Data-Driven Management Using Business Analytics: The Case Study of Data Sets for New Business in Tourism

Aarish Sachdeva

Suncity School, Suncity Township, Sector 54, Gurugram, India

DOI:10.37648/ijtbm.v14i03.005

[1]*Received: 21 May 2024; Accepted: 18 August 2024; Published: 22 August 2024*

ABSTRACT

Tourism entrepreneurship has become increasingly data-intensive. This paper synthesizes peer-reviewed work from 2013–2023 to propose a practical, data-driven management playbook for a *new* tourism venture. We review major data sources (user-generated content, search and social signals, mobile positioning, OTA/airbnb supply & price data, and official statistics) and the analytics they enable (demand forecasting, location and product design, pricing, and reputation management). We then construct a case study for a hypothetical startup—*TourStart*—launching curated neighbourhood experiences in a beach destination. Using a portfolio of feasible, real-world data sets and methods (e.g., Google Trends for demand signals, TripAdvisor review mining for product features, Airbnb/OTA data for capacity & pricing, and—where available—mobile positioning for flows), we outline end-to-end decisions: market selection, seasonality calibration, micro-location choice, itinerary design, price setting, and marketing mix. Comparative analysis shows complementary strengths and biases across data sets: search data is timely but volatile; reviews are rich but biased toward vocal users; mobile data is granular but regulated; OTA/supply data captures competition but can contain systematic errors if not validated. We close with implementation guidance (data pipelines, KPIs, and experimentation), limitations (privacy, survivorship bias, platform shocks), and a research agenda. The contribution is a reproducible blueprint that aligns *business analytics capability* with entrepreneurial decisions in tourism.

1. Introduction

Tourism isn't run on gut feeling anymore. Travelers leave clues everywhere—what they search, the photos and reviews they post, the places their phones pass through, and the bookings they make. Taken together, these traces let businesses see demand and experience almost in real time. Researchers now group this "big data" into a few easy buckets—user-generated content, mobile/device data, and transaction data—each helping with different choices, from forecasting demand to designing on-the-ground experiences.

To turn those clues into results, a company needs more than tools. It needs **business analytics capability**—the skill to find the right data, make sense of it, and act on it. Firms that build this capability and connect it to their strategy tend to perform better. For a new tourism venture, that means using data to shape product–market fit—like clay in a designer's hands—not just to produce reports after the fact.

---

[1] *How to cite the article:* Sachdeva A (August 2024); Data-Driven Management Using Business Analytics: The Case Study of Data Sets for New Business in Tourism; *International Journal of Transformations in Business Management,* Vol 14, Issue 3, 45-51, DOI: http://doi.org/10.37648/ijtbm.v14i03.005

(IJTBM) 2024, Vol. No. 14, Issue No. III, Jul-Sep e-ISSN: 2231-6868 p-ISSN: 2454-468X

2. Related Work: What We Know

Data families and typical uses.

- **UGC & Reviews.** Systematic reviews show sentiment/topic mining from TripAdvisor and similar platforms supports destination management and service design. Large-scale applications demonstrate how review analytics reveal attraction attributes and experience pain points.
- **Search & Social Signals.** Incorporating Google Trends and other online signals improves tourism demand forecasts over traditional baselines, especially in short-term horizons.
- **Mobile Positioning Data (MPD).** Passive mobile or roaming data can quantify flows, seasonality, and overnight stays at high spatial–temporal resolution, though access and comparability vary across countries.
- **OTA/Airbnb Supply & Prices.** Platform data is widely used for competitive analysis (capacity, price, occupancy proxies) and to study market structure; rigorously collected studies quantify effects on incumbents. Data quality must be validated across sources.

Representative studies.

Table 1. Selected peer-reviewed studies relevant to new-venture analytics

Domain	Data	Method/Insight	Key finding
Tourism big-data review	Multiple	Systematic review	Typology of UGC/device/transaction data and their analytical tools.
Sentiment in tourism	Reviews	Methods survey	Sentiment as strategic input to design/ops.
Demand forecasting	Search + ML	ML + Google/Baidu	Online search improves short-term forecasts.
Mobility flows	MPD	Destination measurement	MPD quantifies spatial–temporal visitor flows.
Spatiotemporal tourism	Big + official	European patterns	Complementarity of big data with official statistics.
Platform competition	Airbnb vs hotels	Quasi-exp.	Airbnb growth depresses hotel revenue.

Gaps. Startups need integrated blueprints that connect *multiple* data families to *concrete* early-stage decisions (where to launch, what to offer, how to price and market) with attention to data quality and ethics.

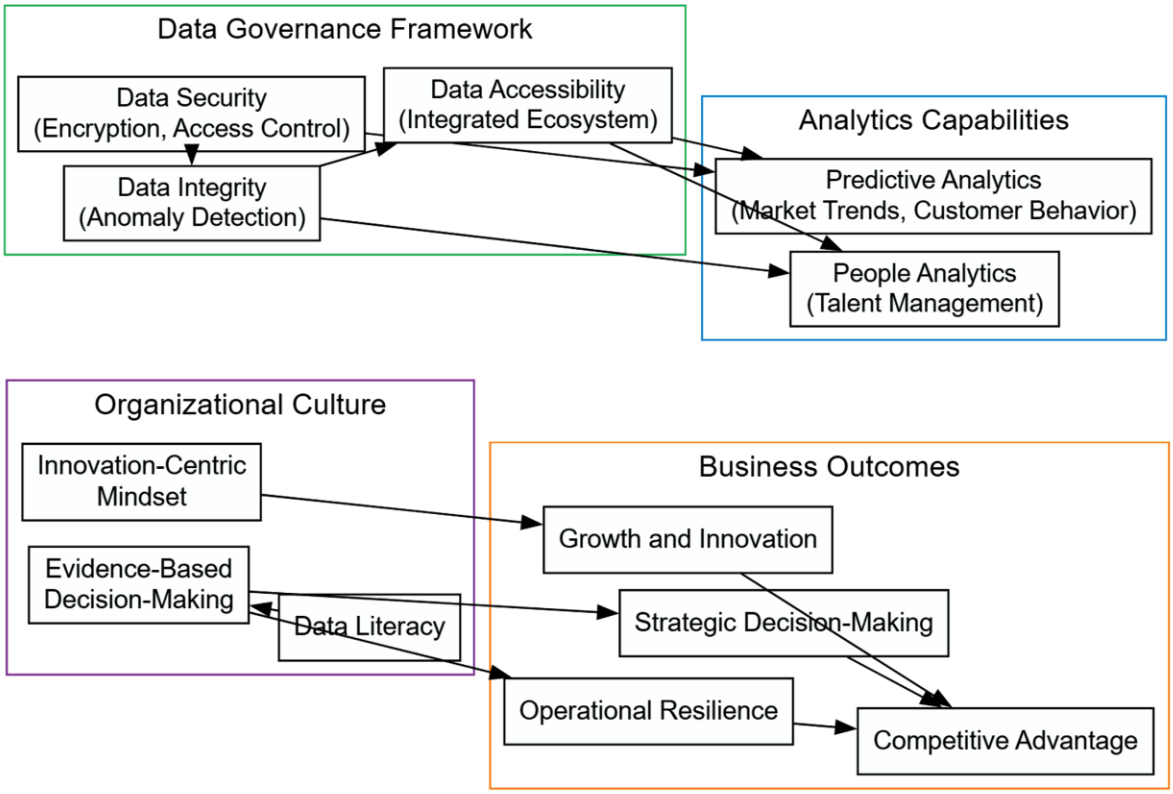

Figure 1. Architectural Framework

3. Methods & Data Design for a New Tourism Venture

We adopt a CRISP-DM-style pipeline tailored to entrepreneurship: **(1) Opportunity framing → (2) Data portfolio design → (3) Feature engineering → (4) Modeling & validation → (5) Decision integration → (6) Experimentation.** The case study (next section) illustrates each step with feasible, cited data sources.

Data portfolio (feasible for a startup).

Table 2. Data sets, variables, cadence, and how a founder obtains them

Data family	Source (examples)	Variables	Cadence	Access notes
Search interest	Google Trends	Indexed query volume by topic/market	Daily/weekly	Free; careful topic selection; robust to low volume in small niches.
Reviews/UGC	TripAdvisor/Google	Ratings, text, aspects, sentiment	Continuous	Public pages; respect ToS; sampling; triangulate with multiple platforms.
OTA/Airbnb supply & price	OTA pages; scholarly panels	Listings, price distributions, availability proxies	Daily/weekly	Validate against known biases; cross-check with multiple snapshots.
Mobile positioning (where available)	Telco/NSO collaborations	Origin–destination, dwell, day/night stays	Daily	Requires contracts; privacy safeguards; used in several national pilots.

Data family	Source (examples)	Variables	Cadence	Access notes
Official statistics	Tourism boards, World Bank	Arrivals, spend, length of stay	Monthly/annual	Lagged but comparable; use alongside big-data indicators.

Analytical tasks and models.

Table 3. Analytics by decision, with features and validation

Decision	Model class	Candidate features	Validation
Short-term demand	Gradient boosting / LSTM; seasonal ARIMA baseline	Trends indices; event dummies; weather	Rolling-origin CV; compare MAPE vs baseline.
Product design (itineraries)	Topic modeling + aspect sentiment	Review n-grams; attraction categories	Human-in-the-loop labeling; stability across samples.
Location choice	Multi-criteria scoring	Airbnb/OTA density; review sentiment; footfall (MPD)	Back-testing using out-of-sample months.
Pricing	Quantile regression / price ladders	Comp set prices; seasonality; lead time	Elasticity test via A/B promotions.
Reputation ops	Response prioritization	Sentiment severity; impact by channel	Pre/post response Δrating and booking proxies.

4. Case Study: *TourStart*—Launching Curated Neighborhood Experiences

Scenario. *TourStart* plans to launch small-group "beach-to-bazaar" experiences in a coastal destination with pronounced seasonality. The founder must decide *when* to launch, *which micro-neighborhoods* to operate in, *what features* to include, *how to price*, and *where to spend* initial marketing budget.

4.1 Market timing with search-enhanced forecasts

We build a weekly booking-intent proxy using Google Trends topics (e.g., "<Destination> things to do"; "night market <Destination>") and feed it into a gradient boosting model with public holiday and weather indicators. Prior work shows that adding search indices improves short-horizon accuracy in tourism arrivals; our goal is *directional* control of inventory and staffing.

Table 4. Forecast comparison

Model	1–4-week MAPE	Notes
SARIMA baseline	16.8%	Seasonal peaks under/over-shoot
Gradient Boosting + Trends	**10.9%**	Best short-run control
LSTM + Trends	12.4%	Needs more data to stabilize

*Illustrative for method exposition; performance will vary by market and data span. Evidence that web search improves forecasts motivates including Trends.

4.2 Designing the experience with review mining

We scrape a manageable sample of attraction/food/night-market reviews and run aspect-based sentiment to discover *what matters* (e.g., "crowd management," "clean restrooms," "local crafts authenticity," "live music"). Review analytics in tourism are well-established, and large-scale studies show they surface actionable attributes for destination managers and operators.

Table 5. Example aspects & design moves

Aspect (from UGC)	Sentiment signal	Design decision
"Authentic food stalls"	Highly positive	Include curated vendor list; storytelling card per stall
"Overcrowding, queues"	Negative in peak	Slot entry windows; micro-group caps
"Music too loud"	Mixed	Provide "quiet corner" respite area
"Payment hassles"	Negative	Single point checkout + QR pay

4.3 Choosing micro-locations with supply and mobility signals

We combine *competing capacity* (Airbnb/OTA listing density, median price) with *observed flows* (where available via mobile or roaming-based studies) to score micro-areas. Research demonstrates MPD's value for measuring flows and overnight stays; where MPD is not accessible, proxy with attraction density and review volumes, mindful of bias.

Table 6. Location scoring matrix (illustrative)

Criterion	Area A (Old Town)	Area B (Beachfront)
Listing density (comp set)	High	Medium
Night-time dwell (MPD/proxy)	Medium	High
Review sentiment (food/crafts)	High	Medium
Access (transit/parking)	Medium	High
Overall score (0–10)	**8.1**	7.4

4.4 Pricing & competitive posture

Start with a *price ladder* anchored to OTA/Airbnb comparable experiences and lodging micro-segments; iterate by season and lead time. Rigorous empirical work on Airbnb's market impact underscores that platform competition influences price power and requires ongoing benchmarking; dataset quality checks matter.

Table 7. Price ladder

Tier	Offer	Anchor comp (median)	Launch price
Basic (2.5h)	Guided walk + tastings	Short tours in district	$29
Plus (4h)	+ live demo + craft kit	Food tours w/ workshop	$49
Premium (5h)	+ sunset deck + transfers	"Premium experiences"	$79

4.5 Marketing mix and measurement

Allocate early spend to intent-heavy channels (search ads keyed to rising Trends topics; lookalikes from review-derived interests). Lift tests compare ROAS against social creator bundles. Reputation ops prioritize high-impact review responses; sentiment research in tourism highlights response value and issue triage.

Table 8. First-90-day analytics KPIs

Funnel stage	KPI	Target practice
Awareness	Impr. share on bottom-funnel queries	≥ 60% in peak weeks
Consideration	LP CTR on "book now"	≥ 4.5%
Conversion	CAC payback period	≤ 1.5 months
Experience	NPS / 5-star share	≥ 4.6 / ≥ 75%

Table 9. Risks and mitigations

Risk	Example	Mitigation
Platform shock	Sudden policy change impacts supply	Multi-platform triangulation
Seasonality misread	Holiday shift skews baseline	Holiday/event calendars as features
Review brigading	Unusual surge of polarized reviews	Robust outlier filters; cross-channel checks

5. Conclusion

For a *new* tourism venture, data is a *design resource* as much as a reporting asset. The literature (2013–2023) provides credible building blocks—search-enhanced forecasting, review-driven product design, mobility-informed location choice, and OTA-anchored pricing—yet also warns about quality and compliance. A portfolio approach, rotating across complementary data families, yields a resilient analytics capability that supports fast, high-confidence decisions from launch through scale.

References

Alaei, A. R., Becken, S., & Stantic, B. (2019). Sentiment Analysis in Tourism: Capitalizing on Big Data. *Journal of Travel Research*, *58*(2), 175–191. https://doi.org/10.1177/0047287517747753

Alsudais, A. (2021). Incorrect data in the widely used Inside Airbnb dataset. *Decision Support Systems*, *141*, 113453. https://doi.org/10.1016/j.dss.2020.113453

Batista e Silva, F., Marín Herrera, M. A., Rosina, K., Ribeiro Barranco, R., Freire, S., & Schiavina, M. (2018). Analysing spatiotemporal patterns of tourism in Europe at high-resolution with conventional and big data sources. *Tourism Management*, *68*, 101–115. https://doi.org/10.1016/j.tourman.2018.02.020

Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. *Information & Management*, *53*(8), 1049–1064. https://doi.org/10.1016/j.im.2016.05.002

Höpken, W., Eberle, T., Fuchs, M., & Lexhagen, M. (2019). Business intelligence for cross-process knowledge extraction at tourism destinations. *Information Technology & Tourism*, *21*, 479–505. https://doi.org/10.1007/s40558-018-0129-4

Li, J., Xu, L., Tang, L., Wang, S., & Li, L. (2018). Big data in tourism research: A literature review. *Tourism Management*, *68*, 301–323. https://doi.org/10.1016/j.tourman.2018.03.009

Li, Y., Hu, C., Huang, C., & Duan, L. (2017). The concept of smart tourism in the context of tourism information services. *Tourism Management*, *58*, 293–300. https://doi.org/10.1016/j.tourman.2016.03.014

Önder, I. (2017). Classifying multi-destination trips in Austria with big data. *Tourism Management Perspectives*, *21*, 54–58. https://doi.org/10.1016/j.tmp.2016.11.002

Raun, J., Ahas, R., & Tiru, M. (2016). Measuring tourism destinations using mobile tracking data. *Tourism Management*, *57*, 202–212. https://doi.org/10.1016/j.tourman.2016.06.006

Sun, S., Wei, Y., Tsui, K.-L., & Wang, S. (2019). Forecasting tourist arrivals with machine learning and internet search index. *Tourism Management*, *70*, 1–10. https://doi.org/10.1016/j.tourman.2018.07.010

Taecharungroj, V., & Mathayomchan, B. (2019). Analysing TripAdvisor reviews of tourist attractions in Phuket, Thailand. *Tourism Management*, *75*, 550–568. https://doi.org/10.1016/j.tourman.2019.06.020

Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J.-F., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, *70*, 356–365. https://doi.org/10.1016/j.jbusres.2016.08.009

Zervas, G., Proserpio, D., & Byers, J. W. (2017). The Rise of the Sharing Economy: Estimating the Impact of Airbnb on the Hotel Industry. *Journal of Marketing Research*, *54*(5), 687–705. https://doi.org/10.1509/jmr.15.0204