FOUNDATION LESSON

WEEK 5 · LESSON 4

Multiple aggregations and transform

Aggregation reduces groups; transform returns a value aligned with each input row.

THE MENTAL MODEL

MENTAL MODEL

The mental model

Aggregation reduces groups; transform returns a value aligned with each input row.

SEVERAL STATISTICS TOGETHER

CODE WALKTHROUGH W05-L04-S03

```
01  import pandas as pd
02  d=pd.DataFrame({"team":["A","A","B"],"sales":[10,30,20]})
03  s=d.groupby("team").agg(total=("sales","sum"),
04                          average=("sales","mean"),
05                          rows=("sales","size"))
06  print(s.to_string())
```

WHAT TO TRACK

Several statistics together

PREDICT THE RESULT

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

SEVERAL STATISTICS TOGETHER · RESULT

CODE WALKTHROUGH W05-L04-S04

```
01  import pandas as pd
02  d=pd.DataFrame({"team":["A","A","B"],"sales":[10,30,20]})
03  s=d.groupby("team").agg(total=("sales","sum"),
04                          average=("sales","mean"),
05                          rows=("sales","size"))
06  print(s.to_string())
```

WHAT TO TRACK

Several statistics together

VALIDATE THE RESULT

```
      total  average  rows
team
A        40     20.0     2
B        20     20.0     1
```

ORDER SHARES WITHIN A REGION

CODE WALKTHROUGH W05-L04-S05

```
01  import pandas as pd
02  d=pd.DataFrame({"region":["N","N","S"],"sales":[10,30,20]})
03  totals=d.groupby("region")["sales"].transform("sum")
04  d["share"]=d["sales"]/totals.where(totals>0)
05  print(d.to_string(index=False))
```

WHAT TO TRACK

Order shares within a region

PREDICT THE RESULT

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

ORDER SHARES WITHIN A REGION · RESULT

CODE WALKTHROUGH W05-L04-S06

```
01  import pandas as pd
02  d=pd.DataFrame({"region":["N","N","S"],"sales":[10,30,20]})
03  totals=d.groupby("region")["sales"].transform("sum")
04  d["share"]=d["sales"]/totals.where(totals>0)
05  print(d.to_string(index=False))
```

WHAT TO TRACK

Order shares within a region

VALIDATE THE RESULT

```
region  sales  share
     N     10   0.25
     N     30   0.75
     S     20   1.00
```

A ZERO-TOTAL GROUP

CODE WALKTHROUGH W05-L04-S07

```
01  import pandas as pd
02  d=pd.DataFrame({"group":["A","A"],"sales":[0,0]})
03  t=d.groupby("group")["sales"].transform("sum")
04  print((d["sales"]/t.where(t>0)).tolist())
```

WHAT TO TRACK

A zero-total group

PREDICT THE RESULT

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

A ZERO-TOTAL GROUP · RESULT

CODE WALKTHROUGH W05-L04-S08

```
01  import pandas as pd
02  d=pd.DataFrame({"group":["A","A"],"sales":[0,0]})
03  t=d.groupby("group")["sales"].transform("sum")
04  print((d["sales"]/t.where(t>0)).tolist())
```

WHAT TO TRACK

A zero-total group

VALIDATE THE RESULT

[nan, nan]

GUIDED PRACTICE

LEARNER PRACTICE

Guided practice

For group A values 2 and 6, attach each row’s group mean with transform.

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

Compare size and count for amounts [5, None] in group A.

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 transform remove repeated group labels? No. Its result remains aligned with the input rows.

RECAP AND NEXT STEP

Recap and next step

Use agg for one summary per group and transform for aligned per-row statistics.
Guard denominators and name counts precisely.

UNDERSTAND APPLY VALIDATE EXPLAIN