OpenAI Model Cheat Sheet: o1 vs o3 vs o4-mini

Summary of Model Strengths

Model	Strengths	Ideal Use Cases	Weaknesses
o1	Fast, conversational, lightweight	Quick summaries, structured answers, basic charts	Shallow reasoning, no live data, approximate values
o3	Deep research, long reasoning chains, real-time search, accurate data	Research, schedules, coding, scientific analysis, data synthesis	Slightly slower, more costly on high settings
o4-mini	Very efficient, smarter than o3-mini and o1, best cost-performance ratio	Mobile/dev work, fast data parsing, code evaluation, CLI tools	Currently has slightly less tool access than o3

Use These Models For:

o3 (High Reasoning + Real-Time)

- Building **itineraries** or schedules from real-world images (e.g. Puy du Fou planning)
- Analyzing **MLB rule changes**, **scientific breakthroughs**, **vehicle performance trends**
- Creating Python **visualizations**, summarizing data across sources
- Real-time fact-checking, **WebGPT-style research**, coding help w/ citations
- Complex workflows (Notion triggers, table mappings, JSON integrations)
- Researching **philosophy, psychology, science**, policy

o4-mini (Efficient Reasoning)

- Terminal + **CLI dev agent work** (Codex CLI)
- Efficient **API-driven projects** or cost-sensitive coding tasks
- Fast local use with low latency and high accuracy
- Code audits, script validation, low-resource environments

⚠ Use o1 only for:

- Fast structured explanations
- Lightweight response formats
- Basic outline + filler content generation

Previously Scanned Conversations – When o3/o4-mini Should've Been Activated

1. Battery Breakthrough Analysis → o3

Should've used o3 for: - Gathering CATL, DOE, IEA data live - Creating actual charts + comparing Leaf vs Tesla

2. MLB Pitch Clock Rule → o3

Should've used o3 for: - Fetching ERA, WHIP, SO9, BB9 from Baseball Reference - Explaining stolen base context (bases + pickoff limits) - Visualizing multi-year trend

3. Puy du Fou Schedule Planning → o3

Already used o3 perfectly: - Extracted show durations + buffers - Built time-efficient plan from 12:00 to end of day

4. Cost vs Performance Graph Interpretation → o3

Used for: - Explaining tradeoffs between AIME score and cost - Understanding why o3 outperforms o1

5. Human Nature Research Prompt → o3 or o4-mini

Ideal for: - Creating a reasoning-based report with empirical citations - Pulling cross-cultural and biological data

6. Security & Privacy Model Audit → o4-mini

Would be useful for: - Efficient scanning of local documents - Terminal-level document reasoning

7. Codex CLI or Self-Awareness YAML Planning → o4-mini

Great fit: - Low-cost agent reasoning directly in dev shell - Reading image data + sketch logic from folder

---

Let me know if you want to append this cheat sheet into Notion, export it as PDF, or generate a ZIP with visual explanations and graphs!