|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| A close up of a sign  Description automatically generated | | | | | | | |
| Data Science Academy Bootcamp | | | | | | | | |
| Midterm exam | | | | | | | | |
| **Name,Surname:** | Şahin Məmmədov | |  | | Date: | 09.25.2022 | | |
| **Exam:** Data Science with Python | |  | | | | | | |
| **Instructions:** You will be given ***2 hours*** to complete examination. There are 2 parts and 4 types of questions including: True/False, Multiple choice, Fill in the blanks and Practical.  **First part** will last ***40 minutes****,* and you must write answers directly on word document. Internet is **not** permitted.  **Second part** is Practical question and will last ***1 hour and 20 minutes***. It is your responsibility to submit your exam answers electronically to [homework@dsa.az](mailto:homework@dsa.az) with **“Python Exam”** subject.  Before you leave exam, you **MUST** contact mentor to confirm that you have submitted your examination.  Good Luck! | | | | | | | | |
| PART 1: TRUE/FALSE | | | | | | | | |
| **Directions: Read each statement below carefully. Place a + on the “True” cell if you think a statement is TRUE, and on the “False” cell if you think the statement is FALSE.** | | | | | | | | |
|  | | | | | | | | |
|  | | | | True | | | False | |
| 1. There is no conceptual limit to the size of a list. (1 point) | | | |  | | | + | |
| 2. These lists return the same outputs: (1 point)  ['a', 'b', 'c']  ['c', 'a', 'b'] | | | |  | | | + | |
| 3. BeautifulSoup, and Scrapy are libraries used for web scraping. (1 point) | | | | + | | |  | |
| 4. A NumPy array is a multidimensional array of objects all the same type.  (1 point) | | | | + | | |  | |
| 5. Supervised Learning uses unlabeled data. (1 point). | | | |  | | | + | |
| 6. The elements of tuple can be changed without any extra operation.  (1 point) | | | |  | | | + | |
| 7. In order to stop “while True:” statement, the keyword “break” should be used at some point in the loop. (1 point) | | | | + | | |  | |
| 8. R Squared always decreases when the number of variables increases.  (1 point) | | | |  | | | + | |
| 9. Pairplot is the plot of Matplotlib library. (1 point) | | | | + | | |  | |
| 10. PCA is used to determine the number of clusters in K-Means. (1 point) | | | | + | | |  | |

|  |
| --- |
| PART 1: MULTIPLE CHOICE |
| **Directions: Read each question carefully, and then place a + next to the answer that best fits the question.** |
|  |
| 1. Which functions will return 4 on the list s = [3, 4, 1, 2]? (2 points)   a. Sum(s), Max(s) b. Len(s), Four(s) +  c. Max(s), Len(s) d. Four(s), Max(s) |
| 1. What will be the output when the following Python code run? (3 points)   def perimeter (a, b):  a = 5  b = 3  return (a+b)\*2  perimeter(20,12)  a. 16 + b. 15  c. Nothing d. 64     1. Suppose random selection of sample values is tested. On which measures outliers have highest impact?   (2 points)  a. Mean + b. Mode  c. Median d. None of above     1. When you change the number of rows and columns which gives a new shape to an object. (2 points)   a. Slicing b. Reshape +  c. Linshape d. Archangel     1. …… is a performance metric for Linear Model. (2 points)   a. R-Squared + b. Precision  c. F-Score d. Gini |

2

|  |
| --- |
| PART 1: FILL IN THE BLANK |
| **Directions: Read each question carefully and provide the answer on the line. Your answer must be in complete sentences.** |
|  |
| 1. Write the correct output of the following string operation: (1 point)   str1 = "Welcome"  print(str1[:6] + " Participants") Output: “Welcom Participants” |
| 1. What is result of following Python code? (1 point)   x = 5  y = x + 3 8  x = x – 1 4  z = 10  x = x + z 14  print(x, y, z)  Output = 14, 8, 10       1. Write about bagging and boosting. (3 points)   Bagging is the way to decrease variance of your prediction by generating extra training data with different combinations.    Boosting is a sequential technique, trains the entire set and models are constructed by fitting the residual errors.       1. Write at least 1 difference between Supervised and Unsupervised Learning. (2 points)   We use labeled and past data for training our supervised model, while unsupervised uses unlabeled data and learns from the user on the go.       1. Which model has the best ROC Curve? Write color and explain why. (2 points)     Blue |
|  |

3