

Quiz 1 – 9/4/2024

Score

/ 10

Instructions. You have 15 minutes to complete this quiz. You may not use any outside materials. This quiz is closed computer, closed book, closed notes, and closed internet. No collaboration allowed.

Problem 1. Suppose you have read in a DataFrame from a csv file using the following code:

```
df = pd.read_csv('data/flights.csv')
```

Which code produces the following output?

```
(5309,8)
```

- A. `df.dimensions`
- B. `df.head()`
- C. `df.shape`
- D. `df.info()`

Problem 2. Suppose each row of the DataFrame `df` represents a company in the 2024 Fortune 500, and that one of the columns contains the company's revenue in millions of dollars. Which code will quickly provide the minimum, maximum, and mean revenue of the companies in the DataFrame?

- A. `df.summarize()`
- B. `df.info()`
- C. `df.describe()`
- D. `df.statistics()`

Your answer:

C

Your answer:

C

Problem 3. Suppose you have evaluated the following line of code:

Your answer:

```
book = ' pride and sensibility '
```

B

Select the code snippet that produces this output:

```
'Pride And Prejudice'
```

(The string method `title()` capitalizes the first letter in each word and makes all other letters in the string lowercase.)

A.

```
book
.strip()
.replace('sensibility', 'prejudice')
.title()
```

B.

```
(
    book
    .strip()
    .replace('sensibility', 'prejudice')
    .title()
)
```

C.

```
book
.title()
.replace('sensibility', 'prejudice')
```

D.

```
book.strip().title().replace('sensibility', 'prejudice')
```

Note that the code in choices A and B are identical, except for the opening and closing parentheses. However, the code in A will result in an error: without the parentheses, Python doesn't know that the code `book` is continued by `.strip()` on the next line.

Scenario for Problems 4-7. Suppose you run the following code to read in a DataFrame and display the first 5 rows:

```
df = pd.read_csv('data/fastfood.csv')
df.head()
```

The first 5 rows are as follows:

	restaurant	item	calories	fat	sodium	carb
0	mcdonalds	big mac	540	28	950	46
1	mcdonalds	filet-o-fish	380	34	640	38
2	chick fil-a	chicken sandwich	440	19	1350	40
3	sonic	jalapeno burger	640	37	930	42
4	dairy queen	crispy chicken wrap	350	30	1250	59

Now consider the following (partial) code snippet which produces a scatterplot of calories versus sodium (in milligrams):

```
alt.Chart(df).mark_A().B(
    alt.X('calories:C'),
    alt.Y('sodium:D'),
    E
)
```

Problem 4. What should go in the place of the letter A in the code?

Your answer:

D

- A. dot
- B. line
- C. bar
- D. point

Problem 5. What should go in the place of the letter B in the code?

Your answer:

B

- A. properties
- B. encode
- C. define
- D. axes

Problem 6. What should go in the place of the letters C and D in the code? (Choose one answer to replace both letters.)

Your answer:

B

- A. N
- B. Q
- C. T
- D. O

Problem 7. Which code, if placed into position E, would color the scatterplot points according to the grams of fat the **vehicle item** has?

Your answer:

A

- A. alt.Color('fat:Q')
- B. rainbow='fat'
- C. alt.Shade('fat:T')
- D. alt.Properties(color='fat')

Scenario for Problems 8-10. Here are the first 5 rows of the DataFrame again:

	restaurant	item	calories	fat	sodium	carb
0	mcdonalds	big mac	540	28	950	46
1	mcdonalds	filet-o-fish	380	34	640	38
2	chick fil-a	chicken sandwich	440	19	1350	40
3	sonic	jalapeno burger	640	37	930	42
4	dairy queen	crispy chicken wrap	350	30	1250	59

Now consider the following (partial) code snippet which produces a bar graph, with one bar for each restaurant in the DataFrame:

```
alt.Chart(df).mark_A().B(  
    alt.Y('restaurant:C'),  
    alt.X('D'),  
)
```

Problem 8. What should go in the place of the letter C in the code?

- A. N
- B. Q
- C. T
- D. o

Your answer:

A

Problem 9. Which code, if placed into position D, would produce a bar graph in which each bar's length corresponds to the number of items from each restaurant?

- A. count()
- B. mean(item)
- C. number(item)
- D. q1(restaurant)

Your answer:

A

Problem 10. Which code, if placed into position D, would produce a bar graph in which each bar's length corresponds to the average number of calories among the items from each restaurant?

- A. count(calories)
- B. mean(calories)
- C. avg(calories)
- D. sd(calories)

Your answer:

B