

3rd Jun 2024, 13:15-16:00

Full Name : _

Student ID: _____

Grade Table (for Lecturer use only)

Question	Points	Score
1	25	
2	25	
3	25	
4	25	
Total:	100	

Instructions for Final Exam

Welcome to the final exam of EEE356 - Data Analytics and good luck!

Please read the following rules and confirm by signing that you have read and understood the rules before you receive your exam:

- The final exam shall be conducted between 13:15 and 16:00. Exam duration is 165 minutes. Students must finalise the exam by delivering it before 16:00. Students are not allowed to leave the exam in the first 30 minutes.
- Student ID cards shall visibly be on the edge of desks till the end of the exam. Students without the student ID cards or Turkish identity cards shall not be participated into the exam.
- This is a closed-book exam which means that students are not allowed to take notes, books, or any other reference material into the exam. Throughout the exam, students shall not possess mobile phones and electronic devices that are capable of storing, receiving or transmitting information or electronic signals, such as computerised watches.
- Students are not allowed to take a glance at the exam questions until told to do so. Students shall not communicate with any other student under any circumstances during the exam period. A student, who cheats, tries to cheat during the exam, or is identified to be cheating after investigating exam documents, is given 0 (zero) for that exam and a disciplinary investigation is opened against the student.
- An incorrect answer to a question is awarded no marks with no consideration of any partial credit. Therefore, no partial credit will be given.

In recognition of and in the spirit of the above rules which constitute Adana Alparslan Türkeş Science and Technology University Honour Code, I certify that I will neither give nor receive unpermitted aid on this examination.

Signature:



3rd Jun 2024, 13:15–16:00

- 1. Download the file named as "PTF-01012016-31122019.csv" in "Electricity Prices of Turkey" data set from Kaggle data sets and answer the following questions by using the downloaded file along with emphasising R codes.
 - (a) **(5 points)** Write down R codes to import "PTF-01012016-31122019.csv" into RStudio as a tibble named as price by using read_delim() function.
 - (b) (5 points) Write down R codes to import "PTF-01012016-31122019.csv" into RStudio as a tibble named as price by using read_delim() function. Check the values in PTF (TL/MWh) column with row numbers 8562, 8580, 8583, and 8584, and report anything wrong. If there is something wrong, explain how you can handle this problem.

(c) **(5 points)** Write down R codes to make required data manipulation in order to plot a similar graph shown below.





3rd Jun 2024, 13:15-16:00

(d) **(10 points)** Write down R codes to make required data manipulation in order to plot a correlation map with Pearson correlation and fill the below table.

Pearson	\mathbf{PTF}
Correlation	(TL/MWh)
PTF (USD/MWh)	
PTF (EUR/MWh)	
Year	
Month	
Day of Month	
Day of Week	
Hour of Day	

- 2. Download the file named as "RU_Electricity_Market_UES_dayahead_price.csv" in "Russian Wholesale Electricity Market" data set from Kaggle data sets and answer the following questions by using the downloaded file along with emphasising R codes.
 - (a) (5 points) Calculate the mean values of each column in the original data set.
 - (b) **(5 points)** Find the number of cells that are smaller than 500 RUB/MWh? Elaborate the numbers for each column separately.
 - (c) (5 points) Use kNN interpolation to impute the cells that are smaller than 500 RUB/MWh.
 - (d) (5 points) Calculate the mean values of each column in the imputed data set.
 - (e) (5 points) Plot correlation maps and determine the strongest correlation for both the original and the imputed data set.



3rd Jun 2024, 13:15–16:00

- 3. Download the file named as "PV_Elec_Gas3.csv" in "Daily Power Production of Solar Panels" data set from Kaggle data sets and answer the following questions by using the downloaded file along with emphasising R codes.
 - (a) (5 points) Omit the first row of the data set and use the rest.
 - (b) (10 points) Mutate a new column named as "DailySolarPower" by utilising "Cumulative_solar_power".

(c) (10 points) Plot a correlation map by excluding "date" column and interpret the obtained results.

4. (25 points) Please note that if you write down "I would prefer to use my HOMEWORK score instead of solving this question", then your homework score will be accepted, otherwise your answers to the question will be evaluated.

Develop a dashboard, web, or a GUI application using Shiny along with emphasising R codes.