MODELING · PARTIAL POOLING · CHEAT SHEET

Hierarchical models: how much to pool

Three ways to estimate nineteen hitters from 45 at-bats each. Complete pooling says everyone is the league. No pooling says everyone is their own start. Partial pooling lets the data decide how far to pull.

Python Compare the three on how well they predict the rest of the season.

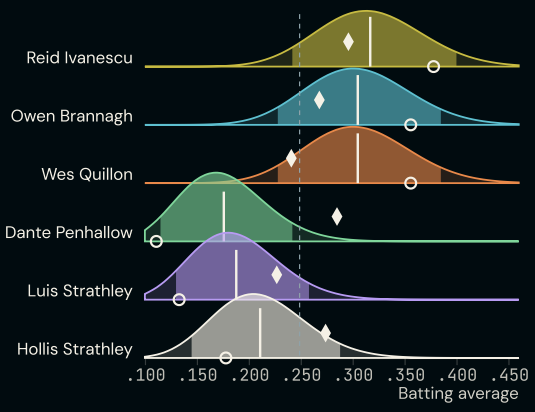
```
import pandas as pd

h = pd.read_csv("data/batting_2026.csv")
league = h.hits.sum() / h.at_bats.sum()
noise = (h.hits / h.at_bats * (1 - h.hits / h.at_bats) /
h.at_bats).mean()
k = league * (1 - league) / max((h.hits / h.at_bats).var()
- noise, 1e-5) - 1

est = h.assign(
    complete=league,
    no_pooling=h.hits / h.at_bats,
    partial=(h.hits + league * k) / (h.at_bats + k),
    rest=h.remaining_hits / h.remaining_at_bats,
)
errors = {c: ((est[c] - est.rest) ** 2).mean() for c in
["complete", "no_pooling", "partial"]}
print(f"prior worth {k:.0f} at-bats")
for name, e in errors.items():
    print(f"{name:<11} mean squared error {e:.5f}")
```

Ran when this sheet was built. Output:

```
prior worth 84 at-bats
complete    mean squared error 0.00127
no_pooling  mean squared error 0.00583
partial     mean squared error 0.00152
```



90% interval · First 45 at-bats (circle) · Rest of season (diamond) · League (dashed)

Each hitter's posterior batting average after 45 at-bats, shaded to the 90% credible interval; the diamond is what they hit the rest of the season. Data: Hot-start hitters (batting_2026.csv). Simulated for practice by AthlyticZ (agents/sim_data.py, seed 20261109). License CC BY 4.0.

Complete pooling
One estimate for everyone. Low variance, ignores real differences.

No pooling
Each unit on its own. Unbiased, far too noisy with small samples.

Partial pooling
A population prior; small samples pull toward the group, large ones keep their own value.

Centered form
Write each unit as drawn from the population. Fine when every unit has plenty of data.

Non-centered form
Write each unit as population mean plus scaled offset. Use it when units have little data and the sampler reports divergences.

From Hierarchical Modeling using PyMC in Baseball. Every code block on this sheet ran when it was built; the data is simulated for practice.