ADANA ALPARSLAN TÜRKEŞ
SCIENCE AND TECHNOLOGY UNIVERSITY

EEE356 - Data Analytics
Online Final Exam
Prepared by Dr Kasım Zor

————————————

7th Jun 2021, 09:15–12:00

Full Name : _____   Student ID: _____

Grade Table (for Lecturer use only)

Question	Points	Score
1	5	
2	5	
3	5	
4	5	
5	5	
6	25	
7	25	
8	25	
Total:	100	

Instructions for Online Final Exam

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

- The final exam will be conducted between 09:15 and 12:00. The system will be closed at 12:00. Students must finalise the exam by saving before 12:00.

- The students are also responsible to have a fully-charged laptop computer and 3G internet connection against an interruption in the electricity or internet services. Please note that, allowed number of entries during the exam is determined as 3 due to the internet disconnections.

- This is an open-book exam that means students are allowed to take notes, books, or any other reference material into the exam.

- The exam must be taken completely alone. Showing it or discussing it with anybody is forbidden, including (but not limited to) the other students in the course in current or previous years. Absolutely no communication is allowed between or among students.

- An incorrect answer to a question is awarded no marks with no consideration of any partial credit. Therefore, no partial credit will be given.

- Please sign the below Honour Code statement.

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: _____

1. **(5 points)** Which of the following functions can be used to create correlation map? Select all that apply.

☐ corrplot::corrplot()
☐ ggstatsplot::ggcorrmat()
☐ PerformanceAnalytics::chart.Correlation()
☐ ggplot2::geom_tile()
☐ GGally::ggpairs()

2. **(5 points)** Which of the following packages can be used to impute missing data? Select all that apply.

☐ VIM
☐ imputeTS
☐ naniar
☐ Amelia
☐ MICE

3. **(5 points)** Which of the following programming languages can be supported by knitr package within R Markdown? Select all that apply.

☐ Python
☐ Julia
☐ C++
☐ MatLab
☐ JavaScript

4. **(5 points)** Which of the following presentation outputs can be generated within R Markdown? Select all that apply.

☐ Powerpoint
☐ Beamer
☐ revealjs
☐ ioslides
☐ slidy

5. **(5 points)** Which of the following data sets has never been used throughout EEE356 Data Analytics (Spring 2021)? Select all that apply.

☐ Fuel economy data from 1999 to 2008 for 38 popular models of cars (mpg)
☐ On-time data for all flights that departed New York City in 2013 (flights)
☐ US gun murders by state for 2010 (murders)
☐ Prices of over 50,000 round cut diamonds (diamonds)
☐ Diabetes data set (Diabetes)

6. 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 linear 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.

ADANA ALPARSLAN TÜRKEŞ
SCIENCE AND TECHNOLOGY UNIVERSITY

EEE356 - Data Analytics
Online Final Exam
Prepared by Dr Kasım Zor

———————————

7th Jun 2021, 09:15–12:00

(e) **(5 points)** Plot correlation maps and determine the strongest correlation for both the original and the imputed data set.

7. Please note that if you write down "I would prefer to use my ASSIGNMENT score instead of solving this question", then your assignment score will be accepted, otherwise your answers to the question will be evaluated.

Download the file named as "people-without-electricity-country.csv" in "Number of people without electricity" data set from Kaggle data sets and answer the following questions by using the downloaded file along with emphasising R codes.

(a) **(5 points)** Plot a graph with red coloured double-sized path geometry that demonstrates the number of people without electricity access in Turkey versus years between 1990 and 2020 in accordance with the followings:
   - Name of y-axis should be "People without Electricity Access",
   - x-axis name should not exist,
   - Title of the plot should be Turkey.

(b) **(10 points)** Find the number of entities in which the number of people without electricity access equals to zero for all years.

(c) **(10 points)** Find the Top25 countries of the year 2016 from the viewpoint of the number of people without electricity access.

8. Please note that if you write down "I would prefer to use my PRESENTATION score instead of solving this question", then your presentation score will be accepted, otherwise your answers to the question will be evaluated.

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.