INTERACTIVE ANALYTICS PLATFORM FOR
INDIAN STARTUP ECOSYSTEM

SHARK CATALYST

A DATA-DRIVEN PROJECT CASE
STUDY OF SHARK TANK INDIA

PROJECT BY :

Sumer Pandey

in | ⬤

DEALS TO DATA

634 Pitches. 360 Deals
₹262.85 Cr Invested
10+ sectors. 7 sharks

EXECUTIVE SUMMARY

PROJECT OVERVIEW

Shark Catalyst analyzes all 4 Shark Tank India seasons, offering deep insights into deals, investors, and founder dynamics. This comprehensive platform transforms raw pitch data into actionable intelligence via visual dashboards, trend analysis, and an equity calculator, blending data storytelling with VC and fintech analytics.

KEY IMPACT

Data Coverage: 634 pitches analyzed across 4 seasons
Investment Landscape: 360 funded deals (56.8%)
Capital Flow: ₹262.85 Cr total investment analyzed
Average Ticket Size: ₹73.6 Lakhs per deal
Sector Insights: 10+ industry verticals mapped

VALUE PROPOSITION

For Founders: Data-driven fundraising strategies with comparable deal analysis and valuation benchmarks
For Investors: Portfolio intelligence with co-investment network mapping and sector concentration analysis
For Ecosystem: Transparent market insights driving informed decision-making

ACTIONABLE

Founders: Match your ask to investor avg tickets and emphasize unit economics.
Investors: Use the co-investor network and HHI metrics to detect concentration and co-partnership opportunities.

WHO IS THIS FOR?

Shark Catalyst translates advanced data engineering into user-centric features, providing founders and investors with actionable market intelligence and shareable ecosystem insights.

VC Analysts & Associates: Investment research and due diligence support

Early-Stage Founders: Fundraising strategy and market positioning

Product Managers: Data product inspiration and technical implementation

Ecosystem Stakeholders: Market transparency and trend analysis

Recruiters: Technical capabilities demonstration for data roles

PRODUCT SUITE

1. Interactive Analytics Platform
- Home Dashboard: Executive KPIs and trend analysis
- Myth Buster: Evidence-based hypothesis testing
- Investor Profiles: Deep-dive analytics with network mapping
- Deal Explorer: Advanced search and comparison tools
- Equity Calculator: Financial modeling and cap table simulation
- Trends & Insights: Market intelligence & concentration metrics
- Methodology Hub: Technical documentation and reproducibility

2. Database & ETL Infrastructure
- Normalized analytics database
- Clean data exports
- ETL orchestration scripts with validation checks

3. SQL Analytics Library
- 10 public SQL files for community use
- 17 internal optimization procedures
- Advanced techniques: CTEs, window functions, network analysis

4. Documentation
- Case Study Repo
- One Pager
- GitHub repository with full source code
- Local deployment instructions

PRODUCT WALKTHROUGH

EXECUTIVE DASHBOARD

Purpose: Strategic overview for C-suite and senior stakeholders

Key Features:

Real-time KPI cards: Pitches, funding success rate, and capital deployed metrics.

Investment trends: Visualizing growth and trajectory

Sector allocation: Capital and concentration analysis

Top investor leaderboard: Activity and performance metrics

Business Impact: Enables quick market pulse assessment for strategic planning and investment committee briefings

MYTH BUSTER

Purpose: Evidence-based hypothesis testing for common market assumptions

Key Features:

Market hypotheses: 20 hypotheses with SQL-driven analysis

Statistical methodology transparency with query visibility

Interactive verdict system: Confidence levels and calculated effect sizes

Strategic Applications: Challenges conventional wisdom with data-driven evidence, initiates ecosystem discussions

INVESTOR INTELLIGENCE

Purpose: Comprehensive investor analytics for co-investment strategy development

Key Features:

Investor scorecards: Individual metrics indicating comprehensive performance
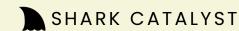
Sector analysis: Investor preference breakdown with diversification data

Pitch recommendations: Personalized insights via investor behavioral analysis

Strategic Value: Enables targeted investor outreach, and provides competitive intelligence for fundraising preparation

Sumer Pandey   SHARK CATALYST

PRODUCT WALKTHROUGH

DEAL EXPLORER

Purpose: Comprehensive deal research and comparable company analysis

Key Features:

Advanced filtering: Multi-parameter system (20+ criterion)

Historical deal timeline: Outcome tracking and follow-up performance

Deal recommendations: Weighted similarity algorithms

Deal details: Comprehensive financial and strategic metrics

Use Cases: Competitive benchmarking, valuation analysis, market sizing, and investor due diligence preparation

MARKET TRENDS

Purpose: Macro market analysis and strategic trend identification

Key Features:

Season-wise Performance: Growth rate for key metrics

HHI Analysis: Herfindahl-Hirschman Index for market concentration assessment

Investment Velocity: Momentum indicators for market timing

Strategic Applications: Challenges conventional wisdom with data-driven evidence, initiates ecosystem discussions

FINANCIAL MODELING SUITE

Purpose: Interactive financial modeling for negotiation preparation and scenario planning

Key Features:

Pre/post-money valuation calculations

Cap table simulation with dilution modeling across multiple rounds

Comparable deal benchmarking for valuation validation

Mathematical Foundation:

Post-money valuation = [Investment Amt.] ÷ [Equity %]

Pre-money valuation = [Post-money Val] - [Investment Amt.]

Founder dilution = [Investment Amt.] ÷ [Post-money Val]

in | ○ Sumer Pandey

SHARK CATALYST

CORE FINDINGS

FUNDING ECOSYSTEM OVERVIEW

The Indian startup ecosystem demonstrates remarkable resilience and growth potential:

Total Capital Deployed: ₹262.85 Crores across analyzed deals

Success Metrics: 56.8% funding success rate exceeding global early-stage averages

Average Deal Size: ₹73.6 Lakhs indicating healthy early-stage market maturity

Investor Participation: 18 active investors with diversified portfolio strategies

INVESTOR BEHAVIOUR INTELLIGENCE

Top Investors by Activity & Capital:
1. Aman Gupta (141 deals, ₹54.54 Cr deployed)
2. Peyush Bansal (₹40.73 Cr across 104 transactions)

Investment Strategy Insights:

Portfolio Construction: Average of 46 deals per active investor annually

Sector Diversification: Top investors maintain exposure across Beauty, Food, Healthcare, Tech sectors

SECTORAL CAPITAL ALLOCATION

Top Sectors:
1. Beauty & Fashion: ₹59.38 Cr
2. Food & Beverage: ₹53.67 Cr
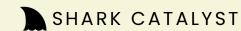3. Healthcare: ₹33.83 Cr
4. Tech: ₹28.87 Cr
5. Manufacturing: ₹16.31 Cr

Strategic Sector Insights:

For Founders: Focus on sectors with 15-25% market share for optimal competition/opportunity balance

For Investors: Diversification opportunities exist in underrepresented verticals with strong fundamentals

LIMITATIONS, ASSUMPTIONS & ETHICAL CONSIDERATIONS

DATA LIMITATIONS & COVERAGE

- Temporal & Scope: Analysis is limited to publicly available TV show data across seasons.
- Financial Gaps: ≈15% of pitch records have incomplete financial data.
- Tracking: Limited follow-up data on post-deal company performance.
- Bias: TV format introduces selection bias, not fully representative of the broader ecosystem.
- Geography: Focuses on the Indian market; limited international comparability.

STATISTICS & METHODOLOGY PRINCIPLES

- Representativeness: Core Assumption: TV participants reasonably proxy early-stage startups. Limitation: Format bias may attract different founder profiles.
- Missing Data: Strategy: Median imputation used for missing financials. Bias Risk: Analytical bias possible if missing data is systematic; all imputation decisions are documented.
- Causation: Findings are correlational, not causal. Recommendations are framed as insights needing external validation due to unobserved factors.

ETHICAL FRAMEWORK & PRIVACY STANDARDS

- Data Handling: Zero PII retained. Analysis restricted to public TV data and original consent context.
- Algorithmic Fairness: Gender/demographic classifications are algorithmic estimates and audited for bias.
- Responsible AI: Methodology is fully documented, reproducible (Open Source), with clear reasoning and confidence levels.
- User Agency: Mechanisms exist for subjects to request data removal and initiate accuracy correction processes.

LEGAL & TECHNICAL CONSTRAINTS

- IP & Copyright: No proprietary show content is used. Fair use, Kaggle licensing, and contributor attribution are fully met. Commercial use is restricted.
- Regulatory Alignment: GDPR and Indian data laws are systematically applied. Comprehensive logging maintains an audit trail.
- Market Assumptions: Assumes temporal and regulatory stability across the study period.
- Technical Limits: SQLite limits concurrent access. No real-time data integration is available (static dataset only).