1. Shape of data:-

Text, letter

Description automatically generated

1. Missing Values - Histogram
2. Number of missing values ranges between 0 - >5000

Chart, histogram

Description automatically generated

1. Missing values – percentage and discretization
2. Converted the missing values to percentage
3. It can be seen that there are almost 175 columns which are having missing values more than 60%. Better to drop this 175 columns
4. Remaining columns we can go for imputation

Chart, histogram

Description automatically generated

1. Created a data dictionary with the question ids and their corresponding questions. Then pointed Important Questions regarding loneliness:-

'ch22o002': 'preloaded variable: age'

'ch22o003': 'preloaded variable: paid job or not'

'ch22o004': 'How would you describe your health, generally speaking?',

'ch22o005': 'Is your health poorer or better, compared to last year?',

'ch22o006': 'How would you rate your chance of living to be 75 years old or older?',

'ch22o007': 'How would you rate your chance of living to be 80 years old or older?',

'ch22o008': 'How would you rate your chance of living to be 85 years old or older?',

'ch22o009': 'How would you rate your chance of living to be 90 years old or older?',

'ch22o010': 'How would you rate your chance of living to be 95 years old or older?',

'ch22o011': 'I felt very anxious',

'ch22o012': 'I felt so down that nothing could cheer me up',

'ch22o013': 'I felt calm and peaceful',

'ch22o014': 'I felt depressed and gloomy',

'ch22o015': 'I felt happy'

'ch22o018': 'Do you suffer from a long-standing disease, affliction or handicap, or do you suffer from the consequences of an accident?',

'ch22o020': 'To what extent did your physical health or emotional problems hinder your daily activities over the past month?',

'ch22o021': 'To what extent did your physical health or emotional problems hinder your social activities over the past month?',

'ch22o022': 'To what extent did your physical health or emotional problems hinder your work over the past month, for instance in your job, the housekeeping, doing volunteer work, or in school?',

'ch22o044': 'performing housekeeping work or maintaining the garden',

'ch22o045': 'taking care of financial affairs, such as paying bills and keeping track of expenditure',

'ch22o076': 'fatigue',

'ch22o077': 'sleeping problems',

'ch22o081': 'a heart attack including infarction or coronary thrombosis or another heart problem including heart failure',

'ch22o082': 'high blood pressure or hypertension',

'ch22o083': 'high cholesterol content in blood',

'ch22o084': 'a stroke or brain infarction or a disease affecting the blood vessels',

'ch22o095': 'Alzheimer, dementia, organic brain syndrome, senility',

'ch22o269': 'COVID-19 (new corona virus)',

'ch22o099': 'How many days during the last month were you unable to go to work, perform housekeeping work or attend school, due to disease?',

'ch22o100': 'At this moment, do you go to work as normal, or do you not or only partly go to work on account of your health?',

'ch22o268': 'For how long have you not been working: number of years',

'ch22o101': 'For how long have you not been working: number of months',

'ch22o102': 'For how long have you not been working: number of weeks',

'ch22o103': 'For how long have you not been working: number of days',

'ch22o104': 'Can you indicate what kind of health problems or what kind of affliction you are suffering from?',

'ch22o126': 'Do you smoke now?',

'ch22o130': 'How many cigarettes (including rolling tobacco) [did/do] you smoke on average per day?'

'ch22o133': 'How often did you have a drink containing alcohol over the last 12 months?'

'cs22o001': 'How satisfied are you with the amount of leisure time that you have?'

'cs22o003': 'a sports club, no connection'

'cs22o008': 'a cultural association or hobby club, no connection'

'cs22o542': 'What kind of illness, handicap or condition [does/did] this person have? Long-term physical disability'

'cs22o121': 'What sport do you practice? - fitness',

'cs22o124': 'How many days do you watch television per week, on average? If you do not watch television, enter a zero (0).'

'cs22o242': 'computer or laptop use, average number of hours per week: at home',

'cs22o280': 'average number of hours per week spent on: chatting, video calling or sending messages (e.g. WhatsApp, Telegram, Signal, Snapchat, Skype, Facebook messenger, etc.)',

'cs22o443': 'average number of hours per week spent on: visiting dating websites (e.g. Relatieplanet, Lexa, Tinder, Grindr or similar services)',

'cs22o436': 'How often did you make use of social media in the past 2 months?'

'cs22o480': 'In general, do you think about what information you share on social media and about who can see this?',

'cs22o602': 'You have indicated previously that you do not use social media. Can you indicate why not? I do not see the point of it',

'cs22o283': 'How satisfied are you with your social contacts?',

'cs22o284': 'I have a sense of emptiness around me',

'cs22o285': 'There are enough people I can count on in case of a misfortune',

'cs22o286': 'I know a lot of people that I can fully rely on',

'cs22o287': 'There are enough people to whom I feel closely connected',

'cs22o288': 'I miss having people around me',

'cs22o289': 'I often feel deserted',

'cs22o290': 'How often do you do the following? Spend an evening with family (other than members of your own household)',

'cs22o291': 'How often do you do the following? Spend an evening with someone from your neighborhood',

'cs22o292': 'How often do you the following things? Spend an evening with someone outside your neighborhood',

'cs22o300': 'Which of these people are very dear to you? person 1: