FOUNDATION LESSON

WEEK 5 · LESSON 1

Missing data and deliberate policies

Missing means unknown until a stated business rule justifies a replacement.

THE MENTAL MODEL

MENTAL MODEL

The mental model

Missing means unknown until a stated business rule justifies a replacement.

COUNT MISSING VALUES

CODE WALKTHROUGH W05-L01-S03

```
01  import pandas as pd
02  sales = pd.Series([100.0, None, 0.0])
03  print(sales.isna().tolist())
04  print("Missing:", sales.isna().sum())
05  print("Observed mean:", sales.mean())
```

WHAT TO TRACK

Count missing values

PREDICT THE RESULT

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

COUNT MISSING VALUES · RESULT

CODE WALKTHROUGH

W05-L01-S04

```
01  import pandas as pd
02  sales = pd.Series([100.0, None, 0.0])
03  print(sales.isna().tolist())
04  print("Missing:", sales.isna().sum())
05  print("Observed mean:", sales.mean())
```

WHAT TO TRACK

Count missing values

VALIDATE THE RESULT

```
[False, True, False]
Missing: 1
Observed mean: 50.0
```

A REQUIRED INVOICE AMOUNT

CODE WALKTHROUGH W05-L01-S05

```
01  import pandas as pd
02  df = pd.DataFrame({"id": ["A", "B", "C"],
03                     "amount": [100, None, 0]})
04  accepted = df.loc[df["amount"].notna()].copy()
05  rejected = df.loc[df["amount"].isna()].copy()
06  print("Accepted:", accepted["id"].tolist())
07  print("Rejected:", rejected["id"].tolist())
08  print("Total:", accepted["amount"].sum())
```

WHAT TO TRACK

A required invoice amount

PREDICT THE RESULT

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

A REQUIRED INVOICE AMOUNT · RESULT

CODE WALKTHROUGH W05-L01-S06

```
01  import pandas as pd
02  df = pd.DataFrame({"id": ["A", "B", "C"],
03                     "amount": [100, None, 0]})
04  accepted = df.loc[df["amount"].notna()].copy()
05  rejected = df.loc[df["amount"].isna()].copy()
06  print("Accepted:", accepted["id"].tolist())
07  print("Rejected:", rejected["id"].tolist())
08  print("Total:", accepted["amount"].sum())
```

WHAT TO TRACK

A required invoice amount

VALIDATE THE RESULT

```
Accepted: ['A', 'C']
Rejected: ['B']
Total: 100.0
```

ALL VALUES ARE ABSENT

CODE WALKTHROUGH W05-L01-S07

```
01  import pandas as pd
02  values = pd.Series([None, None], dtype="float64")
03  print("Observed:", values.count())
04  print("Guarded sum:", values.sum(min_count=1))
```

WHAT TO TRACK

All values are absent

PREDICT THE RESULT

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

ALL VALUES ARE ABSENT · RESULT

CODE WALKTHROUGH W05-L01-S08

```
01  import pandas as pd
02  values = pd.Series([None, None], dtype="float64")
03  print("Observed:", values.count())
04  print("Guarded sum:", values.sum(min_count=1))
```

WHAT TO TRACK

All values are absent

VALIDATE THE RESULT

```
Observed: 0
Guarded sum: nan
```

GUIDED PRACTICE

LEARNER PRACTICE

Guided practice

Use [50, None, 0, 70]. Print missing count and observed mean. Keep the known zero.

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

Create amounts [None, None]. Print a total only if at least one amount is observed; otherwise print No observations.

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 fillna(0) merely change the display, or does it assert that every unknown value is a known zero? It changes the analytical assumption.

RECAP AND NEXT STEP

Recap and next step

Count missing values before changing them. Preserve known zeroes, document exclusions, and guard totals when nothing was observed.

UNDERSTAND APPLY VALIDATE EXPLAIN