**Karthik Senthil Sriranjan S**

**1NT18IS201 1NT18IS159**

**Data Set: Online Shoppers Purchasing Intention**

**Introduction:**

The Dataset has various features describing a session from a visitor on an online shopping website. There are 12,330 sessions(instances) and 18 attributes describing it. These attributes can be used to suggest how likely it is for a visitor to make a purchase.

**Attributes:**

1. **Administrative:** This attribute stores the number of Administrative type pages visited by the visitor in a session.
2. **Administrative Duration:** This attribute stores the amount of time that the visitor totally spends on administrative pages in a session.
3. **Informational:** This attribute stores the number of Informational type pages visited by the visitor in a session.
4. **Informational Duration:** This attribute stores the amount of time that the visitor totally spends on Informational pages in a session.
5. **Product Related:** This attribute stores the number of Product related pages visited by the visitor in a session.
6. **Product Related Duration:** This attribute stores the amount of time that the visitor totally spends on Product related pages in a session.
7. **Bounce Rate:** This stores the percentage of visitors who enter the site from a page and then leave ("bounce") without triggering any other requests during that session.
8. **Exit Rate:** This stores the percentage of the amount of exits / number of pageviews for a particular page. Where an exit is defined as the number of times visitors have left a site from a single page.
9. **Page Value:** The "Page Value" feature represents the average value for a web page that a user visited before completing an e-commerce transaction.
10. **Special Day:** This feature indicates how close a special day is compared to the date on which the visitor visited the site, as there are more chances of a purchase taking place around these days.
11. **Month:** This stores the month in which the visitor visited the website.
12. **Operating System:** Unique ID to identify the operating system of the system that the visitor is using.
13. **Browser:** Unique ID to identify the browser the visitor is using to visit the webpage.
14. **Region:** Unique ID to identify the region that the visitor is from.
15. **Traffic Type:** Unique ID to identify the traffic type(such as Direct, Organic, Referral, Social, Email, Display etc.)
16. **Visitor Type:** Takes 2 values, New Visitor or Returning, used to identify if the visitor is a regular user of the site or is a brand new customer.
17. **Weekend:** Boolean feature that identifies whether the visitor visited the site on a weekend or not.
18. **Revenue:** Boolean feature that identifies whether the visitor performed a transaction on the site i.e whether they purchased item(s). This attribute can be utilised as the class label.