**Project Title : Seller\_enrollment\_analysis**

**Problem Statement:**

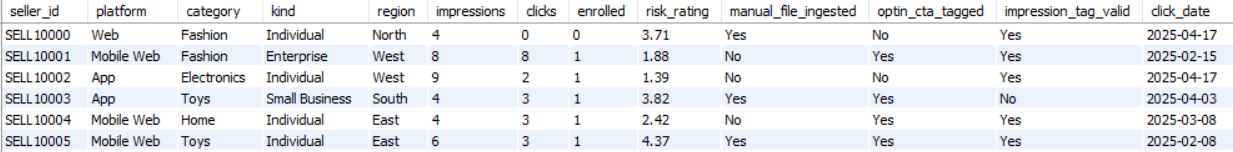
Identify seller enrollment rates by identifying gaps using impression/click tracking data, analyzing seller data , and file-based ingestion events.

**Dataset Overview:**

1. **seller\_id**: Unique identifier for each seller (e.g., SELL10000, SELL10001).
2. **platform**: The platform where the seller interacted with the enrollment campaign.
   * Values: Web, Mobile Web, App
   * Indicates whether the seller used a desktop browser, mobile browser, or mobile app.
3. **category**: The product category the seller belongs to.
   * Values: Fashion, Electronics, Toys, Home, Books
4. **kind**: The type of seller entity.
   * Values: Individual, Small Business, Enterprise.
   * Classifies sellers based on their business size.
5. **region**: The geographical region of the seller.
   * Values: North, South, East, West
6. **impressions**: Number of times the seller viewed the enrollment campaign/ad.
7. **clicks**: Number of times the seller clicked on the enrollment campaign.
   * Measures engagement.
8. **enrolled**: Binary flag (0 or 1) indicating whether the seller enrolled in the program.
   * **0**: Did not enroll.
   * **1**: Enrolled.
9. **risk\_rating**: A numerical score from (0 to 1)representing the perceived risk associated with the seller (e.g., He may be a fraud , He may deliver lately everytime etc).
10. **manual\_file\_ingested**: Indicates whether a seller's Documents/data was **manually reviewed by a human** vs. automated systems based on risk.
11. **optin\_cta\_tagged**: Flags whether the seller received a **targeted Call-to-Action (CTA)** to encourage or improve enrollment.
12. **impression\_tag\_valid**: Whether the impression tracking tag working correctly.
    * Values: Yes, No
13. **click\_date**: Date when the seller clicked the campaign (MM/DD/YYYY).
    * Blank if no click occurred.
14. **enrollment\_date**: Date when the seller enrolled (MM/DD/YYYY).
    * Blank if not enrolled.
15. **product\_opted**: The type of service the seller chose while joining in Ecommerce platform.
16. **campaign\_id**: Identifier for the marketing campaign (e.g., CMP195).
17. **seller\_tenure\_months**: How long the seller has been active on the platform (in months).
18. **conversion\_rate**: Ratio of enrollments to clicks (enrolled÷ clicks).

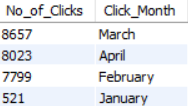
**Analyzing Scarce\_items\_analysis dataset using sql**

**1) Displaying all values in the table**

Select \* from seller\_enrollment; ****

**2) Finding No of clicks from each Month**

select Count(Clicks) as No\_of\_Clicks,Monthname(click\_date)AS Click\_Month from seller\_enrollment group by Click\_Month ;



**Explanation:**  
 Here it shows the no of clicks done by the seller(**represent no of seller clicked and viewed the event ads**) on particular month .Mostly the highest no of clicks done on holiday months only.

**3) Identifying enrollment percentage**

SELECT

(COUNT(CASE WHEN enrolled = 'Yes' THEN 1 END) \* 100.0 / COUNT(\*)) AS yes\_percentage

FROM seller\_enrollment

ORDER BY risk\_rating DESC;



**Explanation:**Total enrollment done by the seller which is very lesser

**4) Finding the seller who are risky**

select seller\_id,risk\_rating from seller\_enrollment where risk\_rating>2.5 ;



**Explanation:**  
The sellers who are risky are mostly not allowed to enroll in new program because they are sent to high verification process before enrollment.

**5) Finding campaign\_id which has highest enrollment rate**

SELECT

campaign\_id,

COUNT(seller\_id) AS total\_sellers,

SUM(enrolled) AS total\_enrollments,

ROUND((SUM(enrolled) / COUNT(seller\_id)) \* 100, 2) AS enrollment\_rate

FROM

seller\_enrollment

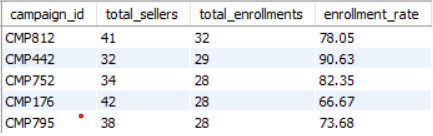
GROUP BY

campaign\_id

ORDER BY

total\_enrollments DESC

LIMIT 5;



**Explanation:**This shows the campaign\_id (promotional technique) which has highest no of enrollment rate .so we need to use this **technique** to convert other sellers to enroll in new event.

**6)** **Finding conversion rate**

SELECT

category,

COUNT(DISTINCT seller\_id) AS total\_sellers,

SUM(impressions) AS total\_impressions,

SUM(clicks) AS total\_clicks,

SUM(enrolled) AS total\_enrollments,

-- Conversion rates

ROUND(SUM(clicks) \* 100.0 / NULLIF(SUM(impressions), 0), 2) AS click\_through\_rate,

ROUND(SUM(enrolled) \* 100.0 / NULLIF(SUM(clicks), 0), 2) AS click\_to\_enroll\_rate,

-- Risk profile

ROUND(AVG(risk\_rating), 2) AS avg\_risk\_rating

-- Platform distribution

FROM

seller\_enrollment

GROUP BY

category

ORDER BY

total\_enrollments DESC;

A screenshot of a computer

AI-generated content may be incorrect.

**Explanation:**Here **Click\_through\_rate** represents no of seller converted from

**viewing(impressions)**to **Clicks(Clicking the event)** and **Click\_to\_enroll\_rate**

represents no of sellers converted from clicks to enrollment**(**not only clicking and

seeing the event and also enrolling it**).**

**7) Finding sellers based on their product\_opted(schemes joined)**

SELECT

product\_opted,

Round( COUNT(seller\_id)\*100/(select count(seller\_id)from seller\_enrollment),2) AS total\_sellers

from

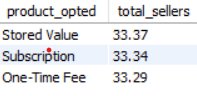
seller\_enrollment

GROUP BY

product\_opted

ORDER BY

total\_sellers DESC;



**Explanation:**  
Here it shows the **product\_opted**(schemes) undertaken by sellers based on their convenience.The product\_opted was taken equally by sellers.

**Product\_opted types:**  
**Stored\_values:**

 Here sellers deposit funds to his seller account for giving commisson to ecommerce platform for displaying and selling his products.

**Subscription:**Seller needs to pay every month for using the platform and accessing premium features like low commisson rates,no fees for listing products ,no shipment costs like that etc.

**One Time fee:**It’s like seller need to pay full amount to get lifetime access for using all features in the platform like low commisson rates etc.

**8) Finding enrollment percentage where seller had an opt-in CTA (Call-to-Action) tagged as 'Yes':**

SELECT

COUNT(DISTINCT seller\_id) AS total\_sellers\_with\_cta,

SUM(enrolled) AS enrolled\_sellers\_with\_cta,

ROUND(SUM(enrolled) \* 100.0 / COUNT(DISTINCT seller\_id), 2) AS enrollment\_percentage

FROM

seller\_enrollment

WHERE

**optin\_cta\_tagged = 'Yes';**



**Explanation:**Here optin\_cta\_tagged = 'Yes' means seller got a strong push to enroll in particular program through promotions conducted by platform owners in holidays like (Big Billion days etc ..).Seller who are got a strong push, has enrolled in this program mostly.

**9) Finding the seller’s risk rating based on seller\_tenure\_months (no of months experience in using platform).  
  
i) Sql query for sellers who are there in this platform for less than 1 year.**

SELECT

CASE

WHEN risk\_rating < 2.0 THEN 'Low Risk (0-2)'

WHEN risk\_rating BETWEEN 2.0 AND 3.5 THEN 'Medium Risk (2-3.5)'

ELSE 'High Risk (>3.5)'

END AS risk\_category,

COUNT(\*) AS seller\_count,

ROUND(COUNT(\*) \* 100.0 / (SELECT COUNT(\*)

FROM seller\_enrollment

WHERE seller\_tenure\_months < 12), 2) AS percentage

FROM

seller\_enrollment

WHERE

seller\_tenure\_months < 12

GROUP BY

risk\_category

ORDER BY

risk\_category;

A screenshot of a computer

AI-generated content may be incorrect.

**ii**) **Sql query for sellers who are there in this platform for more than 2.5 to 3 years**

SELECT

CASE

WHEN risk\_rating < 2.0 THEN 'Low Risk (0-2)'

WHEN risk\_rating BETWEEN 2.0 AND 3.5 THEN 'Medium Risk (2-3.5)'

ELSE 'High Risk (>3.5)'

END AS risk\_category,

COUNT(\*) AS seller\_count,

ROUND(COUNT(\*) \* 100.0 / (SELECT COUNT(\*)

FROM seller\_enrollment

WHERE seller\_tenure\_months between 30 and 36), 2) AS percentage

FROM

seller\_enrollment

WHERE

seller\_tenure\_months between 30 and 36

GROUP BY

risk\_category

ORDER BY

risk\_category;

A screenshot of a computer

AI-generated content may be incorrect.

**Explanation:**In Both **i)** and **ii**) risk\_rating is similar for both(Experienced and Newcomer).

Basically experienced sellers are at low risk .But in this dataset data was total inconsistent.

**10) Finding exact impressions where impression\_tag\_valid='Yes' and grouping them based on kind(Business type).**

select kind,sum(impressions), sum(clicks),sum(enrolled) from seller\_enrollment where **impression\_tag\_valid='Yes'** group by kind ;

A screenshot of a computer

AI-generated content may be incorrect.

**Explanation:**

Here **Kind** means **business type**(seller may be individual, may contain small business or big company where they create products and sell in this platform).

This **impression\_tag\_valid='Yes'** represent the impressions count of seller is correct(**exact**) and tracker is correctly working ,If it is **‘No’** means seller may use ad blocker to avoid ads or any glitch in tracking device for not tracking correctly.

**11) Finding sellers whose manual\_file\_ingested: ‘Yes’**

Select seller\_id ,risk\_rating from seller\_enrollment where **manual\_file\_ingested='Yes';**

A screenshot of a computer screen

AI-generated content may be incorrect.

**Explanation:**

Here **manual\_file\_ingested='Yes' ,**means these sellers are riskier ,they goes under high verification process like verifying the sellers documents,ID’s etc..due to regular complaints on their products like fake products, late shipments and other serious complaints.They are mostly not allowed to enroll.

**Areas to be improved:**

* Eventhough there is more impressions but seller enrolling the program is very less may be the ad is not in creative and not in a catchy way .We need to improve the advertisement methods and add multiple offers to the seller to increase the enrollment.
* We need to remove the high risk sellers as soon as possible after two warnings.
* We need to improve the optin\_cta targeted Call-to-Action (CTA). To improve the make a creative clickbait option like **For Example**: A pop up showing **“Limited spots Enroll Now to get 2 months subscription free”**