**Business Lead Project**

**Aim:**

To understand the concepts of business intelligence using leads generated by digital platforms.

Data cleaning and planning of KBRs (Key Business Requirements) through SQL to provide KBIs (Key Business Indicators) through Tableau Dashboarding.

**Business Problem:**

An education company sells online courses to industry professionals. On any given day, many professionals who are interested in the courses land on their website and browse for courses.

The company markets its courses on several websites and search engines like Google. Once these people land on the website, they might browse the courses or fill up a form for the course or watch some videos, they are known as **leads**

Our goal is to provide the best analytical solution in terms of data relevance and visualization. So that we can find KPIs that would lead into a successful fulfilment of KBRs.

**Learning Outcomes:**

To help learners understand data cleaning through SQL queries.

Version control and other ethics to follow while working with databases in MySQL.

Learn about different terms and concepts of e-commerce business and how leads are generated on a digital platform.

Business Insights that can be derived from leads data visualization through Tableau.

Detailed and powerful dashboard for leads by using it’s various metrices to provide all the hidden insights and success ratio of marketing campaigns/events run by the company.

**Initial Skill Requirements:**

Intermediate skills for MS Excel, SQL filters and logic gates, SQL DDL commands, Tableau parameters and calculated fields, Tableau Dashboarding.

**Data Dictionary:** (These are the features selected by me, although learners are free to decide their own features according to their personal KBRs and KPIs)

| **Variables** | **Description** |
| --- | --- |
| Prospect ID | A unique ID with which the customer is identified. |
| Lead Number | A lead number assigned to each lead procured. |
| Lead Origin | The origin identifier with which the customer was identified to be a lead. Includes API, Landing Page Submission, etc. |
| Lead Source | The source of the lead. Includes Google, Organic Search, Olark Chat, etc. |
| Do Not Email | An indicator variable selected by the customer wherein they select whether of not they want to be emailed about the course or not. |
| Do Not Call | An indicator variable selected by the customer wherein they select whether of not they want to be called about the course or not. |
| Converted | The target variable. Indicates whether a lead has been successfully converted or not. |
| TotalVisits | The total number of visits made by the customer on the website. |
| Total Time Spent on Website | The total time spent by the customer on the website. |
| Page Views Per Visit | Average number of pages on the website viewed during the visits. |
| Last Activity | Last activity performed by the customer. Includes Email Opened, Olark Chat Conversation, etc. |
| Country | The country of the customer. |
| Specialization | The industry domain in which the customer worked before. Includes the level 'Select Specialization' which means the customer had not selected this option while filling the form. |
| How did you hear about X Education | The source from which the customer heard about X Education. |
| What is your current occupation | Indicates whether the customer is a student, unemployed or employed. |
| What matters most to you in choosing this course | An option selected by the customer indicating what is their main motto behind doing this course. |
| Search | Indicating whether the customer had seen the ad in any of the listed items. |
| Magazine |
| Newspaper Article |
| X Education Forums |
| Newspaper |
| Digital Advertisement |
| Through Recommendations | Indicates whether the customer came in through recommendations. |
| Receive More Updates About Our Courses | Indicates whether the customer chose to receive more updates about the courses. |
| Tags | Tags assigned to customers indicating the current status of the lead. |
| Lead Quality | Indicates the quality of lead based on the data and intuition the the employee who has been assigned to the lead. |
| Update me on Supply Chain Content | Indicates whether the customer wants updates on the Supply Chain Content. |
| Get updates on DM Content | Indicates whether the customer wants updates on the DM Content. |
| Lead Profile | A lead level assigned to each customer based on their profile. |
| City | The city of the customer. |
| Asymmetrique Activity Index | An index and score assigned to each customer based on their activity and their profile |
| Asymmetrique Profile Index |
| Asymmetrique Activity Score |
| Asymmetrique Profile Score |
| I agree to pay the amount through cheque | Indicates whether the customer has agreed to pay the amount through cheque or not. |
| a free copy of Mastering The Interview | Indicates whether the customer wants a free copy of 'Mastering the Interview' or not. |
| Last Notable Activity | The last notable acitivity performed by the student. |

Phase: 1

Initially we need to have a good understanding of the project requirement and domain knowledge about how web traffic data is processed, and the terms used for tracking its several metrices.

1. Step 1: Explanation of all the metrices and dimensions recorded in this dataset.
2. Step 2: What is lead and why is it so important for any B2C business model.
3. Step 3: Working with the data in MS Excel, to study and gather initial idea about the data and selection of features with more relevance to potential leads. (Requirement gathering)
4. Step 4: Required Installations: MS Excel, MySQL, Tableau
5. Step 5: Advantages of data cleaning through SQL rather than MS Excel.

Phase: 2

Once the objective and planning of the road map for our task is completed, its time to import the dataset as a table inside a manually created schema for this project in MySQL.

1. Step 1: Data Loading. (Understanding about the different formats of a dataset and their advantages)
2. Step 2: Creating notes for the task required to be completed in data cleaning, while studying the data with the help of DML commands in SQL (eg: Distinct, Desc, date functions, count, groupby)
3. Step 3: Data cleaning in SQL.
4. Step 4: Updating the data, by setting manual values to reduce category.
5. Step 5: Using DDL commands to create different version of datasets, to secure our raw data.
6. Step 6: Need to check each selected/relevant feature and if required make changes in them and update it into the database.

Phase: 3

After cleaning the data, we need connect the database in MySQL’s server to Tableau as the final stage is to achieve the data visualization through various dynamic graphs and an interactive dashboard along with that.

1. Step 1: Connect the database using the MySQL server setup in Tableau
2. Step 2: Identify the business insights to be derived from the cleaned dataset.
3. Step 3: Create each and every graph according to the insights decided in Step 2, by using parameters and calculated fields to customise it according to our own specifications and hence making the chart dynamic.
4. Step 3: Edit the charts, to create visuals that are not directly available in Tableau. (for ex: custom made visuals like doughnut charts, butterfly graph, etc)
5. Use the different types of format and tools available for dashboard in Tableau and customise it to a design that best fits your Insights in terms of visualization.