



## COMSATS University Islamabad, Lahore Campus

Course Title:	Introduction to Data Science			Course Code:	CSC461	Credit Hours:	3(3,0)
Resource Person:	Dr. Muhammad Sharjeel			Programme Name:	BSSE		
Semester:	5 <sup>th</sup>	Batch:	FA21	Section:	C	Max Marks:	10

**SP20-BSE-009**

**Isha Ikhlaq**

Use the dataset file “gender-prediction.csv” available on shared Google Drive folder for this assignment.

Q1: Provide responses to the following questions about the dataset.

**1. How many instances does the dataset contain?**

There are 100 instances in the dataset.

**2. How many input attributes does the dataset contain?**

There are 8 input attributes in the dataset.

**3. How many possible values does the output attribute have?**

The output attribute "gender" seems to have two possible values: male and female.

**4. How many input attributes are categorical?**

The categorical input attributes include "beard," "hair length," "scarf," and "eye color."

**5. What is the class ratio (male vs female) in the dataset?**

To calculate the class ratio, you would need to count the number of instances for each class. The ratio would be the number of instances for one class divided by the total number of instances.