**Data Science Project**

**Proposal**

**Name of Each Group member**

|  |  |
| --- | --- |
| **Name** | **Reg. No.** |
| Muhammad Waleed | L1F15BSCS0066 |

* **Data Source**

|  |  |
| --- | --- |
| **Data set Name** | **Online shoppers intention** |

* **Dataset Selection**

For this we want to provide statistic information based of facts and figures so that people who want to do online shopping based on weekdays and weekends.

* **Questions**

To do shopping or not probabiltity based on different statistics like discounts, end season sales, weekend/holiday discounts and many other.

* **Location Site of Data source**

[**https://www.kaggle.com/roshansharma/online-shoppers-intention**](https://www.kaggle.com/roshansharma/online-shoppers-intention)

* **Dataset information**

No. of Columns : 18

* **NA values or garbage values**

There are 14 missing and many umknown values in data.

* **Formatting Issues**

Yes we will move all numerical values column in one side.

* **Attributes Information**

.Administrative,Administrative\_Duration,Informational,Informational\_Duration,ProductRelated,ProductRelated\_Duration,BounceRates,ExitRates,PageValues,SpecialDay,Month,OperatingSystems,Browser,Region,TrafficType,VisitorType,Weekend,Revenue

* **Machine learning**

We will use decision treee classifier to make a prediction model that will predict wether to shop online or not.