Ed-Tech Customer Lead Prediction

Tool Used: Tableau

Ed-Tech company sells online courses to industry professionals. Many professionals interested in the courses land on their website and browse for courses. This is how campaigns on social media channels, websites, or search engines, such as Google, attract new prospects.

Once people land on the site, they can browse the courses, fill out a form for the course, or watch some videos. If they leave the site again without completing this desired activity, they are brought back with retargeting campaigns and converted to leads. When these people then fill out a form and provide their email address or phone number, they are **classified as prospects**.

Once these leads are acquired, sales team members start making calls, writing emails, etc. Through this process, some of the leads are converted, while most are not. The typical lead conversion rate is around 30%.

For example, if they receive 100 leads in a day, only about 30 of them will convert. To make this process more efficient, the company wants to **identify** the potential leads, also known as "**hot leads**". If they are able to identify these leads, the lead conversion rate should increase. The customers with a higher lead score have a higher chance of conversion, and the customers with a lower lead score have a lower chance of conversion.

The Data

The CSV data set contains information about:

- Whether the requests became customers — the column is called "Converted".
    - Behavior on the website, such as time spent, what content was viewed, etc.

- Information provided in forms
- How the visitor came to the website (search engine, referrer, direct)

The following table shows an extract from the data. Data records of 9240 persons with 37 features are available. The characteristics are stored for each lead. Some are numeric features, such as time spent visiting the website, but there are also many categorical features, such as demographic information or information from web forms.

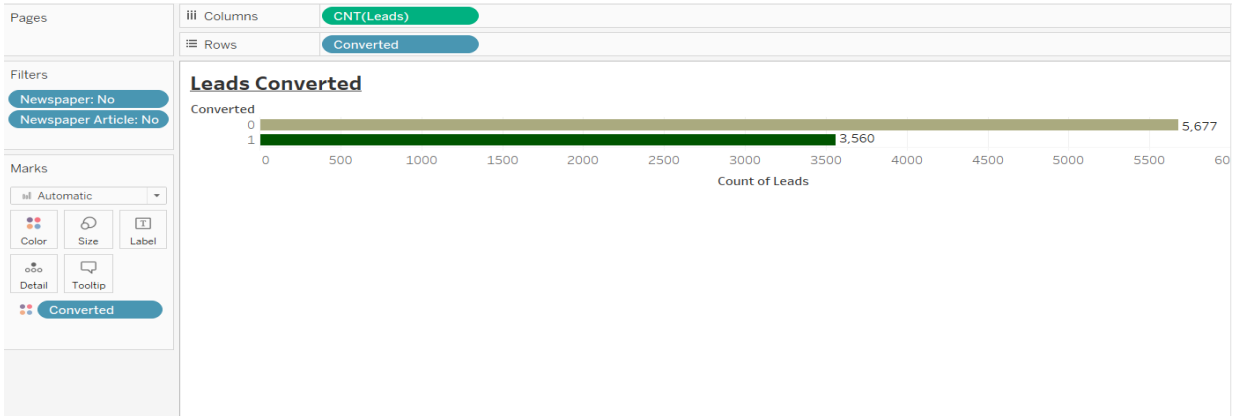
The data can be converted in to Excel file and uploaded in to Tableau canvas.

Data Cleansing

As is often the case in practice, incomplete data sets present a problem that must be solved. Simply removing all incomplete records is usually not a viable approach because too many records are affected. A detailed analysis is necessary.

Exploratory Data Analysis with Tableau

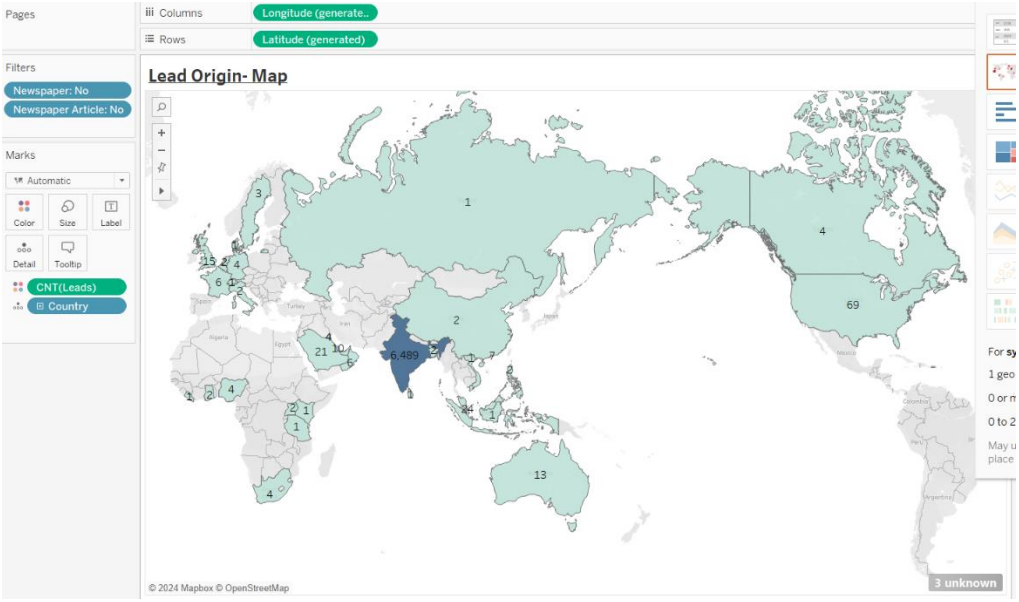
Let's first look at how many leads and non-leads there are by creating a simple bar chart. For this, "Converted" is dragged as a dimension to "rows" and the number of leads to "columns".



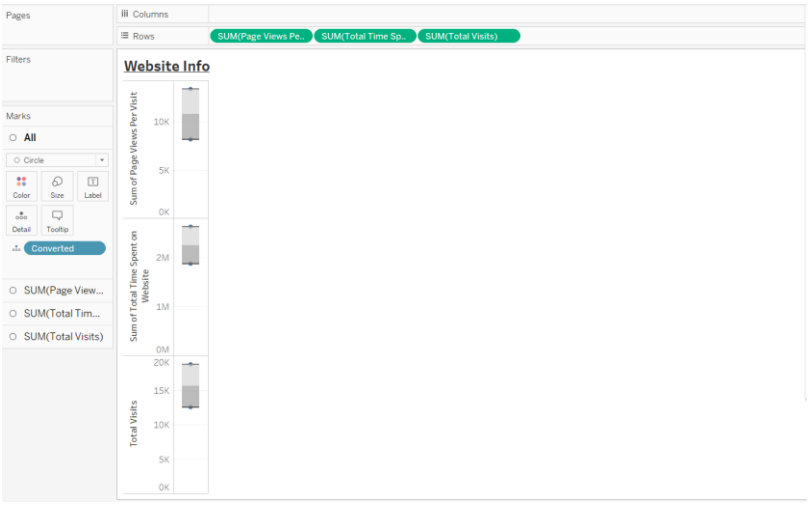
The 'Converted' feature indicates whether a lead has been successfully converted (1) or not (0). Approximately 38% of the customers in our datasets were won.

Now let's look at the regional distribution of the data by displaying a map with colored countries according to the number of records. To do this, we use the "Country" dimension and the number of leads along with the appropriate chart type. Tableau then automatically generates the Latitude and Longitude data and places them in "Rows" and "Columns".

Much of the data comes from India, as shown in the figure.



We now consider the numerical features "Total Visits", "Total Time Spent on Website" and "Page Views Per Visit". We use "box-and-whisker" plots to examine the distribution of the data.



"box-and-whisker" plots of the numerical features

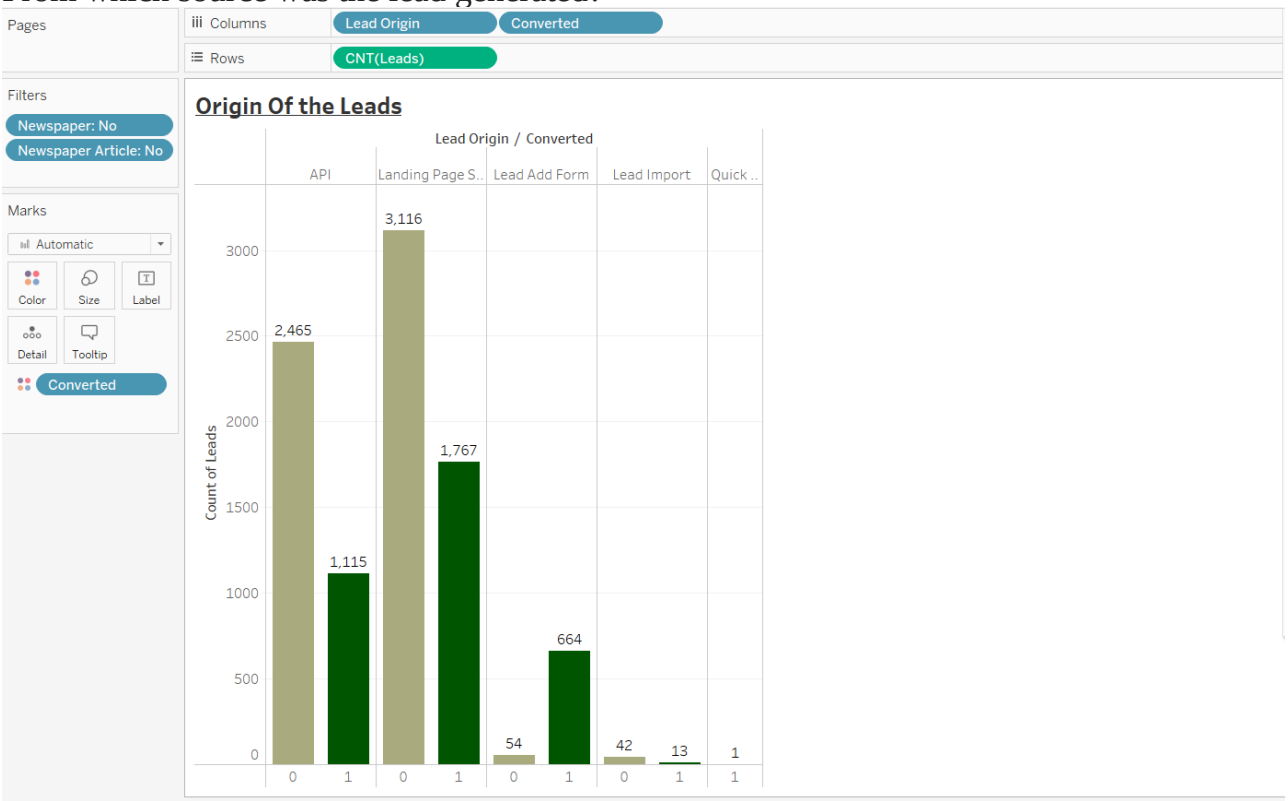
The median for the **number of visits** to the website for converted and non-converted leads is the same. Based on the total number of visits, no conclusive statement can be expected.

People who **spend more time on the website** are more likely to convert. It is recommended to improve the website to make it more helpful for the users and keep them engaged on the website.

The median **number of page views per visit** for converted and non-converted leads is the same. Nothing can be said specifically for lead conversion from the number of page views per visit.

Now we consider categorical features.

From which source was the lead generated?



A bar chart with two dimensions (Lead Origin and Converted) is used to show the lead counts.

- ‘Lead Add Form’ has a very high conversion rate, but the number of leads is not very high.
- ‘API’ and ‘Landing Page Submission’ bring a higher number of leads.
- ‘Lead Import’ generates very few leads

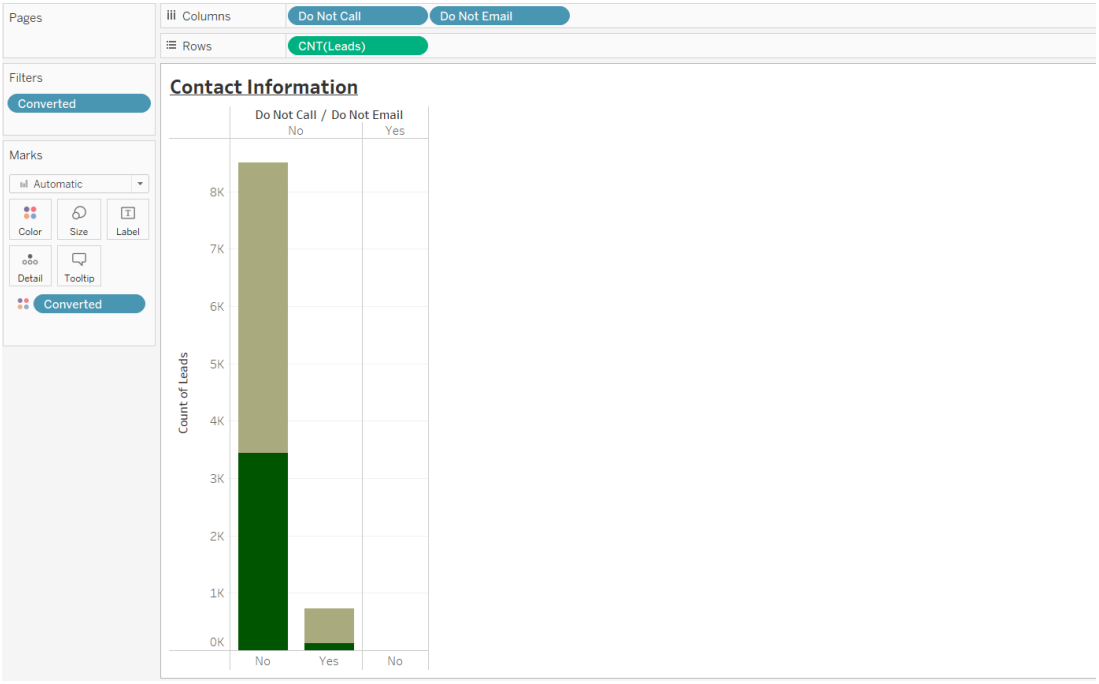
To improve the overall lead conversion rate, we need to focus more on improving the lead conversion of API- and landing page submission origin and generate more leads from the lead add form.

Lead Source

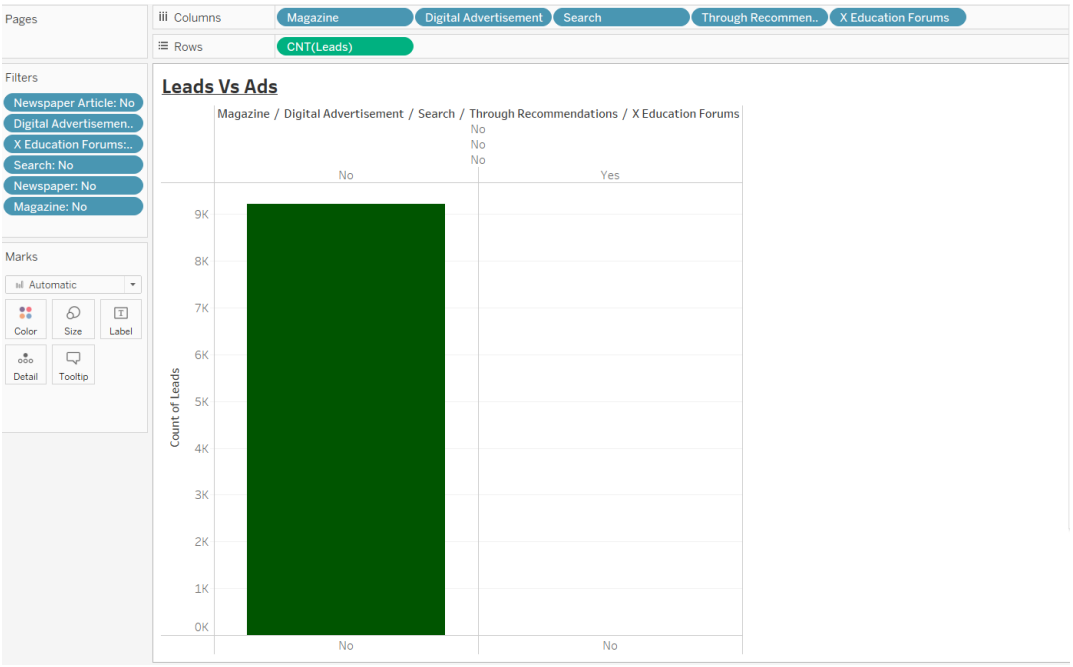
To know the lead from which source has been converted more, drag the Lead source dimension to column and sum of leads converted to rows.

The leads from google has high conversion rate.

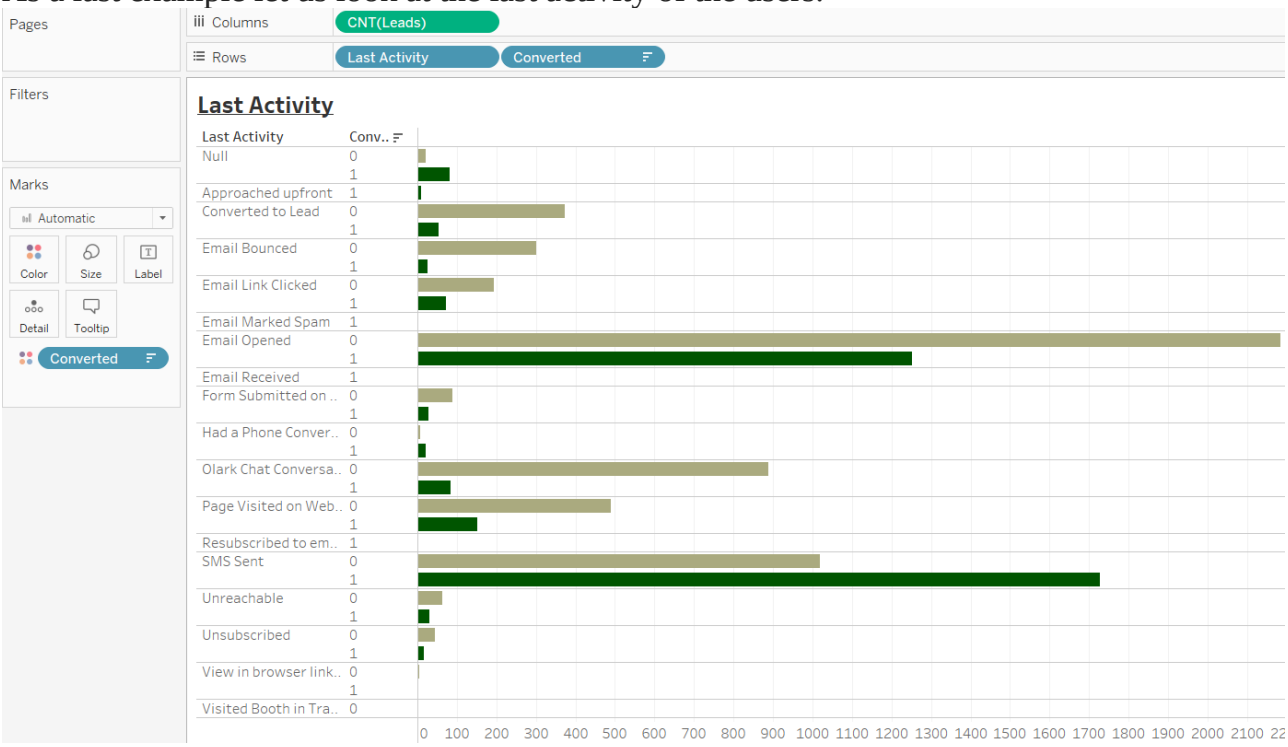
We see that people who don’t want to receive email have a low conversion rate.



Spending on Ads in Magazines, Digital Advertisement, Educational Forum etc. has no impact on conversion of leads. Its visible from the follow Bar diagram that 100% of the leads both converted and non-converted has ‘No’ influence on them.



As a last example let us look at the last activity of the users.



Last activity

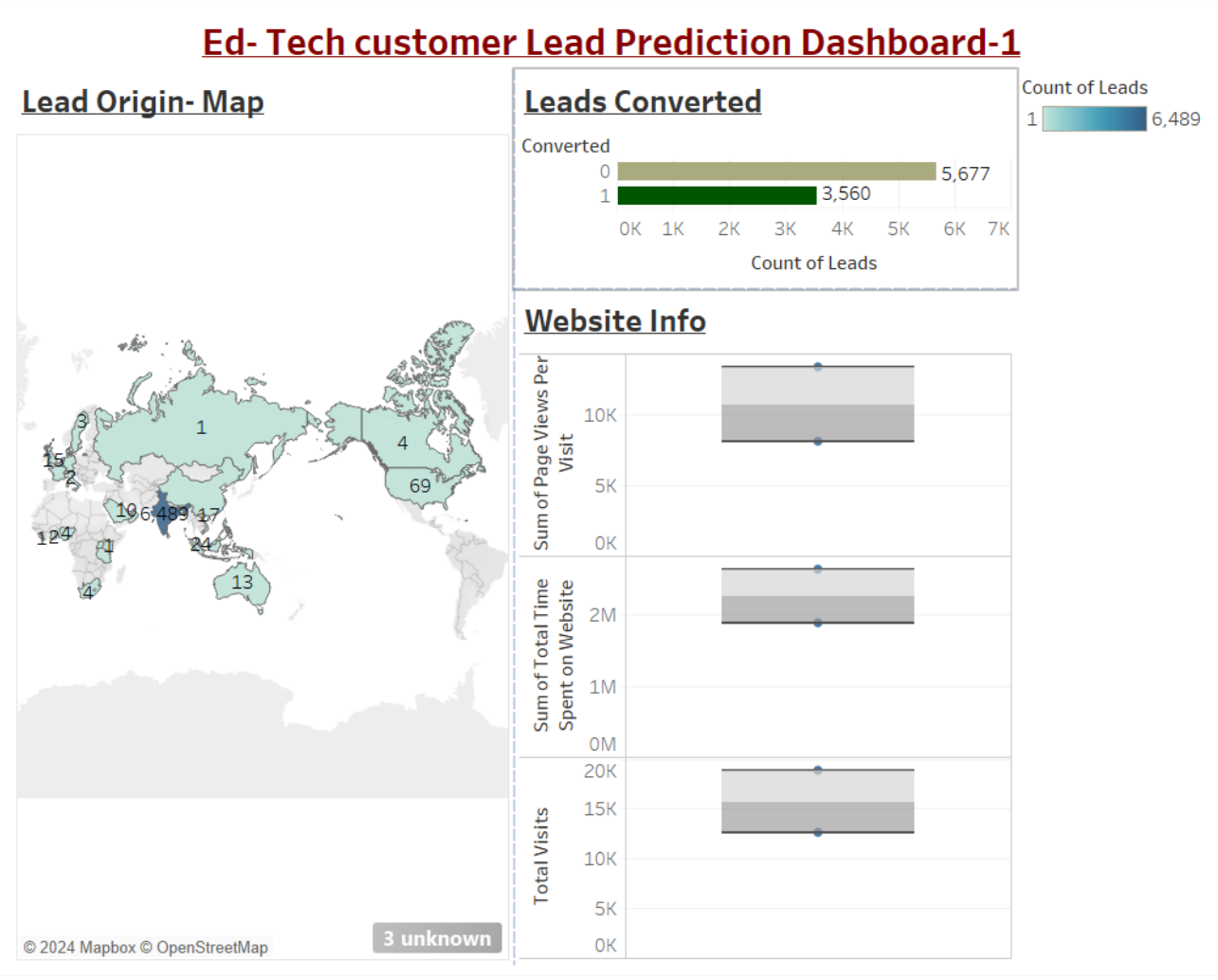
Most of the leads have opened their email as their last activity. **SMS Sent has a high conversion rate** for leads who noted it as their last activity.

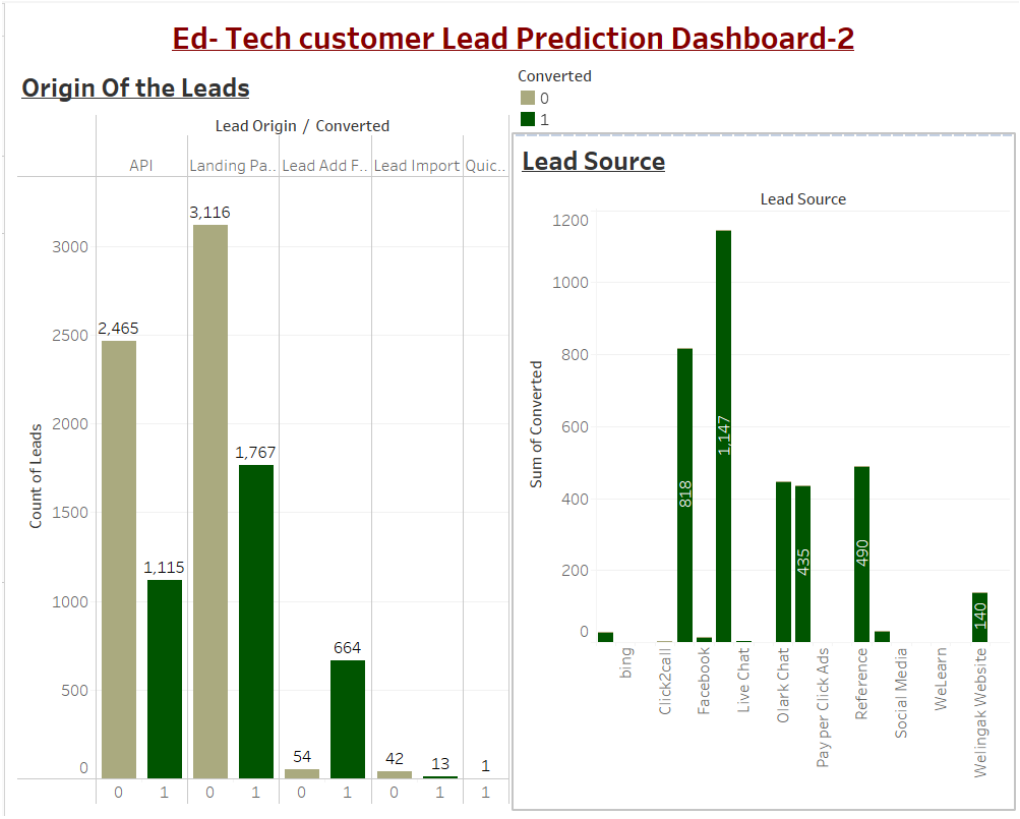
Summary

We have gained the following insights on the data from the analysis:

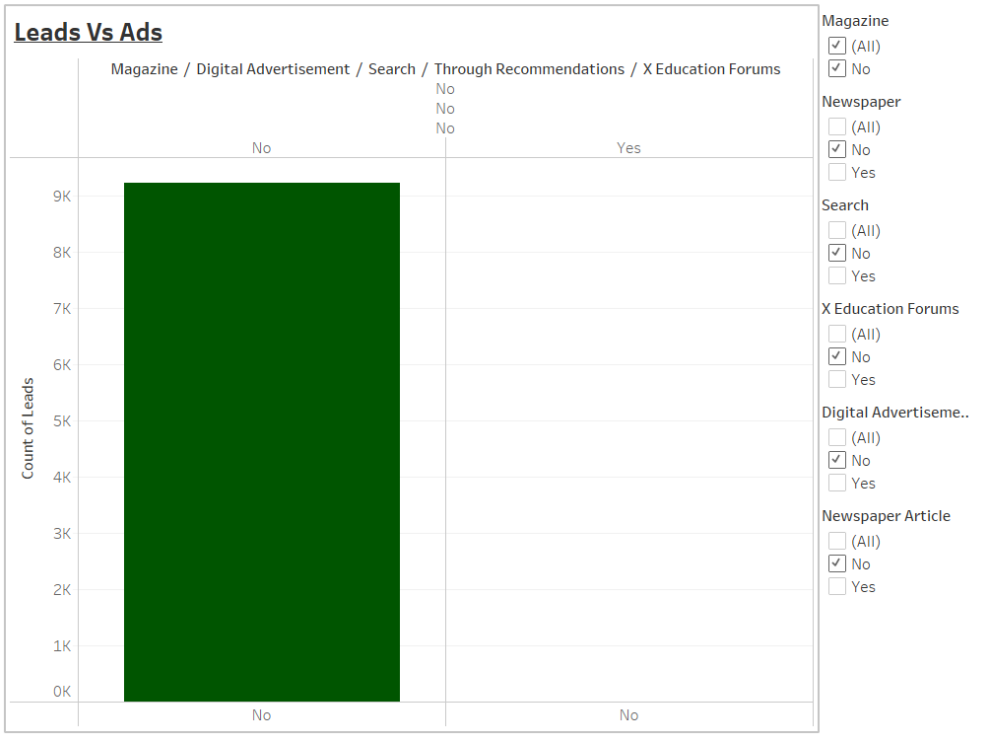
- The time spent on the website is a good indicator of successful completion.
- The “lead sources” had large differences in terms of number of leads and conversion rates.
- The last recorded activities provided important information for prediction.

With the above insights the following dashboards are created.

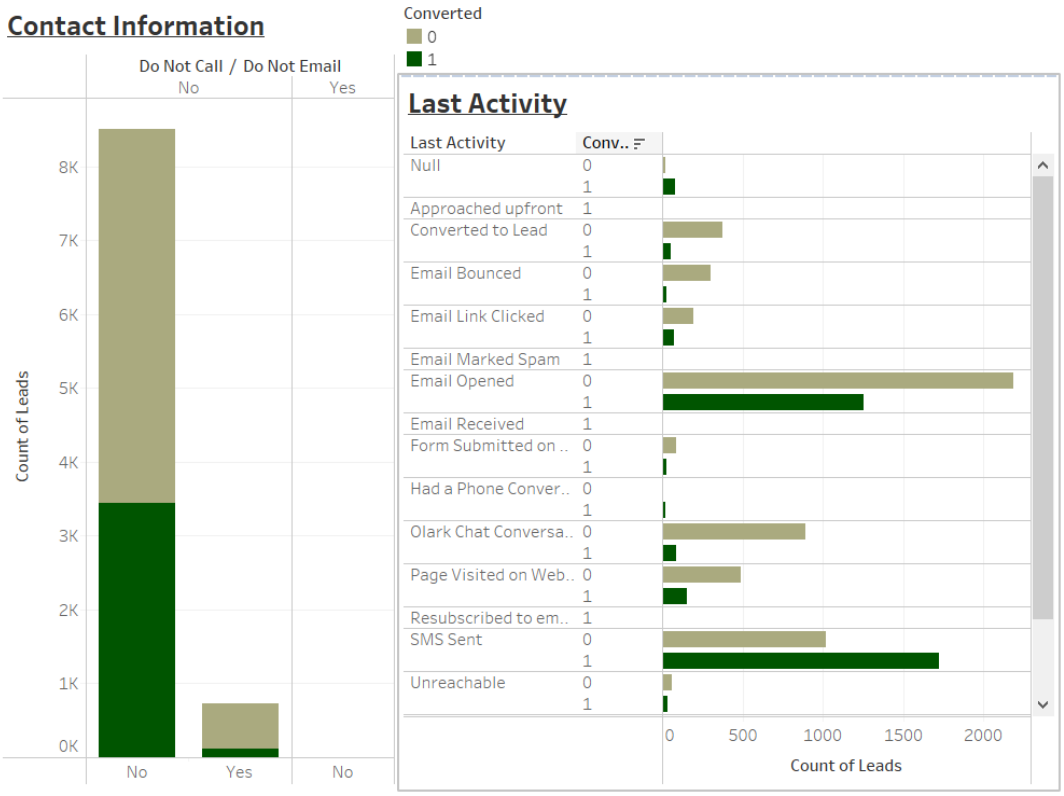




Ed- Tech customer Lead Prediction Dashboard-3



Ed- Tech customer Lead Prediction Dashboard-4



Ed- Tech Customer Lead Conversion Story

nd e	The median for Page views per visit is same for both converted and non converted.	Landing page submission brings higher number of leads.	Google plays vital role in converting the leads.	Magazine, Newspaper,Search, Education Forum, Digital Ads & News paper articles has no	People who don't want email has low conversion rate.	Most of the leads have opened their email but SMS sent has high conversion rate.