Bahria University, Islamabad Campus
Department of Computer Sciences
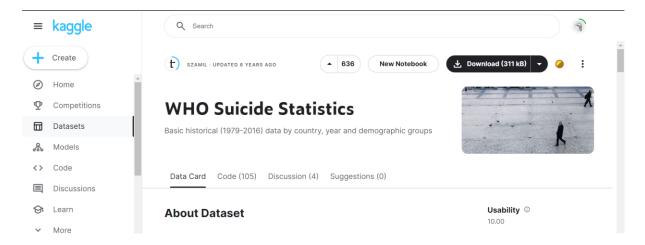### Assignment # 1
Class: BSIT (6th Semester)

Roll No: 01-135201-037   **Name: MARYAM ASHFAQ**

Q1: Choose the dataset from the Kaggle website. Explain how many rows (samples), columns (features), and different categories (classes) are there. Also, explain whether each column contains words or numbers (categorical or numerical). Lastly, count how many different options (unique values) are in each column.

1. **Dataset from the Kaggle: WHO SUICIDE STATISTICS**

2. **Number of rows(samples): 43777**

3. Number of columns(features): 6

I.   Country
II. Year
III. Sex
IV. Age
V.   Suicides number
VI. Population

	A	B	C	D	E	F
1	country	year	sex	age	suicides_no	population
2	Albania	1985	female	15-24 years		277900
3	Albania	1985	female	25-34 years		246800
4	Albania	1985	female	35-54 years		267500
5	Albania	1985	female	5-14 years		298300
6	Albania	1985	female	55-74 years		138700
7	Albania	1985	female	75+ years		34200
8	Albania	1985	male	15-24 years		301400
9	Albania	1985	male	25-34 years		264200
10	Albania	1985	male	35-54 years		296700
11	Albania	1985	male	5-14 years		325800
12	Albania	1985	male	55-74 years		132500
13	Albania	1985	male	75+ years		21100
14	Albania	1986	female	15-24 years		283900
15	Albania	1986	female	25-34 years		252100
16	Albania	1986	female	35-54 years		273200
17	Albania	1986	female	5-14 years		304700
18	Albania	1986	female	55-74 years		141700
19	Albania	1986	female	75+ years		34900

who_suicide_statistics

4. Different categories (classes): There are no yes/no classes.

5. Each column contains words or numbers (categorical or numerical):

I.   Country column contains names of countries(words).
II. Year column contains year numbers from 1979 to 2016(numbers).
III. Sex column contains 2 categories of male and female.
IV. Age column contains 6 numerical ranges from 5-14, 15-24, 25-34, 35-54, 55-74, 75+ years age groups.
V.   Suicides number column have numbers of people who committed suicide. With some blank rows.
VI. Population column have number of all living people in that country.

6. How many different options (unique values) are in each column.

I.   Country. 186 different country names.
II. Year. 37 different years.
III. Sex. 2 genders.
IV. Age. 6 age groups 5-14, 15-24, 25-34, 35-54, 55-74, 75+ years.
V.   Suicides number. This column has minimum value 0 and maximum value 22338. With null values too.
VI. Population. This column has minimum value 259 and maximum value 43805214. With null values too.