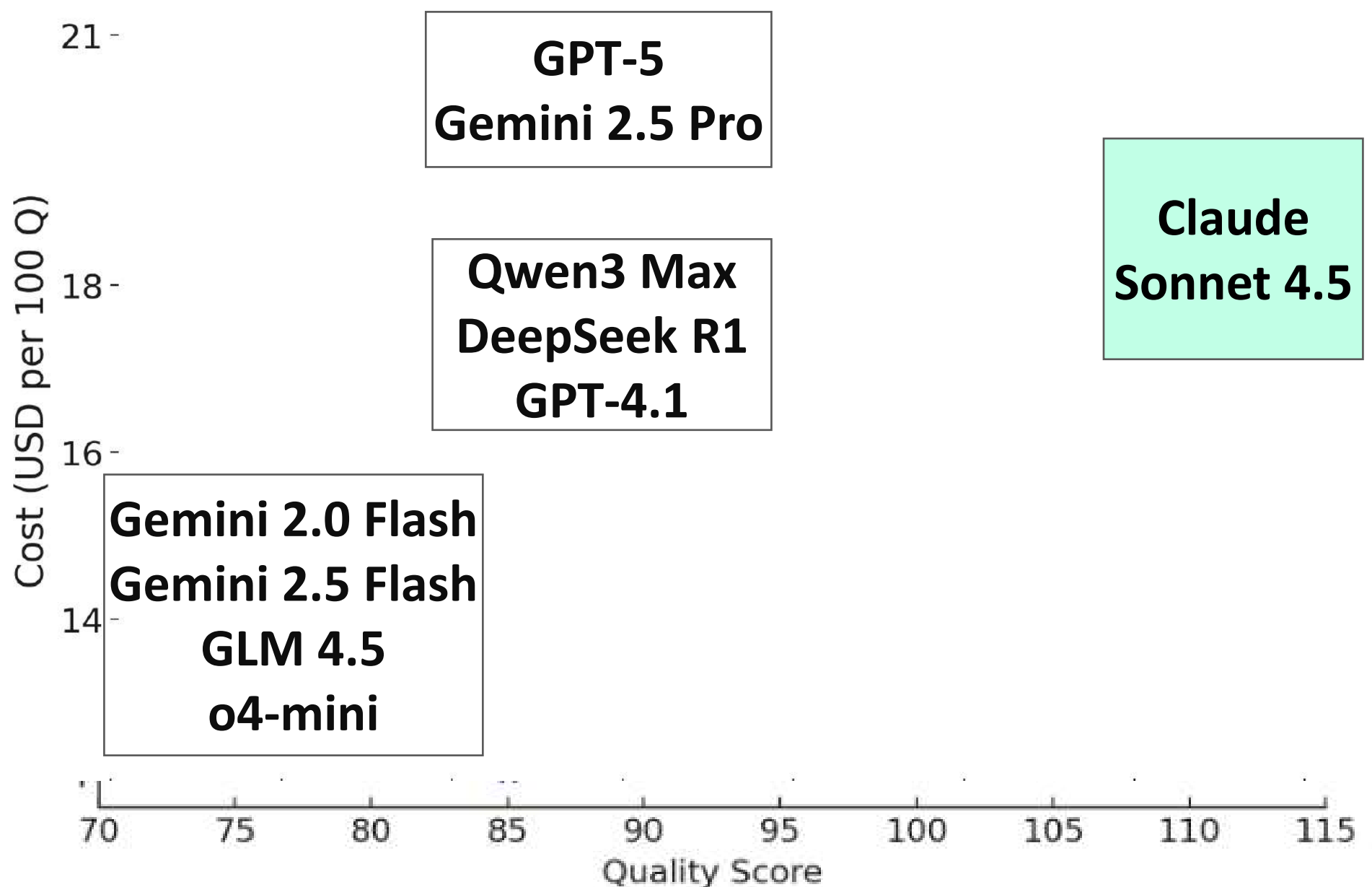


Claude Sonnet 4.5

Cost vs Quality in Database AI

1. Use Claude Sonnet-4.5 for complex reasoning.
2. Gemini 2.5 Flash for value
3. GPT-4.1 / DeepSeek/ Qwen3 Max for balance
4. Avoid GPT-5 and Gemini 2.5 Pro

Reasoning Based Advanced Analysis



~USD per 100 Questions

Direct from DATS-4 app

Live model selector with cost and quality guides

The screenshot shows the DATS-4 Database AI Suite interface. A modal dialog titled "Choose Advanced Analyst Agent Framework" is open, displaying a table of AI models. The table has five columns: Reasoning Model, Type, Quality, Cost, and Select. The "Gemini 2.0 Flash" model is selected, indicated by a blue checkmark in the "Select" column. Below the table, there is a note: "Cost varies based on use case and AI setup. Refer attached price study for more details." and a button labeled "Open Price Guide". At the bottom of the dialog are "Cancel" and "Confirm Selection" buttons.

Reasoning Model	Type	Quality	Cost	Select
Gemini 2.0 Flash	Best Value	75	Lowest	<input checked="" type="checkbox"/>
Claude 4.5 Sonnet	Topmost Quality	115	High	<input type="checkbox"/>
Qwen3 Max	Good	80	Low	<input type="checkbox"/>
GPT-5	Top Quality	95	Volatile	<input type="checkbox"/>
Deepseek-R1-0158	Great Quality	90	Med	<input type="checkbox"/>
Gemini 2.5 Flash	Good	75	Low	<input type="checkbox"/>
GPT-4.1	Great Quality	90	Med	<input type="checkbox"/>
GLM 4.5	High Variances	80	Low	<input type="checkbox"/>
Gemini Pro 2.5	Avoid	85	Highest	<input type="checkbox"/>

Cost varies based on use case and AI setup. Refer attached price study for more details.

Open Price Guide

Cancel **Confirm Selection**

LLM Costs : Advanced Analyst

- In multi-step reasoning-based analysis, execution cost is biggest chunk due to multiple tool calls.
- 1 Question = 7-10 SQL Queries / Tool Calls.
- All executions by GPT-4.1.
- Time: ~2-3 mins per que. Can go upto 10m+. Varies by question.

~USD per 100 Questions

Reasoning Model	Quality Score	Logic USD	Exec. USD	Total USD	Remarks
Gemini Flash 2.0	75	0.25	12.5	12.75	Best value
Gemini Flash 2.5	75	1.75	12.5	14.25	Next after Flash 2.0
Gemini Pro 2.5	85	8.50	12.5	21.00	Avoid. V.High.Cost
Claude Sonnet 4	100	6.50	12.5	19.00	Top Quality
Claude Sonnet 4.5	115	6.50	19.5	19.00	Topmost Quality
DeepSeek R1	85	2.25	12.5	14.75	Great Value
Qwen3 Max	90	2.00	12.5	14.50	Great Value
GLM 4.5	80	1.00	12.5	13.50	High variances
o4-Mini	75	2.75	12.5	15.25	Avoid.
GPT-4.1	90	3.00	12.5	15.50	Great Value
GPT-5	95	15.0	12.5	27.50	Avoid. V. High Cost

Cost Estimates Update : Claude Sonnet 4.5 as of 30th Sep 2025; rest assessed on 9th Sep 2025.

Advanced Analyst : Task Tested

1. Modelling Data Mart creation.

Using 1M Customer Table and a 10M Transaction Table. Summarizing, creating innovative derived variables and merging. Requiring multi-step execution. Tables on AWS RDS MySQL

2. Weighted Score-Based Ranking

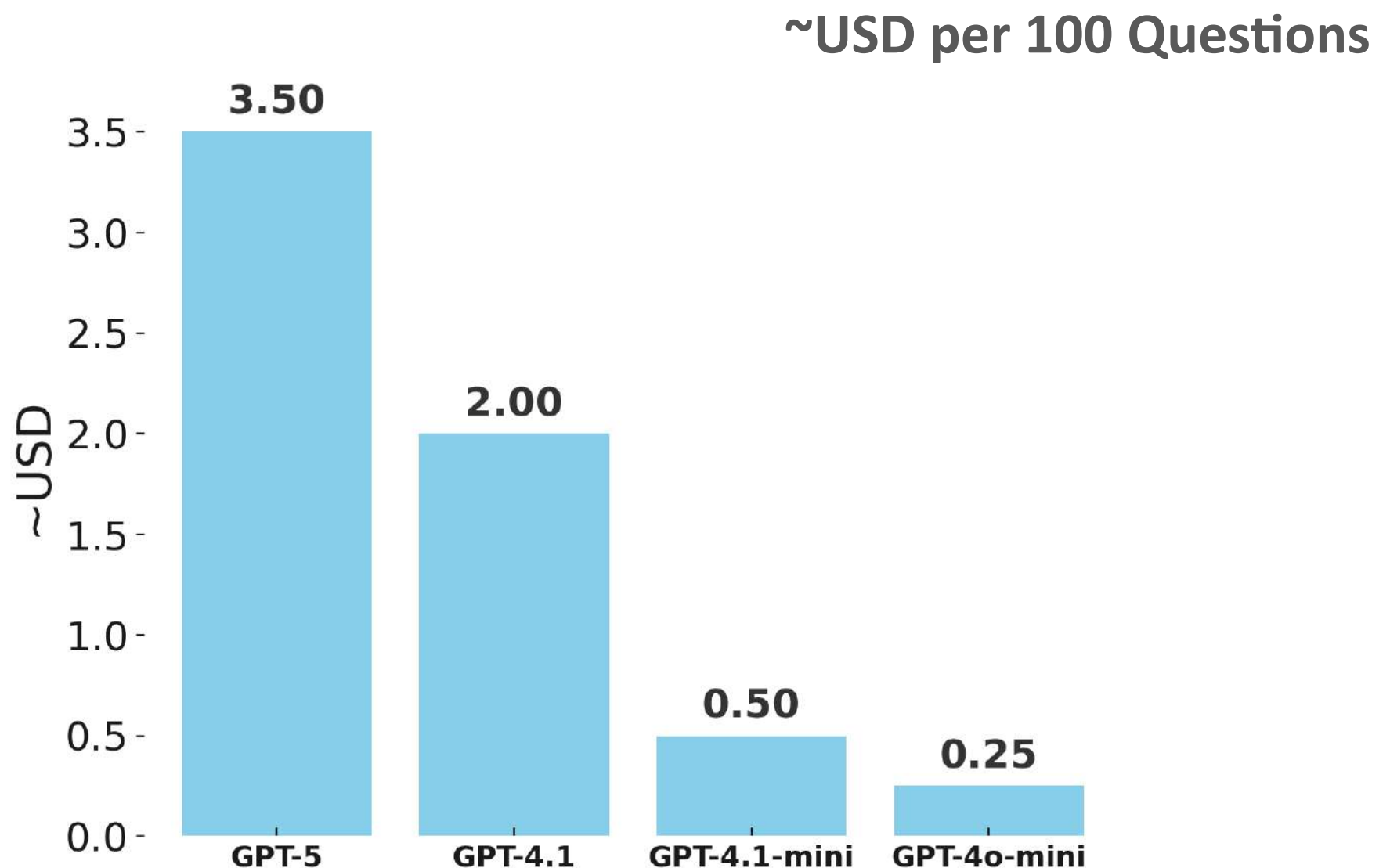
Ranking of Indian credit card issuers - using live RBI Monthly Statistics data. Multi step processing requiring defining scoring methodology, innovative derived variables creation and execution. Table on Supabase

Estimates based on 250+ test run across multi-step advanced analysis and as well as single step general analysis tasks. Always check against your specific use cases.

General Analyst: Single Step SQL Agent

1. OpenAI models most efficient for SQL executions with tool use
2. Avoid GPT-5 – very high cost
3. Choose between GPT-4.1, 4.1-mini and 4o-mini depending upon task complexity

In my own live usage, about 80% of work runs through single-step agents. Use single-step wherever possible. Multi-step - reserve for the tough analysis cases.



LLM Costs: General Analyst

Cost Per 100 Simple Questions

1 Question = 1 SQL Queries/ Tool Call

Single step: no reasoning step, direct execution

~USD per 100 Questions

LLM	~USD	Remarks
GPT-5	3.50	High Cost. Avoid.
GPT-4.1	2.00	Best for complex SQLs
GPT-4.1-mini	0.50	Great for med. complexity SQL
GPT-4o- mini	0.25	Great for simple/ med. complex

General Analyst : Task Example

1. Add new columns
2. Join Tables
3. Summarize
4. Create Chart

- Actual vary by use case and the agent setup. Always estimate for your use cases and compare vs. actuals.
- For all estimates – keep in mind that as context increases the cost goes higher

Cost Estimates Update : Claude Sonnet 4.5 as of 30th Sep 2025; rest assessed on 9th Sep 2025.

DATS-4 : Live

app.tigzig.com

→ Database AI & SQL Apps → DATS-4
Database AI Suite

The screenshot shows the landing page of the app.tigzig.com. The browser address bar displays 'https://app.tigzig.com/landing-preview'. The navigation bar includes a hamburger menu, a 'TIGZIG: Co-Analyst' logo, an 'OSS' badge, and several dropdown menus: 'Analyze', 'Tools', 'GPTs', 'xlwings Lite', 'MCP', and 'Voice'. Below the navigation is a search bar with the placeholder text 'Search tools...'. The main content area is titled 'Recent Updates & Featured Apps' and features three columns of featured content:

- Database AI & SQL Apps**
Complete Database AI Suite
Comprehensive database AI tools: DATS-4, ChatGPT connections, Flowise UI, Voice AI
 - DATS-4 Database AI Suite
 - ChatGPT Database Connections
 - Flowise UI Rapid Deploy
- xlwings Lite: Practice Lab**
Learn xlwings Lite with Practical Examples
A Hands-on guide to get up and running fast with xlwings Lite
 - AI-ready instructions & examples
 - Step-by-step learning guides
 - Interactive workbooks
- Quants & Portfolio Apps**
Run Portfolio Analytics
Quants and Portfolio analysis tools : Performance reports, technical analysis and more
 - Security Performance Report
 - AI Technical Analysis
 - QuantStats Report

- All credentials pass through backend – use test data only
- Free to try
- For live workloads, deploy on your VPN