Data Mining 1 (18-19)

Sep 24th, 2018

Paper #1: Data Quality and Integration Issues in Electronic Health Records
by: Ricardo João Cruz-Correia, Pedro Pereira Rodrigues, Alberto Freitas,
Filipa Canario Almeida, Rong Chen, and Altamiro Costa-Pereira

Question #1:

What are the main kinds of data errors discussed in this paper? Would it be
possible to automatically identify and categorize those errors?

Question #2:

What are missing values? What are the main methods used to handle missing
values?