Data Handling using Pandas – II

Standard difficulty · 20 questions · 40 marks · Time: 39 minutes

Name: _____ Date: _____ Score: ____ / 40

Show your working for Sections B and C. Answers are on the last page — try every question first!

Section A — Choose the correct answer — 8 × 1 = 8 marks

1. Which function gives the middle value of a sorted column?

(a) mode() **(b)** count() **(c)** mean() **(d)** median()

2. DataFrame df:
index | Item | Store | Price | Stock
P1 | Folder | West | 24 | 38
P2 | Scissors | Central | 69 | 62
P3 | Crayons | North | 13 | 44
P4 | Sharpener | Central | 74 | 56
P5 | Glue | Central | 52 | 57
What does print(df['Stock'].median()) show?

(a) 62.0 **(b)** 38.0 **(c)** 56.0 **(d)** 51.4

3. DataFrame df:
index | Name | Dept | Q1 | Q2
E1 | Tara | Support | 84 | 73
E2 | Gita | Design | 82 | 92
E3 | Vikram | Design | 70 | 81
E4 | Asha | Design | 87 | 85
E5 | Rohan | Accounts | 53 | 86
What does print(df['Q1'].mean()) show?

(a) 82.0 **(b)** 87.0 **(c)** 376 **(d)** 75.2

4. DataFrame df:
index | Name | Dept | Q1 | Q2
E1 | Tara | Support | 45 | 78
E2 | Manav | Training | 57 | 86
E3 | Arjun | Support | 54 | 82
E4 | Kabir | Training | 21 | 66
E5 | Isha | Design | 24 | 45
E6 | Sana | Design | 39 | 53
In print(df.sort_values(by='Q2')), which Name is in row 3 of the output?

(a) Arjun **(b)** Sana **(c)** Isha **(d)** Kabir

5. DataFrame df:
index | Name | City | Maths | Science
S1 | Isha | Delhi | 46 | 45
S2 | Zoya | Shimla | 35 | 60
S3 | Pooja | Delhi | 63 | 50
S4 | Farhan | Shimla | 72 | 42
S5 | Imran | Indore | 69 | 74
S6 | Neha | Delhi | 94 | 37
S7 | Karan | Shimla | 88 | 83
What does print(df.groupby('City')['Science'].min()['Shimla']) show?

(a) 74 **(b)** 43 **(c)** 42 **(d)** 37

6. DataFrame df:
index | Item | Store | Price | Stock
P1 | Ruler | West | 60 | 56
P2 | Glue | South | 13 | 57
P3 | Folder | West | — | 15
P4 | Eraser | South | 61 | 88
P5 | Stapler | North | 21 | 77
P6 | Pen | North | 18 | 72
(— marks a missing value; pandas prints it as Not a Number.)
What does print(df['Price'].count()) show?

(a) 5 **(b)** 4 **(c)** 6 **(d)** 1

7. DataFrame df:
index | Name | Dept | Q1 | Q2
E1 | Gita | Accounts | 24 | 26
E2 | Imran | Accounts | — | 77
E3 | Tara | Design | 42 | 44
E4 | Zoya | Support | 25 | 81
E5 | Neha | Support | — | 22
(— marks a missing value; pandas prints it as Not a Number.)
What does print(df['Q1'].isnull().sum()) show?

(a) 0 **(b)** 5 **(c)** 2 **(d)** 3

8. DataFrame df:
index | Name | Dept | Q1 | Q2
E1 | Dev | Accounts | 86 | 55
E2 | Karan | Design | 81 | 32
E3 | Meena | Design | — | 47
E4 | Diya | Accounts | — | 49
E5 | Vikram | Accounts | 87 | 20
E6 | Farhan | Support | 91 | 31
(— marks a missing value; pandas prints it as Not a Number.)
What does print(df.dropna().shape) show?

(a) (4, 3) **(b)** (6, 4) **(c)** (3, 4) **(d)** (4, 4)

Section B — Short answer — 8 × 2 = 16 marks

9. What is descriptive statistics? Name four functions pandas provides for it.

10. DataFrame df:
index | Item | Store | Price | Stock
P1 | Sharpener | Central | 24 | 35
P2 | Folder | Central | 67 | 25
P3 | Scissors | East | 33 | 76
P4 | Ruler | East | 25 | 52
P5 | Notebook | South | 19 | 47
P6 | Stapler | South | 57 | 20
Write the output of:
print(df['Stock'].max() - df['Stock'].min())

11. DataFrame df:
index | Name | City | Maths | Science
S1 | Pooja | Indore | 79 | 39
S2 | Farhan | Pune | 75 | 86
S3 | Meena | Indore | 84 | 55
S4 | Asha | Mysuru | 64 | 73
S5 | Karan | Indore | 73 | 48
S6 | Manav | Indore | 98 | 96
Which values does print(df.sort_values(by='Science').head(2)['Name']) show? Write them in order, separated by commas.

12. DataFrame df:
index | Item | Store | Price | Stock
P1 | Sharpener | South | 81 | 27
P2 | Stapler | North | 55 | 90
P3 | Crayons | West | 37 | 16
P4 | Diary | South | 83 | 28
P5 | Marker | North | 19 | 78
P6 | Pen | West | 26 | 57
P7 | Glue | North | 32 | 50
Write the output of:
print(df.groupby('Store')['Price'].count()['South'])

13. DataFrame df:
index | Name | Dept | Q1 | Q2
E1 | Diya | Training | — | 73
E2 | Farhan | Accounts | — | 53
E3 | Arjun | Accounts | 55 | 60
E4 | Karan | Design | 89 | 37
E5 | Dev | Accounts | 81 | 95
(— marks a missing value; pandas prints it as Not a Number.)
Write the output of:
print(df['Q1'].fillna(0).sum())

14. DataFrame df:
index | Name | Dept | Q1 | Q2
E1 | Zoya | Support | 66 | 45
E2 | Sana | Training | 49 | 67
E3 | Ravi | Sales | — | 65
E4 | Asha | Support | — | 78
E5 | Arjun | Training | 63 | 90
(— marks a missing value; pandas prints it as Not a Number.)
Write the output of:
print(df['Q1'].count())

15. DataFrame df:
index | Item | Store | Price | Stock
P1 | Crayons | Central | 29 | 85
P2 | Eraser | Central | 24 | 17
P3 | Glue | North | 14 | 62
P4 | Pen | North | 80 | 60
P5 | Marker | West | 39 | 80
P6 | Diary | Central | 51 | 78
Write the output of:
print(df.groupby('Store')['Price'].sum())
(Write one "label | value" line for each group.)

16. DataFrame df:
index | Item | Store | Price | Stock
P1 | Crayons | Central | — | 66
P2 | Marker | North | 36 | 57
P3 | Scissors | North | 13 | 80
P4 | Stapler | South | 89 | 33
P5 | Folder | South | 73 | 78
(— marks a missing value; pandas prints it as Not a Number.)
Write the output of:
print(df['Price'].mean())

Section C — Word problems — 4 × 4 = 16 marks

17. Explain any four statistical functions of pandas with examples on a column Marks.

18. DataFrame df:
index | Item | Store | Price | Stock
P1 | Eraser | South | — | 90
P2 | Diary | North | 74 | 50
P3 | Pen | East | 25 | 27
P4 | Notebook | East | 21 | 56
P5 | Folder | East | 60 | 72
P6 | Crayons | North | 45 | 75
(— marks a missing value; pandas prints it as Not a Number.)
Write the output of:
(a) print(df['Price'].count())
(b) print(df['Price'].sum())
(c) print(df.dropna().shape)

19. DataFrame df:

index | Name | Dept | Q1 | Q2
E1 | Isha | Sales | 61 | 79
E2 | Gita | Accounts | 83 | 48
E3 | Rohan | Support | 71 | 64
E4 | Meena | Support | 60 | 93
E5 | Zoya | Accounts | 42 | 72
E6 | Diya | Sales | 45 | 68

Write the output of:

(a) print(df.groupby('Dept')['Q1'].sum())

(b) print(df.groupby('Dept')['Q1'].max())

20. DataFrame df:

index | Item | Store | Price | Stock
P1 | Folder | North | 41 | 46
P2 | Sharpener | East | 54 | 49
P3 | Eraser | East | 30 | 85
P4 | Crayons | North | 83 | 70
P5 | Stapler | South | 27 | 63

Write the output of:

(a) print(df['Price'].max())

(b) print(df['Price'].min())

(c) print(df['Price'].mean())

(d) print(df['Price'].median())

Want another sheet on this chapter? In the Ambiful Worksheets app every sheet is new, with its answer key, and your child never gets the same question twice.
Scan the code to get the free app, or open **worksheets.ambiful.com/get**

Answer key

Data Handling using Pandas – II · AW-12IP-PD2-2026-V1 · For parents — fold back or tear off before handing the sheet to your child.

Q	Answer	Working
1	**(d) median()**	—
2	**(c) 56.0**	sorted: 38, 44, 56, 57, 62; the middle value
3	**(d) 75.2**	376 ÷ 5; mean() always gives a float
4	**(d) Kabir**	ascending order of Q2: Isha, Sana, Kabir, Tara, Arjun, Manav
5	**(c) 42**	the rows with City Shimla
6	**(a) 5**	count() leaves out missing values
7	**(c) 2**	isnull() is True for each missing value; sum() counts the Trues
8	**(d) (4, 4)**	dropna() drops every row that has a missing value
9	**Summarising data by numbers that describe it, such as max(), min(), mean(), median(), mode(), count(), sum(), std(), var() and quantile().**	—
10	**56**	76 − 20
11	**Pooja, Karan**	sorted by Science, lowest first; the first 2 rows
12	**2**	count() over the rows with Store South
13	**225.0**	missing values become 0; the column holds floats, so the sum is a float
14	**3**	missing values are skipped
15	**Central | 104** **North | 94** **West | 39**	one value per group; groupby sorts the group labels
16	**52.75**	missing values are skipped: 211 ÷ 4
17	**df['Marks'].max() gives the highest mark; min() the lowest; mean() the average (sum ÷ count); median() the middle value of the sorted marks; mode() the most frequent; std() the spread around the mean; count() the number of values present.**	one mark for each function
18	**(a) 5; (b) 225.0; (c) (5, 4)**	(a) and (b) skip missing values, and the column is a float column; (c) drops the rows with a missing value
19	**(a)** **Accounts | 125** **Sales | 106** **Support | 131** **(b)** **Accounts | 83**	one line per group, the group labels in sorted order

Q	Answer	Working
	Sales | 61 **Support | 71**	
20	**(a) 83; (b) 27; (c) 47.0; (d) 41.0**	sorted: 27, 30, 41, 54, 83; total 235 over 5 values