SID: 450128

Homework 4

1-(40 pts) Write a MapReduce code to find the average age of all patients. Please take these notes:

•There are patients who have been repeated many times in the system. Remember to take care of them.

•Assume that each patient, has either a unique number representing the age or not a numeric value. It means that one patient, might have been repeated 100 times in the database. However, the age in each row is either not a number or a unique number.

•If there is no numeric age for a patient, then ignore the patient.

•There are patients who do not have any ID (caseid) and hence, we can’t catch them. Then, ignore them as well.

•To help you verify your result, the average is 61.73.

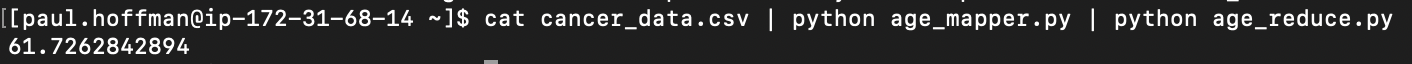
•Don’t worry if your result doesn’t match ours. We are mostly looking for your approach.

Text

Description automatically generated

Text

Description automatically generatedText

Description automatically generated

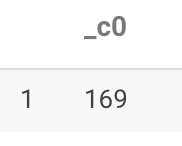
2- (60 pts) After finding the average age (61.73), then find how many unique younger, and older patients have died. By younger and older, we mean younger and older than the average age.

•To illustrate what we are looking for, this should be the result of your code:

o younger 169.0

o older 268.0

•Remember that patients have been repeated in the database. You should be careful about that.

 younger

A picture containing table

Description automatically generatedolder

Graphical user interface, text, application, email

Description automatically generated