ANALYSIS OF TELEMARKETING DATA FOR PEACOCK SOLAR

By: SUKRITI MACKER (sukriti.macker98@gmail.com)

In the following project the objective is to analyze and evaluate the Telemarketing Data which was provided to us.

The major steps to analyze and predict solutions from the data:-

- 1. IMPORTING LIBRARIES AND DATA
- 2. MANIPULATING THE DATA
- 3. DATA VISUALIZATION
- 4. CONCLUSION

1. IMPORTING LIBRARIES AND UNDERSTANDING THE DATA

We will import the pandas library, matplolib library and numpy library as follows:-

```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
```

Now we will import the data file:-

data=pd.read_excel("C:/Users/Sukriti Macker/Peacock Solar Data
Analytics Intern/Telemarketing.xlsx")

	Timestamp	Lead Name	Lead Validation	Calling Date & Time	Picked the phone	Reason Why Phone is not picked	Email (Optional)	City	Lead Interested or Not	Insert Meeting Date & Time Slot	Please state the major highlights of the call.	Select the Reason
0	2019/11/30 11:59:05 AM GMT+5:30	Naveen Garg	Valid Contact	2019-11- 30 10:20:00	Yes	NaN	NaN	Kota	No	NaN	NaN	all read used solar
1	2019/11/23 7:33:23 PM GMT+5:30	Harish	Valid Contact	2019-11- 23 12:10:00	Yes	NaN	NaN	Kota	No	NaN	NaN	all ready solar used
2	2019/11/11 4:27:04 PM GMT+5:30	Shubh Hospital	Valid Contact	2019-11- 11 16:20:00	Yes	NaN	NaN	Jaipur	No	NaN	NaN	all ready used
3	2019/11/11 4:12:52 PM GMT+5:30	Rukmani Birla Hospital CK Birla	Valid Contact	2019-11- 11 15:24:00	Yes	NaN	NaN	Jaipur	No	NaN	NaN	all ready used solar

As we observed there are lot of other irrelevant features which would not contribute towards enhancing the prediction results.

Hence to get rid of them we use the following code:-

```
data.drop(labels=['Timestamp', "Reason Why Phone is not picked", 'Lead
Name', 'Lead Validation', 'Picked the phone', 'Email (Optional)', 'Insert
Meeting Date & Time Slot', 'Please state the major highlights of the
call.'], axis=1, inplace=True)
```

Using the .drop() function would help us drop columns which we find unwanted.

axis=1 will ensure that the column is dropped.

2. MANIPULATING THE DATA

The data will be edited to maintain uniformity and consistency. Also, we will fill up the NaN or None value using mean, median or mode.

Missing values or NaN values are present in the dataset when the customer or a layman is used to fill the form for collection of data. This leads to inconsistency in the data which is collected at last.

- Mean is used when the data is numeric, symmetrical and uniform and does not have any outliers.
- Median is used when data is numeric, left skewed or right skewed and consists of outliers.
- Mode is used when working with categorical data.

We use the following python command to fill NaN or None values, which are also called as missing values: -

```
data["City"]=data["City"].fillna(data["City"].mode()[0])
```

Note that the above code to fill NaN values in the column "City" uses mode, as the data present in the column is categorical.

When the data present in the column is numeric, has certain outliers and is slightly skewed we use median.

The skewness and outliers can be detected by the use of a histogram.

3. DATA VISUALIZATION

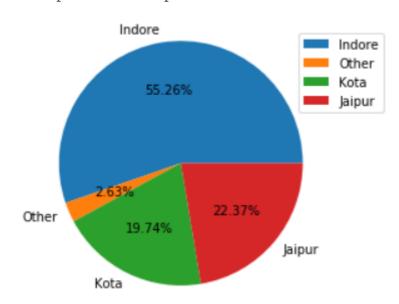
We must visualize the data by plotting graphs as to see the relationship between the top features that effect the target value or the output, which in our case is to analyze the sales of solar panels. Observing the data would lead us to create newer insights and would bring out an elaborate view of which section of customers are more drawn towards investing in Solar Panels.

Through this observation, the company can expand their customers in that specific section.

3.1. City

3.1.1. Cities in which customer showed interest to purchase solar panel

Through the following pie chart we analyze that in which city the customer showed interest to purchase solar panel.



ANALYSING THE LEADS WHICH SHOWED INTEREST

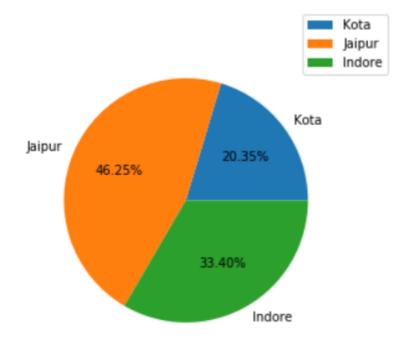
The maximum number of customers are from the city of Indore, with a share of 55.26%.

The second maximum number of customers are from the city of Jaipur, with a share of 22.37%.

The third maximum number of customers are from the city of Kota, with a share of 19.74%

3.1.2. Cities in which customer showed interest to purchase solar panel

Through the following pie chart we analyze that in which city the customer showed the least interest to purchase solar panel.



ANALYSING THE LEADS WHICH DID NOT SHOW INTEREST

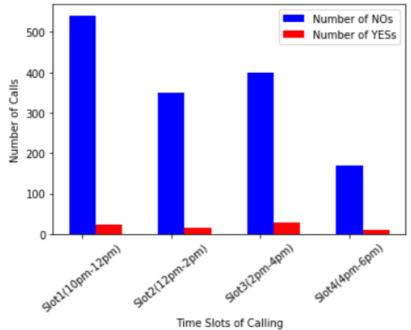
The maximum number of least interested customers are from the city of Jaipur, with a share of 46.25%

The second maximum number of least interested customers are from the city of Indore, with a share of 33.40%

The third maximum number of least interested customers are from the city of Kota, with a share of 20.35%

3.2. Customer Calling Time

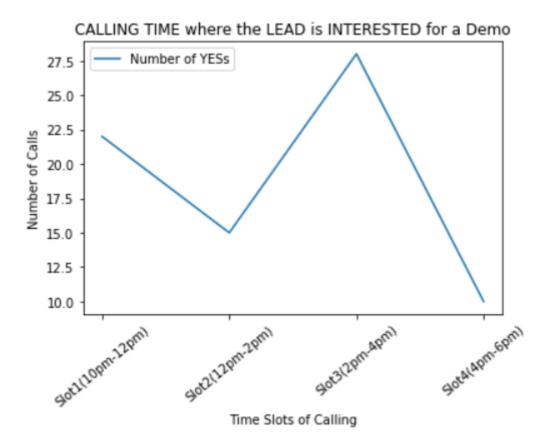
Comparison of whether the LEAD is interested or not based on CALLING TIME for a Demo



When the telemarketing department called the customer in the time slot 3 which is 2pm to 4pm, maximum number of interested customers appeared.

When the telemarketing department called the customer in the time slot 1 which is 10am to 12pm, second maximum number of interested customers appeared.

We can also observe the pattern using a line graph:-

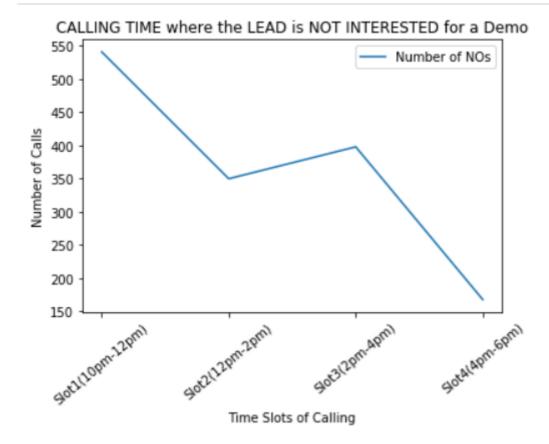


It is clearly established that the slot 3 is the most relevant as it yields the maximum number of customers who are interested in purchasing the solar panel.

ANALYSING THE LEADS WHICH WERE NOT INTRESTED [Blue Bins]

When the telemarketing department called the customer in the time slot 1 which is 10am to 12pm, maximum number of not interested customers appeared.

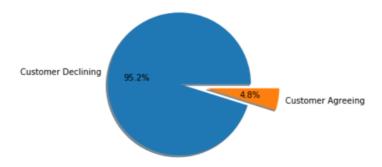
We can also observe the pattern using a line graph:-



It is clearly established that the slot 1 yields the maximum number of customers who are not interested in purchasing the solar panel.

3.3. Denial Reasons

Number of Customers Agreeing and Disagreeing for a Demo based on Reasons recorded by Telemarketing Team

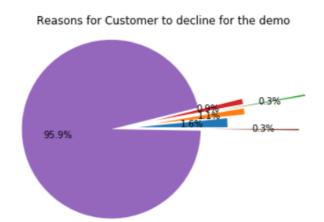


- As analyzed from the above pie chart that the majority of the customers are declining to take a demo by the sales team. The number of customers declining are 95.2%
- A very few portion of the customers actually agreed to take a demo arranged by the telemarketing team which is 4.8% of the customers.

This becomes extremely crucial for us to understand the reasons for denial.

The following pie chart is used to analyze the reasons for denial by the customers:-





Majority customer have not given the reason for the denial. The share of these customers are 95.9% which is approximately 96%

The rest 4% of the customers have given valid reasons for their denial of the demo.

- Out of the 4%
 - 1. 1.6% of the customers say that they already are using solar panels.
 - 2. 1.1% of the customers showed interest in the concept of solar panel but were not interested for the demo.
 - 3. 0.9% of the customers are not able to invest in solar panels as they reside in rented property

4. CONCLUSION

From the Telemarketing Data Analysis, it is observed that:

- 1. The telemarketing department should be aware as to why the customers are not interested in our company Peacock Solar for the installation of solar panels, so that we could understand customer behavior and increase the number of customers who are not interested.
- 2. The reason of denial must be clearly understood and filled by the telemarketing team so that there could be improvements made in the sales pitch for the next customer.

There were certain features that made us understand that customer calling time and city play a vital role in bringing up a lead.