

問1. 以下の問いに答えよ。なお、解答の根拠となるipynb/htmlファイルが両方とも提出されていない場合は、解答を無効とする。(30点)

cs3-mid-1.csvを読み込んで一連の解析を行うノートブックの ____ (1) ____ などの空欄を埋め、また問いに答えよ。空欄を埋めるプログラムは1行で答えること。なお、提出するipynb/htmlファイルには、以下に表示されていないプログラム行が含まれていても構わない。

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

cs3-mid-1.csvの全データをデータフレームdfに読み込む (1)。

```
csv_in = 'cs3-mid-1.csv'
__ (1) __
```

(1)

dfの行数(データ数)と列数 (2)、各列のデータ型と非欠損値数 (3)、dfの最初の5行を表示 (4)、また、データ数 (5) と列数 (6) を答えよ。

```
print( __ (2) __ )
print( __ (3) __ )
display( __ (4) __ )
# num of data: (5)
# num of columns: (6)
```

(2)

(3)

(4)

(5)

(6)

dfからn1列を取り出し、ser_n1に代入 (7)、ser_n1の最小値を表示 (8)、また、その数値を (9) に答えよ。

```
__ (7) __
print( __ (8) __ )
# value: (9)
```

(7)

(8)

(9)

dfの各行を、n1列の値の昇順にソートして、先頭5行を表示 (10)、また、この列の3番目に小さな値を (11) に答えよ。

```
display( __ (10) __ )
# value: (11)
```

(10)

(11)

c列に出現する各値とそれぞれの出現回数の一覧表示 (12)、また、値Aの出現回数を (13) に答えよ。

```
print( __ (12) __ )
# value: (13)
```

(12)

(13)

dfからn5列を削除したデータフレームdf2を作成 (14)、df2のデータを、c列の値ごとにまとめ、残りの列の平均値を表示 (15)、c列の値がCのデータの、n1列の平均値を (16) に答えよ(小数第2位)。

```
__ (14) __
display( __ (15) __ )
# value: (16)
```

(14)

(15)

(16)

dfの各行について、n2列とn4列の値の和を求め、それを新たなn_tot列に格納 (17)、n_totの値について降順にソートした結果をdf3に代入 (18)、df3のn_tot列の先頭行の値を (19) に答えよ(小数第2位)。

```
__ (17) __
__ (18) __
# value: (19)
```

(17)

(18)

(19)

df3のn_tot列のヒストグラムを作成 (20)、bin数は8とし、軸ラベルやタイトル、凡例などの装飾はつけずに、1行で答えること。

```
__ (20) __
plt.show()
```

(20)