

Join card

The four join types with a facility example, and the checks that prove a join.

Join	Keeps	Use it when
inner	Rows that matched on both sides	You only want matched rows and accept silent loss
left	Every left row, missing where the right had nothing	Adding a denominator to a table you trust
right	Every right row	The mirror of left
outer	Everything from both sides	Finding what did not match on either side

Prove the join

```
before = len(services)
```

```
merged = pd.merge(services, pop, on="woreda_code", how="left", validate="many_to_one")
```

```
assert len(merged) == before
```

```
merged["expected_pregnancies"].isna().sum() # unmatched keys
```

```
merged[merged["expected_pregnancies"].isna()]["woreda_code"].unique()
```

- Normalise keys with `str.strip().str.lower()` before joining, and join on codes where they exist.
- Keep any hand mapping in a reviewable file and report how many keys were mapped by hand.
- Never edit the raw export to make a join work.

Based on the ERAS Africa Advanced Data Analysis Using Python course, 2026, WHO surveillance guidance, the EPHI public health emergency management guidelines, the STROBE statement and the pandas, SciPy and statsmodels documentation. Free to copy and adapt with attribution.