

Bibliothèque :

I. Dictionary of Question

Nous avons créé un Dictionnaire afin de regrouper l'ensemble des variables issues du questionnaire initial, accompagnées de leurs descriptions. Il nous permet de mieux comprendre la signification de chaque colonne et de guider notre nettoyage des données. C'est une étape essentielle pour assurer la cohérence et la lisibilité de notre analyse.

Les lettres de C à R correspondent à la numérotation des questions du questionnaire.



C. Analyses des soins de santé

Question: the past 12 months, not counting times you went to an emergency room, how many times did you see a doctor, nurse, or other health professional to get care for yourself?



Légende(FreqGoProvider) :

- -1 = valeur manquante
- 0 = None
- 1 = 1 time = moins fréquent
- 2 = 2-4 time = assez fréquent
- 4 = 5-9 time = plutôt fréquent
- 5 = 10 or more time

Question : Overall, how would you rate the quality of health care you received in the past 12 months?
Légende(qualitycare) :



| hints_code  | | our_code  |
|--|------------------------|--|
| -1 | Missing value | -1 |
| 5 | Poor = mauvaise | 0 |
| 4 | Fair = passable | 1 |
| 3 | Good = Bonne | 2 |
| 2 | Very good = très bonne | 3 |
| 1 | Excellent | 4 |

Question: How much do you trust the health care system (for example, hospitals, pharmacies, and other organizations involved in health care)?
Légende (TrustHCSys):

| hints_code  | | our_code  |
|--|---------------|--|
| -1 | Missing value | -1 |
| 4 | Not at all | 0 |
| 3 | A little | 1 |
| 2 | Some | 2 |
| 1 | A lot | 3 |



G.Test génétiques : Variables

Question: Which of the following types of genetic tests have you had? Ancestry testing to understand where you and your relatives come from ?**Légende(HadTest3_Ancestry2) :**

| hints_code  | | our_code  |
|--|---------------|--|
| -1 | Missing value | -1 → NAN |
| 2 | No | 0 |
| 1 | Yes | 1 |



Question : Which of the following types of genetic tests have you had? Personal trait testing to understand whether you have genes that are linked to certain characteristics - like enjoying cilantro?

Légende(HadTest3_PersonalTrait) :



| hints_code  | | our_code  |
|--|---------------|--|
| -1 | Missing value | -1 → NAN |
| 2 | No | 0 |
| 1 | Yes | 1 |

Question : Which of the following types of genetic tests have you had? Testing for specific diseases to understand your risk of getting certain diseases such as breast cancer, colon cancer, cardiovascular (heart) disease, diabetes, or dementia/Alzheimers?

Légende(HadTest3_SpecificDisease) :



| hints_code  | | our_code  |
|--|---------------|--|
| -1 | Missing value | -1 → NAN |
| 2 | No | 0 |
| 1 | Yes | 1 |

Question : Which of the following types of genetic tests have you had? Prenatal genetic carrier testing to determine the risk that a man and a woman will have a baby with certain diseases such as cystic fibrosis or Tay Sachs?**Légende(HadTest3_Prenatal)**



| hints_code  | | our_code  |
|--|---------------|--|
| -1 | Missing value | -1 → NAN |
| 2 | No | 0 |
| 1 | Yes | 1 |

Question : Which of the following types of genetic tests have you had? Other - Specify?

Légende(HadTest3_Other) :

| hints_code  | | our_code  |
|--|---------------|--|
| -1 | Missing value | -1 → NAN |
| 2 | No | 0 |
| 1 | Yes | 1 |

Question : Which of the following types of genetic tests have you had? Not sure what type of genetic test I've had?**Légende(HadTest3_NotSure)**

| hints_code  | | our_code  |
|--|---------------|--|
| -1 | Missing value | -1 → NAN |
| 2 | No | 0 |
| 1 | Yes | 1 |

Question : Derived variable to categorize responses given in G1 (Had genetic testing)?



Légende(HadTest3_Cat) :

| Original in df_original | Modification in df |
|--|-----------------------------------|
| -1 = valeur manquante | -1 : missing value |
| 8 = Not sure only | 2 = other teste génétique |
| 10 = Multiple genetic test selected | 2 = other teste génétique |
| 15 = Ancestry testing only | 2 = other teste génétique |
| 16 = Personal trait testing only | 2 = other teste génétique |
| 17 = specific disease testing only | 1 = specific disease testing only |
| 18 = Prenatal genetic carrier testing only | 2 = other teste génétique |
| 19 = I have not had any genetic test only | 2 = other teste génétique |
| 91 = Other genetic testing only | 2 = other teste génétique |

Tout les valeurs autre que 17 sont égale à deux , elle correspond au teste génétique autre que des teste génétique pour détecter certaines types de maladie.



Question : What were the reasons you had genetic testing? - Doctors recommendation.

Légende(ReasonTest_DocRec) :

| hints_code  | | our_code  |
|--|---------------|--|
| -1 | Missing value | -1 → NAN |
| 2 | No | 0 |
| 1 | Yes | 1 |



Question : What were the reasons you had genetic testing? - Understand my family ancestry.

Légende(ReasonTest_UnderstandFam)

| hints_code  | | our_code  |
|--|---------------|--|
| -1 | Missing value | -1 → NAN |
| 2 | No | 0 |
| 1 | Yes | 1 |



Question : What were the reasons you had genetic testing? - Learn more about personal traits that may be influenced by genetics ?

Légende(ReasonTest_PersTraits)

| hints_code  | | our_code  |
|--|---------------|--|
| -1 | Missing value | -1 → NAN |
| 2 | No | 0 |
| 1 | Yes | 1 |

Question : What were the reasons you had genetic testing? - Learn more about my risk for certain diseases (for example, cancer or heart disease).

Légende(ReasonTest_DiseaseRisk)

| hints_code  | | our_code  |
|--|---------------|--|
| -1 | Missing value | -1 → NAN |
| 2 | No | 0 |
| 1 | Yes | 1 |

Question : What were the reasons you had genetic testing? - Prenatal testing (for example, carrier testing).

Légende(ReasonTest_Prenatal)

| hints_code  | | our_code  |
|--|--|--|
|--|--|--|

| | | |
|----|---------------|----------|
| -1 | Missing value | -1 → NAN |
| 2 | No | 0 |
| 1 | Yes | 1 |

Question : What were the reasons you had genetic testing?**Légende(ReasonTest_Cat) :**



| Original in df_original | Modification in df |
|---|----------------------------------|
| -1 = <i>valeur manquante</i> | -1 : missing value |
| 1 = Doctor's recommendation only | 1 = Doctor's recommendation only |
| 2 = Understand my family ancestry only | 2 = other reason |
| 3 = Find relatives only | 2 = other reason |
| 4 = Learn more about personal traits only | 2 = other reason |
| 5 = Learn more about disease risk only | 2 = other reason |
| 6 = Understand what might be best for me only | 2 = other reason |
| 7 = Prenatal testing only | 2 = other reason |
| 8 = Received test as a gift only | 2 = other reason |
| 9 = Multiple genetic testing reasons selected | 2 = other reason |
| 91 = Other - Specify only | 2 = other reason |

Toutes les valeurs autre que 1 sont égale à deux , elle correspond au teste génétique autre que des teste génétique pour détecter certaines types de maladie.

Donc:1 = Doctor's recommendation only et 2 = other reason



I.Analyse de la perception de la Santé Global

Question : In general, would you say your health is...? → quel est votre état général ?**Légende:** ('GeneralHealth')

| hints_code  | | our_code  |
|--|------------------------|--|
| -1 | Missing value | -1 |
| 5 | Poor = mauvaise | 0 |
| 4 | Fair = passable | 1 |
| 3 | Good = Bonne | 2 |
| 2 | Very good = très bonne | 3 |
| 1 | Excellent | 4 |

Question :Overall, how confident are you about your ability to take good care of your health?

Légende:(‘OwnAbilityTakeCareHealth’)

| hints_code  | | our_code  |
|--|------------------------------------|--|
| -1 | Missing value | -1 |
| 5 | No confident at all = confiant(e) | 0 |
| 4 | Fair = peu confiant(e) | 1 |
| 3 | Good = Assez confiant(e) | 2 |
| 2 | Very good = très confiant(e) | 3 |
| 1 | Confident = totalement confiant(e) | 4 |

Question: Do any of the following significantly limit your daily activities: deafness or serious difficulty hearing?**Légende:**(‘HealthLimits_Deaf’)

| Original in df_original | Modification in df |
|-------------------------|--------------------|
| -1 : missing value | -1 : missing value |
| 1 : Yes | 1 : Yes |
| 2 : No | 0 : No |

Question: Do any of the following significantly limit your daily activities: visual impairments or blindness?**Légende:**(‘HealthLimits_Blind’)

| Original in df_original | Modification in df |
|-------------------------|--------------------|
| -1 : missing value | -1 : missing value |
| 1 : Yes | 1 : Yes |
| 2 : No | 0 : No |

Question: Do any of the following significantly limit your daily activities: permanent or long term physical or mobility limitations?**Légende:**(‘HealthLimits_Mobility’)

| Original in df_original | Modification in df |
|-------------------------|--------------------|
| -1 : missing value | -1 : missing value |
| 1 : Yes | 1 : Yes |
| 2 : No | 0 : No |

Question: Do any of the following significantly limit your daily activities: chronic pain?**Légende:**(‘HealthLimits_Pain’)

| Original in df_original | Modification in df |
|-------------------------|--------------------|
| -1 : missing value | -1 : missing value |
| 1 : Yes | 1 : Yes |
| 2 : No | 0 : No |

Question: Has a doctor or other health professional ever told you that you had diabetes or high blood sugar?**Légende:**(‘MedConditions_Diabetes’)

| Original in df_original | Modification in df |
|-------------------------|--------------------|
| -1 : missing value | -1 : missing value |
| 1 : Yes | 1 : Yes |
| 2 : No | 0 : No |

Question: Has a doctor or other health professional ever told you that you had high blood pressure or hypertension?**Légende:**(‘MedConditions_HighBP’)

| Original in df_original | Modification in df |
|-------------------------|--------------------|
| -1 : missing value | -1 : missing value |
| 1 : Yes | 1 : Yes |

| | |
|--------|--------|
| 2 : No | 0 : No |
|--------|--------|

Question: Has a doctor or other health professional ever told you that you had a heart condition such as heart attack, angina, or congestive heart failure?**Légende:**
(**'MedConditions_HeartCondition'**)

| Original in df_original | Modification in df |
|-------------------------|--------------------|
| -1 : missing value | -1 : missing value |
| 1 : Yes | 1 : Yes |
| 2 : No | 0 : No |


Question: Has a doctor or other health professional ever told you that you had chronic lung disease, asthma, emphysema, or chronic bronchitis?**Légende:**(**'MedConditions_LungDisease'**)

| Original in df_original | Modification in df |
|-------------------------|--------------------|
| -1 : missing value | -1 : missing value |
| 1 : Yes | 1 : Yes |
| 2 : No | 0 : No |


Question: Has a doctor or other health professional ever told you that you had depression or anxiety disorder?**Légende:**(**'MedConditions_Depression'**)

| Original in df_original | Modification in df |
|-------------------------|--------------------|
| -1 : missing value | -1 : missing value |
| 1 : Yes | 1 : Yes |
| 2 : No | 0 : No |

- Create a new column 'Health_limits' that sums HealthLimits_col for positive values only, not -1 and returns -1 if all values are negative

| Original in df_original | Creation in df |
|---|--|
| No existing  | Health_limits |
| All individual columns only | Sum only positive Health_limits_col = ['HealthLimits_Deaf', 'HealthLimits_Blind', 'HealthLimits_Mobility', 'HealthLimits_Pain'] |
| | -1 : missing value |
| | 0 : None |
| | 1 : 1 Health_limits |
| | 2 : 2 Health_limits |
| | 3 : 3 Health_limits |
| | 4 : 4 Health_limits (i.e : Deaf, Blind, physical or mobility limitations and chronic |

- Create a new column 'Bad_Med_cond' that sums Med_conditions_col for positive values only, not -1 and returns -1 if all values are negative

| Original in df_original | Creation in df |
|--|---|
| No  | Bad_Med_cond |
| All individual | Sum only positive of ['MedConditions_Diabetes', 'MedConditions_HighBP', 'MedConditions_HeartCondition', 'MedConditions_LungDisease', |
| | -1 : missing value |
| | 0 : None |

| | |
|--|--|
| | 1 : 1 Bad_Med_cond (1 of the list) |
| | 2 : 2 Bad_Med_cond |
| | 3 : 3 Bad_Med_con |
| | 4 : 4 Bad_Med_cond |
| | 5 : 5 Bad_Med_cond (i.e diabetes, HighBP, Heart failure, lung disease, Depression) |

Question: About how tall are you without shoes?**Légende:('Height_Feet:')**

- Convert height_feet and height_inches to the new column ['Height_m'] : Height in meter

| Original in df_original | df | df2 |
|-------------------------|--|---------------------------|
| Height_feet | Height_feet | ✗ deleted |
| Height_inches | Height_inches | ✗ |
| ✗ | Height_m = conversion (height_feet+_inches) → height in meters | Height_m |
| | -1 : missing value | -1 : missing value |
| | float64: height in meters | float64: height in meters |

- Convert weight in pounds in kg into the new column ['Weight_kg']

| Original in df_original | df | df2 |
|-------------------------|------------------------|------------------------|
| Weight | Weight | ✗ deleted |
| ✗ | Weight_kg | Weight_kg |
| | -1 : missing value | -1 : missing value |
| | float64 : weight in kg | float64 : weight in kg |

- create new column ['BMI'] : Body mass index = IMC
 $BMI = \text{weight_kg} / (\text{height_m} ** 2)$

| Original in df_original | df |
|-------------------------|------------------------|
| | BMI = IMC |
| ✗ | float64 |
| | -1 : missing value |
| | float64 : weight in kg |

- Drop heigh_feet, height_inches, Weight in pounds (we keep in meters & kg) → at the end_Dorothee_to_do_in_df2

Question: During the past 30 days, how many hours of sleep did you get in a 24 hour period on a weekday?**Légende:('SleepWeekdayHr')**

| | |
|-----------|--------------------------------------|
| -1 | missing value |
| int64 | sleeping hours / day during week day |

Question: During the past 30 days, how many hours of sleep did you get in a 24 hour period on a WEEKEND DAY?**Légende:('SleepWeekendHr2')**

| | |
|-----------|--|
| -1 | missing value |
| int64 | sleeping hours / day during weekend day (saturday, sunday) |

- Creation new column **'SleepWeek'** : Sleeping hours / week

'SleepWeek' = 'SleepWeekdayHr' + 'SleepWeekendHr2'

| Original in df_original | | df |
|-------------------------|--|------------------|
| × | | SleepWeek |
| -1 | missing value | |
| int64 | sleeping hours / day during the all week (from monday to sunday) | |

Question: Over the past 2 weeks, how often have you been bothered by: Little interest or pleasure in doing things?**Légende:('LittleInterest')**

| hints_code × | | our_code ✓ |
|----------------------------|-------------------------|--------------------------|
| -1 | Missing value | -1 |
| 4 | Not at all | 0 |
| 3 | Several days | 1 |
| 2 | More than half the days | 2 |
| 1 | Nearly every day | 3 |

Question: Over the past 2 weeks, how often have you been bothered by: Feeling down, depressed or hopeless?**Légende:('Hopeless')**

| hints_code × | | our_code ✓ |
|----------------------------|-------------------------|--------------------------|
| -1 | Missing value | -1 |
| 4 | Not at all | 0 |
| 3 | Several days | 1 |
| 2 | More than half the days | 2 |
| 1 | Nearly every day | 3 |

Question: Over the past 2 weeks, how often have you been bothered by: Feeling nervous, anxious or on edge?**Légende:('Nervous')**

| hints_code × | | our_code ✓ |
|----------------------------|-------------------------|--------------------------|
| -1 | Missing value | -1 |
| 4 | Not at all | 0 |
| 3 | Several days | 1 |
| 2 | More than half the days | 2 |
| 1 | Nearly every day | 3 |

Question:Over the past 2 weeks, how often have you been bothered by: Not being able to stop or control worrying?**Légende:('Worrying')**

| hints_code × | | our_code ✓ |
|----------------------------|--|--------------------------|
|----------------------------|--|--------------------------|

| | | |
|----|-------------------------|----|
| -1 | Missing value | -1 |
| 4 | Not at all | 0 |
| 3 | Several days | 1 |
| 2 | More than half the days | 2 |
| 1 | Nearly every day | 3 |

Question: Please respond by marking one box per row: I feel left out... **Légende:** ('FeelLeftOut')

→ Feel Left Out = se sentir exclu

| hints_code ❌ | | our_code ✅ |
|--------------|---------------|------------|
| -1 | Missing value | -1 |
| 1 | Never | 0 |
| 2 | Rarely | 1 |
| 3 | Sometimes | 2 |
| 4 | Usually | 3 |
| 5 | Always | 4 |

Question: Please respond by marking one box per row: I feel that people barely know me...

Légende: ('FeelPeopleBarelyKnow')

→ People barely know me = sentiment d'être incompris

| • hints_code ❌ | • | • our_code ✅ |
|----------------|-----------------|--------------|
| • -1 | • Missing value | • -1 |
| • 1 | • Never | • 0 |
| • 2 | • Rarely | • 1 |
| • 3 | • Sometimes | • 2 |
| • 4 | • Usually | • 3 |
| • 5 | • Always | • 4 |

Question: Please respond by marking one box per row: I feel isolated from others... **Légende:** ('FeelIsolated')

→ Feel isolated = sentiment d'être isolé

| hints_code ❌ | | our_code ✅ |
|--------------|---------------|------------|
| -1 | Missing value | -1 |
| 1 | Never | 0 |
| 2 | Rarely | 1 |
| 3 | Sometimes | 2 |
| 4 | Usually | 3 |
| 5 | Always | 4 |

Question : Please respond by marking one box per row: I feel that people are around me but not with me...Légende:('FeelPeopleNotWithMe')

→ sentiment de déconnection social, même en leur présence, comme s'ils étaient physiquement là mais émotionnellement absents. Sentiment de déconnection / de dissociation léger sans perte à la réalité

| hints_code  | | our_code  |
|--|---------------|--|
| -1 | Missing value | -1 |
| 1 | Never | 0 |
| 2 | Rarely | 1 |
| 3 | Sometimes | 2 |
| 4 | Usually | 3 |
| 5 | Always | 4 |

Question: In a typical week, how many days do you do any physical activity or exercise of at least moderate intensity, such as brisk walking, bicycling at a regular pace, and swimming at a regular pace?Légende:('TimesModerateExercise')

| | |
|----|-----------------|
| -1 | Missing value |
| 0 | None |
| 1 | 1 day per week |
| 2 | 2 days per week |
| 3 | 3 days per week |
| 4 | 4 days per week |
| 5 | 5 days per week |
| 6 | 6 days per week |
| 7 | 7 days per week |

Question:On the days that you do any physical activity or exercise of at least moderate intensity, how long do you typically do these activities?Légende:('HowLongModerateExerciseMinutes')

| | |
|-------|---------------------------|
| -1 | missing value |
| int64 | minutes of exercise / day |

Question: In a typical week, outside of your job or work around the house, how many days do you do leisure-time physical activities specifically designed to strengthen your muscles such as lifting weights or circuit training?Légende:('TimesStrengthTraining')

| | |
|----|-----------------|
| -1 | Missing value |
| 0 | None |
| 1 | 1 day per week |
| 2 | 2 days per week |
| 3 | 3 days per week |
| 4 | 4 days per week |
| 5 | 5 days per week |
| 6 | 6 days per week |
| 7 | 7 days per week |

- Creation of a new column : 'Moderate_exercise_min_week' : Moderate exercise in minutes per week

'Moderate_exercise_min_week' = TimesModerateExercise * HowLongModerateExerciseMinutes

| Original in df_original | df | |
|----------------------------|-----------|--|
| × | SleepWeek | |
| | -1 | missing value |
| | int64 | sleeping hours / day during the all week (from monday to sunday) |

J.Exploration des données environnementale et indicateur de Santé

Question:How much do you think climate change will harm your health?**Legend:**
(' ClimateChgHarmHealth')

Ce qu'il faut comprendre ici c'est l'importance qu'accorde les personnes au changement climatique sur leur santé

| hints_code × | | our_code ✓ |
|---------------------|--------------------|-------------------|
| -1 | Missing value | -1 |
| 5 | Don't know | -1 |
| 4 | Not at all | 0 |
| 3 | A little = un peu | 1 |
| 2 | Some = moyennement | 2 |
| 1 | A lot = beaucoup | 3 |

Cette variable permet d'étudier les variations sociodémographiques (âge région, statut socio-économique) dans la perception du risque climatique, à croiser avec le stress et la "santé mentale"

Question: Thinking about your neighborhood, over the past 12 months, how much was it affected by extreme weather events (such as severe storms, droughts, floods, heat waves, cold snaps, etc.)?

Legend : ('ExtremeWeatherEvents')

| hints_code × | ExtremeWeatherEvents | our_code ✓ |
|---------------------|----------------------|-------------------|
| -1 | Missing value | -1 |
| 4 | Not at all | 0 |
| 3 | A little = un peu | 1 |
| 2 | Some = moyennement | 2 |
| 1 | A lot = beaucoup | 3 |

Peut être croisée avec l'état de santé.

Question:How much do you worry that outdoor air pollution will harm your health? Dans quelle mesure craignez-vous que la pollution de l'air extérieur nuise à votre santé. **Legende:** ('HarmHealth_OutdoorAir')

Cette question évalue le niveau d'inquiétude des répondants quant aux effets potentiellement néfastes de la pollution de l'air extérieur sur leur santé.

Sur cette celle-ci je ne regrouperais pas le "un peu" et le "un petit peu".

| hints_code ✖ | HarmHealth_OutdoorAir | our_code ✔ |
|--------------|-----------------------|------------|
| -1 | Missing value | -1 |
| 4 | Not at all | 0 |
| 3 | A little = un peu | 1 |
| 2 | Some = moyennement | 2 |
| 1 | A lot = beaucoup | 3 |

Peut être croisée avec certains facteurs de l'état de santé.

Croiser avec ClimateChgHarmHealth pour voir si les personnes inquiètes pour leur santé à cause de la pollution le sont aussi à cause du changement climatique.

Question:How much do you worry that indoor air pollution will harm your health? **Legende:** ('HarmHealth_IndoorAir')

Sur celle-ci je ne regrouperais pas le "un peu" et le "un petit peu" non plus.

| hints_code ✖ | HarmHealth_IndoorAir | our_code ✔ |
|--------------|----------------------|------------|
| -1 | Missing value | -1 |
| 4 | Not at all | 0 |
| 3 | A little = un peu | 1 |
| 2 | Some = moyennement | 2 |
| 1 | A lot = beaucoup | 3 |

Piste : Voir si ceux qui s'inquiètent beaucoup ont des antécédents de santé respiratoire.

Question:During the past 12 months, how many times have you had a sunburn (even a small part of your skin turns red or hurts for 12 hours or more) from too much sun exposure? **Legend :** ('TimesSunburned')

Cette question mesure la fréquence à laquelle les répondants ont subi un ou plusieurs coups de soleil liés à une exposition excessive au soleil durant les 12 derniers mois.

Il n'y a pas de codage prédéfini comme dans les questions précédentes ; les réponses sont généralement 0, 1, 2, 3...

Ici c'est un chiffre exact et entier pour analyser les risques de **cancer de la peau** (notamment les mélanomes).

Legend : (nombre de coup de soleil à l'année)

- -1 Missing values
- 0 Sunburns in past 12 months

- 1 Sunburns in past 12 months
- 2 Sunburns in past 12 months
- 3 Sunburns in past 12 months



Jusqu'à

- 99 Sunburns in past 12 months

La période sur 12 mois permet également d'établir une fréquence d'exposition pertinente.

Question: Were you drinking alcohol at any of the times when you were sunburned? **Legend:** ('Sunburned_Alcohol2')

Question: Were you using marijuana or a marijuana product at any of the times when you were sunburned? **Legende:** ('Sunburned_Marijuana')

| hints_code  | Sunburned_Alcohol2 | our_code  |
|--|--------------------|--|
| -1 | Missing value | -1 |
| 2 | No | 0 |
| 1 | Yes | 1 |


K. Santé et Stress Financier

Question : In the past 12 months how often was the following true - Someone in your household cut the size of meals or skipped meals because

there wasn't enough money for food?


- La légende est applicable pour toutes les réponses*

Légende :

| hints_code  | |
|--|-----------------|
| -1 | Missing value |
| 0 | None |
| 1 | Very often True |
| 2 | Often True |
| 3 | Sometimes True |
| 4 | Rarely True |
| 5 | Never True |

L. Analyses de la Consommation Nutritionnel

Question : About how many cups of fruit (including 100% pure fruit juice) do you eat or drink each day? **Légende:** ('Fruit2') :

| hints_code  | |
|--|-----------------|
| -1 | Missing value |
| 0 | None |
| 1 | 1/2 cup or less |

| | |
|---|----------------------------|
| 2 | 1/2 cup to less than 1 cup |
| 3 | 1 to less than 2 cups |
| 4 | 2 to less than 3 cups |
| 5 | 3 to less than 4 cups |
| 6 | 4 or more cups |

Question : About how many cups of vegetables (including 100% pure vegetable juice) do you eat or drink each day?

Légende('Vegetables2') :

| hints_code ✗ | |
|---------------------|----------------------------|
| -1 | Missing value |
| 0 | None |
| 1 | 1/2 cup or less |
| 2 | 1/2 cup to less than 1 cup |
| 3 | 1 to less than 2 cups |
| 4 | 2 to less than 3 cups |
| 5 | 3 to less than 4 cups |
| 6 | 4 or more cups |

M.Analyses de la Consommation d'Alcool

- DrinkDaysPerMonth: M1. During the past 30 days, on how many days did you have at least one drink of any alcoholic beverage?**

→ Monthly number of day with ≥ 1 alcohol drink → nombre de jours mensuel avec au moins 1 verre d'alcool

| | |
|-----------|--|
| -1 | Missing value |
| 0 | Zero / Never |
| 1 to 30 | number of day with ≥ 1 alcohol drink (i.e. 30 = the person has at least 1 drink each day) |

- DrinksPerDay2: M2. During the past 30 days, on the days when you drank alcohol, about how many alcoholic drinks did you drink on average?**

| | |
|-----------|--|
| -1 | Missing value |
| 0 | Zero / Never |
| 1 to 20 | number of day with ≥ 1 alcohol drink (i.e. 30 = the person has at least 1 drink each day) |


- Create a new column 'Drink_nb_PerMonth' for positive values only in DrinkDaysPerMonth AND DrinksPerDay2, not -1 . returns -1 if one of the variables values is negative

| Original in df_original | Creation in df |
|-------------------------|-------------------|
| No existing ✗ | Drink_nb_PerMonth |

| | |
|-----------------------------|--|
| All individual columns only | ['DrinkDaysPerMonth'] *['DrinksPerDay2'] |
| | -1 : missing value |
| | 0 : None |
| | int64 : nb of drink per month |

N. Analyses des Comportements de Consommation: Tabac et Cannabis

Question: Have you smoked at least 100 cigarettes in your entire life?**Légende(Smoke100) :**

| hints_code | | our_code  |
|------------|---------------|--|
| -1 | Missing value | -1 |
| 3 | Not at all | 0 |
| 2 | Some days | 1 |
| 1 | Everyday | 2 |

Question: Have you smoked at least 100 cigarettes in your entire life?**Légende(SmokeNow):**

Question: Have you ever used any of the following tobacco products? (- E-cigarettes)**Légende(EverUsed_ECig) :**

Question: Have you ever used any of the following tobacco products? - Cigars, including cigars, cigarillos, and little filtered cigars.**Légende(EverUsed_Cigars) :**

Question: Have you ever used any of the following tobacco products? - Hookah (also known as Waterpipe or shisha).**Légende(EverUsed_Hookah) :**

Question: Have you ever used any of the following tobacco products? - Smokeless tobacco, including chewing tobacco, snus, and snuff ? **Légende(EverUsed_Smokeless) :**

Question: Have you ever used any of the following tobacco products? - Nicotine pouches?**Légende(EverUsed_NicPouch) :**

Question: Have you ever used any of the following tobacco products? - Nicotine pouches?**Légende(EverUsed_NicPouch) :**

Question: Have you ever used any of the following tobacco products? - Heated tobacco products, like IQOS? **Légende(EverUsed_HeatTob) :**

Question: Have you ever used any of the following tobacco products? - Modified Risk Tobacco Products, like General Snus or VLN KING, VLN? **Légende(EverUsed_ModRisk) :**

Question: Have you ever used any of the following tobacco products? - I have never used any of these tobacco products? **Légende(EverUsed_None) :**

Question: Have you ever used any of the following tobacco products? - I have never used any of these tobacco products? **Légende(EverUsed_Cat) :**

Question: Do you now use any of the following tobacco products? - E-cigarettes (also known as vapes, vape-pens, tanks, mods or pod-mods)? **Légende(NowUse_ECig) :**

Question: Do you now use any of the following tobacco products? - Cigars, including cigars, cigarillos, and little filtered cigars? **Légende(NowUse_Cigars) :**

Question: Do you now use any of the following tobacco products? - Hookah (also known as Waterpipe or shisha)? **Légende(NowUse_Hookah) :**

Question: Do you now use any of the following tobacco products? - Smokeless tobacco, including chewing tobacco, snus, and snuff? **Légende(NowUse_Smokeless) :**

Question: Do you now use any of the following tobacco products? - Nicotine pouches

Question: Do you now use any of the following tobacco products? - Heated tobacco products, like IQOS? **Légende(NowUse_HeatTob) :**

Question: Do you now use any of the following tobacco products? - Modified Risk Tobacco Products, like General Snus or VLN KING, VLN? **Légende(NowUse_ModRisk) :**

Question: Do you now use any of the following tobacco products? - I do not currently use any of these products? **Légende(NowUse_None) :**

Question: Derived variable to categorize responses given in N7 (Now use - tobacco products). **Légende(NowUse_Cat) :**

Question: When you used marijuana during the past 12 months was it usually? **Légende(marijuanausereason) :**



O. Analyses des Pratiques de Dépistages du cancer

Question: At any time in the past year, did a doctor or other health professional talk with you about having a low-dose CT (LDCT) scan to check for lung cancer?

Légende('DocTalkLDCT'):



Cette question cherche à savoir si un professionnel de santé a parlé avec la personne d'un dépistage du cancer du poumon par scanner à faible dose (LDCT) dans l'année écoulée. Elle permet de mesurer si les personnes à risque (comme les fumeurs ou ex-fumeurs) reçoivent des informations et des recommandations pour prévenir le cancer du poumon.

Legend :

| hints_code  | | | our_code  |
|--|--------------------------|----|--|
| -1 | Missing value | -1 | -1 |
| 1 | Never heard of this test | 2 | 0 |
| 2 | Yes | 1 | 1 |
| 3 | No | 2 | 0 |
| 4 | Don't know | -1 | -1 |



Question: Has a doctor or other health professional ever told you there are a few different tests to detect colorectal cancer? **Légende :** ('DocTellColorectalTests2')

Cette question vise à savoir si un professionnel de santé a **informé le répondant qu'il existe plusieurs options de dépistage** pour le **cancer colorectal**

| hints_code  | | our_code  |
|--|--------------------------|--|
| -1 | Missing value | -1 |
| 1 | Yes | 1 |
| 2 | No | 0 |
| 3 | Never heard of this test | 0 |

Q. Analyse des Cas de Cancer Déclarés



Question: Have you ever been diagnosed as having cancer? **Légende(EverHadCancer) :**

| hints_code  | | our_code  |
|--|---------------|--|
| -1 | Missing value | -1 → NAN |
| 2 | No | 0 |
| 1 | Yes | 1 |



Question: What type(s) of cancer have you been diagnosed with?(La légende est applicable pout tout les Cancer)

['EverHadCancer', 'CaBladder', 'CaBone', 'CaBreast', 'CaBrain', 'CaCervical', 'CaColon', 'CaEndometrial', 'CaEye', 'CaHeadNeck', 'CaLeukemia', 'CaLiver', 'CaLung', 'CaHodgkins', 'CaNonHodgkin', 'CaMelanoma', 'CaMultMyeloma', 'CaOral', 'CaOvarian', 'CaPancreatic', 'CaPharyngeal', 'CaProstate', 'CaRectal', 'CaRenal', 'CaSkin', 'CaStomach', 'CaTesticular', 'CaThyroid', 'CaOther', 'CaOther_OS', 'Cancer_Cat', 'WhenDiagnosedCancer', 'FamilyEverHadCancer2']

Question: Have any of your first- or second-degree biological relatives (parents, brothers and sisters, children, grandparents, aunts and uncles, nieces and nephews) ever had cancer?legend : cf Cancer_type

| hints_code  | | our_code  |
|--|---------------|--|
| -1 | Missing value | -1 → NAN |
| 2 | No / not sure | 0 |
| 1 | Yes | 1 |

Question: What type(s) of cancer have you been diagnosed with? **Légende(CaOther_OS) :**

| hints_code  | | our_code  |
|--|---------------|--|
| -1 | Missing value | -1 → NAN |
| 2 | No / not sure | 0 |
| 1 | Yes | 1 |

R.Analyses du Contexte Familiales

Question: What is your age? **Legende ('Age')**

| -1 | Missing value |
|-------------------|---------------|
| Int64 : 18 to 102 | Current Age |

Question: What sex were you assigned at birth, on your original birth certificate?

Legende('BirthSex')

| Hints 7 | | Our code |
|---------|---------------------------------------|----------|
| -1 | Missing value | -1 |
| 1 | Female | 1 |
| 2 | Male | 2 |
| 3 | Don't know → Ne souhaite pas répondre | 0 |

Question: About how many hours do you work per week at all of your jobs and businesses combined?**Legende:('WorkHrsPerWeek')**

Nombre d'heures de travail du répondant / semaine

| -1 | Missing value |
|------------------|------------------------|
| 0 | Zero |
| Int64 : 1 to 168 | Hours of work per week |

Question: Which of the following best describe your current occupational status? - Employed, including self-employed/ Legende : ('Occupation2_Employed')

| Original in df_original | Modification in df |
|-------------------------|---|
| -1 : missing value | -1 : missing value |
| 1 : Yes | 0 : No |
| 2 : No | 1 : Yes employe or self-employed (employé ou auto-entrepreneur) |

Question: Which of the following best describe your current occupational status? - Out of work for 1 year or more Legende: ('Occupation2_OOW1Y')

| Original in df_original | Modification in df |
|-------------------------|---|
| -1 : missing value | -1 : missing value |
| 1 : Yes | 0 : No |
| 2 : No | 1 : Yes, ne travaille pas depuis > 1 an |

Question: Which of the following best describe your current occupational status? - Out of work for less than 1 year. Legende: ('Occupation2_OOWLess1Y')

| Original in df_original | Modification in df |
|-------------------------|---|
| -1 : missing value | -1 : missing value |
| 1 : Yes | 0 : No |
| 2 : No | 1 : Yes, ne travaille pas depuis < 1 an |

Question: Which of the following best describe your current occupational status? - A homemaker/caregiver /Legende: ('Occupation2_Homemaker')

| Original in df_original | Modification in df |
|-------------------------|----------------------------|
| -1 : missing value | -1 : missing value |
| 1 : Yes | 0 : No |
| 2 : No | 1 : Yes, personne au foyer |

Question: Which of the following best describe your current occupational status? Legende: ('Occupation2_Student')

| Original in df_original | Modification in df |
|-------------------------|--------------------|
| -1 : missing value | -1 : missing value |
| 1 : Yes | 0 : No |
| 2 : No | 1 : Yes, student |

Question: Which of the following best describe your current occupational status? **Legende:** ('Occupation2_Retired:')

| Original in df_original | Modification in df |
|-------------------------|--------------------|
| -1 : missing value | -1 : missing value |
| 1 : Yes | 0 : No |
| 2 : No | 1 : Yes, retired |

Question: Which of the following best describe your current occupational status?

Legende: ('Occupation2_Disabled')

| Original in df_original | Modification in df |
|-------------------------|--------------------|
| -1 : missing value | -1 : missing value |
| 1 : Yes | 0 : No |
| 2 : No | 1 : Yes, invalide |

Question: Which of the following best describe your current occupational status?

Legende: ('Occupation2_Other')

| Original in df_original | Modification in df |
|-------------------------|--------------------|
| -1 : missing value | -1 : missing value |
| 1 : Yes | 0 : No |
| 2 : No | 1 : Yes, invalide |

Question: Which of the following best describe your current occupational status? **Legende:** ('Occupation2_Other_OS')

| Original in df_original | Modification in df |
|-------------------------|--|
| -1 : missing value | -1 : missing value |
| str | <i>Do : still have to think about it</i> |

Question: Derived variable to categorize responses given in R5 (Occupational status) ?**Legende** ('Occupation2_Cat')

| | |
|----|--|
| -1 | Missing data (Web partial - Question Never Seen) |
| 1 | Employed, including self-employed only |
| 2 | Out of work for 1 year or more only |
| 3 | Out of work for less than 1 year only |
| 4 | A homemaker/caregiver only |
| 5 | A student only |
| 6 | Retired only |
| 7 | Disabled only |
| 8 | Multiple Occupation statuses selected |
| 91 | Other Occupation only |

Question: Derived variable, categorizing combined responses given in R4 (Total hours worked per week) and R5 (Occupation status) **Legende :** (' FullTimeOcc2_Cat')

| | |
|----|-----------------------------|
| -1 | missing value |
| 1 | Employed full time |
| 2 | Employed part time |
| 3 | Homemaker/Caregiver |
| 4 | Student |
| 5 | Retired |
| 6 | Unable to work (Disabled) |
| 7 | Unemployed less than 1 year |
| 8 | Unemployed 1 year or more |
| 9 | Other |

Question: In the past 30 days, my work hours included... - I did not work in the past 30 days?

Legende : (' WorkHrs_DidNotWork')

| Original in df_original | Modification in df |
|-------------------------|-----------------------------------|
| -1 : missing value | -1 : missing value |
| 1 : Yes | 0 : No |
| 2 : No | 1 : Yes, didn't work past 30 days |

Question: In the past 30 days, my work hours included... - Early morning hours (4am - 8am) ?

Legende ('WorkHrs_EarlyAM')

| Original in df_original | Modification in df |
|-------------------------|---|
| -1 : missing value | -1 : missing value |
| 1 : Yes | 0 : No |
| 2 : No | 1 : Yes, early morning hours 4am to 8am |

Question: R6. In the past 30 days, my work hours included... - Morning hours (8am - 12pm)?

Legende ('WorkHrs_AM')

| Original in df_original | Modification in df |
|-------------------------|------------------------------------|
| -1 : missing value | -1 : missing value |
| 1 : Yes | 0 : No |
| 2 : No | 1 : Yes, morning hours 8am to 12pm |

Question: R6. In the past 30 days, my work hours included... - Afternoon hours (12pm - 6pm)?

Legende ('WorkHrs_Afternoon')

| Original in df_original | Modification in df |
|-------------------------|--------------------------------------|
| -1 : missing value | -1 : missing value |
| 1 : Yes | 0 : No |
| 2 : No | 1 : Yes, afternoon hours 12pm to 6pm |

Question: In the past 30 days, my work hours included... - Evening hours (6pm - 11pm)?

Legende (WorkHrs_Evening')

| Original in df_original | Modification in df |
|-------------------------|--|
| -1 : missing value | -1 : missing value |
| 1 : Yes | 0 : No |
| 2 : No | 1 : Yes, evening hours 6pm - 11pm |

Question: R6. In the past 30 days, my work hours included... - Night hours (11pm - 4am)?

Legende: ('WorkHrs_Night')

| Original in df_original | Modification in df |
|-------------------------|----------------------------------|
| -1 : missing value | -1 : missing value |
| 1 : Yes | 0 : No |
| 2 : No | 1 : Yes, night hours 11pm to 4am |

Question: Derived variable to categorize responses given in R6 (Working hours)?

Legende: ('WorkHrs_Cat')

| | |
|-----------|------------------------------|
| -1 | Missing data |
| 1 | I did not work... only |
| 2 | Early morning hours only |
| 3 | Morning jours only |
| 4 | Afternoon hours only |
| 5 | Evening hours only |
| 6 | Night hours only |
| 7 | Multiple work hours selected |

Question: What is your marital status?Legende: ('MaritalStatus)

| | |
|-----------|---|
| -1 | Missing data |
| 1 | Married |
| 2 | Living as married or living with a romantic partner |
| 3 | Divorced |
| 4 | Widowed |
| 5 | Separated |
| 6 | Single, never been married |

Question: What is the highest grade or level of schooling you completed?Legende: ('Education')

| | |
|-----------|--|
| -1 | Missing data |
| 1 | Less than 8 years |
| 2 | 8 through 11 years |
| 3 | 12 years or completed high school = lycée |
| 4 | Post high school training other than college (vocational or technical) = BTS / DUT |
| 5 | Some college = études supérieures universitaires licence, BTS |
| 6 | College graduate = BAC + 3 |
| 7 | Postgraduate = école supérieure BAC + 5 |

Question: Are you of Hispanic, Latino/a, or Spanish origin? - No, not of Hispanic, Latino/a, or Spanish origin? **Legende: ('NotHisp')**

Question: Are you of Hispanic, Latino/a, or Spanish origin? - Yes, Mexican, Mexican American, Chicano/a ? **Legende: ('Mexican')**

Question: Are you of Hispanic, Latino/a, or Spanish origin? - Yes, Puerto Rican? **Legende: ('PuertoRican')**

Question: Are you of Hispanic, Latino/a, or Spanish origin? - Yes, Cuban? **Legende: ('Cuban')**

Question: Are you of Hispanic, Latino/a, or Spanish origin? - Yes, another Hispanic, Latino/a, or Spanish origin? **Legende: ('OthHisp')**

Question: What is your race? - White? **Legende: ('White')**

Question: What is your race? - Black or African American? **Legende: ('Black')**

Question: What is your race? - American Indian or Alaska Native? **Legende: ('AmerInd')**

Question: What is your race? - Asian Indian? **Legende: ('AsInd')**

Question: What is your race? - Chinese? **Legende: ('Chinese')**

Question: What is your race? - Filipino? **Legende: ('Filipino')**

Question: What is your race? - Japanese? **Legende: ('Japanese')**

Question: What is your race? - Korean? **Legende: ('Korean')**

Question: What is your race? - Vietnamese? **Legende: ('Vietnamese')**

Question: What is your race? - Other Asian? **Legende: ('OthAsian')**

Question: What is your race? - Other Pacific Islander? **Legende: ('OthPacIsl')**

| Original in df_original | Modification in df |
|-------------------------|--------------------|
| -1 : missing value | -1 : missing value |
| 1 : Yes | 0 : No |
| 2 : No | 1 : Yes |

Question: Derived variable to categorize responses given in R9 (Hispanic ethnicity)? **Legende: ('Hisp_Cat')**

| | |
|----|--|
| -1 | Missing data |
| 10 | Not hispanic only |
| 21 | Mexican only |
| 22 | Puerto Rican only |
| 23 | Cuban Only |
| 24 | Other Hispanic only |
| 25 | Multiple Hispanic ethnicities selected |

Question: Derived variable to categorize responses given in R10 (Race)? **Legende: ('Race_Cat2')**

| | |
|----|---------------------------------------|
| -1 | Missing data |
| 11 | White only |
| 12 | Black only |
| 14 | American Indian or Alaska Native only |
| 16 | Multiple races selected |
| 31 | Asian Indian only |
| 32 | Chinese only |
| 33 | Filipino only |

| | |
|----|-----------------------------|
| 34 | Japanese only |
| 35 | Korean only |
| 36 | Vietnamese |
| 37 | Other Asian only |
| 54 | Other Pacific Islander only |

Question: How many children < 18 live in your household?**Legende:** ('ChildrenInHH')

Cette question vise à connaître **le nombre d'enfants mineurs** vivant **dans le même logement** que le répondant. Elle permet de mieux comprendre la **composition familiale** du foyer.

| -1 | Missing data |
|--------------|-------------------------|
| int64 | |
| from 0 to 32 | Number of children < 18 |

Permet de **croiser facilement avec d'autres variables** (revenu, santé, stress ou les comportements de prévention). En résumé, cette question sert à mieux cerner la structure du foyer, ce qui est essentiel pour comprendre les contextes de vie et les facteurs sociaux qui influencent la santé et le bien-être.

Question: Thinking about members of your family living in this household, what is your combined annual income, meaning the total pre-tax income from all sources earned in the past year?**Legende:** ('IncomeRanges')

Cette question demande aux répondants de déclarer le **revenu annuel total du foyer (avant impôts)** en incluant tous les membres et toutes les sources de revenus.

| -1 | Missing data |
|----|------------------------|
| 1 | \$0 to \$9,999 |
| 2 | \$10,000 to \$14,999 |
| 3 | \$15,000 to \$19,999 |
| 4 | \$20,000 to \$34,999 |
| 5 | \$35,000 to \$49,999 |
| 6 | \$50,000 to \$74,999 |
| 7 | \$75,000 to \$99,999 |
| 8 | \$100,000 to \$199,999 |
| 9 | \$200,000 or more |

Sert à étudier les **inégalités sociales**, l'**accès aux soins**, ou les **facteurs de risque en santé publique**.

Permet des **regroupements statistiques** (ex. : seuils de pauvreté, classe moyenne, hauts revenus). Peut être croisée avec de nombreuses autres variables : alimentation, santé perçue, couverture médicale, etc.

Question: Which one of these comes closest to your own feelings about your household income?**Legende:** ('IncomeFeelings')

Cette question mesure comment les personnes ressentent leur situation financière, c'est-à-dire si elles estiment pouvoir vivre confortablement, s'en sortir, ou rencontrer des difficultés avec leur revenu actuel.

| | |
|----|---|
| -1 | Missing data |
| 1 | Multiple responses selected in error |
| 2 | Living comfortably on present income |
| 3 | Getting by on present income |
| 4 | Finding it difficult on present income |
| 5 | Finding it very difficult on present income |

Points à explorer: Les liens entre difficultés financières, "santé mentale" **renoncement aux soins, comportements de santé et accès à l'information.**