research report. Meeting minimum requirements will not get you a high grade. You are expected to analyze and present results and discussion beyond what has been asked here.

##

Starting Guide:

1) There will be a discussion on the problem statement and how to approach it. Make sure to attend the lecture and take notes, as appropriate.
2) You can use ideas from the two case studies that will be discussed in the lecture.

Instructions for submission:

You need to submit **two** files:

1. A copy of your Jupyter Notebook (.**ipynb** file) - it should run without any errors and should include all steps (including required python package installation).
2. A **pdf** file consisting of the code and related output.

At the start of the report, add a section (including names and KU IDs of all the group members, and title of the report), a ToC (Table of Contents) and Table of Figures, etc.

Make sure that you have informed the module leader of your team and got approval before starting to work on the report.

Assessment and Feedback:

You will be assessed primarily on (but not necessarily limited to) the following: completeness of the minimum requirements (see above), coding quality (proper commenting, etc.), implementation of appropriate functionalities, executability of code, suitability and quality of visualization plots addressing the identified research questions. See the below rubric for a division of marks.

You will be provided feedback on your work by the deadline mentioned on the *Canvas Assessment and Feedback Schedule* page here.

Rubric

Content	Marks
Quality of practical reports (formatting, appropriate headings and subheadings, proper commenting, code execution, etc.)	**10**
Basic IDA and EDA	**25**
Suitability and quality of visualization plots addressing the identified research questions	**40**
Interpretation and observations for the identified RQs; recommendations/conclusions?	**25**

University Policy and Regulations

Any submission found/suspected of academic misconduct including but not limited to collusion, plagiarism and Purchasing or commissioning will NOT be marked and all the members of the group will be reported to the University for further action. Make sure you read the below guidelines available here:

Academic Misconduct (including Plagiarism)

To apply for extensions and mitigating circumstance, please refer to the guidelines here:

Extensions and Mitigating Circumstances Policy

Viva: We reserve the right to withhold your marks and invite you to a *viva* where you will be asked questions about your work to demonstrate ownership and understanding of your own submission. Whether you are asked to *viva* or not will depend on:

- Your engagement during scheduled classes
- Your engagement with the team
- Your formative submissions (i.e. the work you do for learning purposes rather than marks)
- The work you submit (if it's an interesting submission and we want to ask some questions – this is a GOOD thing, as you may end up with more marks!)

If you are asked to do a *viva*, then your final mark will be subjected to a multiplier factor as follows:

Multiplier	Viva assessment
1.2	Ownership and understanding of the artefact submitted was excellent, and gave an insight into not only their own work as-is but also possible alternative approaches, demonstrating more understanding of development techniques than the work alone provides.

1	Ownership and understanding of the artefact submitted is clearly established. The student responded to questions about their work with authority.
0.6	The student was able to answer most questions about their work, but there were some gaps in understanding of the artefact submitted.
0.3	The student was able to answer some questions and demonstrate only a limited understanding of their own submission. However, there was a clear disconnect between the quality of artefact and the students' understanding of the technology and principles encapsulated within.
0	The student was not able to provide any meaningful responses to questions about their work. Extremely limited, or no understanding of the artefact submitted was apparent. Alternatively, the student did not respond to the invitation to viva and/or did not attend the viva at the appointed time.

Please note that, as they contribute to an assessment event, *vivas* will be recorded to provide sight of such assessment events to moderators and external examiners. Where appropriate, the recordings will be used in other university regulatory processes.

Contact

Please contact your module leader, Dr Nabajeet Barman (n.barman@kingston.ac.uk) or your Course Leader Dr James Denholm-Price (J.Denholm-Price@kingston.ac.uk) if you have any questions or concerns.