



# Session 19: Spark SQL

---

## Assignment 2

**ACADGILD**

## Session 19: Spark SQL

### Assignment 2

#### Table of Contents

1. Introduction .....	3
2. Objective .....	3
3. Prerequisites .....	3
4. Associated Data Files .....	3
5. Problem Statement .....	3
6. Approximate Time to Complete Task .....	<b>Error! Bookmark not defined.</b>

## 1. Introduction

In this assignment, you have to implement the concepts of the session.

## 2. Objective

This assignment will help you to understand the concepts learnt in the session.

## 3. Prerequisites

None

## 4. Associated Data Files

[Link](#)

## 5. Problem Statement

Using udfs on dataframe

1. Change firstname, lastname columns into  
Mr.first\_two\_letters\_of\_firstname<space>lastname

for example - michael, phelps becomes Mr.mi phelps

2. Add a new column called ranking using udfs on dataframe, where :  
gold medalist, with age  $\geq 32$  are ranked as pro  
gold medalists, with age  $\leq 31$  are ranked amateur  
silver medalist, with age  $\geq 32$  are ranked as expert  
silver medalists, with age  $\leq 31$  are ranked rookie

## 6. Expected output

Query with screenshots