

# Lab - 5

**Deadline:** Friday, 10 March 2023, 11:59 PM

## Assignment

*Sample Customer Master data frame:*

id	name	dob	validity_start	validity_end
1	Harsha	20-08-1990	01-01-1970	12-12-9999
2	Goldie	11-02-1990	01-01-1970	12-12-9999
3	Divya	25-12-1990	01-01-1970	12-12-9999

*Update to be performed:*

name	updated_dob
Harsha	05-09-1990

- 1) Write PySpark code to implement SCD Type II on the given sample customer master data frame.
- 2) Write SparkSQL code to implement SCD Type II on the given sample customer master data frame.