1.read heading, and find about incorrect format and correct

2.find out missing data and fill it with median (for number) or mode (for string), or just drop those rows.