FOUNDATION LESSON

WEEK 5 · LESSON 5

Merging and stacking tables safely

A merge matches keys; concat stacks compatible rows.

THE MENTAL MODEL

MENTAL MODEL

The mental model

A merge matches keys; concat stacks compatible rows.

A MANY-TO-ONE LOOKUP

CODE WALKTHROUGH W05-L05-S03

```
01  import pandas as pd
02  orders=pd.DataFrame({"customer":["A","A","B"],"amount":[10,20,3
↪   0]})
03  customers=pd.DataFrame({"customer":["A","B"],"region":["N","S"]
↪   })
04  joined=orders.merge(customers,on="customer",how="left",
05                      validate="many_to_one",indicator=True)
06  print(joined.to_string(index=False))
07  print("Total:", joined["amount"].sum())
```

WHAT TO TRACK

A many-to-one lookup

PREDICT THE RESULT

Predict the output.
Explain each operation.
Pause before continuing.

A MANY-TO-ONE LOOKUP · RESULT

CODE WALKTHROUGH W05-L05-S04

```
01  import pandas as pd
02  orders=pd.DataFrame({"customer":["A","A","B"],"amount":[10,20,3
↪   0]})
03  customers=pd.DataFrame({"customer":["A","B"],"region":["N","S"]
↪   })
04  joined=orders.merge(customers,on="customer",how="left",
05                      validate="many_to_one",indicator=True)
06  print(joined.to_string(index=False))
07  print("Total:", joined["amount"].sum())
```

WHAT TO TRACK

A many-to-one lookup

VALIDATE THE RESULT

```
customer  amount region
_merge
       A      10      N
both
       A      20      N
both
       B      30      S
both
Total: 60
```

STACK TWO DELIVERY BATCHES

CODE WALKTHROUGH W05-L05-S05

```
01  import pandas as pd
02  a=pd.DataFrame({"id":["X"],"amount":[10]})
03  b=pd.DataFrame({"id":["Y"],"amount":[20]})
04  both=pd.concat([a,b],ignore_index=True)
05  print(both.to_string(index=False))
06  print("Rows:", len(both))
```

WHAT TO TRACK

Stack two delivery batches

PREDICT THE RESULT

Predict the output.
Explain each operation.
Pause before continuing.

STACK TWO DELIVERY BATCHES · RESULT

CODE WALKTHROUGH W05-L05-S06

```
01  import pandas as pd
02  a=pd.DataFrame({"id":["X"],"amount":[10]})
03  b=pd.DataFrame({"id":["Y"],"amount":[20]})
04  both=pd.concat([a,b],ignore_index=True)
05  print(both.to_string(index=False))
06  print("Rows:", len(both))
```

WHAT TO TRACK

Stack two delivery batches

VALIDATE THE RESULT

```
id  amount
 X      10
 Y      20
Rows: 2
```

A DUPLICATED LOOKUP KEY

CODE WALKTHROUGH W05-L05-S07

```
01  import pandas as pd
02  o=pd.DataFrame({"customer":["A"],"amount":[10]})
03  c=pd.DataFrame({"customer":["A","A"],"region":["N","S"]})
04  o.merge(c,on="customer",validate="many_to_one")
```

WHAT TO TRACK

A duplicated lookup key

PREDICT THE RESULT

Predict the output.
Explain each operation.
Pause before continuing.

A DUPLICATED LOOKUP KEY · RESULT

CODE WALKTHROUGH W05-L05-S08

```
01 import pandas as pd
02 o=pd.DataFrame({"customer":["A"],"amount":[10]})
03 c=pd.DataFrame({"customer":["A","A"],"region":["N","S"]})
04 o.merge(c,on="customer",validate="many_to_one")
```

WHAT TO TRACK

A duplicated lookup key

VALIDATE THE RESULT

```
pandas.errors.MergeError:
Merge keys are not unique in
right dataset; not a
many-to-one merge
```

GUIDED PRACTICE

LEARNER PRACTICE

Guided practice

Left-merge orders for A and Z with a lookup containing only A. Use indicator and print the unmatched customer.

PREDICT · MODIFY · RUN

01 Restate the supplied rules.

02 Predict the result.

03 Modify and run the code.

04 Compare actual output.

05 Explain any difference.

INDEPENDENT PRACTICE

LEARNER PRACTICE

Independent practice

Concatenate two one-row batches with the same invoice id, then detect the repeated id.

PREDICT · MODIFY · RUN

01 Restate the supplied rules.

02 Predict the result.

03 Modify and run the code.

04 Compare actual output.

05 Explain any difference.

CHECK YOUR UNDERSTANDING

MENTAL MODEL

Check your understanding

Diagnostic: Does a successful left merge guarantee every key matched? No. Inspect the merge indicator.

RECAP AND NEXT STEP

Recap and next step

Declare join cardinality, inspect unmatched rows, and reconcile totals. Use concat to stack batches with compatible schemas.

UNDERSTAND APPLY VALIDATE EXPLAIN