Dataset difficulty ranking and characteristics

Most difficult: nasa-FD002
Easiest: healthcare
Avg difficulty: 0.220

- healthcare — #5 — 0.872 / 0.829
- financial — #4 — 0.858 / 0.824
- eligibility — #3 — 0.791 / 0.707
- nasa-FD001 — #2 — 0.696 / 0.629
- nasa-FD002 — #1 — 0.680 / 0.567

Legend:
- Average performance
- Ease score (1 - difficulty)
- Performance variability

X-axis: Score (0-1 scale)