Most common problems found in the datasets:

* inaccurate data -

the data in some of the datasets doesn't fit with the rest and is out of place,

like in the taarif dataset where some of the bonuses are unrealistically high.

* wrong dtypes -

the dtypes in many columns don't fit the data that they hold, like in the date columns where the dtype is string and not datetime.

* ambiguous data -

The format of the data is not explicit, like in the vetek columns where the format could be either in years or in months.

* inconsistent data -

The format of the data is not clear and changes between records, i.e the gender columns have multiple formats for writing the gender.

* too much data -

in many of the datasets there are unused columns which do not provide any extra information

* missing data -

both missing complete rows and important data that should not be missing,

for example the basic taarif for drives, it doesn't make sense for it to be no taarif for drives.