

## Exercise 2 – Reading CSV File

---

In this exercise we'll read a csv file that contains data about sleeping habits. The data is taken from a clinical research.

Open the Sleep\_health.csv file and investigate it's structure. Note at the columns names and values.

Write a spark program that performs the following tasks –

1. Read the csv file into a Dataframe.
2. Count how many rows the dataframe contains.
3. Count how many male patients are there.
4. For each occupation category, count how many patients are there.
5. For each occupation category, show how many patients had a sleep disorder, and sort the results. Which is the most sleep disorder prone occupation?