

# Assignment

This week's assignment will be fun and interesting to work on as it is an open assignment, where you can apply all your creativity. There are no restrictions with the Datasets that are used for the assignment. You can feel free to choose a dataset of your choice and explore.

## Question 1.

Research more on Uber mode & create an in detail write-up for the same..

## Question 2.

Explore multi-node cluster by navigating to the resource manager in the Labs. Create an extensive document with proper screenshots of your observations.

## Question 3.

Select a relevant dataset of your choice and write some queries to demonstrate the following :

- Running total, Grouping aggregates and various Window functions like rank, dense\_rank, row\_num, lead, lag.
- Also try creating a pivot view.

## Question 4.

How to deal with nulls in Apache Spark? Give a detailed explanation with examples wherever necessary.

## Process to Submit the Assignment -

You need to create a Google Document consisting of answers to all the above 4 questions. Name the Google Document as **yourname\_week8\_assignment**

**Please upload your solution by filling the following form -**

<https://forms.gle/uCYtWKzvU1YroXhp9>

Top 5 answers will be selected and they will be compiled into a solution document and added to the Learning portal. Exceptional solutions will be chosen and you will receive a Trendytech T-shirt as a token of appreciation.

### **Why an open assignment this week?**

1. We would like to understand how good you are at researching and exploring various things as this is an important aspect when working on real life projects.
2. With this we would be able to analyse your understanding on when to apply what. Create a proper use case and then apply the relevant functions.

