Fraud Example

Average transaction size: $750

Total fraud losses: $760,500

All Data

Fraud	500	5%
Not Fraud	9,500	95%
Total	10,000	100%

Fraud	350	4%
Not Fraud	8,200	96%
Total	8,550	100%

Small Transactions

Large Transactions

Fraud	150	10%
Not Fraud	1,300	90%
Total	1,450	100%

Average transaction size: $30

Total fraud losses: $10,500

Average transaction size: $5,000
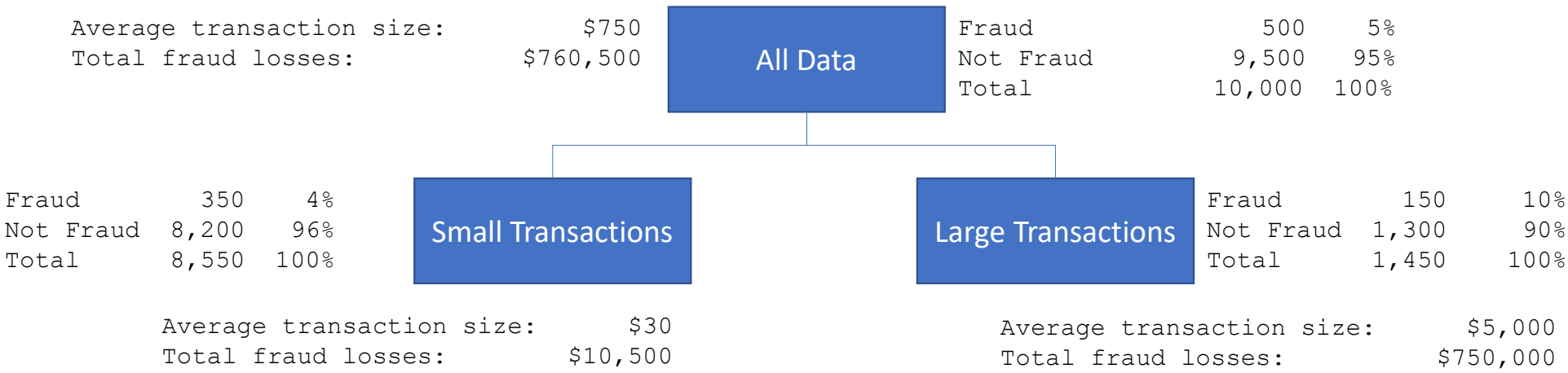
Total fraud losses: $750,000

Cost to review potential fraud case = $150

Baseline:
Benefit from fraud capture: $ 760,500
Cost to review all cases: $ 1,500,000

Net Benefit: $ -739,500

Large Transactions:
Benefit from fraud capture: $ 750,000
Cost to review all cases: $ 217,500

Net Benefit: $ +532,500

Entropy increases in the large transaction node.
Profitability occurs post-split.
Information value is insufficient.

Root causes:
--- Imbalanced dataset
--- Financial model not 1:1 with purity of nodes