Pandas and friends

Data cleaning

Steps

1. import all needed libraries and data
2. check for shape, info, data types
3. check for duplicates
4. check the amount of null in columns and remove the ones with >75% null
5. check the info in each column and drop the ones which don't seem relevant
6. check whether the info in each case number column is the same and drop the useless ones
7. clean the column Sex, Age and Fatal making them easier to analyse

Observations

1. calculating variation or outliers was pointless since most infos were objects
2. not that easy to group by age