**BDAD Project**

**Requirements Gathering**

**TEAM H:**

Akshay Saxena (BDA 01)

Divij Bhutani (BDA 02)

Chandan Kamal (BDA 03)

**Aim:** Understanding customer feedback through twitter data and knowledge graph to decide various strategies and changes a brand can make in their product.

**Platform:**

Jupyter Notebook

**For Tweet Extraction:**

Python

Twint

PySpark

**For Knowledge Graphs:**

Neo4j

**The advantage of Twint is that you don’t need Twitter’s API to make TWINT work -Twint utilizes Twitter’s search operators to let you:**

-scrape Tweets from specific users

-scrape Tweets relating to certain topics

-hashtags & trends

-or sort out sensitive information from Tweets like e-mail and phone numbers.

**Some of the benefits of using Twint:**

-Can fetch almost all Tweets (Twitter API limits to last 3200 Tweets only);

-Fast initial setup;

-Can be used anonymously and without Twitter sign up;

-No rate limitations.