Primary and Secondary Analysis Framework

Primary Analysis Questions (Based on Available Data)

9. Digital Readiness Trends by City Tier

- How have literacy rates, smartphone penetration, and internet penetration trended from 2019 to 2024 across different city tiers (Tier 1, Tier 2, Tier 3)?

- Which tier shows the fastest year-over-year improvement in digital readiness?

10. Language Impact on Print Circulation

- How does net circulation differ between Hindi and English editions across cities from 2019 to 2024?

- Has the circulation gap between Hindi and English widened or narrowed in states like Uttar Pradesh and Maharashtra?

11. Monthly vs. Quarterly Circulation Fluctuations

- What are the seasonal patterns in copies sold and net circulation on a monthly basis (*fact_print_sales.csv*)?

- How do these align with quarterly ad revenue trends (*fact_ad_revenue.csv*)?

12. Ad Revenue Currency Analysis

- How much ad revenue was generated in each currency (INR, USD, EUR) from 2019–2024?

- What is the year-over-year growth after converting to INR (USD=80, EUR=90)?

- Which categories contribute the most in foreign currencies?

13. Pilot Platform Performance Comparison (2021)

- Which digital pilot platforms (e.g., PDF WhatsApp Push, E-paper Mobile Web) had:

  o The lowest average bounce rates?

  o The highest downloads_or_accesses per users_reached?

- How does this vary by ad category?

14. Cost Efficiency of Digital Pilots (2021)

- What is the cost per user reached = (dev_cost + marketing_cost) ÷ users_reached?

- Which cities and platforms show the best ROI?

15. Print Returns Rate Analysis

- What is the average returns rate (copies_returned ÷ copies_sold) per city and edition from 2019–2024?

- Which states/tiers have the highest return rates, and how has this impacted net circulation?

16. Ad Category Group Performance

- How has ad revenue evolved by category group (Public Sector, Commercial Brands, Private Sector) from 2019–2024 (*dim_ad_category.csv*)?

- Which group correlates most strongly with print circulation declines?

17. Feedback Themes in Digital Pilots

- From *fact_digital_pilot.csv* cumulative_feedback, what are the top recurring issues (e.g., load times, font size)?

- How do these themes correlate with bounce rates?

18. State-Level Aggregation of Circulation & Revenue

- What is the total net circulation and ad revenue per state (aggregated from city data) from 2019–2024?

- Which states show the strongest positive correlation between circulation growth and revenue growth?

19. Year-Over-Year Decline in Specific Editions

- Which editions (e.g., ED1001, ED1005) saw the sharpest net circulation declines from 2020–2024?

- How do these declines align with changes in their ad revenue?

20. Digital Readiness vs. Pilot Engagement (2021)

- What is the correlation coefficient between city digital readiness (literacy_rate + smartphone_penetration + internet_penetration) and digital pilot metrics (users_reached, downloads_or_accesses)?

- Which cities deviate most from the trend?

21. Ad Revenue Comments Analysis

- From *fact_ad_revenue.csv* comments (e.g., "New FMCG client onboarded"), which ad categories/quarters show:

    o   Positive sentiments?

    o   Negative sentiments?

- How do these align with revenue spikes/drops?

22. Tier-Based Digital Pilot Costs (2021)

- How do dev_cost and marketing_cost vary by tier?

- Which tier has the highest cost per download_or_access, and what does this suggest about scalability?

23. Monthly Circulation Volatility

- What is the standard deviation of monthly net circulation per edition (2019–2024)?

- Which cities/languages show the highest volatility, and how does this affect ad revenue predictability?

24. Ad Revenue per Ad Category per City

- Using edition_id and city_id mappings, what is ad revenue per category per city (2019–2024)?

- Which cities dominate categories like Real Estate or Automobile?

25. Print vs. Digital Overlap (2021)

- In pilot cities, how does print net circulation compare with users_reached?

- Which cities show potential cannibalization of print by digital?

26. Readiness Metrics Quarterly Changes

- What are quarter-over-quarter % changes in smartphone_penetration & internet_penetration (2019-Q1 → 2024-Q4)?

- Which cities dropped most during COVID quarters (2020-Q1 → 2020-Q4)?

27. Example Brands Impact on Revenue

- From *dim_ad_category.csv* example_brands and comments (e.g., "New FMCG client"), how do new client onboardings correlate with Commercial Brands revenue growth?

28. Bounce Rate Trends by Launch Month (2021)

- How did avg_bounce_rate change month-over-month per pilot platform?

- Which platforms improved most, and in which cities?

29. Net Circulation Growth Rates

- What is the CAGR of net circulation per city (2019–2024)?

- Which Tier 3 cities (e.g., Ranchi) show positive growth despite overall declines?

30. Ad Revenue Share by Category

- What % of total revenue does each category contribute year-over-year (2019–2024)?

- How has the share shifted (e.g., toward/away from Government ads)?

31. Pilot Feedback vs. Readiness (2021)

- In high-literacy cities (>80%), how do feedback themes differ from low-literacy cities?

- Do these differences explain bounce rate variations?

32. Edition-Specific Waste Gap

- Per edition_ID, what is the avg gap between copies_sold and net_circulation (2019–2024)?

- Which Hindi vs. English editions have the largest waste gaps?

33. Currency Conversion Impact on Trends

- After converting revenues to INR, how do overall trends shift (2019–2024)?

- Which quarters show artificial inflation due to FX?

34. Marketing Cost Efficiency (2021)

- What is marketing_cost ÷ users_reached for pilots?

- Which categories/cities show the best efficiency?

35. State × Tier Aggregation

- By state and tier, which combos (e.g., Tier 2 in Uttar Pradesh) have:
    - Highest net circulation?
    - Lowest digital readiness?

36. Pilot Platform Success by Ad Category (2021)

- Which ad_category_id had the highest downloads_or_accesses ÷ users_reached?
- How does this vary across platforms like Mobile App Beta?

37. Returns Rate vs. Readiness (2019–2024)

- Is there a negative correlation between city internet_penetration and print returns?
- How has this relationship evolved?

38. Quarterly Ad Revenue Volatility

- What is the variance in ad_revenue per quarter (2019–2024)?
- Which categories (e.g., Real Estate) show highest volatility?

39. City Tier vs. Performance

- By tier, how does per-city ad revenue correlate with net circulation?
- Do Tier 1 cities outperform Tier 2/3 in ROI?

40. Digital Pilot Cumulative Metrics (2021)

- What is the total users_reached and downloads_or_accesses across all pilots?
- Which months/cities contributed most to growth?

Secondary Analysis & Recommendations

6. Optimizing Print Waste Reduction → Target demand forecasting AI to reduce return rates.
7. Phased Digital Rollout → Prioritize low-readiness cities with PDF WhatsApp Push.
8. Advertiser Retention → Personalized outreach for declining categories (e.g., Real Estate).
9. Content Customization → Audio summaries for low-literacy areas to boost engagement.
10. Hybrid Subscription Models → Bundle digital access with print subscriptions.
11. Influencer Partnerships → Leverage local influencers in low-engagement states.

12. **Pilot Cost-Benefit Analysis** → Budget scaling toward Tier 2 cities with best ROI.

13. **Loyalty Programs** → Rewards linked to feedback themes to improve retention.

14. **Cross-Category Bundling** → Package weak categories with stable ones (Govt + FMCG).

15. **Regional Journalist Empowerment** → Train reporters with digital tools for local trust.

16. **Paywall Strategies** → Dynamic pay-per-article/freemium for high-readiness metros.

17. **Seasonal Marketing** → Festive campaign pushes to stabilize Q4 revenue.

18. **Data-Driven Expansion** → Use readiness data to expand pilots into Tier 3.

19. **Feedback Loop** → Rapid content iteration using customer feedback.

20. **Tech Partnerships** → Collaborate with firms to optimize load times.

21. **Revenue Diversification** → Explore e-commerce & sponsored content beyond ads.

22. **Training Programs** → Bridge readiness gaps in low-internet cities.

23. **Competitor Benchmarking** → Track recovery against industry peers.

24. **Sustainability in Print** → Eco-friendly printing to cut costs in high-returns cities.

25. **Multi-Language Digital Strategy** → Tailor Hindi vs. English for regions like UP.

26. **ROI Forecasting Models** → Predictive models for city-level digital ROI.

27. **Crisis Response Planning** → Apply COVID-era lessons for future resilience.

28. **User Segmentation** → Segment pilot users by inferred demographics for targeting.

29. **Brand Rebuilding Campaigns** → PR efforts to regain advertiser trust in weak markets.

30. **Unified Metrics** → Combine print & digital KPIs for holistic tracking.