|  |  |  |
| --- | --- | --- |
| This question paper consists  of 3 printed pages, each is identified by the Code Number LUBS5309M |  |  |
| **© UNIVERSITY OF LEEDS**  **(Semester 2, 2023/2024)**  **Assessed Coursework**  **LUBS5309 Forecasting and Advanced Business Analytics**  **100% Assignment** | | |

|  |
| --- |
| The assignment contains two equally weighted parts. You must submit your work as a report describing the work that you have done in Microsoft Word (.docx) format. The submission should answer both questions and may be generated either from R Markdown or manually. Additionally, you **must** also submit the code that was used in your analysis, either an .RMD file if that is used to generate the word document or in the format of .R if the report is prepared separately from the analysis.  **Part 1:** Using the US seasonally-adjusted personal consumption expenditures (PCE.csv) compare the predictive ability of the following three models:   1. One of the four simple forecasting methods (average, naïve, seasonal naïve, or drift) 2. An Exponential smoothing model (ses, holt, hw/a, hw/m) 3. An ARIMA model   Your goal is to identify the best performing model.  In you report you should clearly present and discuss the following:   * Describe the steps and decisions made for your analysis (missing data, dataset split etc.). * Present and interpret the criteria and the results for the selection of the best model. * Plot the prediction of the models against the real values in one graph. * Based on your findings make an estimation of the personal consumption expenditures for the **October of 2024.** * Repeat and present **only** the models’ comparison using this time one-step ahead rolling forecasting without re-estimation of the parameters.   **Part 2:** The file HotelData.csv contains customer online reviews for hotels and their corresponding ratings (Likert scale from 1/low to 5/high satisfaction). Randomly, select a sample of 2,000 reviews as follows:  Use the **sample\_n()** function from dplyr package, but to ensure reproducibility use the function **set.seed(XXX)** before you use the sample\_n() function where XXX should be the last 3 digits of your student ID. For example, if your student ID is 2017XX123 then your code could have the following form:  **set.seed(123)**  **test<-sample\_n(reviews, 2000)**  Based on the new sample you are asked to perform topic modelling to identify (separately) the factors (topics) that are discussed in the positive and negative reviews.  **In your report you should clearly address and report the following:**   * Explain the criterion used to classify positive and negative reviews. * Describe all the steps that are followed for your analysis (for example, tokenization, removal of stop words, etc.). * Present the criteria for the selection of the number of topics. * Identify and label the topics. * Discuss the three top factors that affect the satisfaction of the customers. * Discuss the three top factors that affect the dissatisfaction of the customers.   Your report should use tables and figures as appropriate as well as text to present your findings. No appendices are required, and any material provided in appendices will not contribute to the final mark. The number of words in the Word file should not exceed 3000 words in total for both parts **(the length of the r code will not be considered)**. The coursework will be marked using the following marking criteria:  **Weighting Criteria**  Presentation: 10%  Reproducibility of the results: 10%  Code Efficiency and Novelty: 20%  Critical Thinking: 20%  Analysis: 40% |

**Assignments should be a maximum of 3000 words in length.**

All coursework assignments that contribute to the assessment of a module are subject to a word limit, as specified on the assessment brief. **The word limit is an extremely important aspect of good academic practice and must be adhered to.** Unless stated otherwise in the relevant module handbook (if one has been provided), the word count includes EVERYTHING (i.e. all text in the main body of the assignment including summaries, subtitles, contents pages, tables, supportive material whether in footnotes or in-text references) except the main title, reference list and/or bibliography and any appendices. It is not acceptable to present matters of substance, which should be included in the main body of the text, in the appendices (“appendix abuse”). It is not acceptable to attempt to hide words in graphs and diagrams; only text which is strictly necessary should be included in graphs and diagrams.

You are required to adhere to the word limit specified and state an accurate word count on the cover page of your assignment brief. Your declared word count must be accurate, and should not mislead. Making a fraudulent statement concerning the work submitted for assessment could be considered academic malpractice and investigated as such. If the amount of work submitted is higher than that specified by the word limit or that declared on your word count, this may be reflected in the mark awarded and noted through individual feedback given to you.

**The deadline date for this assignment is 12:00:00 noon on Monday 13th May 2024.**

An electronic copy of the assignment must be submitted to the Assignment Submission area within the module resource on the Blackboard MINERVA website no later than 12:00:00 noon prompt on the deadline date.

Faxed, emailed or hard copies of the assignment will not be accepted.

Failure to meet this initial deadline will result in a reduction of marks, details of which can be found at the following place:

<https://students.business.leeds.ac.uk/assessment/code-of-practice-on-assessment/>

**SUBMISSION**

Please ensure that you leave sufficient time to complete the online submission process, as upload times can vary. Accessing the submission link before the deadline does **NOT** constitute completion of submission.You **MUST** click the ‘**CONFIRM**’ button before 12:00:00 noon for your assignment to be classed as submitted on time, if not you will need to submit to the Late Area and your assignment will be marked as late. It is your responsibility to ensure you upload the correct file to the MINERVA, and that it has uploaded successfully.

**It is important that any file submitted follows the conventions stated below**:

**FILE NAME**

The name of the file that you upload must be your student ID only.

**ASSIGNMENT TITLE**

During the submission process the system will ask you to enter the title of your submission. This should also be your student ID only.

**FRONT COVER**

The first page of your assignment should always be the Assessed Coursework Coversheet (individual), which is available to download from the following location:

<https://students.business.leeds.ac.uk/forms-guidance-and-coversheets/>

**STUDENT NAME**

You should **NOT** include your name anywhere on your assignment

**END**