Modules
Grades
Inbox 30
Discuss
Calc

Assignment
Week 1 Graded Assignment 1.1
Assignment
Week 1 Graded Assignment 1.2
Assignment
Week 2
Week 3
Week 4
Week 5
Week 6
Week 7
Week 8
Week 9
Week 10

Week 1 Graded Assignment 1.1

The due date for submitting this assignment has passed.
Due on 2024-02-04, 23:59 IST.

You may submit any number of times before the due date. The final submission will be considered for grading.

You have last submitted on: 2024-01-27, 19:28 IST

Click here for the dataset required to solve below questions

Shape of the dataset is (10000, 12) [Please note that if you do not obtain this shape, it may indicate that the data was not uploaded correctly to the Colab environment.]
For simplicity please consider all the unknown values ("**?**") as null values for this assignment.
Consider each question as individual question unless instructed otherwise and use the given csv data to compute all the solutions.

---

1) How many unknown ("?) values are present in the dataset? Remove/Delete unknown ("?) values present in the dataset to make it null value.
Remove/Delete - means it will show NAN in place of "?"
Note: If there is no value present in the dataset it is represented as **NAN**(read pandas documentation for all the other ways to represent null values) in data

5548

No, the answer is incorrect.
Score: 0

Accepted Answers:

(Type: Numeric) 1823

2 points

2) What is the shape of the dataset ? ***2 points***

- A) (10000, 12)
- B) (8985, 13)
- C) (100043, 14)
- D) (100000, 12)

Yes, the answer is correct.
Score: 2

Accepted Answers:

A) (10000, 12)

3) What is the value present at the 692th indexed row and 0th indexed column in the data ? ***2 points***

dataframe[692,0] (simpy saying this in a matrix)
rows/ columns starts indexing from zero(0) in python

- A) 2009-11-16
- B) 2009
- C) 2009-11-17
- D) 2010-10-16

Yes, the answer is correct.
Score: 2

Accepted Answers:

A) 2009-11-16

4) What is the value present at the 546th indexed row and 7th indexed column in the data ?
dataframe[546,7] (simpy saying this in a matrix)
rows/ columns starts indexing from zero(0) in python

3

Yes, the answer is correct.
Score: 2

Accepted Answers:

(Type: Numeric) 3

2 points

5) What are the unique values present in the Locality feature of the dataset? ***2 points***

- A) [Detached House, Duplex, Triplex, Fourplex]
- B) [Bridgeport, Fairfield, Greenwich, Norwalk,Stamford,Waterbury, West Hartford]
- C) [Boston, California, Greenwich, Stamford]

D) [West Hartford, Waterbury, Norwalk]

Yes, the answer is correct.
Score: 2

Accepted Answers:

B) [Bridgeport, Fairfield, Greenwich, Norwalk,Stamford,Waterbury, West Hartford]

6) Which of the following features have missing(NaN) values present in the dataset? (Note: compute after removing "?") **2 points**

- [ ] A) Year
- [x] B) Estimated Value
- [ ] C) Sale Price
- [x] D) Locality
- [x] E) Property
- [ ] F) property_tax_rate

Yes, the answer is correct.
Score: 2

Accepted Answers:

B) Estimated Value
D) Locality
E) Property

7) Which of the following feature has most missing(NaN) values present in the dataset? (Note: compute after removing "?") **2 points**

- ( ) A) Locality
- ( ) B) Estimated Value
- (•) C) Property
- ( ) D) carpet_area

Yes, the answer is correct.
Score: 2

Accepted Answers:

C) Property

8) Drop all the samples(rows) with missing values strictly greater than 2. How many samples remains after that ?

9917

Yes, the answer is correct.
Score: 2

Accepted Answers:

(Type: Numeric) 9917

2 points

9) Drop all the samples(rows) with missing values in the original dataframe . How many samples remains after that ?

5449

Yes, the answer is correct.
Score: 2

Accepted Answers:

(Type: Numeric) 5449

2 points