The Atlas

How we mapped what people do, need and depend on

Aishwary Singh, Opankitaab Private Limited
19 October 2026, public edition, 30-minute read

Abstract

The atlas is a method for counting what people in India do, need and depend on, built in six layers, then corrected and given a mathematical spine. It is the evidence base under The Distance.

The six layers were drafted in sequence. Each was checked against live sources and corrected where it was wrong. Eight recurring errors were found, the most serious a population figure that would have inflated every worker-based count by more than half.

The layers build on each other. The first counts people and places. The second adds what changes behaviour and how to count it without mixing units. The third adds relationships, time and institutions. The fourth treats daily life as a network. The fifth turns everything into an engine with a field plan. The sixth asks why any action begins at all.

On top of the six layers sits a body of mathematics: the chain law, queueing, agreement under liars, comparison traps and the arithmetic of persuasion. That mathematics is what turned a catalogue of behaviour into a thesis about where the distance between wanting and happening actually lives.

This paper records how the atlas was made, presents each layer in refined form, and ends with one consolidated specification: one notation shared with The Distance, one data schema, and one build order. It also states plainly what no atlas of this kind can know.

Why an atlas

Counting people tells you who exists. Counting person-events tells you what happens. The atlas counts the second.

A population table says that 24.3 crore people live in Uttar Pradesh. It says nothing about how often a farmer there chases a payment, a mother arranges a pickup, or a trader waits on a supplier. Those events, not the head count, are where time, money and trust are lost.

The unit: one person, one occurrence

The atlas measures person-events. The same person generates very different counts across activities:

Activity	Typical frequency for a participant
Checking a phone	About 200 times a day
Filing GST, where required	12 times a year
Buying a train ticket	About 2 times a year
Visiting a dentist	About once every 2 years

These rates are illustrative, not measurements. They show why frequency must be estimated per activity, never assumed per person.

Five things kept apart

Most data blurs five different things. The atlas keeps them separate from the first row:

1. What people **do**.
2. What they **need**.
3. What they **say**.
4. What they **intend**.
5. What **actually happens**.

A survey answer is something said. A payment log is something done. Neither is an outcome. Confusing them is the root of most bad behavioural statistics.

Three universes of activity

Every activity falls into one of three universes: things people do themselves, things they do for others, and things they need others to do. The atlas maps all three without favouring any. The Distance then showed, with the chain law, that the third universe is where the gap between wanting and happening is longest.

What the atlas is for

- **For Kitaab:** where coordination stalls, for whom, and how often.
- **For business:** where demand sits, by place, income and occupation.
- **For economics:** how much time and money a society spends on coordination itself.

How it was built

The atlas grew out of one proof. A question about whether society is bad led to the chain law, and the chain law is what made a behaviour atlas worth building.

Drafting the six layers was the quick part. The slow part was checking them: each was compared with a running corrections log, its figures verified where possible, its arithmetic re-run and its symbols reconciled before it was kept.

Layer 1 — Who lives where

Layer 1 describes a person's world in seven dimensions, names sixteen domains of activity, and defines the one count every later layer refines.

Seven dimensions of a lived environment

Dimension	Segments	Count
Settlement	Metro, large city, small city, town, village, remote habitation	6
Age	0–14, 15–24, 25–34, 35–44, 45–59, 60+	6
Economic position	Destitute, low-income, lower-middle, middle, affluent, ultra-wealthy	6
Livelihood	Salaried, self-employed, farming, informal labour, unpaid care, student, unemployed, retired	8
Technology access	Offline, shared phone, basic smartphone, digitally active, power user	5
Household	Alone, nuclear, joint, multigenerational, migrant	5
Institutional access	Banked, documented, insured, formally employed (each yes or no)	2^4 = 16

Multiplied out, that is $6 \times 6 \times 6 \times 8 \times 5 \times 5 \times 16 = 691{,}200$ possible combinations. The number is a warning, not a plan. Most combinations are nearly empty, and one person can hold several livelihoods at once: a tribal woman may be a teacher, a farmer and an elected village representative. Roles are therefore weights a person carries, not boxes they sit in.

Sixteen domains of activity

Domain	Examples
Survival and body	Sleeping, eating, water, washing, cooking
Livelihood and labour	Farming, construction, driving, coding, selling
Money and transactions	Cash, UPI, borrowing, chit funds, remittances, taxes
Mobility and logistics	Buses, trains, autos, trucking, parcels, migration
Family and relationships	Marriage, childcare, eldercare, weddings, inheritance
Health	Pharmacies, hospitals, vaccination, mental healthcare
Government and civic life	ID documents, ration cards, land records, voting, courts
Digital behaviour	Messaging, video, shopping, payments, AI conversations
Education and learning	School, tuition, exams, vocational training
Culture, faith and community	Worship, festivals, community councils, volunteering
Leisure	Cricket, music, cinema, restaurants, travel
Property and infrastructure	Rent, land, repairs, electricity, water, internet
Commerce and enterprise	Kirana shops, vendors, factories, wholesale, startups
Conflict and illegality	Fraud, corruption, harassment, litigation, mediation
Social coordination	Favours, chasing payments, appointments, hiring, following up
Inner life	Worry, grief, planning, ambition, prayer, boredom

The core count

For an activity, a district and a population segment, expected occurrences in a period are:

$$E = N \times q \times f$$

Events equal eligible people N, times the share of them who take part q, times how often each participant does it f. The draft used p for participation; this paper uses q, because p already means a party's reliability in The Distance.

Every estimate is stored as one row: geography, population, behaviour, frequency, channels, dependencies and evidence.

What was refined in Layer 1

- **Base population.** The draft leaned on Census 2011. The atlas now uses the official 2026 projections (Appendix A): 142.94 crore nationally. International estimates run higher, up to about 147 crore, so every national figure carries a band of roughly 3%.
- **Scale.** State totals are too coarse for N. District-level populations are the next data to pull.
- **Order.** The full activity catalogue comes before any district rates. Precise numbers on an incomplete map would be false precision.

Layer 2 — What changes behaviour, and how to count it

Layer 2 adds seventeen forces that shape behaviour, nine units that must never be mixed, and nine rhythms of frequency, and it forbids the shortcut of multiplying dimensions as if they were independent.

Seventeen variables that modify behaviour

Age and life stage; sex and gender; income and assets; occupation; urbanisation; education and literacy; language; household composition; migration; religion and culture; caste and community; physical and mental health; disability; digital access and autonomy; season and climate; local infrastructure; and institutional trust and safety.

These explain patterns across groups. They are never rules about individuals: nobody's religion predicts their spending, and nobody's state predicts their job.

Nine units, never mixed

Unit	Example
People	Number of people who work in construction
Households	Households with an unpaid electricity bill
Enterprises	Kiranas keeping customer credit
Events	Dental appointments booked
Transactions	Payments made
Time	Minutes spent caring for family
Relationships	Active caregiver and dependant pairs
Open undertakings	Things someone said they would do, still open
Value	Rupees transferred, earned or lost

One person can have ten open undertakings; one household five earners; one enterprise hundreds of transactions. A model that divides one unit by another without a stated bridge produces nonsense, however good its population numbers are.

Nine rhythms of frequency

Rhythm	Examples	Measured as
Continuous	Sleep, ongoing pain	Minutes or hours
Daily	Meals, commuting, phone use	Events per day
Weekly	Groceries, worship, weekly markets	Events per week
Monthly	Rent, salaries, loan instalments	Events per month
Seasonal	Harvest, festivals, monsoon preparation	Events per season
Annual	Tax returns, insurance renewal, exams	Events per year
Life-stage	School admission, marriage, retirement	Events per eligible lifetime
Triggered	Emergencies, disputes, accidents	Events per exposure
Latent	Fear, desire, private intention	Survey or diary measures

The count, summed across groups

$$E_{d,a} = \sum_{s} N_{d,s}\, q_{d,s,a}\, f_{d,s,a}$$

The events of activity a in district d are the sum, across population segments s, of eligible people times participation share times frequency.

The forbidden shortcut

The dimensions are correlated, so their shares cannot be multiplied. Suppose, purely as an illustration, that 35% of people are urban and 40% are power users of technology. Independence would predict that 14% of all people are urban power users, and that power users are as common in remote villages as in cities. Both are false. The atlas uses observed joint distributions from survey microdata, and statistical estimation where district samples are thin.

Where the data starts

The draft named three national surveys as the base: the Time Use Survey 2024, the Household Consumption Expenditure Survey 2023–24, and the Annual Survey of Unincorporated Sector

Enterprises 2023–24, which it cites at about 7.34 crore non-farm informal businesses. For sparse or hidden behaviour, every estimate carries a low, base and high value, and a note on how reliable its evidence is.

What was refined in Layer 2

- **"Obligations" became "open undertakings".** The draft counted "promises awaiting completion". Not every open matter is a promise, so the unit is named for what it counts.
- **The working-population error was caught here first.** The draft called 148 crore India's "working population". It is a high estimate of total population (see "What we refined").
- **The 7.34 crore figure is the draft's.** It has not been independently checked and is marked that way.

Layer 3 — Relationships, time and institutions

Layer 3 recognises that activity happens inside relationships, rules and finite time, and that much of life's effort goes into coordinating rather than doing.

What travels through a relationship

Information, resources, labour, authority, reputation, risk, emotion and obligation. A wholesaler extends credit to a retailer because of ten years of dealings; a landlord rents to someone a neighbour vouched for. Listed prices and individual preferences do not explain these choices.

Three intellectual lenses describe the same fact: Ronald Coase on the costs of transacting, Mark Granovetter on economic life embedded in social ties, and Elinor Ostrom on how communities govern shared resources.

Four kinds of trust

Trust	Rests on
Personal	I know you
Reputational	People I know vouch for you
Institutional	Rules and organisations protect this
Evidence-based	There is independently checkable proof

Where institutions are weak, people lean on personal and reputational trust. Software can lower the cost of evidence, but it cannot replace affection, duty or legitimate authority.

Time is finite; open matters are not

Everyone has 1,440 minutes a day. Open matters accumulate by a simple rule:

$$Q_{t+1} = \max\left(0,\ Q_t + A_t - C_t\right)$$

Tomorrow's backlog is today's backlog, plus what arrived, minus what was completed, never below zero. If arrivals A outpace completions C over the long run, the backlog grows without limit. Little's Law, in The Distance, adds the steady-state version: open matters equal arrival rate times closing time.

The hidden coordination tax

A refrigerator repair may take 45 minutes of work. Around it sit search, calls, waiting, supervision, payment and recovery if it fails again:

$$C_{\text{coord}} = C_{\text{search}} + C_{\text{communication}} + C_{\text{waiting}} + C_{\text{monitoring}} + C_{\text{verification}} + C_{\text{recovery}}$$

Measured in person-minutes or money, without double-counting waiting that overlaps with work. How large this tax is nationally is an open empirical question.

Six institutional worlds

The same ₹20,000 owed is governed differently in each: the household (affection, authority), the community (reciprocity, reputation), the market (price, contract), the organisation (hierarchy,

rules), the state (law, entitlement) and the digital network (protocols, identity). A payment receipt proves money moved; it does not prove the underlying obligation was met.

Intention is not action

The draft's invented dental funnel: 10,000 people with a need; 6,500 intend to seek care; 4,200 contact a provider; 3,000 book; 2,400 attend; 1,700 complete treatment. That is 17%. A funnel is a chain: the end rate is the product of the stage rates, exactly the p^k structure with a different p at each step.

Thirteen steps of action

Need, discover, choose, request, negotiate, authorise, accept, execute, coordinate, observe, verify, close and recover. Appointments, weddings, shipments and government applications all combine some of these. Similar structure does not make activities interchangeable: an emergency transfusion and a cake delivery carry different consent, evidence and safety rules.

What was refined in Layer 3

- **Two step names changed.** The draft's "commit" is "accept", which is what its own example describes, and its "settle" is "close": paying is only one way a matter closes.
- **Labour figures were checked as arithmetic.** Per 100,000 adults, the draft's 2025 rates give 57,400 workers: 32,259 self-employed, 13,546 salaried and 11,595 casual. The sums are right; the source rates are the draft's and not independently verified.
- **Payment volume was verified.** UPI processed 24.51 billion (2,451 crore) transactions worth ₹29.82 lakh crore in August 2026, about 791 million a day. These are transactions, not people and not completed obligations.
- **The rate model is a choice to test.** The draft multiplies a base rate by seasonality, institutions and shocks. That form is a hypothesis, not a law.

Layer 4 — The network underneath daily life

Layer 4 treats daily life as a changing network of people and institutions, and asks what actually passes between them.

Twelve things that pass between people

Payload	Example
Information	Tell me when the train arrives
Attention	Listen to my problem
Advice	Which doctor should I see?
Permission	Can I use your vehicle?
Undertaking	I'll deliver it tomorrow
Labour	Please repair the motor
Goods	Send twenty bags of rice
Money	Transfer ₹5,000
Care	Look after my father
Evidence	Send the payment receipt
Influence	Introduce me to the officer
Emotion	Forgive me, comfort me, celebrate with me

One call can carry several at once. A son's call to his mother may hold reassurance, money news, a health update and a plan to visit. Reducing it to one transaction loses most of it.

One household, one day

The draft's fictional household (a shopkeeper, a teacher, two children and a grandparent) generated eight events in a day: milk bought, school pickup confirmed, shop stock ordered and pending, a clinic visit scheduled, a customer's credit unresolved, the electricity bill paid, homework done, and a family trip still undecided. Three closed that day, two were in motion, and three stayed open. Some outcomes, like emotional support, have no binary "done" at all.

The network, formally

$$G_t = (V,\ \mathcal{L}_t,\ R)$$

The network at time t is the set of actors V, the links active at that time $\mathscr{L}$, and the kinds of relationship R (kinship, employment, commerce, care, authority). Each link records two actors, the relationship, the activity, the time and its state. Events with three or more participants, such as a wedding, a construction job or a clinical referral, fit better as one event with many actors than as many pairs.

Frequency is not importance

Each activity carries a profile, not one number: frequency f, time consumed τ, money value v, risk ρ_{risk}, the number of parties who must come through k, and the person's own control κ. Scrolling happens hundreds of times a day; a property deal happens once. The first is frequent; the second is consequential.

The people a model misses

Unpaid domestic work, informal deals, people without documents, people facing disability barriers, seasonal workers, hidden harm and remote settlements. Some are unrecorded; others are under-reported out of stigma or fear. The response is never invented precision. Every estimate carries: observed count, coverage correction, uncertainty interval, source quality and date.

$$\text{true count} \approx \frac{\text{observed count}}{\text{coverage}}$$

If a source captures 60% of informal transactions, the true count is about 1.67 times the observed one. If coverage is only known to lie between 50% and 70%, the true count lies between 1.43 and 2 times the observed one, and that range travels with the number.

Engine version 0

Eight steps: reconcile state and district populations to one national total; build household and individual profiles from survey joint distributions; assign daily schedules from time-use data; add occupation, household and seasonal activities; generate interactions; track open, completed, abandoned and delayed matters; calibrate against independent administrative and transaction data; and report by place, segment, activity and period, with uncertainty. A simulation that looks plausible but fails calibration is corrected, however elegant its theory.

What was refined in Layer 4

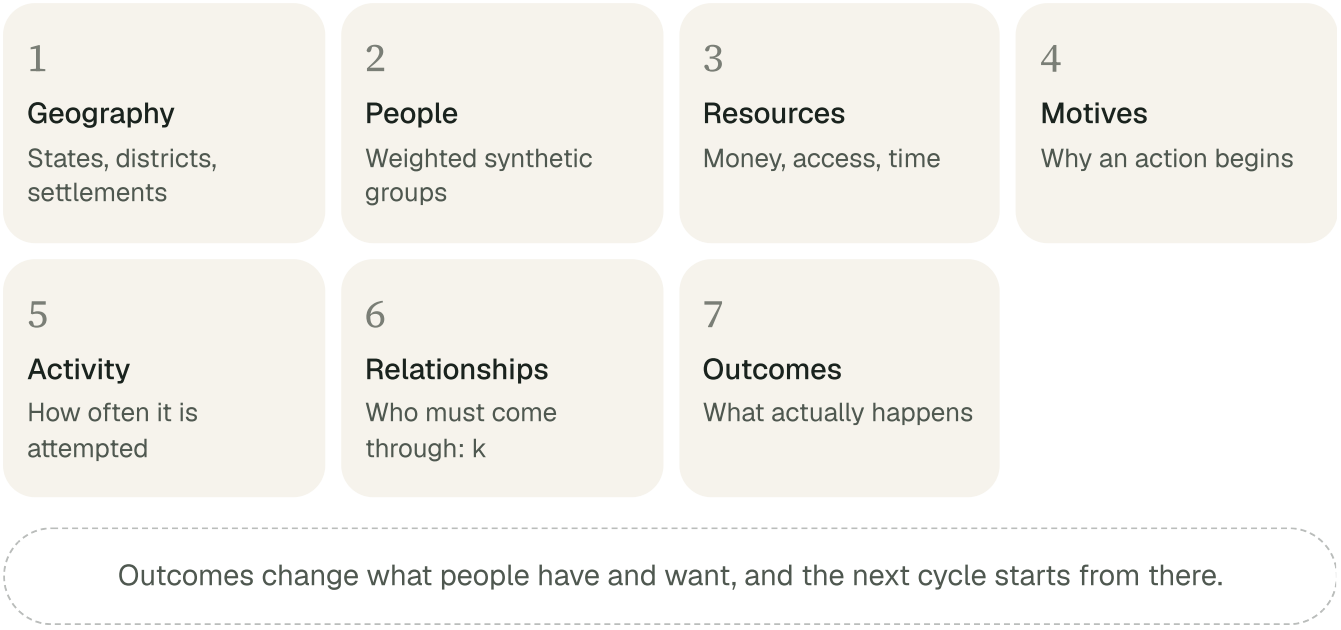
- **"Commitment" became "undertaking"** in the payload list, for the same reason as Layer 2.
- **Letters were renamed.** The draft used E for links (its third meaning of E) and r for both relationship type and risk. Links are now $\mathscr{L}$, relationship types R, and risk ρ_{risk}.

- **The profile gained k.** The draft's "dependency complexity" is exactly the number of parties in the chain law, so it now carries that name.

Layer 5 — The engine, and the field

Layer 5 turns the first four layers into a working engine with two separate models, one for how often things are attempted and one for how often they succeed, and a plan to measure both in the field.

The engine runs in seven stages: geography, people, resources, motives, activity, relationships and outcomes. It works on weighted synthetic groups that stand for real populations, never on real identities.



The engine. Stages one to five set how many attempts happen; stage six is where the chain law enters.

Stages one to five drive the count of attempts. Stage six is where the chain law enters: every additional party who must come through lowers the chance of the outcome in stage seven.

Two models, not one

The count model says how many attempts happen:

$$\mathbb{E}[E_{a,d,t}] = \sum_{g} W_{d,g,t}\, q_{a,d,g,t}\, f_{a,d,g,t}$$

The expected number of events is the sum over synthetic groups g of the people each group represents (W), times their participation share, times their frequency.

The success model says how many of those attempts work:

$$P(\text{success} \mid a, d, g, t) = \int P(\text{success} \mid Z, a)\, \pi(Z \mid d, g, t)\, dZ$$

It averages the chance of success over every situation a person might be in: Z is the person's conditions (money, access, counterparties, timing), and π(Z) is how common each condition is in that place, group and time. This is the law of total probability, so the formula is exact; the assumption is that place and group matter only through Z.

Separating the models separates three questions: how many people need something, how many attempt it, and how many accomplish it.

Sixteen fields per activity

Field	Question it answers
Eligibility	Who could experience or do it?
Need rate	How often does the need arise?
Awareness	Is the need recognised?
Intent rate	How often is action chosen?
Attempt rate	How often is it started?
Frequency	Attempts per participant
Duration	Time consumed
Participants	People and institutions involved
Dependencies	Whose cooperation is required (k)
Money	Value exchanged or committed
Authority	Who may legally or socially decide
Channel	In person, phone, paper, digital or mixed
Outcome	Success, failure, withdrawal, ongoing, unknown
Delay	Time until progress or resolution
Risk	Financial, physical, emotional, social, legal
Confidence	How well the estimate is supported

Five reminders about one unpaid invoice are not five invoices. A wedding with hundreds of guests and thousands of messages is one event.

Three funnel models

The draft's illustrative dental funnel runs 10,000 with a need, 9,000 intend (90%), 7,200 book (80% of those), and 6,480 attend (90% of those): 64.8%. Raising intent from 90% to 99% gives three different answers depending on the model:

Model	What raising intent to 99% does
Multiplied stages	Attendance rises 10%
Bottleneck (say 5,000 clinic slots)	No change; capacity binds
Dependent stages (new intenders book at 20%)	Attendance rises about 2.5%

Every activity must declare which model it uses before anyone decides where to act.

The data stack

Source	What it reveals	Main limit
Time Use Survey 2024	Paid work, care, learning, leisure, the daily clock	Not a record of intentions
NFHS-5 district data	Health, household conditions, autonomy	Mostly 2019–21
India Human Development Survey	Change in households over time	Older waves
All India Debt and Investment Survey	Assets and debts	Reference year mostly 2018–19
Telecom survey 2025	Phone and internet access	Access is not autonomy
Village network datasets	Real ties, favours, diffusion	Specific villages, older
NABARD self-help group data	Savings and credit groups	Programme coverage only

No existing survey measures the thing this research cares most about: what people intended, who they needed, and why it stalled. A one-week consented diary study fills that gap.

Eight places to measure first

District	What it shows
Bengaluru Urban	Metropolitan work, services, digital coordination
Mumbai Suburban	Density, migration, commuting
Patna	Education, informal trade, institutions
Varanasi	Pilgrimage, tourism, trade, public services
Coimbatore	Manufacturing, suppliers, family enterprise
Ernakulam	Services, healthcare, migration-linked homes
Kamrup Metropolitan	Regional commerce and transport
Papum Pare	Urban-rural links, terrain and access

Eight mechanisms to test

1. **Scarcity:** do income shocks increase delays and short-term borrowing?
2. **Access:** does less travel or waiting raise completed visits?
3. **Network:** does information spread further through well-connected households?
4. **Authority:** does financial independence change what is attempted and completed?
5. **Coordination:** do multi-person tasks fail more than similar solo ones? This is the chain law, as a field test.
6. **Institutions:** do clear rules and reliable timelines reduce abandoned applications?
7. **Planning:** do explicit when-and-how plans beat reminders alone?
8. **Trust:** does verified history make people engage with strangers?

Each result is reported with its effect size, confidence interval and whether it generalises beyond the sample.

Society as distributed computation, with a boundary

People are nodes with private knowledge; conversations are messages; laws and norms are protocols; consent is permission; nobody sees everything; waiting is delay; forgetting and refusal are failures; reattempts and escalation are recovery; receipts and witnesses are evidence. The analogy has a hard limit. Humans are not servers. They have dignity, change their minds and have the right to refuse, and silence is never consent.

Six layers of reality

Existence, condition, motivation, interaction, execution and consequence, and then the cycle repeats. The economy, politics, family life and faith all run inside this cycle. It gives one set of questions with which to study them all.

What was refined in Layer 5

- **"1.48 billion people" appeared again.** The category is right this time (total population), but the figure is the top of the range. The official 2026 projection is 142.94 crore.
- **Letters were untangled.** The draft's N meant people in one layer and events in another, and p meant both participation and the probability of a condition. People are N or W, events are E, participation is q, and the condition density is π.
- **The binding-constraint claim was corrected** into the three models above.
- **One boundary to verify before fieldwork:** whether Itanagar's capital area now sits in a separate district from Papum Pare.

Layer 6 — Motivation, and how a desire is observed

Layer 6 asks why any action begins, and how anyone could know. Its answer is that desire leaves only partial traces, so evidence must be graded.

Ten kinds of motivation

The final layers separate things usually lumped together as "wants": biological drives; psychological needs such as autonomy, competence and belonging; acquired tastes; instrumental goals pursued for something else; values held for their own sake; desires shaped by comparison and norms; obligations felt regardless of preference; habits and compulsions; behaviour under coercion; and play or exploration with no stated goal.

Five levels of evidence for a desire

Evidence	What it establishes	What it cannot establish
Self-report	What someone says they want	Their unconscious motives
Behavioural choice	Which option they picked	Every reason for picking it
Willingness to pay	Value under stated conditions	What it means to them
Time and effort spent	What they commit to the pursuit	Whether it was freely chosen
Completed action	What actually happened	Whether it satisfied them

Someone may say saving matters while a money diary shows emergencies eating every rupee. That is constrained saving, not a lack of discipline.

Three populations

The eligible could do something; the motivated value it; the acting actually do it. The gaps between them say more about friction and opportunity than any attitude survey.

A study of motivation

Across contrasting districts, consenting adults answer five questions, with follow-up weeks later: what matters, what is possible, who influences it, what happened, and what changed. Task and choice measures sit beside the questions, because what people choose often differs from what they say.

Five propositions to test

1. **The future people can picture shapes what they want now.**
2. **An observed action says less about motive than it appears to.** A ₹10,000 transfer fits rent, a parent, a loan, a gift, a bribe or an investment.
3. **Agency is distributed.** A person can be fully motivated and still unable to act alone.
4. **Reward is not pleasure.** Wanting, liking and learning are different systems.
5. **Social systems both enable and constrain.**

The strongest evidence in the set

The first proposition rests on an Indian randomised experiment. A 1993 law reserved village council leadership for women in randomly chosen villages. In West Bengal, 495 villages and 8,453 surveys

later, villages with a woman leader for two terms had closed the gender gap in adolescents' career and education aspirations by 32%, and in parents' aspirations by 20 to 25%. The gap in actual schooling disappeared (Beaman, Duflo, Pande and Topalova, Science, 2012). One visible role model changed what a village thought possible.

What was refined in Layer 6

- **Acting is not always a subset of wanting.** Obligation, coercion and habit produce action without desire, so acting is never computed as wanting times conversion for those activities.
- **The parent figure is a range.** The published abstract reports 25%; the author manuscript reports 20%.
- **"1.48 billion" appeared twice more** and was corrected as before.

The mathematics added on top

The six layers describe the world; the mathematics explains why it behaves as it does. Every result below is a theorem or exact arithmetic, with its assumptions stated. Full derivations for the first group are in The Distance, Appendix B.

Results shared with The Distance

Result	What it says	Status
Chain law, S = p^k	Success decays exponentially with the number of parties	Exact, if failures are independent
The 99.95% bound	Twenty-person chains that work 99% of the time need each person at 99.95%	Exact, given the observation
Breakers versus rent-takers	1% chain-breakers would cut a 99% chain to 81%; rent-takers tax rather than break	Exact, given the model
Little's Law, L = λW	Open matters equal arrival rate times closing time	Theorem
Overload, ρ ÷ (1 – ρ)	At 90% load, 9 items wait; at 95%, 19	Assumption: random arrivals
Retries	13 steps at 90%: 25% succeed without follow-up, nearly all with it, for 11% more effort	Exact
One-third law, n ≥ 3f + 1	Without proof, agreement fails if a third lie; with signatures it survives any number	Theorem
Silence (FLP)	You cannot tell slow from gone without a time limit	Theorem
Hub law	200 people: 19,900 pairwise links, or 199 through one hub	Exact
Friendship paradox	A random person's contacts are better connected than the person	Theorem
Simpson's paradox	A district can be better in every group and worse overall	Exact

The good and bad actor model

A party in a chain completes its part with probability:

$$p = (1 - b)(1 - \varepsilon) + b\,u$$

The honest share (1 − b) completes unless it slips, at rate ε, and the bad share b completes with probability u. Chain-breakers have u near 0 and are tightly bounded by how reliably everyday chains work. Rent-takers have u near 1: they complete, but take a cut. A rent-taker's income is roughly the cut, times the chains through his seat, times p^k, so it depends on everyone else's honesty. The damage a bad actor does grows with centrality: one registrar sits on thousands of chains.

Results particular to the atlas

1. An action reveals little about motive (Bayes' rule). Take a ₹10,000 transfer and assumed starting beliefs: rent 30%, a parent 25%, an investment 20%, a loan 15%, a gift 5%, a bribe 5%. Add a strong clue: it was sent on the 1st to 5th of the month.

Motive	Before	After the clue
Rent	30%	52%
Parent	25%	19%
Loan	15%	15%
Investment	20%	12%
Gift	5%	1%
Bribe	5%	1%

Even with a strong clue, the best guess is barely better than a coin toss. Payment logs cannot reveal motive; people's own words can. The starting beliefs and clue strengths are illustrative.

2. One visible success doubles what seems possible (rule of succession). With no other information, the perceived chance of an outcome is (successes + 1) ÷ (cases seen + 2). After 50 girls and no scientist, that is 1 ÷ 52, about 1.9%. After one success in 51, it is 2 ÷ 53, about 3.8%. One role model doubles a village's sense of the possible, the shape of the effect the leadership study measured.

3. One person can flip a crowd (Granovetter's threshold model, 1978). A hundred people each join once enough others have; their thresholds are 0, 1, 2 and so on to 99. Person 0 starts, triggering person 1, and so on until all 100 join. Change one threshold from 1 to 2 and only person 0 ever joins. Averages cannot predict cascades; distributions can.

4. Why wedding spending spirals (a prisoner's dilemma). Two families score each outcome, higher being better:

	Other spends modestly	Other spends lavishly
You spend modestly	3, 3	1, 4
You spend lavishly	4, 1	2, 2

Whatever the other family does, spending lavishly pays you more, so both do and each gets 2 instead of 3. The cure is a binding shared rule, Ostrom's point. The payoffs are illustrative.

5. Correcting for what a source misses. The true count is about the observed count divided by coverage. Coverage of 50% to 70% means the truth lies between 1.43 and 2 times what was observed, and that range must travel with the number.

What we refined

Every layer passed through the same five checks before it was kept, and those checks caught eight recurring errors and six weaker modelling choices.

The five checks applied to each layer

1. **Absorb:** extract every number, symbol and claim.
2. **Compare:** check the layer against the running corrections log.
3. **Verify:** open the source page for any figure that can be checked live.
4. **Recompute:** re-run every piece of arithmetic in the draft.
5. **Reconcile:** map each symbol onto one shared notation before any formula is combined with another.

Errors found

#	Error	First seen	How it was caught	Correction
1	"148 crore working population"	Layer 2, repeated through Layer 6	Checked against the official projection and the age structure	Total population is 142.94 crore (2026 projection); working age 15–59 is about 90–93 crore
2	Census 2011 used as current	Layer 1	Compared with 2026 projections	Use the Technical Group's 2026 state projections
3	Unreconciled state table with mixed-year economic shares	Layer 3, about an earlier draft table	The draft flagged its own table	One reconciled projection table; output figures from one base year
4	S, E, N, r and p each used for two or three different things	Layers 3 to 5	Symbol mapping during reconcile	One notation across both papers (Appendix B)
5	Payment counts read as people or outcomes	Layer 3	Unit check	Transactions stay transactions
6	Two funnel models mixed	Layer 5	Recomputed the draft's own funnel	Each activity declares its model
7	Acting treated as a subset of wanting	Layer 6	Tested against obligation and coercion cases	Never multiply wanting by conversion for those
8	A study's result quoted as one number	Layer 6	Opened two versions of the paper	Cite the 20–25% range for parents

Weaker choices that were sharpened

- **Participation renamed q,** so p keeps one meaning: a party's reliability.
- **"Obligations" and "commitment" became "open undertakings" and "undertaking",** because not every open matter is a pledge.
- **Two action steps renamed:** "commit" became "accept", and "settle" became "close".

- **Dependency complexity became k**, linking every activity profile to the chain law.
- **The draft's multiplicative rate model is labelled a hypothesis**, not a law.
- **Unverified draft figures are marked as such**, including the informal-enterprise count and the labour-force rates.

The refined atlas specification

After refinement, the atlas is one notation, six tables and an eight-step build, with every estimate carrying its uncertainty and source.

Six tables

Table	Holds	Key fields added in refinement
`geography`	States, districts, boundaries, settlement types	Population year and source; boundary version
`population_cells`	Weighted synthetic groups by the seven dimensions	Role weights instead of exclusive boxes
`activity_catalog`	Every activity, its definition, unit and eligibility	Unit (one of nine); frequency rhythm; funnel model type
`activity_rates`	Participation q, frequency f, duration, uncertainty	Low, base and high values; coverage estimate
`dependency_patterns`	Who must take part, in what order	Parties k; handoffs; evidence level reached
`sources_and_versions`	Survey, year, method, confidence, revisions	Opened-and-checked flag; date checked

Around them sit two supporting structures: origin-destination matrices for migration, remittances and travel, and an event calendar of festivals, harvests, exams, elections and weather.

Three tiers of confidence

- **Directly observed:** transaction counts, administrative events, surveyed time use, verified outcomes.
- **Statistically estimated:** weighted survey rates, projections, small-area estimates with intervals.
- **Weakly observed:** private intentions, hidden behaviour, unreported harm, informal undertakings.

The build, in order

1. **Reconcile population** to the 142.94 crore national projection, then pull district-level figures.
2. **Build synthetic groups** from observed joint distributions, never from multiplied margins.
3. **Lay down the daily clock** from time-use data.
4. **Catalogue activities** with their unit, rhythm and funnel model declared.
5. **Attach dependencies:** the parties k, the handoffs and the evidence each step can produce.
6. **Run the success model** over conditions Z, not just the count model.
7. **Calibrate** against independent data: UPI volumes, labour-force rates, time use, service use. Correct anything that fails, however elegant.
8. **Report** by district, segment, activity and period, always with intervals.

Where to start

Begin with two kinds of activity. High-frequency daily behaviour anchors calibration, because it is the best measured. Activities with many parties (k of 3 or more) carry the thesis of The Distance, because that is where the chain law predicts the largest gap between wanting and happening.

What the atlas cannot know

The atlas can estimate how often things happen across groups. It cannot read minds, see hidden lives, or predict any single person, and it should never pretend to.

- **Private motive.** As the Bayes example shows, actions underdetermine reasons. Only people's own words, given freely, reveal why.
- **Hidden behaviour.** Stigma, fear and illegality suppress reporting. Coverage corrections widen the interval; they do not remove the blind spot.
- **Individuals.** Synthetic groups describe populations. Nothing in the atlas may be used to predict or judge a real person.

- **Causes, without experiments.** Patterns across districts are not causes. Mechanisms H1 to H8 need randomised or carefully identified designs before anyone claims an effect.
- **The present moment.** Surveys lag reality by years. The atlas is a map, not a live feed.
- **Comparisons on totals.** Simpson's paradox means districts and countries are compared within conditions, or not at all.

Privacy is part of the method

The atlas is built from anonymised survey microdata, aggregate statistics and synthetic people. It never needs identifiable records to reach its estimates, and it never infers private motives from surveillance or payment logs. Sensitive topics in any primary research need careful consent and protection from identification.

Appendix A — India's 2026 state and UT populations

The atlas's base populations: the official 2026 projections of the Technical Group on Population Projections, built on Census 2011. Figures in millions, from StatisticsTimes, which reproduces the official report. These are projections, not counts; international estimates for India run about 3% higher.

State or UT	2011 census	2026 projection	Growth in 2026 (% a year)
India	1,210.85	1,429.40	0.84
Uttar Pradesh	199.81	243.47	0.91
Bihar	104.10	132.85	1.38
Maharashtra	112.37	129.58	0.72
West Bengal	91.28	100.63	0.43
Madhya Pradesh	72.63	89.97	1.10
Rajasthan	68.55	83.88	0.99
Tamil Nadu	72.15	77.58	0.24
Gujarat	60.44	74.34	1.13
Karnataka	61.10	69.07	0.58
Andhra Pradesh	49.39	53.74	0.29
Odisha	41.97	47.22	0.57
Jharkhand	32.99	41.11	1.19
Telangana	35.19	38.67	0.43
Assam	31.21	36.82	0.88
Kerala	33.41	36.24	0.35
Haryana	25.35	31.41	1.13
Punjab	27.74	31.37	0.58
Chhattisgarh	25.55	31.31	1.06
Delhi (UT)	16.79	22.67	1.78
Jammu and Kashmir (UT)	12.27	13.93	0.69

State or UT	2011 census	2026 projection	Growth in 2026 (% a year)
Uttarakhand	10.09	12.03	0.97
Himachal Pradesh	6.86	7.59	0.44
Tripura	3.67	4.27	0.85
Meghalaya	2.97	3.45	0.88
Manipur	2.86	3.32	0.88
Nagaland	1.98	2.30	0.88
Puducherry (UT)	1.25	1.77	2.25
Arunachal Pradesh	1.38	1.61	0.88
Goa	1.46	1.60	0.50
Dadra and Nagar Haveli and Daman and Diu (UT)	0.59	1.59	7.71
Mizoram	1.10	1.28	0.87
Chandigarh (UT)	1.06	1.27	0.87
Sikkim	0.61	0.71	0.85
Andaman and Nicobar Islands (UT)	0.38	0.41	0.25
Ladakh (UT)	0.27	0.31	0.66
Lakshadweep (UT)	0.06	0.07	1.45

The five largest states hold about half the country; the ten largest about 74%; the twelve smallest states and UTs together about 1%.

Appendix B — Unified notation across both papers

One symbol, one meaning, across The Distance and The Atlas. Letters written as subscripts (a, d, g, s, t) are indices: activity, district, group, segment and time. They never stand alone as quantities.

Symbol	Meaning
N	Eligible people in a segment
W	People a synthetic group stands for
q	Share of eligible people who take part
f	Occurrences per participant in a period
E	Expected number of events
Σ	Add up over every segment or group
∫	Add up over a continuous range
Z	A person's conditions: money, access, counterparties, timing
π(Z)	How common each condition is
p	Chance one party does its part
k	Parties who must all do their part
$S = p^k$	Chance the whole chain succeeds
D	The distance: chance a wanted thing does not happen
b	Share of bad actors in a chain
ε	Honest slip rate
u	Chance a bad actor still completes
Q_t	Open matters at time t
A_t, C_t	Matters arriving and completed in a period
L	Average open matters
λ	Arrival rate of new matters
W (queueing)	Average time to close a matter; context separates it from group weight

Symbol	Meaning
ρ	Load: arrivals over capacity
ρ_{risk}	Risk or consequence of an activity
G_t	The network at time t
V	Actors in the network
$\mathscr{L}$	Links active at a time
R	Kinds of relationship
τ	Time an activity consumes
v	Money value of an activity
κ	A person's own control over an activity
n, f (agreement)	Group size and number of faulty parties

Two letters still carry a second meaning that context resolves: W (group weight or closing time) and f (frequency or faulty parties). Each is defined where it is used.

Appendix C — Sources

Pages opened and checked on 19 October 2026, followed by figures taken from the drafts that were not independently verified.

Source	Used for
StatisticsTimes, reproducing the Technical Group on Population Projections	2026 state and UT populations; 142.94 crore national projection
Free Press Journal, citing NPCI, 1 September 2026	24.51 billion UPI transactions worth ₹29.82 lakh crore in August 2026; 791 million a day
Beaman, Duflo, Pande and Topalova, Science, 2012 (MIT author manuscript)	495 villages, 8,453 surveys; aspiration gap closed 32% for adolescents and 20% for parents in the manuscript; education gap erased
The Distance, companion paper	Derivations of the shared mathematics; product thesis

From the drafts, not independently verified

- About 7.34 crore non-farm unincorporated establishments (ASUSE 2023–24).
- The 2025 labour-force rates: 57.4% worker-population ratio; 56.2% self-employed, 23.6% salaried, 20.2% casual.
- Dataset descriptions and their date limits in the Layer 5 data stack.
- All illustrative funnels, frequencies and households, which the drafts themselves marked as invented.

Classic references

- Coase, R. (1937). The nature of the firm. Economica.
- Granovetter, M. (1985). Economic action and social structure: the problem of embeddedness. American Journal of Sociology.
- Granovetter, M. (1978). Threshold models of collective behavior. American Journal of Sociology.
- Ostrom, E. (1990). Governing the Commons. Cambridge University Press.
- Little, J. D. C. (1961). A proof for the queuing formula L = λW. Operations Research.
- Lamport, L., Shostak, R. and Pease, M. (1982). The Byzantine Generals Problem. ACM Transactions on Programming Languages and Systems.
- Fischer, M., Lynch, N. and Paterson, M. (1985). Impossibility of distributed consensus with one faulty process. Journal of the ACM.
- Feld, S. (1991). Why your friends have more friends than you do. American Journal of Sociology.

Cite this paper

Singh, A. (2026). *The Atlas: How we mapped what people do, need and depend on*. Opankitaab Private Limited. kitaab.live/research/the-atlas