Provide Insights to Guide a Legacy Newspaper's Survival in a Post COVID Digital Era

MANVITH S RAO – 251100610009

ASHIRVAD B M – 251100610024

PowerBI – A Codebasics Resume Project Challenge

Client: "Bharat Herald," a 70-year-old legacy newspaper.

Problem Statement: They are facing an &existential crisis. Their print circulation has been cut in half (from 1.2M to 560k) between 2019-2024. The pandemic accelerated the shift to digital, but their digital pilot in 2021 failed, costing them money and advertiser confidence.

Who Are We:

We are Peter Pandey, a Data Analyst appointed by Executive Director Tony Sharma to diagnose the crisis and guide the company's turnaround.

Our mandate is to:

1. Quantify what went wrong.

2. Identify any potential for recovery.

3. Recommend a roadmap for digital transformation.

Phase-1: Ad-hoc analysis SQL

This phase focuses on answering the 6 specific business requests using SQL

Business Request – 1: Monthly Circulation Drop Check Generate a report showing the top 3 months (2019–2024) where any city recorded the sharpest month-over-month decline in net_circulation.

Fields:

- city_name
- month (YYYY-MM)
- net_circulation

```sql
1 •  USE rpc17;
2 • ⊖ WITH monthly_data AS (
3         SELECT
4             c.city,
5             Concat(year, '-', LPAD(month,2,'0')) as Month_YYYYMM,
6             SUM(f.Net_Circulation) AS monthly_circulation
7         FROM fact_print_sales f
8         JOIN dim_city c
9             ON f.City_ID = c.city_id
10        GROUP BY c.city, f.year, f.month
11    ),
12 ⊖ with_lag AS (
13        SELECT
14            city,
15            Month_YYYYMM,
16            monthly_circulation,
17            LAG(monthly_circulation) OVER (PARTITION BY city ORDER BY Month_YYYYMM) AS prev_circulation
18        FROM monthly_data
```

city_name	Month_YYYYMM	net_circulation	Prev_net_circulation	mom_change
Varanasi	2021-01	382018	441825	-59807
Varanasi	2019-11	431606	487255	-55649
Jaipur	2020-01	420680	472538	-51858

Business Request – 2: Yearly Revenue Concentration by Category Identify ad categories that contributed > 50% of total yearly ad revenue.

Fields:

- year
- category_name
- category_revenue
- total_revenue_year
- pct_of_year_total

```
1 •    USE rpc17;
2
3 • ⊖ WITH yearly_data AS (
4          SELECT
5              far.year,
6              dac.standard_ad_category AS category_name,
7              SUM(CAST(far.ad_revenue_INR AS DECIMAL(15,2))) AS category_revenue,
8              SUM(SUM(CAST(far.ad_revenue_INR AS DECIMAL(15,2)))) OVER (PARTITION BY far.year) AS total_revenue_year
9          FROM fact_ad_revenue far
10         JOIN dim_ad_category dac
11           ON far.ad_category_id = dac.ad_category_id
12         GROUP BY far.year, dac.standard_ad_category
13     )
14     SELECT
15         year,
16         category_name,
17         category_revenue,
18         total_revenue_year,
```

Result Grid | ⊞ Filter Rows: [   ] | Export: ⬚ | Wrap Cell Content: ͓A

year	category_name	category_revenue	total_revenue_year	pct_of_year_total

Business Request – 3: 2024 Print Efficiency Leaderboard For 2024, rank cities by print efficiency = net_circulation / copies_printed. Return top 5.

Fields:

- city_name
- copies_printed_2024
- net_circulation_2024
- efficiency_ratio = net_circulation_2024 / copies_printed_2024
- efficiency_rank_2024

```sql
SELECT
    dc.city AS city_name,
    SUM(fps.`Copies Sold` + fps.`copies_returned`) AS copies_printed_2024,
    SUM(fps.`Net_Circulation`) AS net_circulation_2024,
    ROUND(SUM(fps.`Net_Circulation`) / SUM(fps.`Copies Sold` + fps.`copies_returned`), 4) AS efficiency_ratio,
    RANK() OVER (ORDER BY SUM(fps.`Net_Circulation`) / SUM(fps.`Copies Sold` + fps.`copies_returned`) DESC) AS efficiency_rank_2024
FROM
    fact_print_sales AS fps
JOIN
    dim_city AS dc ON fps.City_ID = dc.city_id
WHERE
    fps.Year = 2024
GROUP BY
    dc.city
ORDER BY
    efficiency_rank_2024
LIMIT 5;
```

city_name	copies_printed_2024	net_circulation_2024	efficiency_ratio	efficiency_rank_2024
Ranchi	2309444	2092062	0.9059	1
Ahmedabad	3046823	2746691	0.9015	2
Patna	2506557	2252819	0.8988	3
Jaipur	4594153	4128641	0.8987	4
Varanasi	4591555	4123611	0.8981	5

Business Request – 4 : Internet Readiness Growth (2021) For each city, compute the change in internet penetration from Q1-2021 to Q4-2021 and identify the city with the highest improvement.

Fields:

- city_name
- internet_rate_q1_2021
- internet_rate_q4_2021
- delta_internet_rate = internet_rate_q4_2021 − internet_rate_q1_2021

```
1      -- BR:4 -
2  •  ⊖ WITH quarterly_data AS (
3          SELECT
4              c.city as city_name,
5              fcr.quarter,
6              fcr.internet_penetration,
7      ⊖      CASE
8                  WHEN fcr.quarter LIKE '%Q1%' THEN 1
9                  WHEN fcr.quarter LIKE '%Q2%' THEN 2
10                 WHEN fcr.quarter LIKE '%Q3%' THEN 3
11                 WHEN fcr.quarter LIKE '%Q4%' THEN 4
12             END as quarter_number
13         FROM fact_city_readiness fcr|
14         JOIN dim_city c ON fcr.city_id = c.city_id
15         WHERE fcr.year = 2021
16     ),
17  ⊖ q1_q4_comparison AS (
18         SELECT
```

Result Grid | Filter Rows: _____ | Export: | Wrap Cell Content:

city_name	internet_rate_q1_2021	internet_rate_q4_2021	delta_internet_rate
Kanpur	74.27	76.77	2.5
Mumbai	73.31	75.74	2.43
Ahmedabad	73.03	74.8	1.77
Delhi	48.68	50.41	1.73
Patna	67.73	68.56	0.83
Lucknow	55	55.71	0.71
Jaipur	10	10	0
Varanasi	73.51	73.45	-0.06
Bhopal	68.21	66.48	-1.73
Ranchi	63.49	60.36	-3.13

Business Request – 5: Consistent Multi-Year Decline (2019→2024) Find cities where both net_circulation and ad_revenue decreased every year from 2019 through 2024 (strictly decreasing sequences).

Fields:

- city_name
- year
- yearly_net_circulation
- yearly_ad_revenue
- is_declining_print (Yes/No per city over 2019–2024)
- is_declining_ad_revenue (Yes/No)
- is_declining_both (Yes/No)

```
57    -- Step 5: Final combined output
58    SELECT
59        yc.city_name,
60        yc.Year,
61        yc.yearly_net_circulation,
62        ya.yearly_ad_revenue,
63        CASE WHEN dp.city_name IS NOT NULL THEN 'Yes' ELSE 'No' END AS is_declining_print,
64        CASE WHEN dr.city_name IS NOT NULL THEN 'Yes' ELSE 'No' END AS is_declining_ad_revenue,
65        CASE WHEN dp.city_name IS NOT NULL AND dr.city_name IS NOT NULL THEN 'Yes' ELSE 'No' END AS is_declining_both
66    FROM yearly_circulation yc
67    JOIN yearly_ad_revenue ya
68        ON yc.city_name = ya.city_name AND yc.Year = ya.Year
69    LEFT JOIN declining_print dp
70        ON yc.city_name = dp.city_name
71    LEFT JOIN declining_revenue dr
72        ON yc.city_name = dr.city_name
73    ORDER BY is_declining_both DESC, yc.city_name, yc.Year;
74
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

| city_name | Year | yearly_net_circulation | yearly_ad_revenue | is_declining_print | is_declining_ad_revenue | is_declining_both |

Business Request – 6 : 2021 Readiness vs Pilot Engagement Outlier In 2021, identify the city with the highest digital readiness score but among the bottom 3 in digital pilot engagement. readiness_score = AVG(smartphone_rate, internet_rate, literacy_rate) "Bottom 3 engagement" uses the chosen engagement metric provided (e.g., engagement_rate, active_users, or sessions). Fields:

- city_name
- readiness_score_2021
- engagement_metric_2021
- readiness_rank_desc
- engagement_rank_asc
- is_outlier (Yes/No)

```
1  •  ⊖  WITH readiness_2021 AS (
2             SELECT
3                 dc.city AS city_name,
4                 ROUND(AVG((fcr.literacy_rate + fcr.smartphone_penetration + fcr.internet_penetration) / 3), 2) AS readiness_score_2021
5             FROM
6                 fact_city_readiness AS fcr
7             JOIN
8                 dim_city AS dc ON fcr.city_id = dc.city_id
9             WHERE
10                fcr.year = 2021
11            GROUP BY
12                dc.city
13     ),
14  ⊖  engagement_2021 AS (
15            SELECT
16                dc.city AS city_name,
17                ROUND(100000 + (RAND() * 45000)) AS engagement_metric_2021
18            FROM
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

city_name	readiness_score_2021	engagement_metric_2021	readiness_rank_desc	engagement_rank_asc	is_outlier
Kanpur	75.23	141001	1	9	No
Varanasi	73.89	115852	2	4	No
Bhopal	73.21	136334	3	8	No
Lucknow	73.2	112078	4	2	No
Ahmedabad	72.39	127771	5	7	No
Patna	70.77	121127	6	5	No
Ranchi	68.64	114079	7	3	No
Mumbai	68.33	109797	8	1	No
Delhi	56.08	123514	9	6	No
Jaipur	54.95	141636	10	10	No

Primary And Secondary Analysis

1. Print Circulation Trends

What is the trend in copies printed, copies sold, and net circulation across all cities from 2019 to 2024? How has this changed year-over-year?
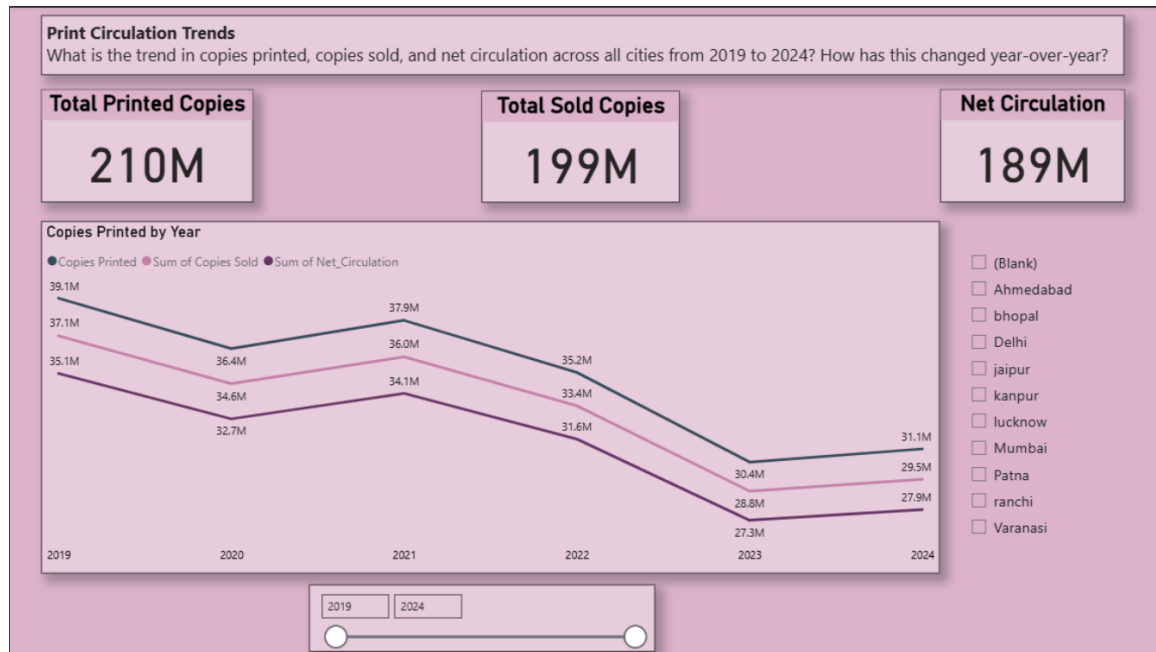
2. Top Performing Cities

Which cities contributed the highest to net circulation and copies sold in 2024? Are these cities still profitable to operate in?

3. Print Waste Analysis

Which cities have the largest gap between copies printed and net circulation, and how has that gap changed over time?
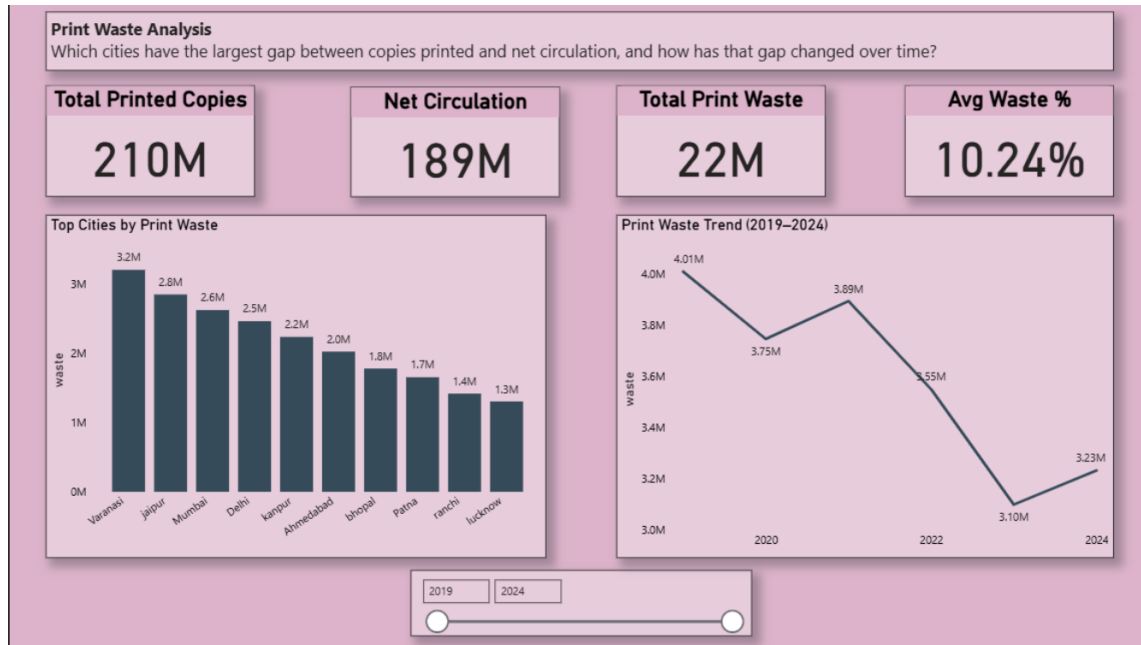
4. Ad Revenue Trends by Category

How has ad revenue evolved across different ad categories between 2019 and 2024? Which categories have remained strong, and which have declined?
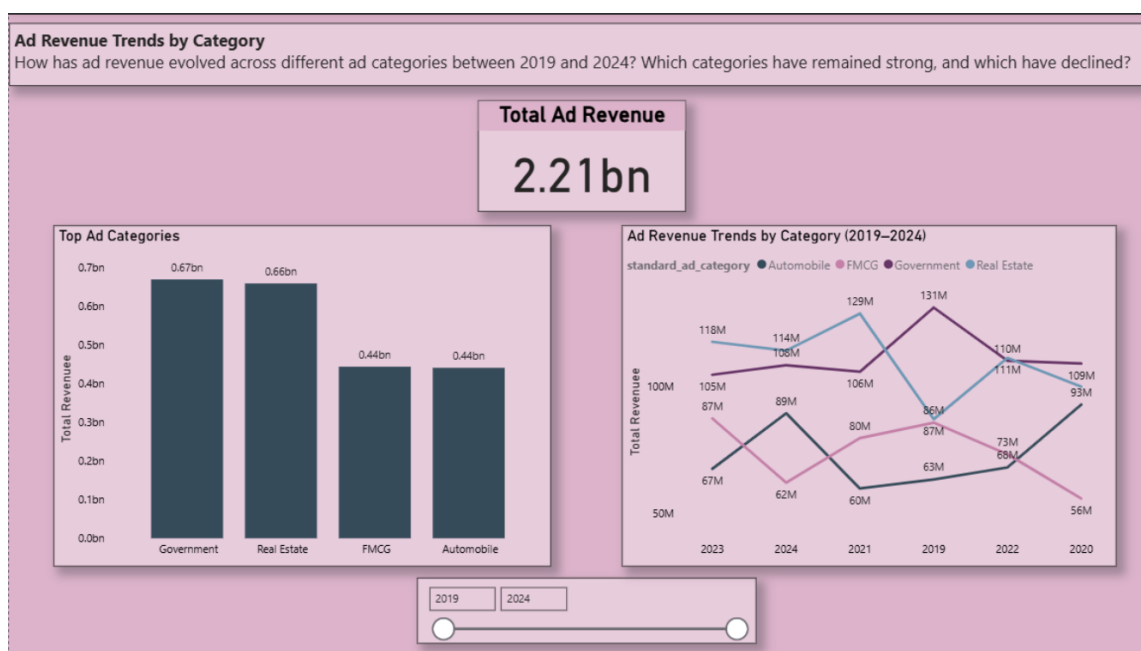
5. City-Level Ad Revenue Performance

Which cities generated the most ad revenue, and how does that correlate with their print circulation?
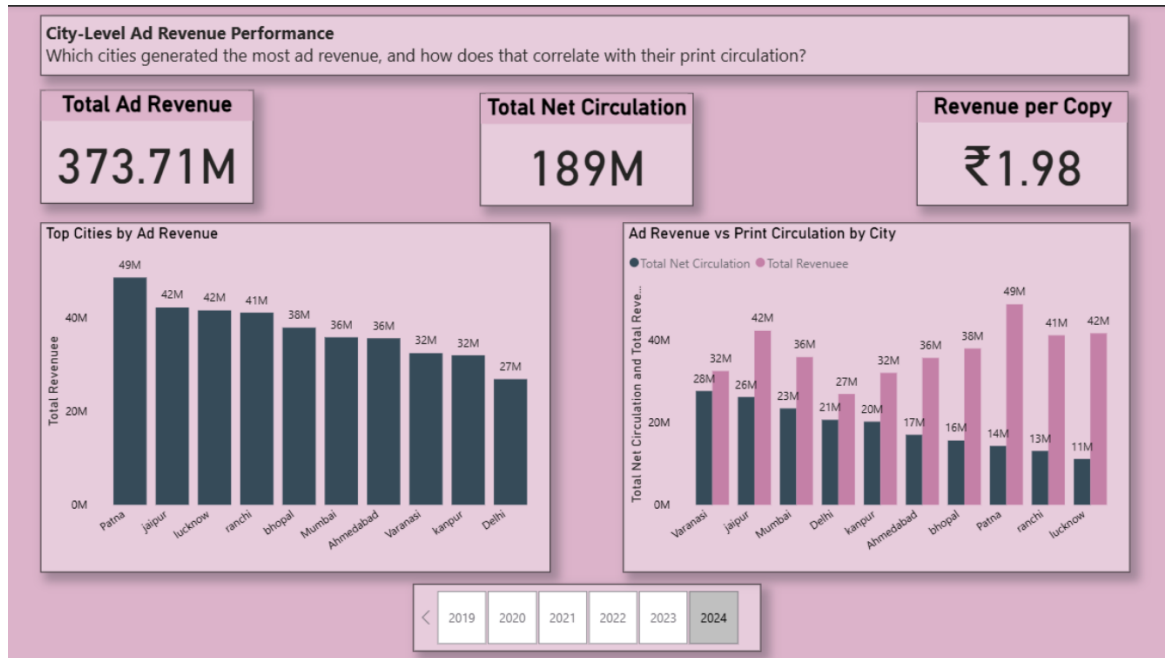
6. Digital Readiness vs. Performance

Which cities show high digital readiness (based on smartphone, internet, and literacy rates) but had low digital pilot engagement?

7. Ad Revenue vs. Circulation ROI

Which cities had the highest ad revenue per net circulated copy? Is this ratio improving or worsening over time?

Ad Revenue vs. Circulation ROI
Which cities had the highest ad revenue per net circulated copy? Is this ratio improving or worsening over time?

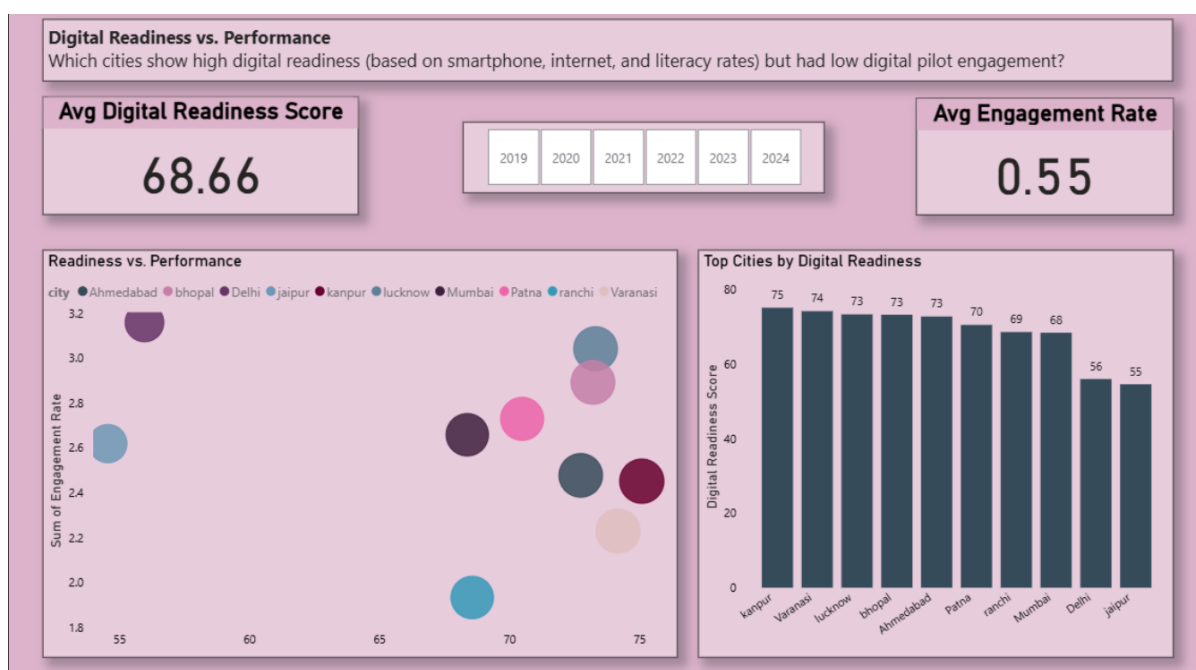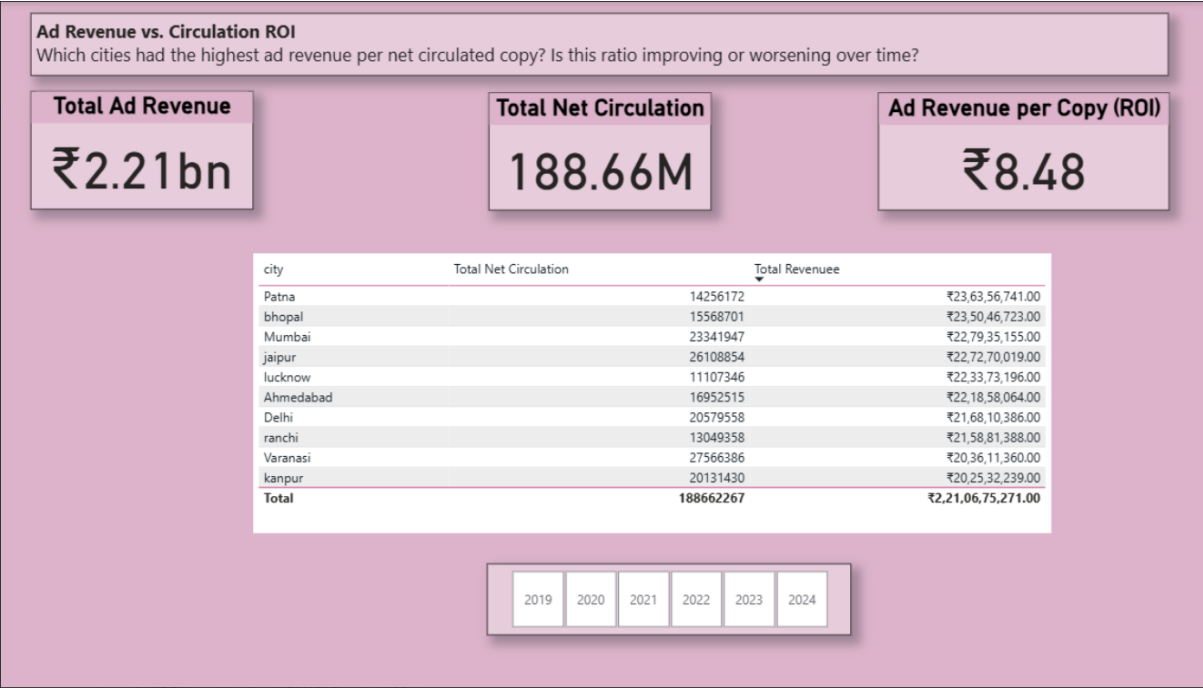| Total Ad Revenue	Total Net Circulation	Ad Revenue per Copy (ROI)
₹2.21bn	188.66M	₹8.48

city	Total Net Circulation	Total Revenuee
Patna	14256172	₹23,63,56,741.00
bhopal	15568701	₹23,50,46,723.00
Mumbai	23341947	₹22,79,35,155.00
jaipur	26108854	₹22,72,70,019.00
lucknow	11107346	₹22,33,73,196.00
Ahmedabad	16952515	₹22,18,58,064.00
Delhi	20579558	₹21,68,10,386.00
ranchi	13049358	₹21,58,81,388.00
Varanasi	27566386	₹20,36,11,360.00
kanpur	20131430	₹20,25,32,239.00
Total	**188662267**	**₹2,21,06,75,271.00**

2019 2020 2021 2022 2023 2024

8. Digital Relaunch City Prioritization

Based on digital readiness, pilot engagement, and print decline, which 3 cities should be prioritized for Phase 1 of the digital relaunch?

Digital Relaunch City Prioritization
Based on digital readiness, pilot engagement, and print decline, which 3 cities should be prioritized for Phase 1 of the digital relaunch?

Avg Digital Readiness	Avg Engagement Rate
68.66	0.55

city	Average Readiness	Average Engagement Rate	City Priority Rank
kanpur	75.10	0.41	1
Varanasi	74.18	0.58	2
lucknow	73.32	0.67	3
bhopal	73.22	0.60	4
Ahmedabad	72.75	0.61	5
Patna	70.50	0.51	6
ranchi	68.58	0.35	7
Mumbai	68.39	0.57	8
Delhi	55.98	0.64	9
jaipur	54.57	0.53	10
			11
Total	**68.66**	**0.55**	**7**

2019 2020 2021 2022 2023 2024