

Take-home Assignment

Assume you are a Data Analyst in an EdTech company. Your company is focused on accelerating its growth by increasing the number of enrolled users.

Therefore, you have been asked to analyse various aspects of customer acquisition to see the status of new users' growth in your company. The insights you discover will help your business team in designing a better marketing strategy for your company.

Expected Outcome

1. Brainstorm and identify the right metrics and frame proper questions for analysis. Your analysis should help your :
 - a. Business team to understand the lead's journey and stages with scope for improvement
 - b. Business heads to understand their team performance
 - c. Managers to understand their target areas
2. In case you identify any outliers in the data set, make a note of them and exclude them from your analysis.
3. Build the best suitable dashboard presenting your insights.

Your recommendations must be backed by data insights and professional visualizations to help your business team design road maps, strategies, and action items to achieve their goals.

We are looking for in-depth analysis of the data and highly meaningful outcomes from the dashboard.

Tools to Use

- Google Data Studio (preferred), Tableau, Power Bi or any other visualization tool
- You can use BigQuery SQL if you wish, not mandatory

Overview of the Dataset

Dataset Description:

You are given a month's data for your analysis. This dataset contains the details of the leads in various stages of the customer acquisition flow.

Customer Acquisition Flow (flowchart)

Customer Acquisition Key Stages:

Lead - Awareness - Consideration - Conversion

Dataset provided contains:

- Basic details of leads
- Hierarchy structure of sales managers and their assigned leads
- Lead interaction details

- Demo watched details of leads
 - Reason for not being interested
1. **lead_basic_details**: Contains details of the leads.
 2. **sales_managers_assigned_leads_details**: Contains the details of the senior and junior sales managers and their assigned leads.
 - Each senior sales manager is assigned to 4 junior sales managers.
 - Each junior sales manager is assigned few leads in each cycle (a cycle is approximately a week).
 3. **leads_interaction_details**: Contains the details of call interactions of junior sales managers with the leads.
 - A lead can drop out at any stage of the flow. If there is no call by the junior sales manager to the lead after a certain stage then the lead is considered as dropped at that stage.
 4. **leads_demo_watched_details**: Contains the details of the demo session watched by the leads.
 5. **leads_reasons_for_no_interest**: Contains the details of the reasons given by the leads for their lack of interest.

Tables Description

lead_basic_details

- **lead_id**: unique id of the lead [string]
- **age**: age of the lead [int]
- **gender**: gender of the lead [string]
- **current_city**: city of residence of the lead [string]
- **current_education**: current education details of the lead [string]
- **parent_occupation**: occupation of the parent of the lead [string]
- **lead_gen_source**: source from which the lead is generated [string]

sales_managers_assigned_leads_details

- **snr_sm_id**: unique id of the senior sales manager [string]
- **jnr_sm_id**: unique id of the junior sales manager [string]
- **assigned_date**: date at which certain leads are assigned to junior sales manager [date]
- **cycle**: cycle in which the lead is assigned [string]
- **lead_id**: unique id of the lead [string]

leads_interaction_details:

- **jnr_sm_id**: unique id of the junior sales manager [string]
- **lead_id**: unique id of the lead [string]
- **lead_stage**: stage of the lead when contacted by junior sales manager [string]
 - Stage can be “lead”, “awareness”, “consideration”, “conversion”
- **call_done_date**: date of call done to lead by junior sales manager [date]
- **call_status**: status of the call made to the lead [string]
 - If lead answers the call then “successful”
 - In all other cases “unsuccessful”
- **call_reason**: reason for calling the lead [string]
 - The reason depends on the stage lead is in.

- If lead is in “lead” stage, then reasons could be
 - `lead_introduction`
 - `demo_scheduled`
 - `demo_not_attended`
- If lead is in “awareness” stage, then reasons could be
 - `after_demo_followup`
 - `followup_for_consideration`
- if lead is in “consideration” stage, then reasons could be
 - `interested_for_conversion`
 - `followup_for_conversion`
- if lead is in “conversion” stage, then reasons could be
 - `successful_conversion`

leads_demo_watched_details

- **lead_id**: unique id of the lead [string]
- **demo_watched_date**: date at which demo session is watched by the lead [date]
- **language**: language in which the demo session is watched by the lead [string]
 - It can be “English”, “Telugu” or “Hindi”
- **watched_percentage**: percentage of the session watched by the lead (out of 100) [float]

leads_reasons_for_no_interest

- **lead_id**: unique id of the lead [string]
- **reasons_for_not_interested_in_demo**: the reason stated by the lead for their lack of interest in watching the demo session. [string]
- **reasons_for_not_interested_in_consideration**: the reason stated by the lead for not considering the product as a solution. [string]
- **reasons_for_not_interested_in_conversion**: the reason stated by the lead for not converting.

Download the files using the link provided - <https://www.kaggle.com/datasets/nxtwaveda/data-analyst/download?datasetVersionNumber=2>