

Lead Scoring Case Study

Shivangi Saraswat

Chandra Babu S. N.

Sunil Tandel



Agenda

Reading And Understanding Data

Data Cleaning

Exploratory Data Analysis

Building Predictive Model

Assessing the Model

Conclusion







Reading And Understanding Data & Cleaning

First step to clean the dataset we chose was to drop the variables having unique values.

• Then, there were few columns with value 'Select' which means the leads did not choose any given option. We changed those values to Null values.

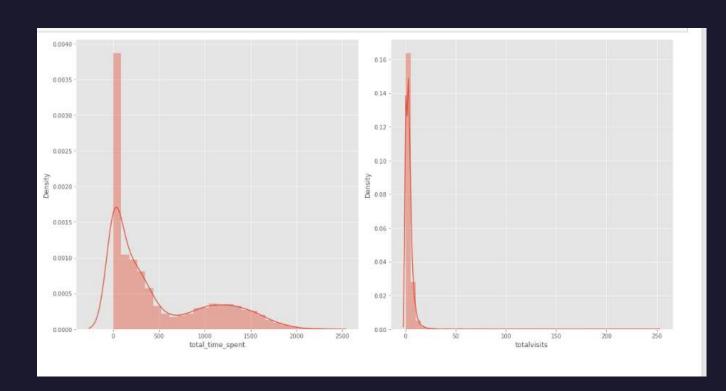
We dropped the columns having NULL values greater than
50%.
Changed NA values to Others for country names

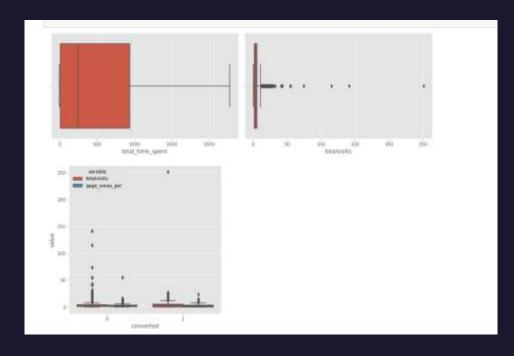


Exploratory Data Analysis

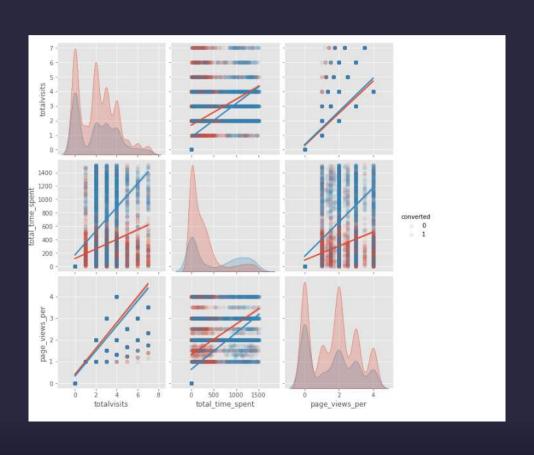
totalvisits total time spent page views per lead origin landing page submission lead origin lead import lead origin quick add form lead source google lead source organic search lead source others lead_source_reference lead source referral sites lead source wellingak website do not email yes do not call yes country others. country_unknown search yes newspaper article yes x education forums yes newspaper yes digital advertisement yes through recommendations yes lead quality low in relevance lead quality might be lead quality not sure lead quality worst -02 city_mumbai oby other cities of maharashtra city_other_metro_cities city thane & outskirts city tier ii cities city unknown asymmetrique activity index low asymmetrique activity index medium asymmetrique profile index low asymmetrique profile index medium a free copy yes last notable activity email bounced last_notable_activity_email_link_clicked last notable activity email marked spam last notable activity email opened last notable activity email received last notable activity form submitted on website last notable activity had a phone conversation last notable activity modified last notable activity olark chat conversation last notable activity page visited on website last notable activity resubscribed to emails last notable activity sms_sent last notable activity unreachable last notable activity unsubscribed last notable activity view in browser link clicked

Univariant Analysis

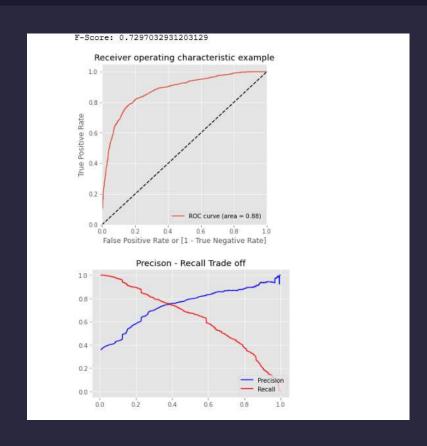




Bivariant Analysis



Measuring Model Performance



Conclusion:

67]:				
		lead_score	predicted_outcome	actual_outcome
	0	6	0	0
	1	58	1	1
	2	86	1	1
	3	12	0	0
	4	19	0	0

- The lead score calculated in the test set of data shows the conversion rate of 83% on the final predicted model which clearly meets the expectation of CEO has given a ballpark of the target lead conversion rate to be around 80%.
- Good value of sensitivity of our model will help to select the most promising leads.
- Do not focus on unemployed leads. They might not have a budget to spend on the course
- Do not focus on students, since they are already studying as it will be burden for them to pay Money
- Focus on Employees who are looking for upskilling and growth in Career

Thank You



