analysis

July 22, 2024

1 Question: Did people's dietary and eating habits get better over the decade?

We are going to look at NHANES fasting questionnaire data in order to answer this question.

[4]:
```
import os
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import re
```

[5]:
```
filenames = [x for x in os.listdir() if re.search('.XPT',x)]
filenames
```

[5]:
```
['FASTQX_D.XPT',
 'FASTQX_E.XPT',
 'FASTQX_F.XPT',
 'FASTQX_G.XPT',
 'FASTQX_H.XPT',
 'FASTQX_I.XPT',
 'FASTQX_J.XPT',
 'PH.XPT',
 'PH_B.XPT',
 'PH_C.XPT',
 'P_FASTQX.XPT']
```

[6]:
```
total_df = pd.DataFrame()
for i in range(len(filenames)):
    if filenames[i] == 'FASTQX_E.XPT':
        year = 2008
    elif filenames[i] == 'PH.XPT':
        year = 2000
    elif filenames[i] == 'FASTQX_D.XPT':
        year = 2006
    elif filenames[i] == 'PH_B.XPT':
        year = 2002
```

```
    elif filenames[i] == 'PH_C.XPT':
        year = 2004
    elif filenames[i] == 'FASTQX_F.XPT':
        year = 2010
    elif filenames[i] == 'FASTQX_G.XPT':
        year = 2012
    elif filenames[i] == 'FASTQX_H.XPT':
        year = 2014
    elif filenames[i] == 'FASTQX_I.XPT':
        year = 2016
    elif filenames[i] == 'FASTQX_J.XPT':
        year = 2018
    elif filenames[i] == 'P_FASTQX.XPT': # this is 2017 - 2020
        year = 2020
    df = pd.read_sas(filenames[i])
    df['year'] = year
    total_df = pd.concat([df,total_df],ignore_index=True,sort=False)

    print(df.head())

    print(total_df.tail())
```

```
      SEQN  PHQ020  PHACOFHR  PHACOFMN  PHQ030  PHAALCHR  PHAALCMN  PHQ040  \
0  31128.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
1  31129.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
2  31130.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
3  31131.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
4  31132.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0

   PHAGUMHR  PHAGUMMN  PHQ050  PHAANTHR  PHAANTMN  PHQ060  PHASUPHR  \
0       NaN       NaN     2.0       NaN       NaN     2.0       NaN
1       NaN       NaN     2.0       NaN       NaN     2.0       NaN
2       NaN       NaN     2.0       NaN       NaN     2.0       NaN
3       NaN       NaN     2.0       NaN       NaN     2.0       NaN
4       NaN       NaN     2.0       NaN       NaN     1.0       1.0

       PHASUPMN  PHAFSTHR  PHAFSTMN       PHDSESN  year
0           NaN      14.0       2.0  5.397605e-79  2006
1           NaN       3.0      24.0  2.000000e+00  2006
2           NaN      10.0       3.0  5.397605e-79  2006
3           NaN      14.0       9.0  5.397605e-79  2006
4  5.397605e-79      11.0      29.0  5.397605e-79  2006
         SEQN  PHQ020  PHACOFHR  PHACOFMN  PHQ030  PHAALCHR  PHAALCMN  PHQ040  \
9435  41469.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
9436  41471.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
9437  41472.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
9438  41473.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
```

```
9439  41474.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0

      PHAGUMHR  PHAGUMMN  PHQ050  PHAANTHR  PHAANTMN  PHQ060  PHASUPHR  \
9435       NaN       NaN     2.0       NaN       NaN     2.0       NaN
9436       NaN       NaN     2.0       NaN       NaN     2.0       NaN
9437       NaN       NaN     1.0       2.0      39.0     1.0      12.0
9438       NaN       NaN     2.0       NaN       NaN     2.0       NaN
9439       NaN       NaN     2.0       NaN       NaN     2.0       NaN

      PHASUPMN  PHAFSTHR  PHAFSTMN       PHDSESN  year
9435       NaN       5.0      54.0  2.000000e+00  2006
9436       NaN      11.0       1.0  5.397605e-79  2006
9437      39.0      13.0      38.0  5.397605e-79  2006
9438       NaN       4.0      23.0  2.000000e+00  2006
9439       NaN       9.0      39.0  5.397605e-79  2006
      SEQN  PHQ020  PHACOFHR  PHACOFMN  PHQ030  PHAALCHR  PHAALCMN  PHQ040  \
0  41475.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
1  41476.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
2  41477.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
3  41478.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
4  41479.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0

   PHAGUMHR  PHAGUMMN  PHQ050  PHAANTHR  PHAANTMN  PHQ060  PHASUPHR  PHASUPMN  \
0       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
1       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
2       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
3       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
4       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN

   PHAFSTHR      PHAFSTMN       PHDSESN  year
0       7.0  5.600000e+01  1.000000e+00  2008
1      12.0  5.100000e+01  5.397605e-79  2008
2       2.0  3.100000e+01  1.000000e+00  2008
3       5.0  5.397605e-79  2.000000e+00  2008
4      14.0  3.000000e+01  5.397605e-79  2008
         SEQN  PHQ020  PHACOFHR  PHACOFMN  PHQ030  PHAALCHR  PHAALCMN  \
18742  41469.0     2.0       NaN       NaN     2.0       NaN       NaN
18743  41471.0     2.0       NaN       NaN     2.0       NaN       NaN
18744  41472.0     2.0       NaN       NaN     2.0       NaN       NaN
18745  41473.0     2.0       NaN       NaN     2.0       NaN       NaN
18746  41474.0     2.0       NaN       NaN     2.0       NaN       NaN

       PHQ040  PHAGUMHR  PHAGUMMN  PHQ050  PHAANTHR  PHAANTMN  PHQ060  \
18742     2.0       NaN       NaN     2.0       NaN       NaN     2.0
18743     2.0       NaN       NaN     2.0       NaN       NaN     2.0
18744     2.0       NaN       NaN     1.0       2.0      39.0     1.0
18745     2.0       NaN       NaN     2.0       NaN       NaN     2.0
18746     2.0       NaN       NaN     2.0       NaN       NaN     2.0
```

```
       PHASUPHR  PHASUPMN  PHAFSTHR  PHAFSTMN       PHDSESN  year
18742       NaN       NaN       5.0      54.0  2.000000e+00  2006
18743       NaN       NaN      11.0       1.0  5.397605e-79  2006
18744      12.0      39.0      13.0      38.0  5.397605e-79  2006
18745       NaN       NaN       4.0      23.0  2.000000e+00  2006
18746       NaN       NaN       9.0      39.0  5.397605e-79  2006
      SEQN  PHQ020  PHACOFHR  PHACOFMN  PHQ030  PHAALCHR  PHAALCMN  PHQ040  \
0  51624.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
1  51625.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
2  51626.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
3  51627.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
4  51628.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0

   PHAGUMHR  PHAGUMMN  PHQ050  PHAANTHR  PHAANTMN  PHQ060  PHASUPHR  PHASUPMN  \
0       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
1       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
2       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
3       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
4       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN

       PHAFSTHR  PHAFSTMN       PHDSESN  year
0  5.397605e-79      15.0  1.000000e+00  2010
1  1.000000e+00      35.0  1.000000e+00  2010
2  1.000000e+00      11.0  2.000000e+00  2010
3  1.500000e+01       7.0  5.397605e-79  2010
4  3.000000e+00      28.0  2.000000e+00  2010
          SEQN  PHQ020  PHACOFHR  PHACOFMN  PHQ030  PHAALCHR  PHAALCMN  \
28577  41469.0     2.0       NaN       NaN     2.0       NaN       NaN
28578  41471.0     2.0       NaN       NaN     2.0       NaN       NaN
28579  41472.0     2.0       NaN       NaN     2.0       NaN       NaN
28580  41473.0     2.0       NaN       NaN     2.0       NaN       NaN
28581  41474.0     2.0       NaN       NaN     2.0       NaN       NaN

       PHQ040  PHAGUMHR  PHAGUMMN  PHQ050  PHAANTHR  PHAANTMN  PHQ060  \
28577     2.0       NaN       NaN     2.0       NaN       NaN     2.0
28578     2.0       NaN       NaN     2.0       NaN       NaN     2.0
28579     2.0       NaN       NaN     1.0       2.0      39.0     1.0
28580     2.0       NaN       NaN     2.0       NaN       NaN     2.0
28581     2.0       NaN       NaN     2.0       NaN       NaN     2.0

       PHASUPHR  PHASUPMN  PHAFSTHR  PHAFSTMN       PHDSESN  year
28577       NaN       NaN       5.0      54.0  2.000000e+00  2006
28578       NaN       NaN      11.0       1.0  5.397605e-79  2006
28579      12.0      39.0      13.0      38.0  5.397605e-79  2006
28580       NaN       NaN       4.0      23.0  2.000000e+00  2006
28581       NaN       NaN       9.0      39.0  5.397605e-79  2006
      SEQN  PHQ020  PHACOFHR  PHACOFMN  PHQ030  PHAALCHR  PHAALCMN  PHQ040  \
```

```
0  62161.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
1  62162.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
2  62163.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
3  62164.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
4  62165.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0

   PHAGUMHR  PHAGUMMN  PHQ050  PHAANTHR  PHAANTMN  PHQ060  PHASUPHR  PHASUPMN  \
0       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
1       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
2       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
3       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
4       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN

   PHAFSTHR      PHAFSTMN       PHDSESN  year
0      14.0  3.700000e+01  5.397605e-79  2012
1       1.0  5.397605e-79  2.000000e+00  2012
2      17.0  5.500000e+01  1.000000e+00  2012
3      11.0  6.000000e+00  5.397605e-79  2012
4      12.0  1.100000e+01  5.397605e-79  2012
          SEQN  PHQ020  PHACOFHR  PHACOFMN  PHQ030  PHAALCHR  PHAALCMN  \
37533  41469.0     2.0       NaN       NaN     2.0       NaN       NaN
37534  41471.0     2.0       NaN       NaN     2.0       NaN       NaN
37535  41472.0     2.0       NaN       NaN     2.0       NaN       NaN
37536  41473.0     2.0       NaN       NaN     2.0       NaN       NaN
37537  41474.0     2.0       NaN       NaN     2.0       NaN       NaN

       PHQ040  PHAGUMHR  PHAGUMMN  PHQ050  PHAANTHR  PHAANTMN  PHQ060  \
37533     2.0       NaN       NaN     2.0       NaN       NaN     2.0
37534     2.0       NaN       NaN     2.0       NaN       NaN     2.0
37535     2.0       NaN       NaN     1.0       2.0      39.0     1.0
37536     2.0       NaN       NaN     2.0       NaN       NaN     2.0
37537     2.0       NaN       NaN     2.0       NaN       NaN     2.0

       PHASUPHR  PHASUPMN  PHAFSTHR  PHAFSTMN       PHDSESN  year
37533       NaN       NaN       5.0      54.0  2.000000e+00  2006
37534       NaN       NaN      11.0       1.0  5.397605e-79  2006
37535      12.0      39.0      13.0      38.0  5.397605e-79  2006
37536       NaN       NaN       4.0      23.0  2.000000e+00  2006
37537       NaN       NaN       9.0      39.0  5.397605e-79  2006
      SEQN  PHQ020  PHACOFHR  PHACOFMN  PHQ030  PHAALCHR  PHAALCMN  PHQ040  \
0  73557.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
1  73558.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
2  73559.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
3  73560.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
4  73561.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0

   PHAGUMHR  PHAGUMMN  PHQ050  PHAANTHR  PHAANTMN  PHQ060  PHASUPHR  PHASUPMN  \
0       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
```

```
1       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
2       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
3       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
4       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN

   PHAFSTHR  PHAFSTMN       PHDSESN  year
0       3.0      47.0  1.000000e+00  2014
1       3.0      14.0  2.000000e+00  2014
2      15.0      51.0  5.397605e-79  2014
3       2.0      35.0  1.000000e+00  2014
4      14.0      42.0  5.397605e-79  2014
          SEQN  PHQ020  PHACOFHR  PHACOFMN  PHQ030  PHAALCHR  PHAALCMN  \
46955  41469.0     2.0       NaN       NaN     2.0       NaN       NaN
46956  41471.0     2.0       NaN       NaN     2.0       NaN       NaN
46957  41472.0     2.0       NaN       NaN     2.0       NaN       NaN
46958  41473.0     2.0       NaN       NaN     2.0       NaN       NaN
46959  41474.0     2.0       NaN       NaN     2.0       NaN       NaN

       PHQ040  PHAGUMHR  PHAGUMMN  PHQ050  PHAANTHR  PHAANTMN  PHQ060  \
46955     2.0       NaN       NaN     2.0       NaN       NaN     2.0
46956     2.0       NaN       NaN     2.0       NaN       NaN     2.0
46957     2.0       NaN       NaN     1.0       2.0      39.0     1.0
46958     2.0       NaN       NaN     2.0       NaN       NaN     2.0
46959     2.0       NaN       NaN     2.0       NaN       NaN     2.0

       PHASUPHR  PHASUPMN  PHAFSTHR  PHAFSTMN       PHDSESN  year
46955       NaN       NaN       5.0      54.0  2.000000e+00  2006
46956       NaN       NaN      11.0       1.0  5.397605e-79  2006
46957      12.0      39.0      13.0      38.0  5.397605e-79  2006
46958       NaN       NaN       4.0      23.0  2.000000e+00  2006
46959       NaN       NaN       9.0      39.0  5.397605e-79  2006
      SEQN  PHQ020  PHACOFHR  PHACOFMN  PHQ030  PHAALCHR  PHAALCMN  PHQ040  \
0  83732.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
1  83733.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
2  83734.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
3  83735.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
4  83736.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0

   PHAGUMHR  PHAGUMMN  PHQ050  PHAANTHR  PHAANTMN  PHQ060  PHASUPHR  PHASUPMN  \
0       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
1       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
2       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
3       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
4       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN

   PHAFSTHR  PHAFSTMN       PHDSESN  year
0       3.0      50.0  2.000000e+00  2016
```

```
1      12.0       2.0  5.397605e-79  2016
2      10.0      27.0  5.397605e-79  2016
3       2.0      17.0  1.000000e+00  2016
4      10.0      35.0  5.397605e-79  2016
          SEQN  PHQ020  PHACOFHR  PHACOFMN  PHQ030  PHAALCHR  PHAALCMN  \
56120  41469.0     2.0       NaN       NaN     2.0       NaN       NaN
56121  41471.0     2.0       NaN       NaN     2.0       NaN       NaN
56122  41472.0     2.0       NaN       NaN     2.0       NaN       NaN
56123  41473.0     2.0       NaN       NaN     2.0       NaN       NaN
56124  41474.0     2.0       NaN       NaN     2.0       NaN       NaN

       PHQ040  PHAGUMHR  PHAGUMMN  PHQ050  PHAANTHR  PHAANTMN  PHQ060  \
56120     2.0       NaN       NaN     2.0       NaN       NaN     2.0
56121     2.0       NaN       NaN     2.0       NaN       NaN     2.0
56122     2.0       NaN       NaN     1.0       2.0      39.0     1.0
56123     2.0       NaN       NaN     2.0       NaN       NaN     2.0
56124     2.0       NaN       NaN     2.0       NaN       NaN     2.0

       PHASUPHR  PHASUPMN  PHAFSTHR  PHAFSTMN       PHDSESN  year
56120       NaN       NaN       5.0      54.0  2.000000e+00  2006
56121       NaN       NaN      11.0       1.0  5.397605e-79  2006
56122      12.0      39.0      13.0      38.0  5.397605e-79  2006
56123       NaN       NaN       4.0      23.0  2.000000e+00  2006
56124       NaN       NaN       9.0      39.0  5.397605e-79  2006
      SEQN  PHQ020  PHACOFHR  PHACOFMN  PHQ030  PHAALCHR  PHAALCMN  PHQ040  \
0  93703.0     NaN       NaN       NaN     NaN       NaN       NaN     NaN
1  93704.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
2  93705.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
3  93706.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
4  93707.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0

   PHAGUMHR  PHAGUMMN  PHQ050  PHAANTHR  PHAANTMN  PHQ060  PHASUPHR  PHASUPMN  \
0       NaN       NaN     NaN       NaN       NaN     NaN       NaN       NaN
1       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
2       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
3       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
4       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN

   PHAFSTHR  PHAFSTMN  PHDSESN  year
0       NaN       NaN      2.0  2018
1       1.0      38.0      1.0  2018
2       4.0      32.0      2.0  2018
3      18.0      25.0      1.0  2018
4       3.0      16.0      2.0  2018
          SEQN  PHQ020  PHACOFHR  PHACOFMN  PHQ030  PHAALCHR  PHAALCMN  \
64486  41469.0     2.0       NaN       NaN     2.0       NaN       NaN
64487  41471.0     2.0       NaN       NaN     2.0       NaN       NaN
64488  41472.0     2.0       NaN       NaN     2.0       NaN       NaN
```

```
64489  41473.0     2.0       NaN       NaN     2.0       NaN       NaN
64490  41474.0     2.0       NaN       NaN     2.0       NaN       NaN

       PHQ040  PHAGUMHR  PHAGUMMN  PHQ050  PHAANTHR  PHAANTMN  PHQ060  \
64486     2.0       NaN       NaN     2.0       NaN       NaN     2.0
64487     2.0       NaN       NaN     2.0       NaN       NaN     2.0
64488     2.0       NaN       NaN     1.0       2.0      39.0     1.0
64489     2.0       NaN       NaN     2.0       NaN       NaN     2.0
64490     2.0       NaN       NaN     2.0       NaN       NaN     2.0

       PHASUPHR  PHASUPMN  PHAFSTHR  PHAFSTMN       PHDSESN  year
64486       NaN       NaN       5.0      54.0  2.000000e+00  2006
64487       NaN       NaN      11.0       1.0  5.397605e-79  2006
64488      12.0      39.0      13.0      38.0  5.397605e-79  2006
64489       NaN       NaN       4.0      23.0  2.000000e+00  2006
64490       NaN       NaN       9.0      39.0  5.397605e-79  2006
   SEQN  PHQ020  PHACOFHR  PHACOFMN  PHQ030  PHAALCHR  PHAALCMN  PHQ040  \
0   1.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
1   2.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
2   3.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
3   4.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
4   5.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0

   PHAGUMHR  PHAGUMMN  PHQ050  PHAANTHR  PHAANTMN  PHQ060  PHASUPHR  PHASUPMN  \
0       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
1       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
2       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
3       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
4       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN

   PHAFSTHR  PHAFSTMN       PHDSESN  year
0      13.0      15.0  5.397605e-79  2000
1      12.0      47.0  5.397605e-79  2000
2       3.0       1.0  5.397605e-79  2000
3       6.0      10.0  1.000000e+00  2000
4      12.0      44.0  5.397605e-79  2000
          SEQN  PHQ020  PHACOFHR  PHACOFMN  PHQ030  PHAALCHR  PHAALCMN  \
73318  41469.0     2.0       NaN       NaN     2.0       NaN       NaN
73319  41471.0     2.0       NaN       NaN     2.0       NaN       NaN
73320  41472.0     2.0       NaN       NaN     2.0       NaN       NaN
73321  41473.0     2.0       NaN       NaN     2.0       NaN       NaN
73322  41474.0     2.0       NaN       NaN     2.0       NaN       NaN

       PHQ040  PHAGUMHR  PHAGUMMN  PHQ050  PHAANTHR  PHAANTMN  PHQ060  \
73318     2.0       NaN       NaN     2.0       NaN       NaN     2.0
73319     2.0       NaN       NaN     2.0       NaN       NaN     2.0
73320     2.0       NaN       NaN     1.0       2.0      39.0     1.0
73321     2.0       NaN       NaN     2.0       NaN       NaN     2.0
```

```
73322     2.0       NaN       NaN     2.0       NaN       NaN     2.0

       PHASUPHR  PHASUPMN  PHAFSTHR  PHAFSTMN       PHDSESN  year
73318       NaN       NaN       5.0      54.0  2.000000e+00  2006
73319       NaN       NaN      11.0       1.0  5.397605e-79  2006
73320      12.0      39.0      13.0      38.0  5.397605e-79  2006
73321       NaN       NaN       4.0      23.0  2.000000e+00  2006
73322       NaN       NaN       9.0      39.0  5.397605e-79  2006
     SEQN  PHQ020  PHACOFHR  PHACOFMN  PHQ030  PHAALCHR  PHAALCMN  PHQ040  \
0  9966.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
1  9967.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
2  9968.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
3  9969.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
4  9970.0     NaN       NaN       NaN     NaN       NaN       NaN     NaN

   PHAGUMHR  PHAGUMMN  PHQ050  PHAANTHR  PHAANTMN  PHQ060  PHASUPHR  PHASUPMN  \
0       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
1       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
2       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
3       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
4       NaN       NaN     NaN       NaN       NaN     NaN       NaN       NaN

   PHAFSTHR  PHAFSTMN       PHDSESN  year
0      12.0      43.0  5.397605e-79  2002
1      12.0      50.0  1.000000e+00  2002
2      15.0      47.0  5.397605e-79  2002
3      13.0      33.0  5.397605e-79  2002
4       NaN       NaN  5.397605e-79  2002
          SEQN  PHQ020  PHACOFHR  PHACOFMN  PHQ030  PHAALCHR  PHAALCMN  \
83247  41469.0     2.0       NaN       NaN     2.0       NaN       NaN
83248  41471.0     2.0       NaN       NaN     2.0       NaN       NaN
83249  41472.0     2.0       NaN       NaN     2.0       NaN       NaN
83250  41473.0     2.0       NaN       NaN     2.0       NaN       NaN
83251  41474.0     2.0       NaN       NaN     2.0       NaN       NaN

       PHQ040  PHAGUMHR  PHAGUMMN  PHQ050  PHAANTHR  PHAANTMN  PHQ060  \
83247     2.0       NaN       NaN     2.0       NaN       NaN     2.0
83248     2.0       NaN       NaN     2.0       NaN       NaN     2.0
83249     2.0       NaN       NaN     1.0       2.0      39.0     1.0
83250     2.0       NaN       NaN     2.0       NaN       NaN     2.0
83251     2.0       NaN       NaN     2.0       NaN       NaN     2.0

       PHASUPHR  PHASUPMN  PHAFSTHR  PHAFSTMN       PHDSESN  year
83247       NaN       NaN       5.0      54.0  2.000000e+00  2006
83248       NaN       NaN      11.0       1.0  5.397605e-79  2006
83249      12.0      39.0      13.0      38.0  5.397605e-79  2006
83250       NaN       NaN       4.0      23.0  2.000000e+00  2006
83251       NaN       NaN       9.0      39.0  5.397605e-79  2006
```

```
      SEQN  PHQ020  PHACOFHR  PHACOFMN  PHQ030  PHAALCHR  PHAALCMN  PHQ040  \
0  21005.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
1  21006.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
2  21007.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
3  21008.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
4  21009.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0

   PHAGUMHR  PHAGUMMN  PHQ050  PHAANTHR  PHAANTMN  PHQ060  PHASUPHR  PHASUPMN  \
0       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
1       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
2       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
3       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
4       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN

   PHAFSTHR  PHAFSTMN       PHDSESN  year
0      10.0       2.0  5.397605e-79  2004
1      12.0      22.0  5.397605e-79  2004
2      20.0      55.0  1.000000e+00  2004
3       4.0      35.0  1.000000e+00  2004
4       6.0      15.0  1.000000e+00  2004
         SEQN  PHQ020  PHACOFHR  PHACOFMN  PHQ030  PHAALCHR  PHAALCMN  \
92426  41469.0     2.0       NaN       NaN     2.0       NaN       NaN
92427  41471.0     2.0       NaN       NaN     2.0       NaN       NaN
92428  41472.0     2.0       NaN       NaN     2.0       NaN       NaN
92429  41473.0     2.0       NaN       NaN     2.0       NaN       NaN
92430  41474.0     2.0       NaN       NaN     2.0       NaN       NaN

       PHQ040  PHAGUMHR  PHAGUMMN  PHQ050  PHAANTHR  PHAANTMN  PHQ060  \
92426     2.0       NaN       NaN     2.0       NaN       NaN     2.0
92427     2.0       NaN       NaN     2.0       NaN       NaN     2.0
92428     2.0       NaN       NaN     1.0       2.0      39.0     1.0
92429     2.0       NaN       NaN     2.0       NaN       NaN     2.0
92430     2.0       NaN       NaN     2.0       NaN       NaN     2.0

       PHASUPHR  PHASUPMN  PHAFSTHR  PHAFSTMN       PHDSESN  year
92426       NaN       NaN       5.0      54.0  2.000000e+00  2006
92427       NaN       NaN      11.0       1.0  5.397605e-79  2006
92428      12.0      39.0      13.0      38.0  5.397605e-79  2006
92429       NaN       NaN       4.0      23.0  2.000000e+00  2006
92430       NaN       NaN       9.0      39.0  5.397605e-79  2006
       SEQN  PHQ020  PHACOFHR  PHACOFMN  PHQ030  PHAALCHR  PHAALCMN  PHQ040  \
0  109263.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
1  109264.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
2  109265.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
3  109266.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
4  109269.0     2.0       NaN       NaN     2.0       NaN       NaN     2.0
```

```
   PHAGUMHR  PHAGUMMN  PHQ050  PHAANTHR  PHAANTMN  PHQ060  PHASUPHR  PHASUPMN  \
0       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
1       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
2       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
3       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN
4       NaN       NaN     2.0       NaN       NaN     2.0       NaN       NaN

   PHAFSTHR      PHAFSTMN       PHDSESN  year
0      15.0  3.700000e+01  5.397605e-79  2020
1      10.0  2.400000e+01  5.397605e-79  2020
2       6.0  5.397605e-79  2.000000e+00  2020
3      16.0  5.100000e+01  1.000000e+00  2020
4      12.0  1.800000e+01  5.397605e-79  2020
           SEQN  PHQ020  PHACOFHR  PHACOFMN  PHQ030  PHAALCHR  PHAALCMN  \
106198  41469.0     2.0       NaN       NaN     2.0       NaN       NaN
106199  41471.0     2.0       NaN       NaN     2.0       NaN       NaN
106200  41472.0     2.0       NaN       NaN     2.0       NaN       NaN
106201  41473.0     2.0       NaN       NaN     2.0       NaN       NaN
106202  41474.0     2.0       NaN       NaN     2.0       NaN       NaN

        PHQ040  PHAGUMHR  PHAGUMMN  PHQ050  PHAANTHR  PHAANTMN  PHQ060  \
106198     2.0       NaN       NaN     2.0       NaN       NaN     2.0
106199     2.0       NaN       NaN     2.0       NaN       NaN     2.0
106200     2.0       NaN       NaN     1.0       2.0      39.0     1.0
106201     2.0       NaN       NaN     2.0       NaN       NaN     2.0
106202     2.0       NaN       NaN     2.0       NaN       NaN     2.0

        PHASUPHR  PHASUPMN  PHAFSTHR  PHAFSTMN       PHDSESN  year
106198       NaN       NaN       5.0      54.0  2.000000e+00  2006
106199       NaN       NaN      11.0       1.0  5.397605e-79  2006
106200      12.0      39.0      13.0      38.0  5.397605e-79  2006
106201       NaN       NaN       4.0      23.0  2.000000e+00  2006
106202       NaN       NaN       9.0      39.0  5.397605e-79  2006
```

[7]:
```
total_df
```

[7]:
```
            SEQN  PHQ020  PHACOFHR  PHACOFMN  PHQ030  PHAALCHR  PHAALCMN  \
0       109263.0     2.0       NaN       NaN     2.0       NaN       NaN
1       109264.0     2.0       NaN       NaN     2.0       NaN       NaN
2       109265.0     2.0       NaN       NaN     2.0       NaN       NaN
3       109266.0     2.0       NaN       NaN     2.0       NaN       NaN
4       109269.0     2.0       NaN       NaN     2.0       NaN       NaN
…            …     …         …         …       …         …         …
106198   41469.0     2.0       NaN       NaN     2.0       NaN       NaN
106199   41471.0     2.0       NaN       NaN     2.0       NaN       NaN
106200   41472.0     2.0       NaN       NaN     2.0       NaN       NaN
106201   41473.0     2.0       NaN       NaN     2.0       NaN       NaN
```

```
106202   41474.0     2.0       NaN       NaN     2.0       NaN       NaN

        PHQ040  PHAGUMHR  PHAGUMMN  PHQ050  PHAANTHR  PHAANTMN  PHQ060  \
0          2.0       NaN       NaN     2.0       NaN       NaN     2.0
1          2.0       NaN       NaN     2.0       NaN       NaN     2.0
2          2.0       NaN       NaN     2.0       NaN       NaN     2.0
3          2.0       NaN       NaN     2.0       NaN       NaN     2.0
4          2.0       NaN       NaN     2.0       NaN       NaN     2.0
…          …         …         …       …         …         …       …
106198     2.0       NaN       NaN     2.0       NaN       NaN     2.0
106199     2.0       NaN       NaN     2.0       NaN       NaN     2.0
106200     2.0       NaN       NaN     1.0       2.0      39.0     1.0
106201     2.0       NaN       NaN     2.0       NaN       NaN     2.0
106202     2.0       NaN       NaN     2.0       NaN       NaN     2.0

        PHASUPHR  PHASUPMN  PHAFSTHR      PHAFSTMN      PHDSESN  year
0            NaN       NaN      15.0  3.700000e+01  5.397605e-79  2020
1            NaN       NaN      10.0  2.400000e+01  5.397605e-79  2020
2            NaN       NaN       6.0  5.397605e-79  2.000000e+00  2020
3            NaN       NaN      16.0  5.100000e+01  1.000000e+00  2020
4            NaN       NaN      12.0  1.800000e+01  5.397605e-79  2020
…            …         …        …     …             …             …
106198       NaN       NaN       5.0  5.400000e+01  2.000000e+00  2006
106199       NaN       NaN      11.0  1.000000e+00  5.397605e-79  2006
106200      12.0      39.0      13.0  3.800000e+01  5.397605e-79  2006
106201       NaN       NaN       4.0  2.300000e+01  2.000000e+00  2006
106202       NaN       NaN       9.0  3.900000e+01  5.397605e-79  2006

[106203 rows x 20 columns]
```

1.0.1 Variable (feature) explanations:

SEQN - Respondent sequence number PHQ020 - Coffee or tea with cream or sugar? - binary yes (1) or no (2) PHACOFHR - Coffee/tea fast time (hours) PHACOFMN - Coffee/tea fast time (minutes) PHQ030 - Alcohol, such as beer, wine, or liquor? - binary yes (1) or no (2) PHAALCHR - Alcohol fast time (hours) PHAALCMN - Alcohol fast time (minutes) PHQ040 - Gum, mints, lozenges or cough drops - binary yes (1) or no (2) PHAGUMHR - Gum, mints cough drops fast time (hours) PHAGUMMN - Gum, mints, cough fast time (minutes) PHQ050 - Antacids, laxatives, or anti-diarrheals? - binary yes (1) or no (2) PHAANTHR - Antacids, laxatives fast time (hours) PHAANTMN - Antacids, laxatives fast time (minutes) PHQ060 - Dietary supplements? - binary yes (1) or no (2) PHASUPHR - Dietary supplements fast time (hours) PHASUPMN - Dietary supplements fast time (minutes) PHAFSTHR - Total length of "food fast," hours PHAFSTMN - Total length of "food fast," minutes PHDSESN - Session in which SP was examined

[8]:
```
nums = 0
for i in filenames:
    num = pd.read_sas(i).shape[0]
```

```
    nums += num
print(nums)
```

```
106203
```

[9]:
```
total_df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 106203 entries, 0 to 106202
Data columns (total 20 columns):
 #   Column    Non-Null Count   Dtype  
---  ------    --------------   -----  
 0   SEQN      106203 non-null  float64
 1   PHQ020    102064 non-null  float64
 2   PHACOFHR  848 non-null     float64
 3   PHACOFMN  848 non-null     float64
 4   PHQ030    102064 non-null  float64
 5   PHAALCHR  335 non-null     float64
 6   PHAALCMN  335 non-null     float64
 7   PHQ040    102064 non-null  float64
 8   PHAGUMHR  2488 non-null    float64
 9   PHAGUMMN  2488 non-null    float64
 10  PHQ050    102064 non-null  float64
 11  PHAANTHR  389 non-null     float64
 12  PHAANTMN  389 non-null     float64
 13  PHQ060    102064 non-null  float64
 14  PHASUPHR  1001 non-null    float64
 15  PHASUPMN  1001 non-null    float64
 16  PHAFSTHR  102057 non-null  float64
 17  PHAFSTMN  102057 non-null  float64
 18  PHDSESN   106202 non-null  float64
 19  year      106203 non-null  int64  
dtypes: float64(19), int64(1)
memory usage: 16.2 MB
```

[10]:
```
total_df.groupby('year')['PHACOFHR'].describe()
```

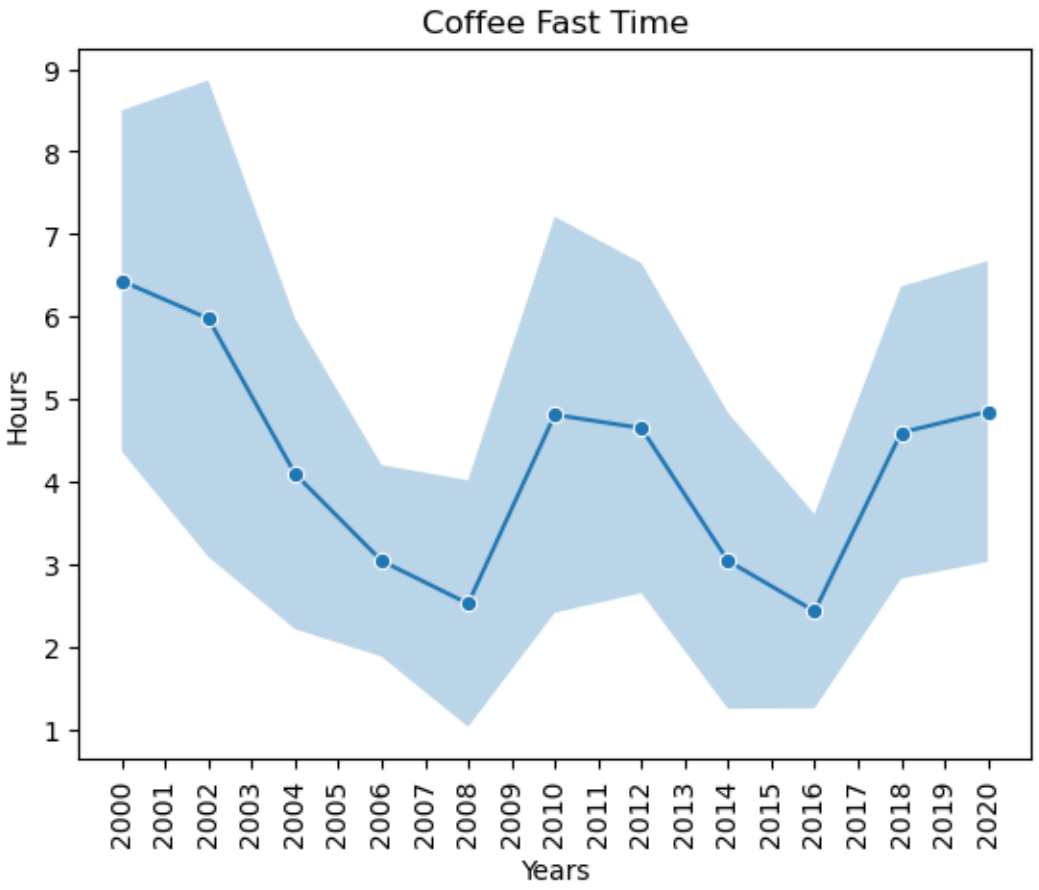
[10]:

year	count	mean	std	min	25%	50%	75%	max
2000	44.0	6.431818	4.133796	5.397605e-79	3.0	5.5	10.00	15.0
2002	123.0	5.975610	5.777945	5.397605e-79	2.0	5.0	8.00	49.0
2004	64.0	4.093750	3.765840	5.397605e-79	2.0	3.0	5.00	21.0
2006	26.0	3.038462	2.323459	5.397605e-79	1.0	2.0	4.75	10.0
2008	25.0	2.520000	2.987753	5.397605e-79	1.0	1.0	3.00	12.0
2010	21.0	4.809524	4.802281	1.000000e+00	2.0	2.0	4.00	19.0
2012	71.0	4.647887	4.000352	5.397605e-79	1.0	3.0	6.50	14.0
2014	129.0	3.038760	3.589053	5.397605e-79	1.0	2.0	4.00	18.0

```
2016   68.0  2.426471  2.351983  5.397605e-79  1.0  2.0   4.00  10.0
2018  118.0  4.593220  3.542748  5.397605e-79  2.0  4.0   5.00  17.0
2020  159.0  4.849057  3.647764  5.397605e-79  2.0  4.0   6.00  17.0
```

[11]:
```
sns.lineplot(data=total_df.groupby('year')['PHACOFHR'].
 ↪describe(),x='year',y='mean',marker="o")
upper = total_df.groupby('year')['PHACOFHR'].describe()['mean'] + total_df.
 ↪groupby('year')['PHACOFHR'].describe()['std']/2
lower = total_df.groupby('year')['PHACOFHR'].describe()['mean'] - total_df.
 ↪groupby('year')['PHACOFHR'].describe()['std']/2
plt.fill_between(total_df.groupby('year')['PHACOFHR'].describe().index,
                 lower,
                 upper,
                 alpha=.3)
plt.title('Coffee Fast Time')
plt.ylabel('Hours')
plt.xlabel('Years')
plt.xticks(np.arange(2000,2021,step=1), rotation=90)
plt.show()
```

```
C:\Users\brian\Anaconda3\Lib\site-packages\seaborn\_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed in a
future version. Convert inf values to NaN before operating instead.
  with pd.option_context('mode.use_inf_as_na', True):
C:\Users\brian\Anaconda3\Lib\site-packages\seaborn\_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed in a
future version. Convert inf values to NaN before operating instead.
  with pd.option_context('mode.use_inf_as_na', True):
```



We are looking at coffee fasting time in hours. It seems like the average is around 4 hours between each coffee. It looks like the coffee fast time decreased until 2008 and then increased slightly to fall back again. It is increasing as of late.

[12]:
```
total_df.groupby('year')['PHACOFMN'].describe()
```

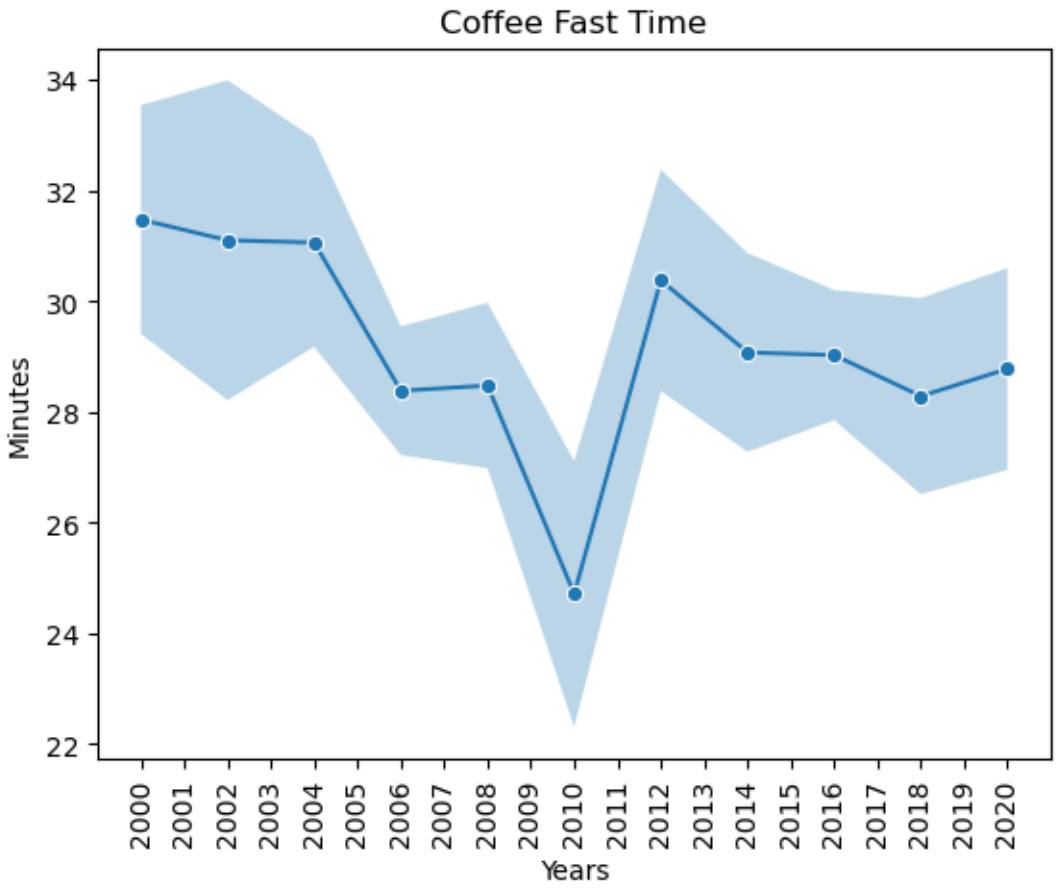
[12]:

year	count	mean	std	min	25%	50%	75%	max
2000	44.0	31.477273	16.391275	2.000000e+00	18.50	36.0	43.25	58.0
2002	123.0	31.105691	17.542834	5.397605e-79	15.00	34.0	46.00	59.0
2004	64.0	31.062500	17.704699	5.397605e-79	14.00	32.5	46.00	59.0
2006	26.0	28.384615	16.603799	3.000000e+00	17.00	27.0	45.00	54.0
2008	25.0	28.480000	15.697983	2.000000e+00	19.00	29.0	37.00	58.0
2010	21.0	24.714286	14.993808	1.000000e+00	16.00	23.0	33.00	54.0
2012	71.0	30.380282	19.331962	5.397605e-79	10.00	32.0	48.50	59.0
2014	129.0	29.077519	17.895798	5.397605e-79	14.00	30.0	46.00	59.0
2016	68.0	29.029412	18.242946	5.397605e-79	12.75	31.0	46.00	58.0
2018	118.0	28.288136	18.377384	5.397605e-79	10.25	30.0	43.00	59.0
2020	159.0	28.779874	18.380868	5.397605e-79	11.00	31.0	44.50	59.0

[13]:
```
sns.lineplot(data=total_df.groupby('year')['PHACOFMN'].
 ↪describe(),x='year',y='mean',marker="o")
upper = total_df.groupby('year')['PHACOFMN'].describe()['mean'] + total_df.
 ↪groupby('year')['PHACOFHR'].describe()['std']/2
lower = total_df.groupby('year')['PHACOFMN'].describe()['mean'] - total_df.
 ↪groupby('year')['PHACOFHR'].describe()['std']/2
plt.fill_between(total_df.groupby('year')['PHACOFMN'].describe().index,
                 lower,
                 upper,
                 alpha=.3)

plt.title('Coffee Fast Time')
plt.ylabel('Minutes')
plt.xlabel('Years')
plt.xticks(np.arange(2000,2021,step=1), rotation=90)
plt.show()
```

```
C:\Users\brian\Anaconda3\Lib\site-packages\seaborn\_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed in a
future version. Convert inf values to NaN before operating instead.
  with pd.option_context('mode.use_inf_as_na', True):
C:\Users\brian\Anaconda3\Lib\site-packages\seaborn\_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed in a
future version. Convert inf values to NaN before operating instead.
  with pd.option_context('mode.use_inf_as_na', True):
```

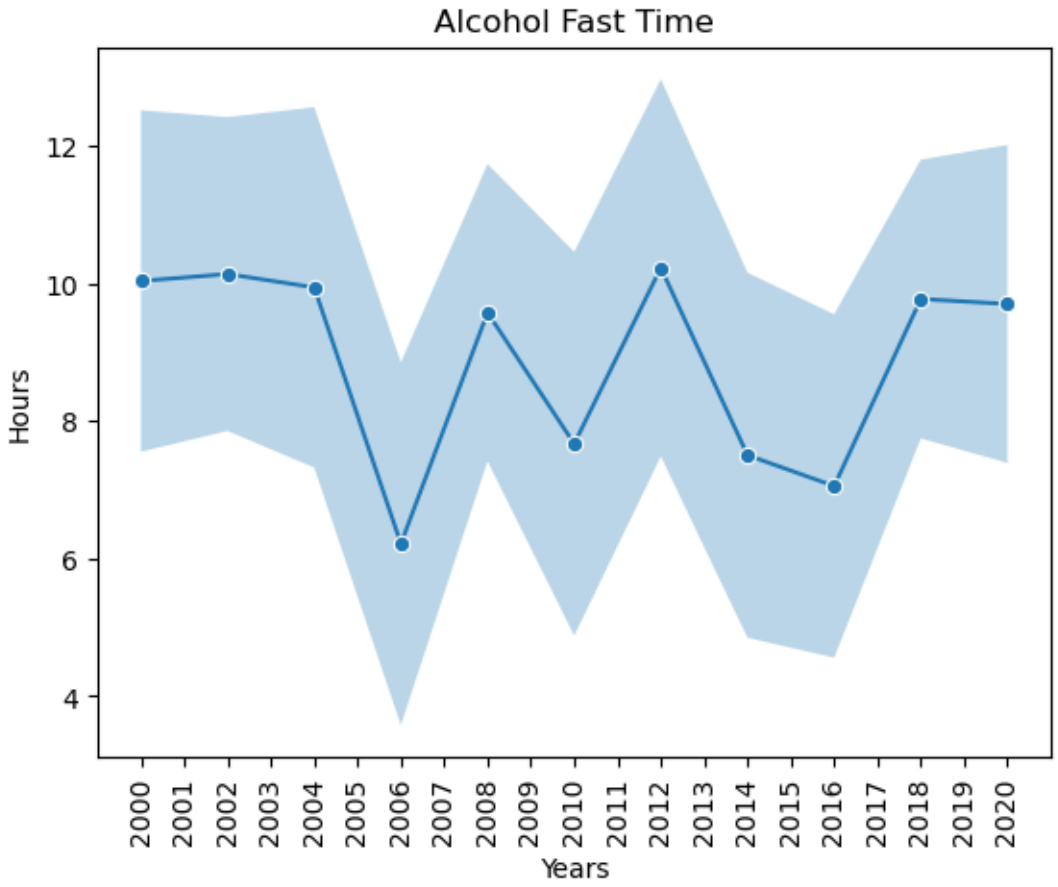


Unlike the data above, the fasting times between minutes for coffee is averaged at around 28 minutes. Also, the fasting time has decreased until 2010 and then increased to 30 minutes to stabilize there.

[14]:
```
sns.lineplot(data=total_df.groupby('year')['PHAALCHR'].
 ↪describe(),x='year',y='mean',marker="o")
upper = total_df.groupby('year')['PHAALCHR'].describe()['mean'] + total_df.
 ↪groupby('year')['PHAALCHR'].describe()['std']/2
lower = total_df.groupby('year')['PHAALCHR'].describe()['mean'] - total_df.
 ↪groupby('year')['PHAALCHR'].describe()['std']/2
plt.fill_between(total_df.groupby('year')['PHAALCHR'].describe().index,
                 lower,
                 upper,
                 alpha=.3)

plt.title('Alcohol Fast Time')
plt.ylabel('Hours')
plt.xlabel('Years')
plt.xticks(np.arange(2000,2021,step=1), rotation=90)
plt.show()
```

```
C:\Users\brian\Anaconda3\Lib\site-packages\seaborn\_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed in a
future version. Convert inf values to NaN before operating instead.
  with pd.option_context('mode.use_inf_as_na', True):
C:\Users\brian\Anaconda3\Lib\site-packages\seaborn\_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed in a
future version. Convert inf values to NaN before operating instead.
  with pd.option_context('mode.use_inf_as_na', True):
```



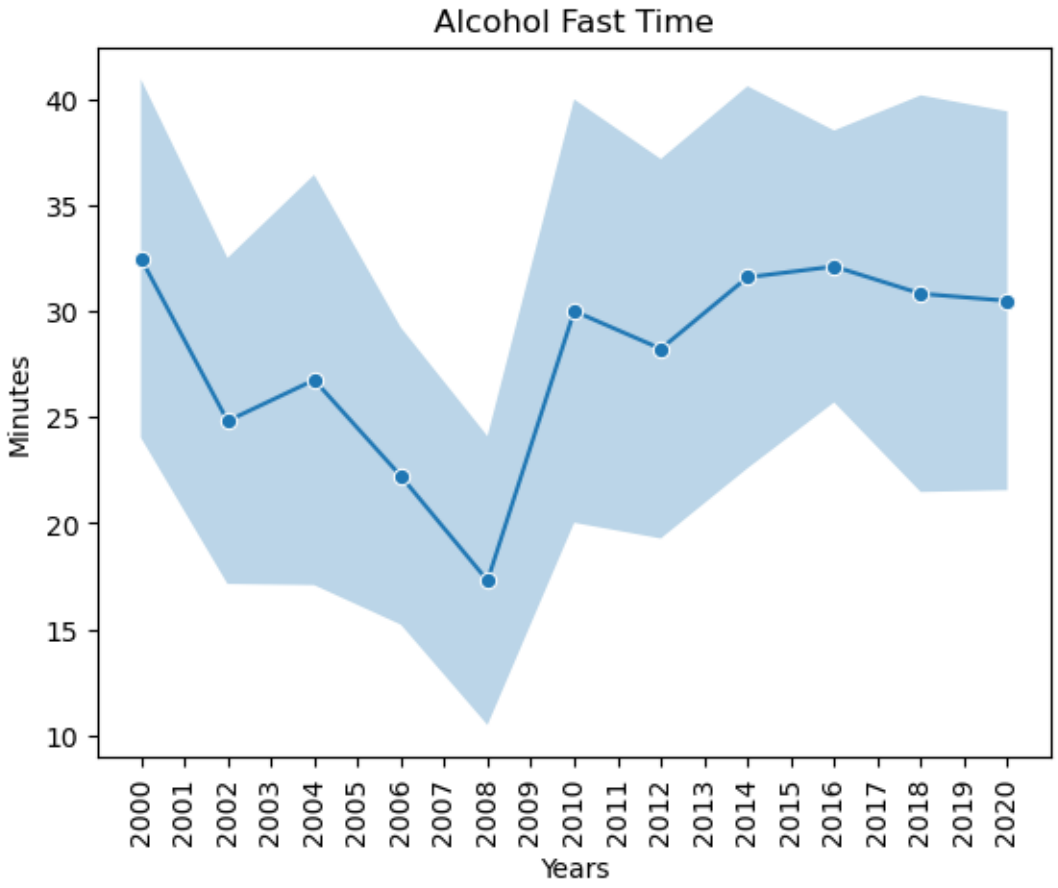
It looks like the alcohol is relatively stable without much increase or decrease (although there is some "noise" or deviation). It looks like the average fasting time is at around 9-10 hours.

[15]:
```
sns.lineplot(data=total_df.groupby('year')['PHAALCMN'].
  ↪describe(),x='year',y='mean',marker="o")
upper = total_df.groupby('year')['PHAALCMN'].describe()['mean'] + total_df.
  ↪groupby('year')['PHAALCMN'].describe()['std']/2
lower = total_df.groupby('year')['PHAALCMN'].describe()['mean'] - total_df.
  ↪groupby('year')['PHAALCMN'].describe()['std']/2
plt.fill_between(total_df.groupby('year')['PHAALCMN'].describe().index,
```

```
                lower,
                upper,
                alpha=.3)

plt.title('Alcohol Fast Time')
plt.ylabel('Minutes')
plt.xlabel('Years')
plt.xticks(np.arange(2000,2021,step=1), rotation=90)
plt.show()
```

```
C:\Users\brian\Anaconda3\Lib\site-packages\seaborn\_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed in a
future version. Convert inf values to NaN before operating instead.
  with pd.option_context('mode.use_inf_as_na', True):
C:\Users\brian\Anaconda3\Lib\site-packages\seaborn\_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed in a
future version. Convert inf values to NaN before operating instead.
  with pd.option_context('mode.use_inf_as_na', True):
```

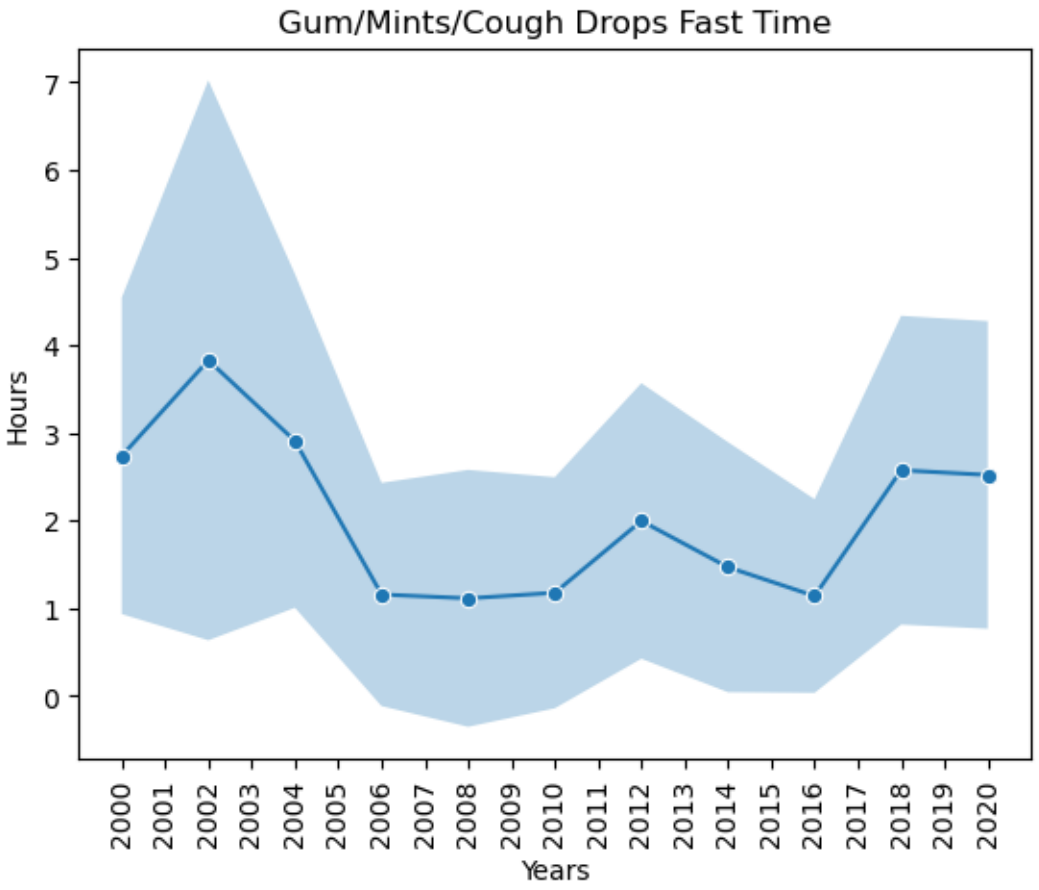


Unlike the data above, the alcohol fasting time in minutes is decreasing to 15 minutes until 2008, and then it increases to stabilize at around 30 minutes.

[16]:
```
sns.lineplot(data=total_df.groupby('year')['PHAGUMHR'].
 ↪describe(),x='year',y='mean',marker="o")
upper = total_df.groupby('year')['PHAGUMHR'].describe()['mean'] + total_df.
 ↪groupby('year')['PHAGUMHR'].describe()['std']/2
lower = total_df.groupby('year')['PHAGUMHR'].describe()['mean'] - total_df.
 ↪groupby('year')['PHAGUMHR'].describe()['std']/2
plt.fill_between(total_df.groupby('year')['PHAGUMHR'].describe().index,
                 lower,
                 upper,
                 alpha=.3)

plt.title('Gum/Mints/Cough Drops Fast Time')
plt.ylabel('Hours')
plt.xlabel('Years')
plt.xticks(np.arange(2000,2021,step=1), rotation=90)
plt.show()
```

```
C:\Users\brian\Anaconda3\Lib\site-packages\seaborn\_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed in a
future version. Convert inf values to NaN before operating instead.
  with pd.option_context('mode.use_inf_as_na', True):
C:\Users\brian\Anaconda3\Lib\site-packages\seaborn\_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed in a
future version. Convert inf values to NaN before operating instead.
  with pd.option_context('mode.use_inf_as_na', True):
```

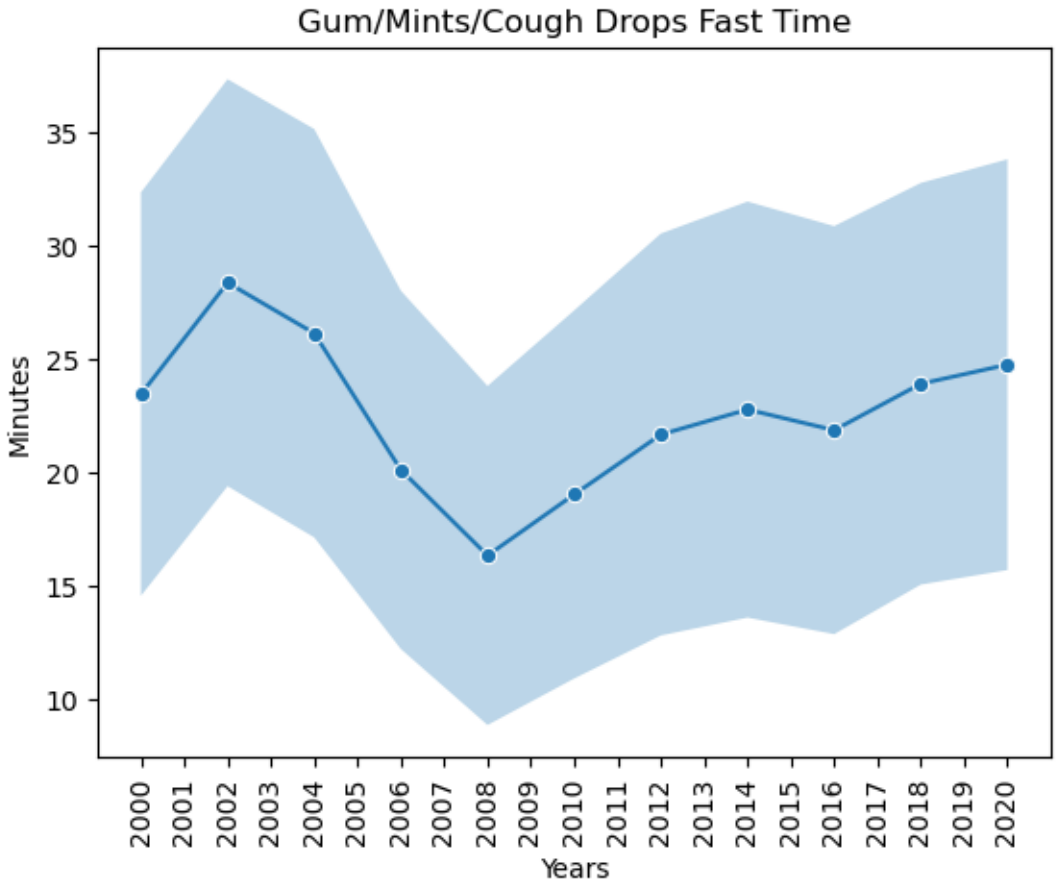


It seems to be stable, averaging at around 2 minutes without much deviation other than at the beginning of 2001.

[17]:
```
sns.lineplot(data=total_df.groupby('year')['PHAGUMMN'].
 ↪describe(),x='year',y='mean',marker="o")
upper = total_df.groupby('year')['PHAGUMMN'].describe()['mean'] + total_df.
 ↪groupby('year')['PHAGUMMN'].describe()['std']/2
lower = total_df.groupby('year')['PHAGUMMN'].describe()['mean'] - total_df.
 ↪groupby('year')['PHAGUMMN'].describe()['std']/2
plt.fill_between(total_df.groupby('year')['PHAGUMMN'].describe().index,
                 lower,
                 upper,
                 alpha=.3)

plt.title('Gum/Mints/Cough Drops Fast Time')
plt.ylabel('Minutes')
plt.xlabel('Years')
plt.xticks(np.arange(2000,2021,step=1), rotation=90)
plt.show()
```

```
C:\Users\brian\Anaconda3\Lib\site-packages\seaborn\_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed in a
future version. Convert inf values to NaN before operating instead.
  with pd.option_context('mode.use_inf_as_na', True):
C:\Users\brian\Anaconda3\Lib\site-packages\seaborn\_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed in a
future version. Convert inf values to NaN before operating instead.
  with pd.option_context('mode.use_inf_as_na', True):
```



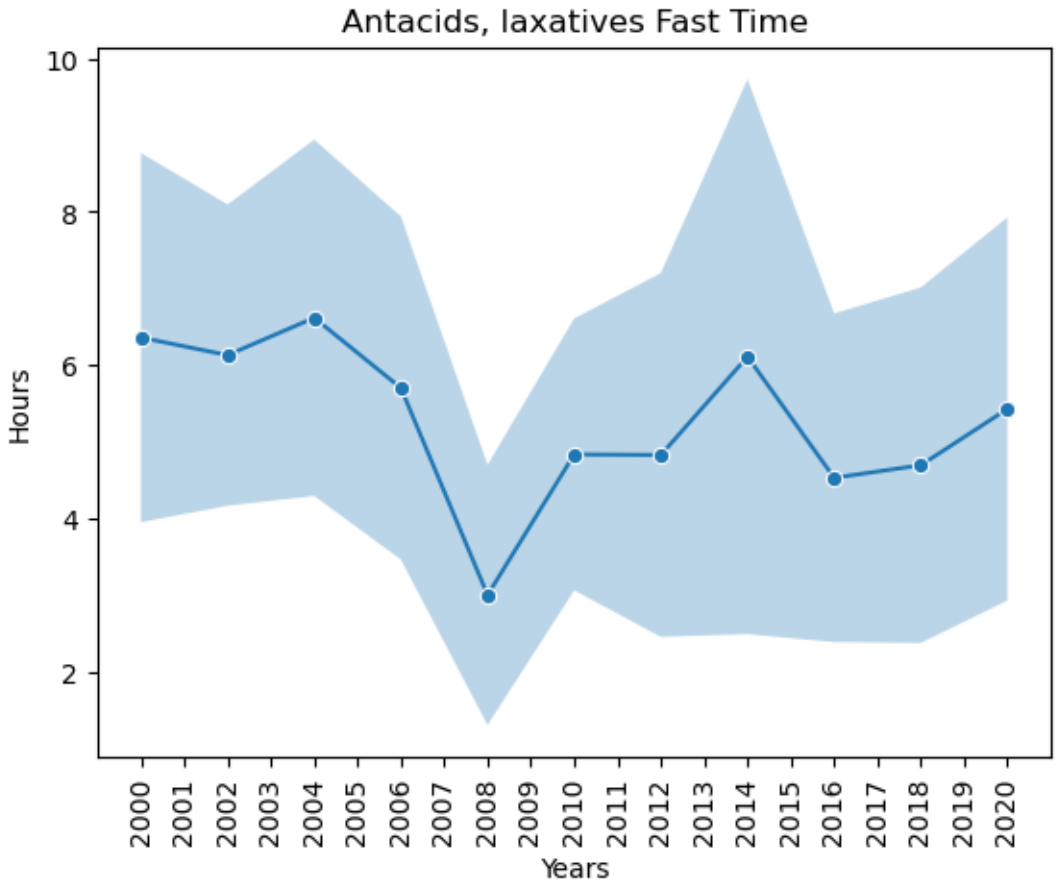
Unlike the hourly fasting information, the fasting of Gum/Mints/Cough peaks at 30 minutes in 2002, and then decreases to 15 minutes until 2008, and is slowly increasing.

[18]:
```
sns.lineplot(data=total_df.groupby('year')['PHAANTHR'].
  ↪describe(),x='year',y='mean',marker="o")
upper = total_df.groupby('year')['PHAANTHR'].describe()['mean'] + total_df.
  ↪groupby('year')['PHAANTHR'].describe()['std']/2
lower = total_df.groupby('year')['PHAANTHR'].describe()['mean'] - total_df.
  ↪groupby('year')['PHAANTHR'].describe()['std']/2
plt.fill_between(total_df.groupby('year')['PHAANTHR'].describe().index,
```

```
                lower,
                upper,
                alpha=.3)

plt.title('Antacids, laxatives Fast Time')
plt.ylabel('Hours')
plt.xlabel('Years')
plt.xticks(np.arange(2000,2021,step=1), rotation=90)
plt.show()
```

```
C:\Users\brian\Anaconda3\Lib\site-packages\seaborn\_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed in a
future version. Convert inf values to NaN before operating instead.
  with pd.option_context('mode.use_inf_as_na', True):
C:\Users\brian\Anaconda3\Lib\site-packages\seaborn\_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed in a
future version. Convert inf values to NaN before operating instead.
  with pd.option_context('mode.use_inf_as_na', True):
```

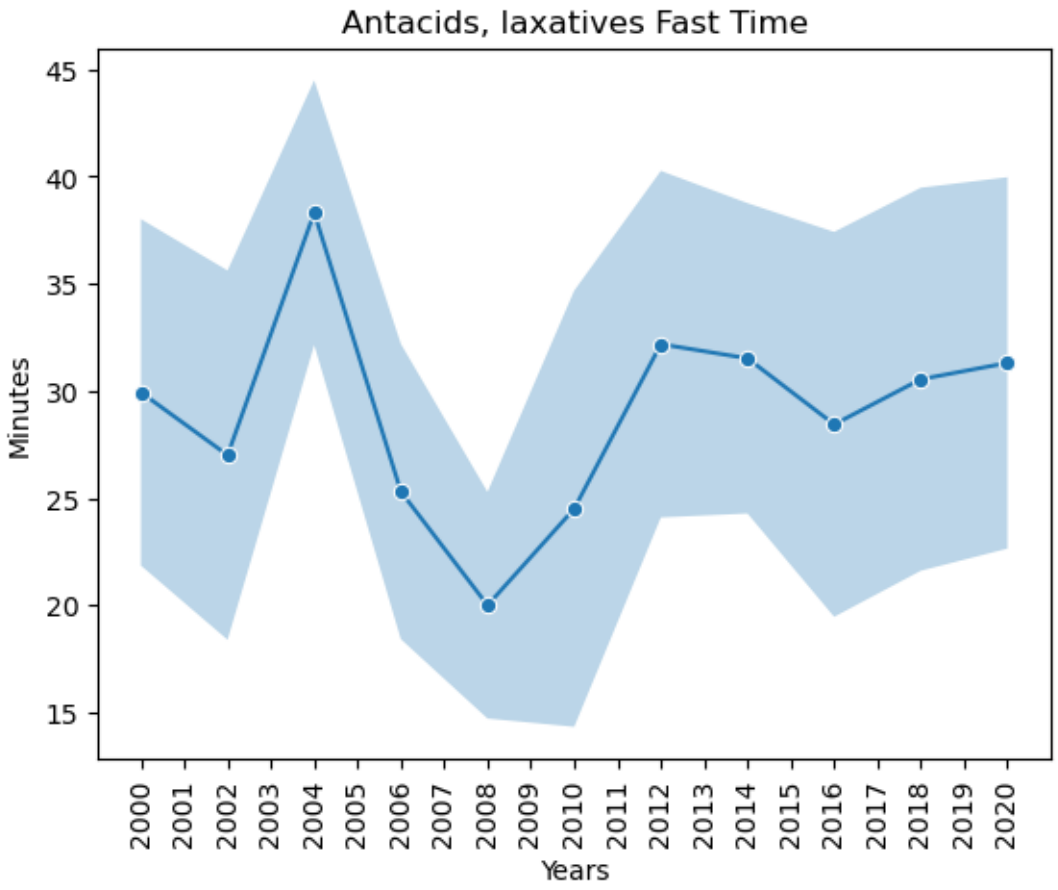


Antacids and laxatives are pretty stable at the average of 5 hours, with a dip at 3 hours in 2008.

[19]:
```
sns.lineplot(data=total_df.groupby('year')['PHAANTMN'].
 ↪describe(),x='year',y='mean',marker="o")
upper = total_df.groupby('year')['PHAANTMN'].describe()['mean'] + total_df.
 ↪groupby('year')['PHAANTMN'].describe()['std']/2
lower = total_df.groupby('year')['PHAANTMN'].describe()['mean'] - total_df.
 ↪groupby('year')['PHAANTMN'].describe()['std']/2
plt.fill_between(total_df.groupby('year')['PHAANTMN'].describe().index,
                 lower,
                 upper,
                 alpha=.3)

plt.title('Antacids, laxatives Fast Time')
plt.ylabel('Minutes')
plt.xlabel('Years')
plt.xticks(np.arange(2000,2021,step=1), rotation=90)
plt.show()
```

```
C:\Users\brian\Anaconda3\Lib\site-packages\seaborn\_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed in a
future version. Convert inf values to NaN before operating instead.
  with pd.option_context('mode.use_inf_as_na', True):
C:\Users\brian\Anaconda3\Lib\site-packages\seaborn\_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed in a
future version. Convert inf values to NaN before operating instead.
  with pd.option_context('mode.use_inf_as_na', True):
```

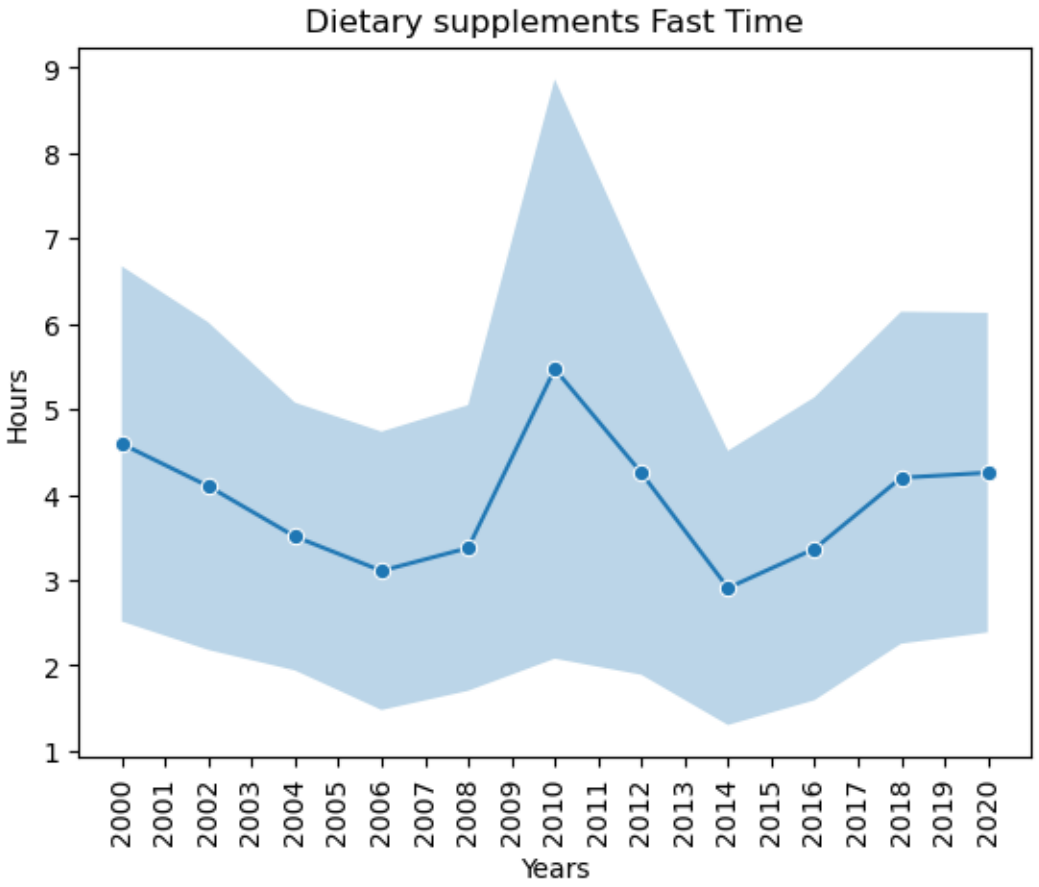


Unlike the hourly fasting data, the antacids and laxatives fasting time in minutes peak at around 40 minutes to decrease to 20 minutes in 2008, and increasing steeply.

[61]:
```
sns.lineplot(data=total_df.groupby('year')['PHASUPHR'].
 ↪describe(),x='year',y='mean',marker="o")
upper = total_df.groupby('year')['PHASUPHR'].describe()['mean'] + total_df.
 ↪groupby('year')['PHASUPHR'].describe()['std']/2
lower = total_df.groupby('year')['PHASUPHR'].describe()['mean'] - total_df.
 ↪groupby('year')['PHASUPHR'].describe()['std']/2
plt.fill_between(total_df.groupby('year')['PHASUPHR'].describe().index,
                 lower,
                 upper,
                 alpha=.3)

plt.title('Dietary supplements Fast Time')
plt.ylabel('Hours')
plt.xlabel('Years')
plt.xticks(np.arange(2000,2021,step=1), rotation=90)
plt.show()
```

```
C:\Users\brian\Anaconda3\Lib\site-packages\seaborn\_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed in a
future version. Convert inf values to NaN before operating instead.
  with pd.option_context('mode.use_inf_as_na', True):
C:\Users\brian\Anaconda3\Lib\site-packages\seaborn\_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed in a
future version. Convert inf values to NaN before operating instead.
  with pd.option_context('mode.use_inf_as_na', True):
```



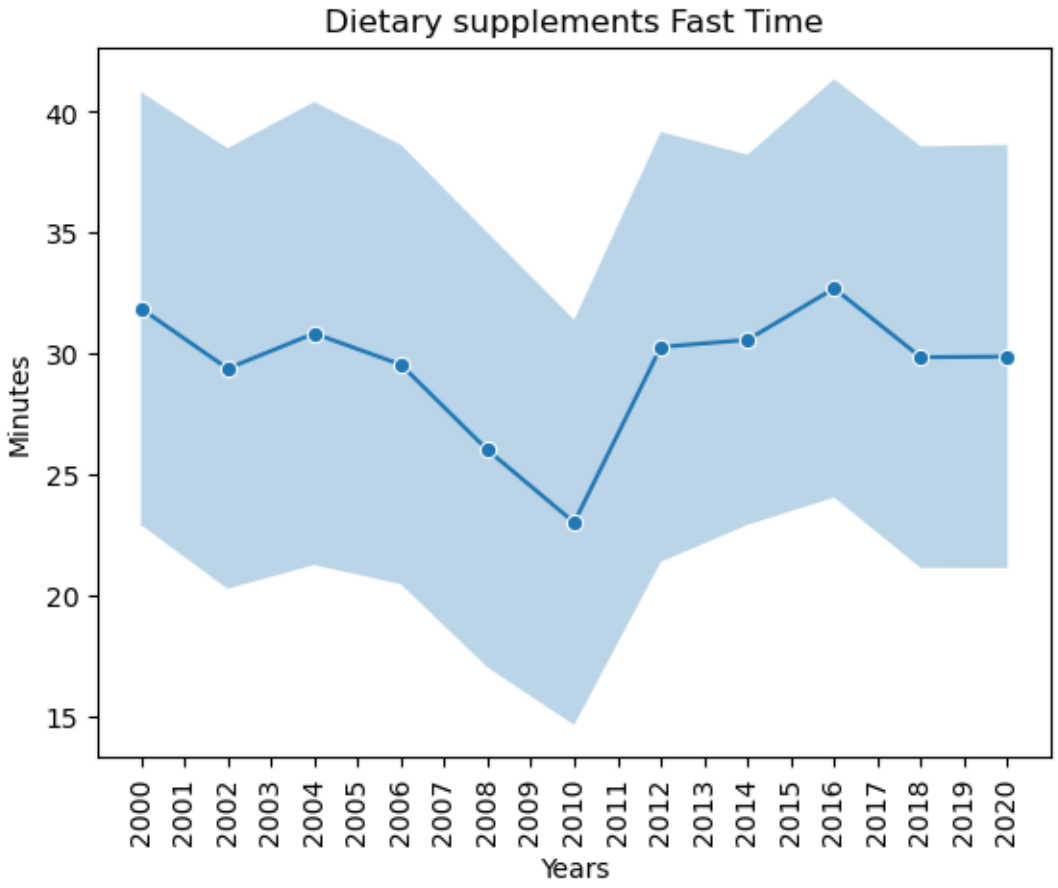
The dietary supplemenets are pretty stable at around 4 hours except the sharp peak at 5.5 hours in 2010.

[62]:
```
sns.lineplot(data=total_df.groupby('year')['PHASUPMN'].
  ↪describe(),x='year',y='mean',marker="o")
upper = total_df.groupby('year')['PHASUPMN'].describe()['mean'] + total_df.
  ↪groupby('year')['PHASUPMN'].describe()['std']/2
lower = total_df.groupby('year')['PHASUPMN'].describe()['mean'] - total_df.
  ↪groupby('year')['PHASUPMN'].describe()['std']/2
plt.fill_between(total_df.groupby('year')['PHASUPMN'].describe().index,
```

```
                lower,
                upper,
                alpha=.3)

plt.title('Dietary supplements Fast Time')
plt.ylabel('Minutes')
plt.xlabel('Years')
plt.xticks(np.arange(2000,2021,step=1), rotation=90)
plt.show()
```

```
C:\Users\brian\Anaconda3\Lib\site-packages\seaborn\_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed in a
future version. Convert inf values to NaN before operating instead.
  with pd.option_context('mode.use_inf_as_na', True):
C:\Users\brian\Anaconda3\Lib\site-packages\seaborn\_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed in a
future version. Convert inf values to NaN before operating instead.
  with pd.option_context('mode.use_inf_as_na', True):
```

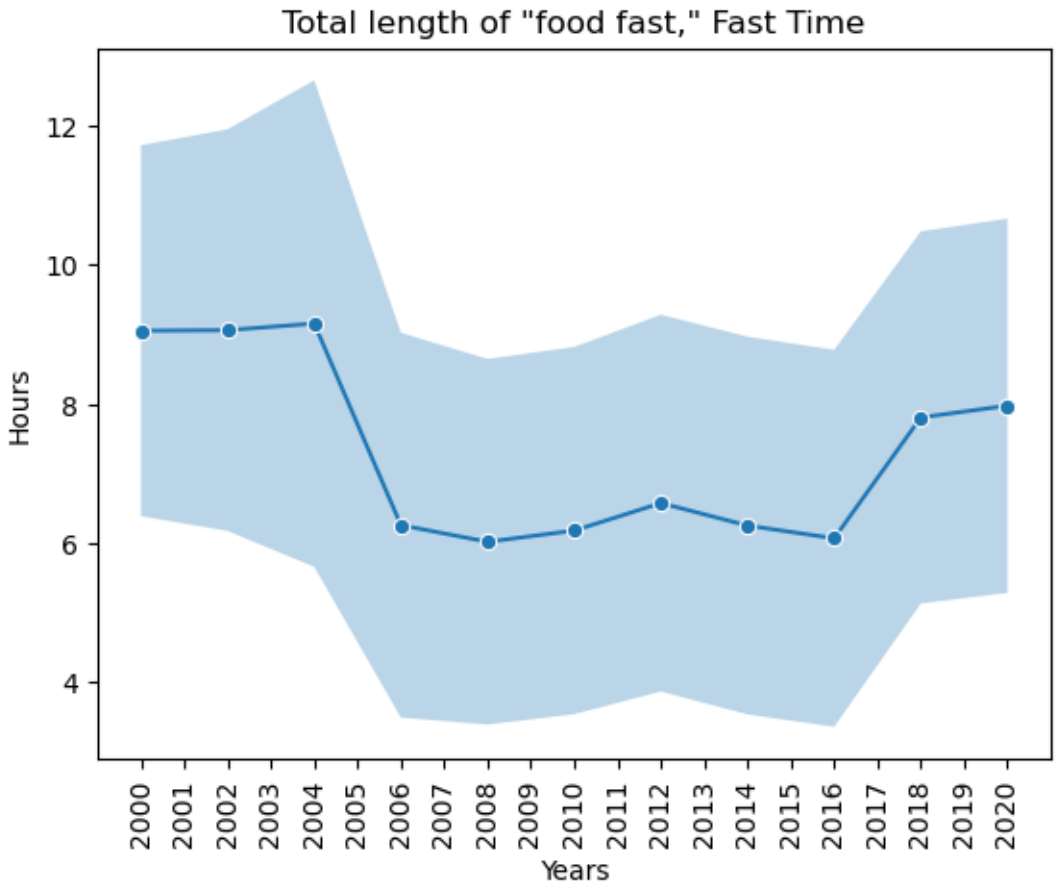


Unlike the data above, the minutely data shows a sharp dip at 2010. Other than that, it is pretty stable at around 30 min.

[64]:
```
sns.lineplot(data=total_df.groupby('year')['PHAFSTHR'].
 ↪describe(),x='year',y='mean',marker="o")
upper = total_df.groupby('year')['PHAFSTHR'].describe()['mean'] + total_df.
 ↪groupby('year')['PHAFSTHR'].describe()['std']/2
lower = total_df.groupby('year')['PHAFSTHR'].describe()['mean'] - total_df.
 ↪groupby('year')['PHAFSTHR'].describe()['std']/2
plt.fill_between(total_df.groupby('year')['PHAFSTHR'].describe().index,
                 lower,
                 upper,
                 alpha=.3)

plt.title('Total length of "food fast," Fast Time')
plt.ylabel('Hours')
plt.xlabel('Years')
plt.xticks(np.arange(2000,2021,step=1), rotation=90)
plt.show()
```

```
C:\Users\brian\Anaconda3\Lib\site-packages\seaborn\_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed in a
future version. Convert inf values to NaN before operating instead.
  with pd.option_context('mode.use_inf_as_na', True):
C:\Users\brian\Anaconda3\Lib\site-packages\seaborn\_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed in a
future version. Convert inf values to NaN before operating instead.
  with pd.option_context('mode.use_inf_as_na', True):
```

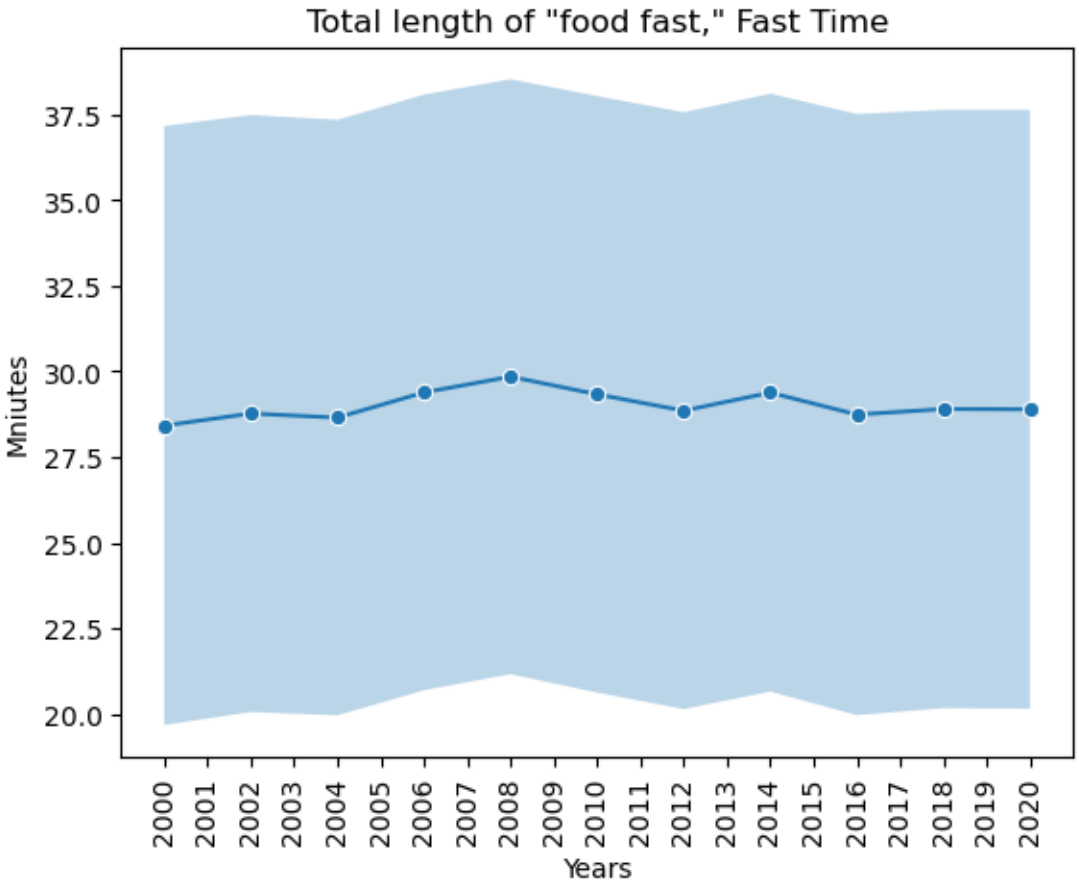


Interestingly, the food fasting time is stable at 9 hours until the decrease in 2006, which stays stable again at 6 hours, and then is increasing after 2016.

[65]:
```
sns.lineplot(data=total_df.groupby('year')['PHAFSTMN'].
 ↪describe(),x='year',y='mean',marker="o")
upper = total_df.groupby('year')['PHAFSTMN'].describe()['mean'] + total_df.
 ↪groupby('year')['PHAFSTMN'].describe()['std']/2
lower = total_df.groupby('year')['PHAFSTMN'].describe()['mean'] - total_df.
 ↪groupby('year')['PHAFSTMN'].describe()['std']/2
plt.fill_between(total_df.groupby('year')['PHAFSTMN'].describe().index,
                 lower,
                 upper,
                 alpha=.3)

plt.title('Total length of "food fast," Fast Time')
plt.ylabel('Mniutes')
plt.xlabel('Years')
plt.xticks(np.arange(2000,2021,step=1), rotation=90)
plt.show()
```

```
C:\Users\brian\Anaconda3\Lib\site-packages\seaborn\_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed in a
future version. Convert inf values to NaN before operating instead.
  with pd.option_context('mode.use_inf_as_na', True):
C:\Users\brian\Anaconda3\Lib\site-packages\seaborn\_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed in a
future version. Convert inf values to NaN before operating instead.
  with pd.option_context('mode.use_inf_as_na', True):
```



Interestingly again, the food fasting time is very stable at around 28 minutes without a dip or sharp peak.

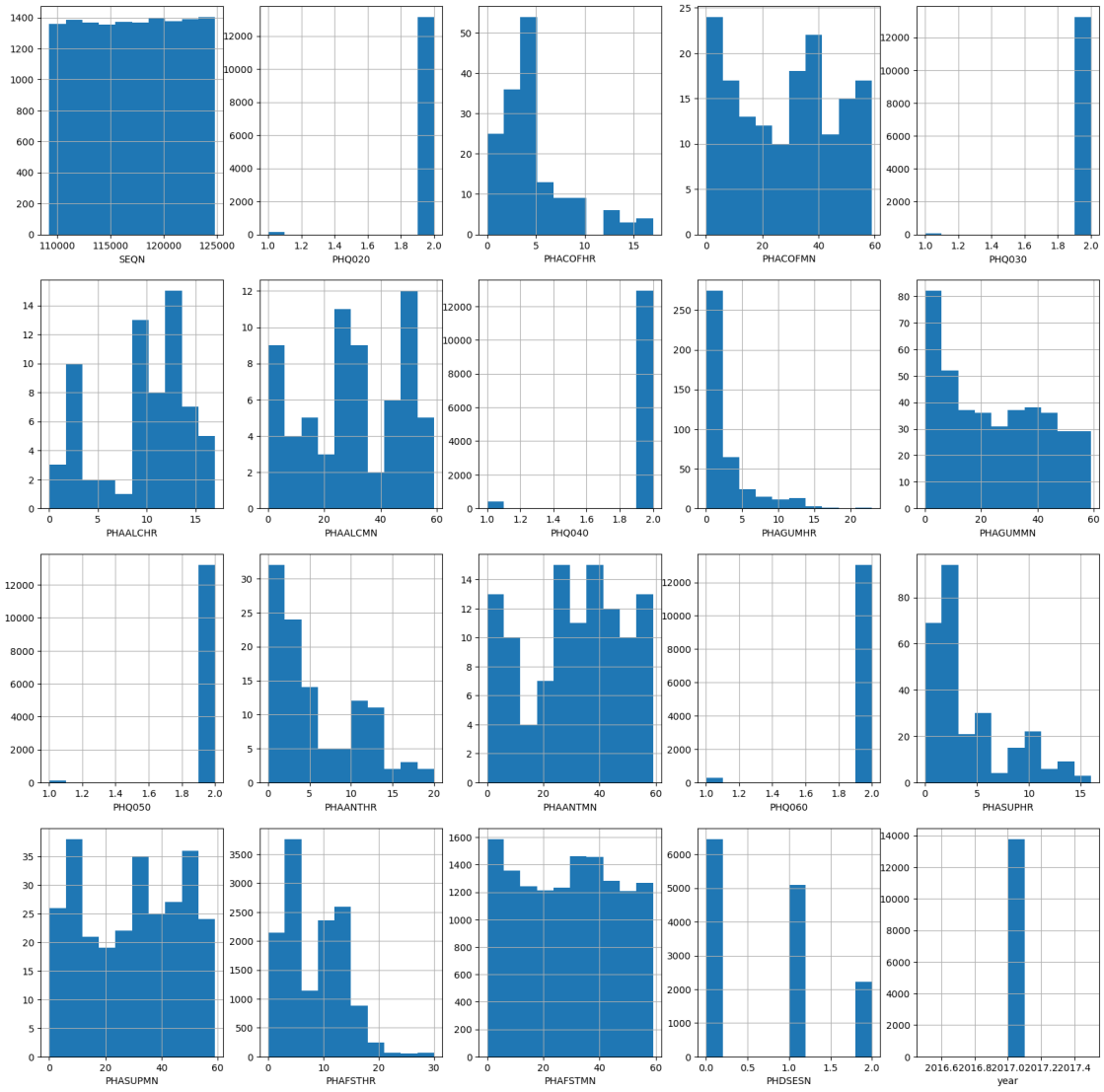
[7]:
```
np.sort(total_df['year'].unique())
```

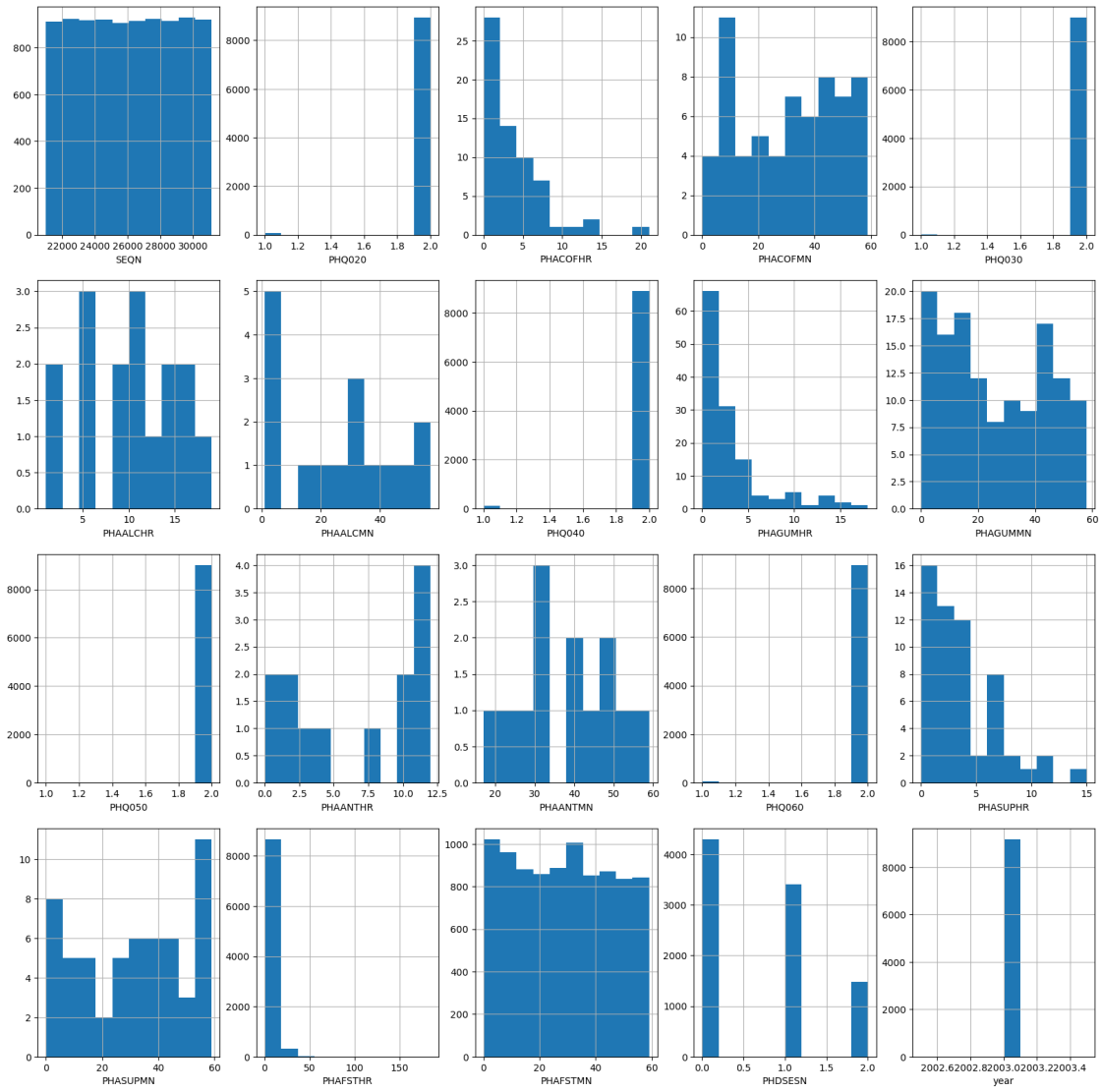
[7]:
```
array([1999, 2001, 2003, 2005, 2007, 2009, 2011, 2013, 2015, 2017],
      dtype=int64)
```

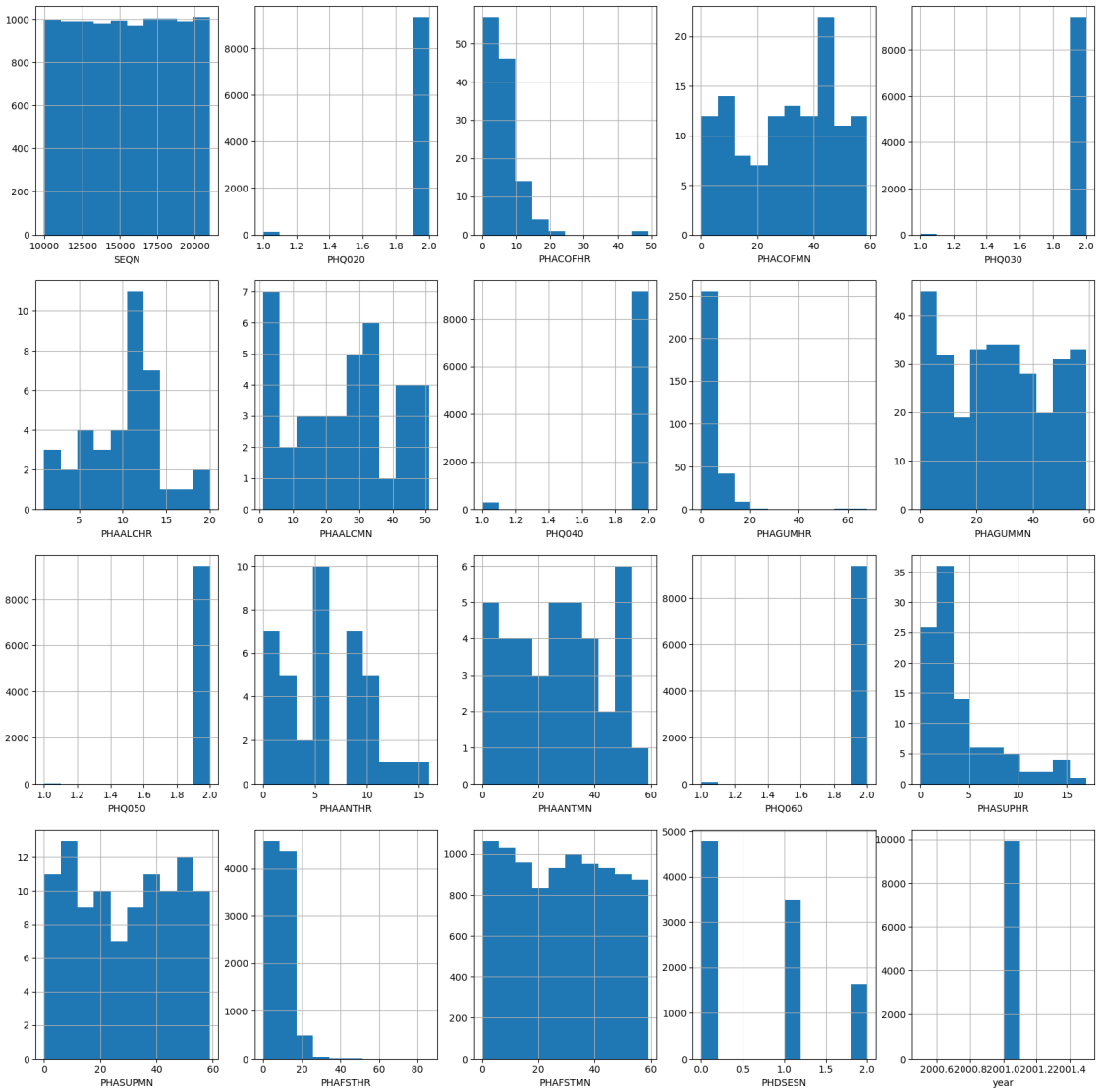
[8]:
```
total_df[total_df['year']==1999]['PHAALCHR'].describe()
```

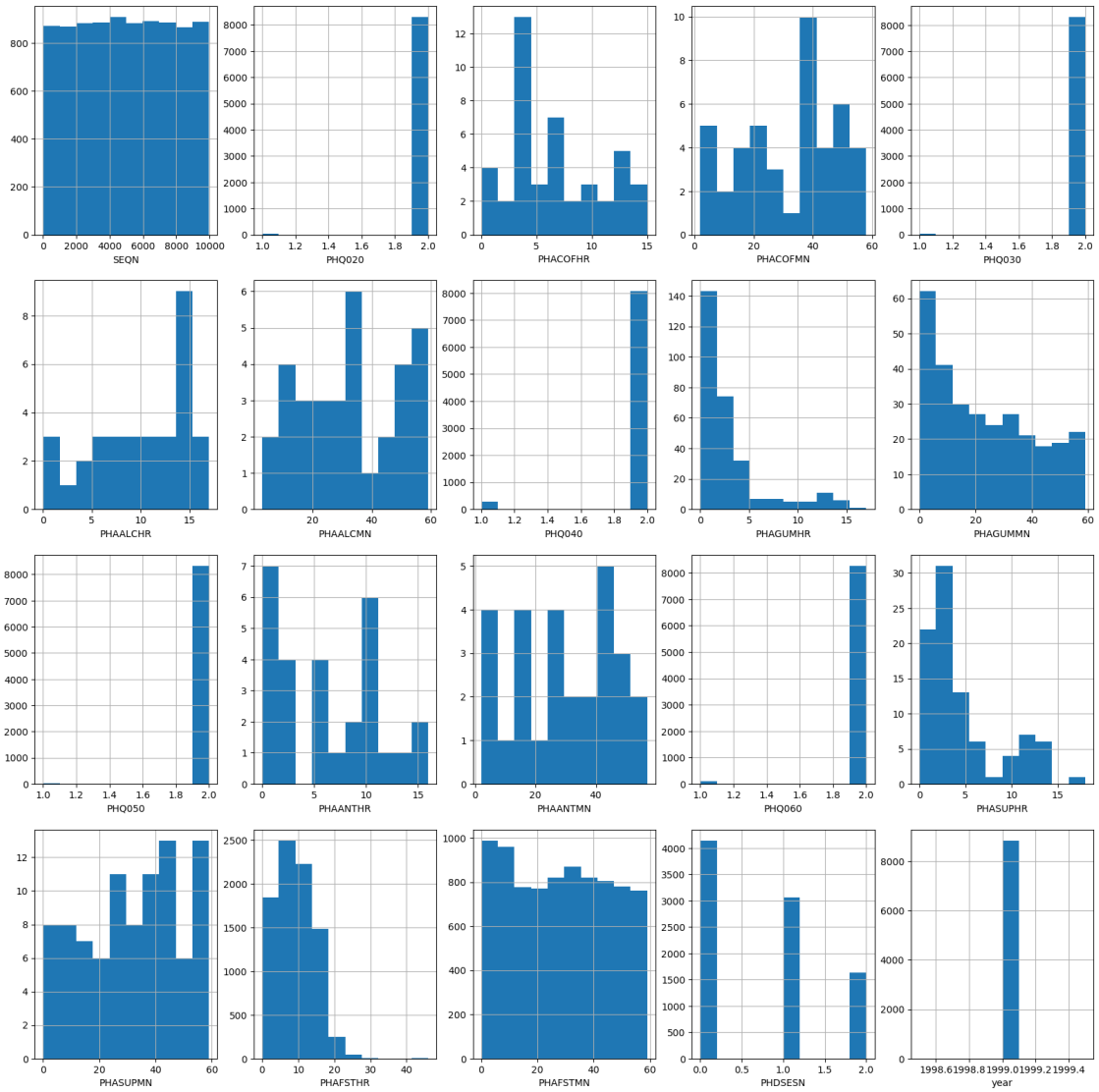
[8]:
```
count    3.300000e+01
mean     1.003030e+01
std      4.959113e+00
min      5.397605e-79
25%      6.000000e+00
50%      1.100000e+01
75%      1.400000e+01
max      1.700000e+01
Name: PHAALCHR, dtype: float64
```

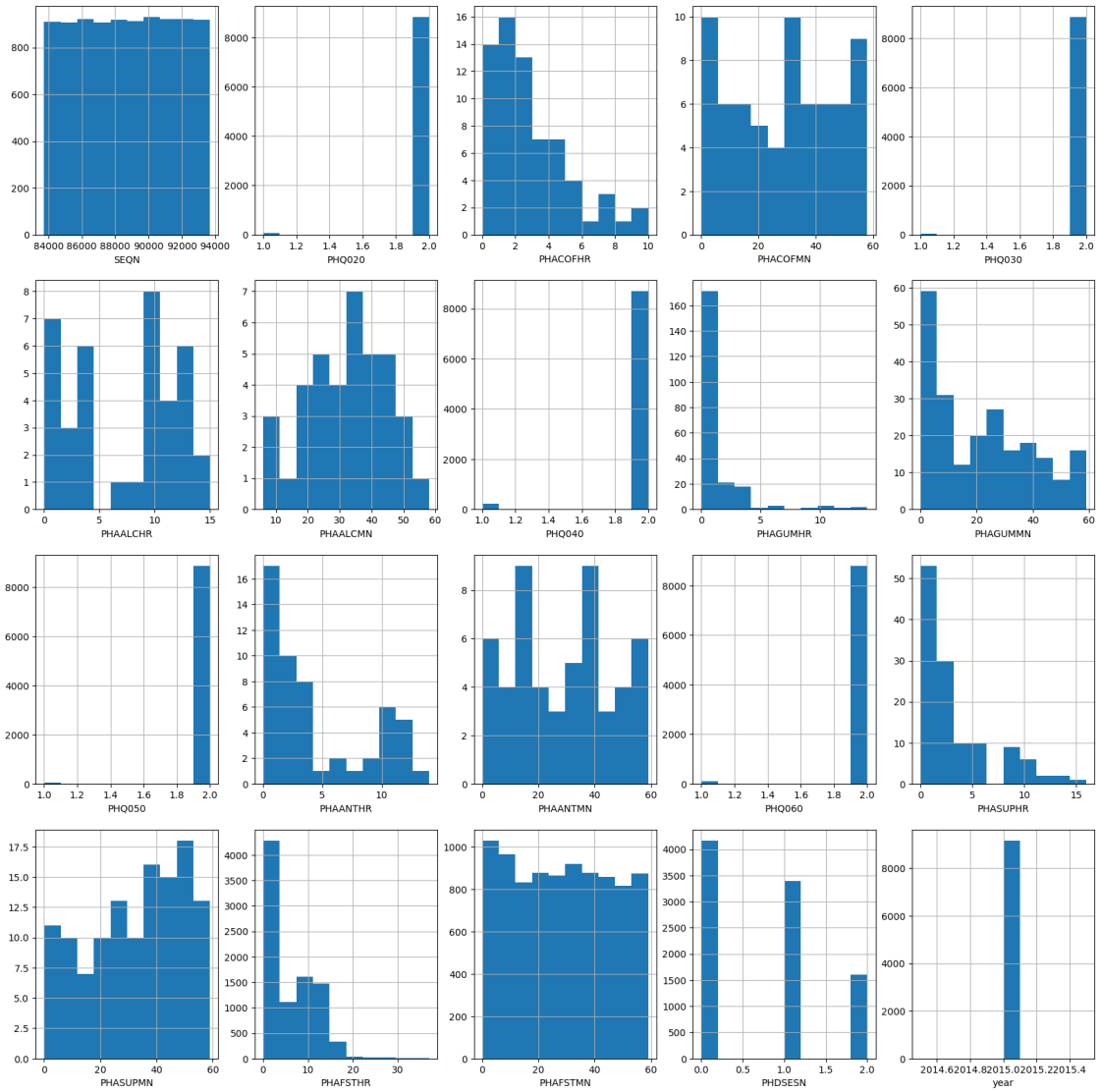
[9]:
```
for i in range(len(total_df['year'].unique())):
    fig, ax = plt.subplots(4,int(len(total_df.columns)/4),figsize=(20,20))
    for k in range(len(total_df.columns)):
        if k < 5:
            total_df[total_df['year']==total_df['year'].unique()[i]][total_df.
 ↪columns[k]].hist(ax=ax[0,k])
            ax[0,k].set_xlabel(total_df.columns[k])
        elif 5 <= k < 10:
            total_df[total_df['year']==total_df['year'].unique()[i]][total_df.
 ↪columns[k]].hist(ax=ax[1,k-5])
            ax[1,k-5].set_xlabel(total_df.columns[k])
        elif 10 <= k < 15:
            total_df[total_df['year']==total_df['year'].unique()[i]][total_df.
 ↪columns[k]].hist(ax=ax[2,k-10])
            ax[2,k-10].set_xlabel(total_df.columns[k])
        elif 15 <= k < 20:
            total_df[total_df['year']==total_df['year'].unique()[i]][total_df.
 ↪columns[k]].hist(ax=ax[3,k-15])
            ax[3,k-15].set_xlabel(total_df.columns[k])
    fig.suptitle(total_df['year'].unique()[i])
    plt.show()
```

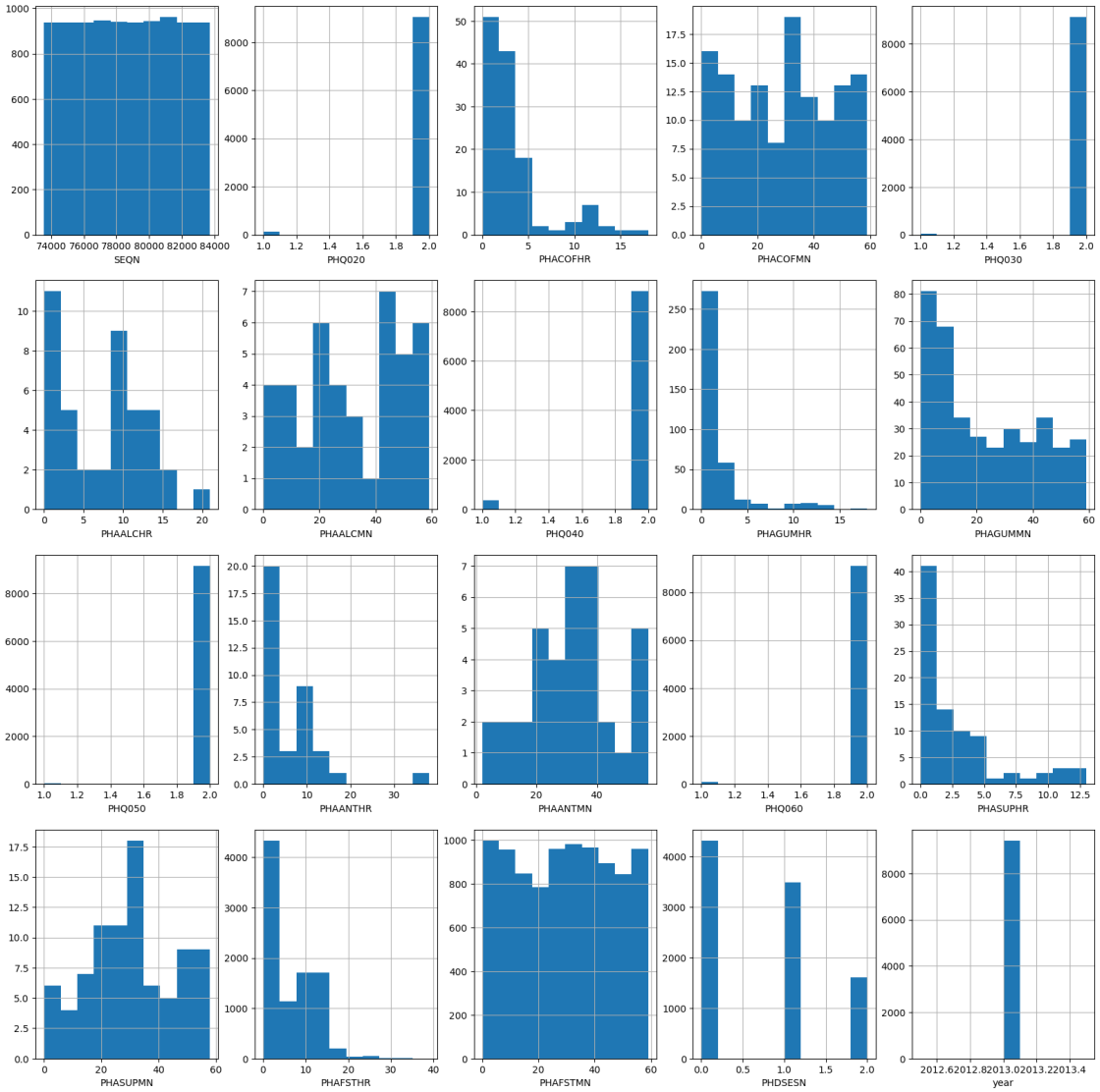

SEQN
PHQ020
PHACOFHR
PHACOFMN
PHQ030
PHAALCHR
PHAALCMN
PHQ040
PHAGUMHR
PHAGUMMN
PHQ050
PHAANTHR
PHAANTMN
PHQ060
PHASUPHR
PHASUPMN
PHAFSTHR
PHAFSTMN
PHDSESN
year

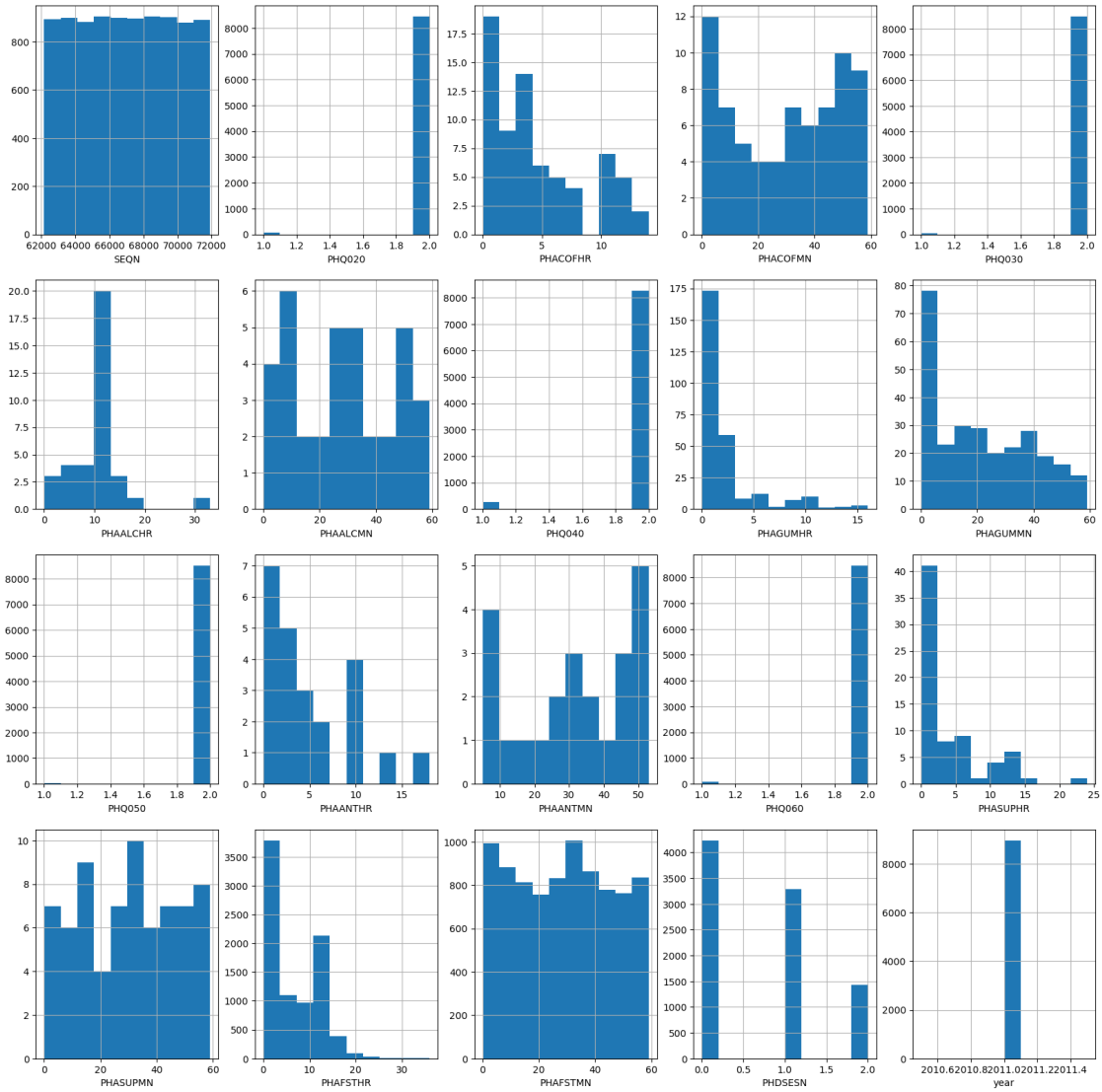

SEQN
PHQ020
PHACOFHR
PHACOFMN
PHQ030
PHAALCHR
PHAALCMN
PHQ040
PHAGUMHR
PHAGUMMN
PHQ050
PHAANTHR
PHAANTMN
PHQ060
PHASUPHR
PHASUPMN
PHAFSTHR
PHAFSTMN
PHDSESN
year

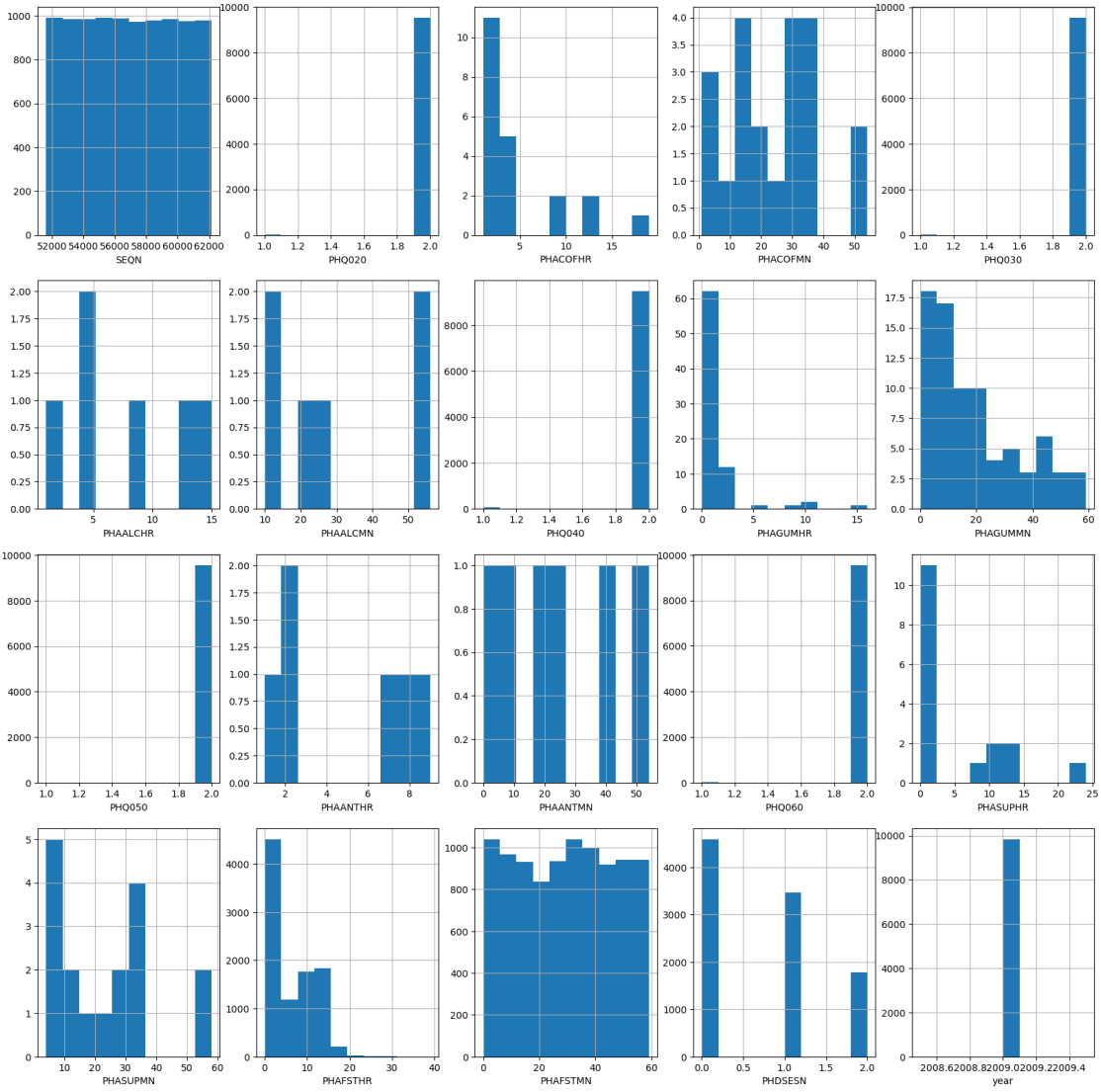

SEQN
PHQ020
PHACOFHR
PHACOFMN
PHQ030
PHAALCHR
PHAALCMN
PHQ040
PHAGUMHR
PHAGUMMN
PHQ050
PHAANTHR
PHAANTMN
PHQ060
PHASUPHR
PHASUPMN
PHAFSTHR
PHAFSTMN
PHDSESN
year

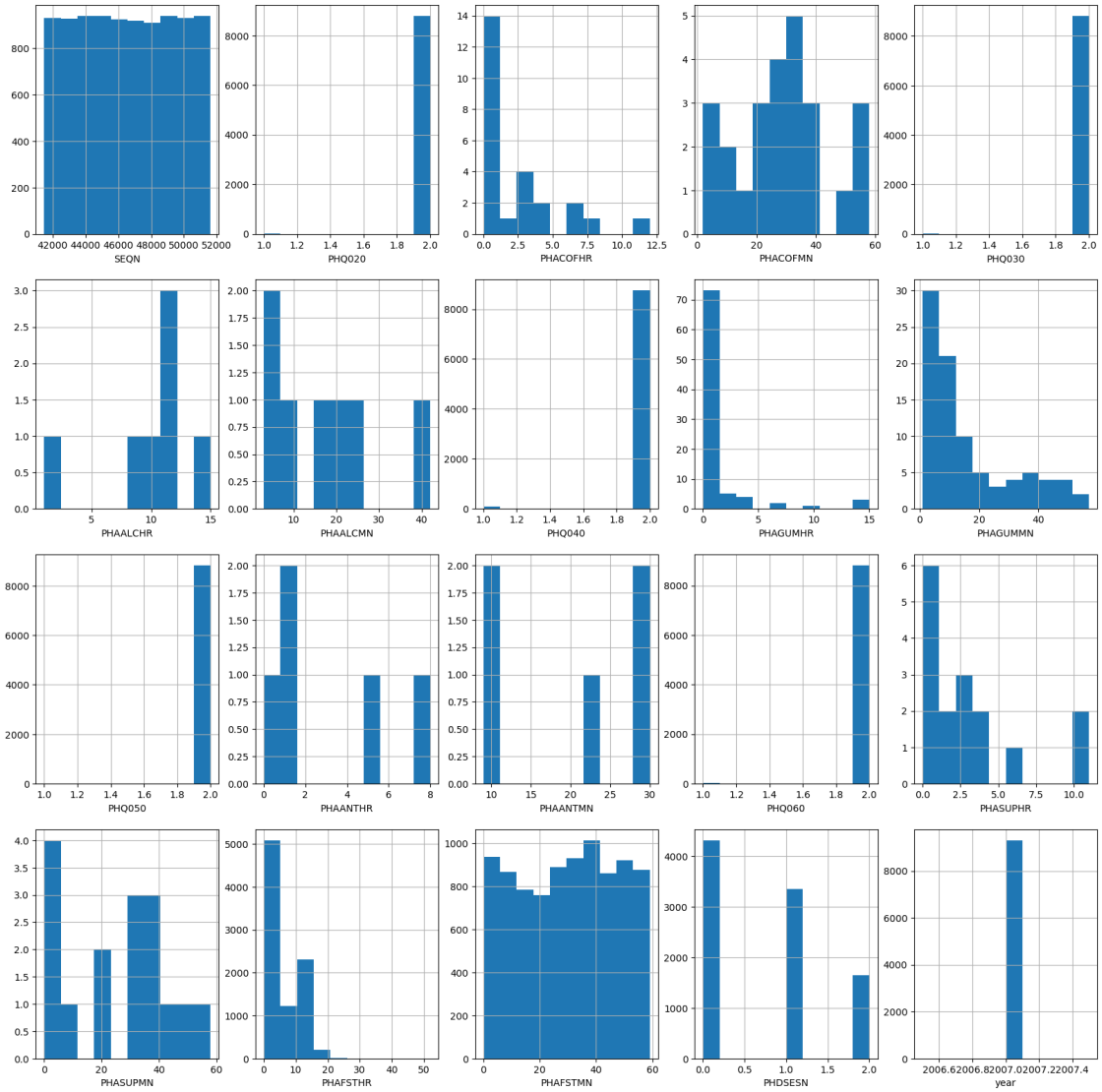

SEQN
PHQ020
PHACOFHR
PHACOFMN
PHQ030
PHAALCHR
PHAALCMN
PHQ040
PHAGUMHR
PHAGUMMN
PHQ050
PHAANTHR
PHAANTMN
PHQ060
PHASUPHR
PHASUPMN
PHAFSTHR
PHAFSTMN
PHDSESN
year


SEQN
PHQ020
PHACOFHR
PHACOFMN
PHQ030
PHAALCHR
PHAALCMN
PHQ040
PHAGUMHR
PHAGUMMN
PHQ050
PHAANTHR
PHAANTMN
PHQ060
PHASUPHR
PHASUPMN
PHAFSTHR
PHAFSTMN
PHDSESN
year

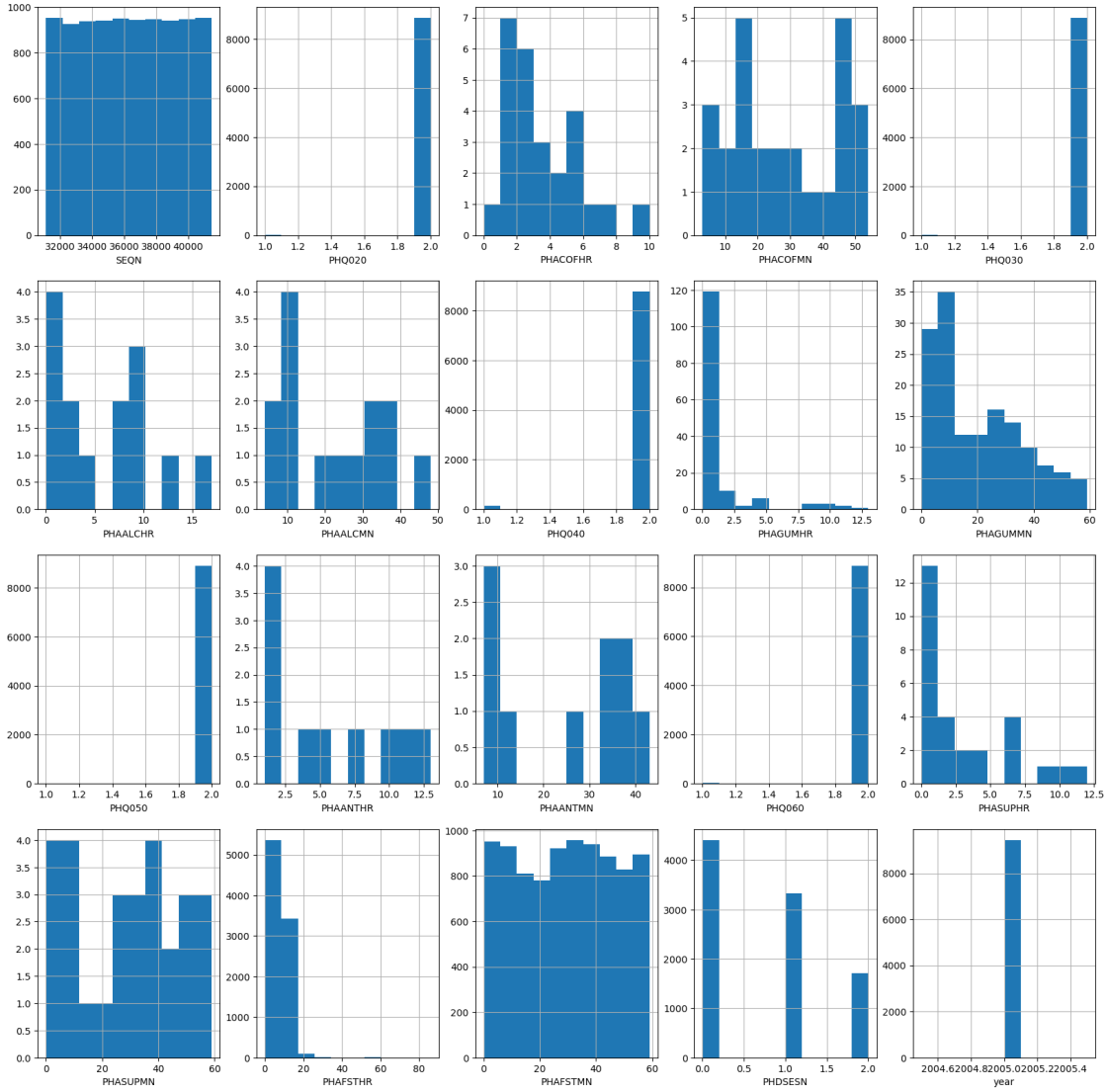

SEQN
PHQ020
PHACOFHR
PHACOFMN
PHQ030
PHAALCHR
PHAALCMN
PHQ040
PHAGUMHR
PHAGUMMN
PHQ050
PHAANTHR
PHAANTMN
PHQ060
PHASUPHR
PHASUPMN
PHAFSTHR
PHAFSTMN
PHDSESN
year


SEQN
PHQ020
PHACOFHR
PHACOFMN
PHQ030
PHAALCHR
PHAALCMN
PHQ040
PHAGUMHR
PHAGUMMN
PHQ050
PHAANTHR
PHAANTMN
PHQ060
PHASUPHR
PHASUPMN
PHAFSTHR
PHAFSTMN
PHDSESN
year


SEQN
PHQ020
PHACOFHR
PHACOFMN
PHQ030
PHAALCHR
PHAALCMN
PHQ040
PHAGUMHR
PHAGUMMN
PHQ050
PHAANTHR
PHAANTMN
PHQ060
PHASUPHR
PHASUPMN
PHAFSTHR
PHAFSTMN
PHDSESN
year


SEQN
PHQ020
PHACOFHR
PHACOFMN
PHQ030
PHAALCHR
PHAALCMN
PHQ040
PHAGUMHR
PHAGUMMN
PHQ050
PHAANTHR
PHAANTMN
PHQ060
PHASUPHR
PHASUPMN
PHAFSTHR
PHAFSTMN
PHDSESN
year

2005



[10]:
```
total_df.columns
```

[10]:
```
Index(['SEQN', 'PHQ020', 'PHACOFHR', 'PHACOFMN', 'PHQ030', 'PHAALCHR',
       'PHAALCMN', 'PHQ040', 'PHAGUMHR', 'PHAGUMMN', 'PHQ050', 'PHAANTHR',
       'PHAANTMN', 'PHQ060', 'PHASUPHR', 'PHASUPMN', 'PHAFSTHR', 'PHAFSTMN',
       'PHDSESN', 'year'],
      dtype='object')
```

[11]:
```
np.sort(total_df['year'].unique())
```

[11]:
```
array([1999, 2001, 2003, 2005, 2007, 2009, 2011, 2013, 2015, 2017],
      dtype=int64)
```

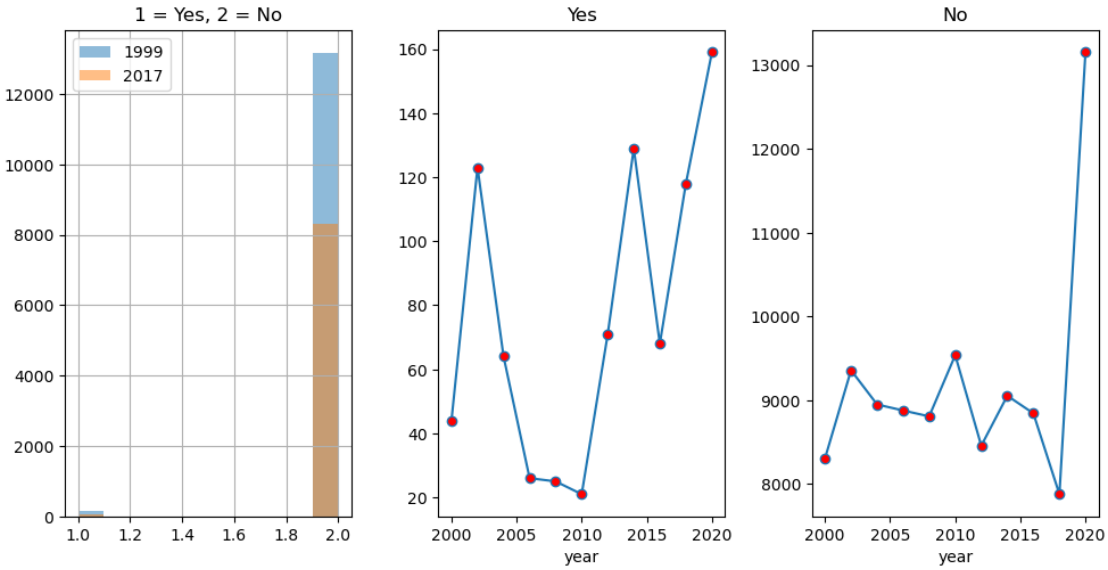
2 Coffee or tea with cream or sugar? [Include milk or non-dairy creamers.]

2.0.1 Have you had any of the following since the fasting time in variable PHQ010. PHQ010: When was the last time you ate or drank anything other than plain water? [Do not include diet soda, black coffee or tea with saccharine or Equal.] (Note: Variable PHQ010 was not released as part of this data set because of disclosure issues)

[69]:
```
fig, ax = plt.subplots(1,3,figsize=(10,5))

total_df[total_df['year']==2020]['PHQ020'].hist(alpha=0.5,ax=ax[0])
total_df[total_df['year']==2000]['PHQ020'].hist(alpha=0.5,ax=ax[0])

total_df[total_df['PHQ020']==1].groupby(['year']).count()['PHQ020'].
 ↪plot(ax=ax[1],marker='o',mfc = 'r')
total_df[total_df['PHQ020']==2].groupby(['year']).count()['PHQ020'].
 ↪plot(ax=ax[2],marker='o',mfc = 'r')
plt.tight_layout()
ax[0].legend(['1999','2017'])
ax[0].set_title('1 = Yes, 2 = No')
ax[1].set_title('Yes')
ax[2].set_title('No')
plt.show()
```



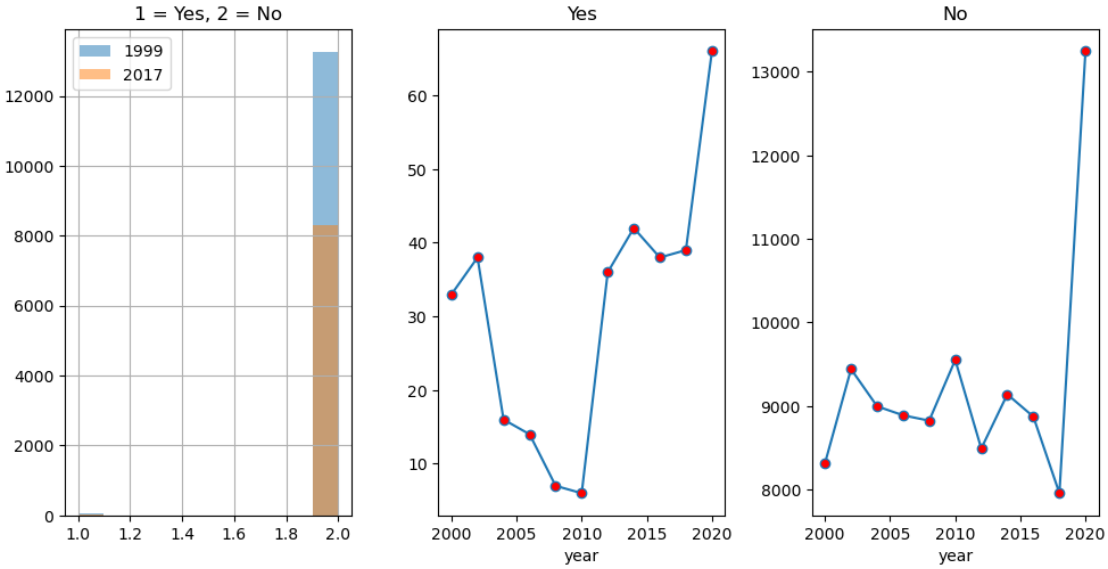
3 Alcohol, such as beer, wine, or liquor?

3.0.1 Have you had any of the following since the fasting time in variable PHQ010. PHQ010: When was the last time you ate or drank anything other than plain water? [Do not include diet soda, black coffee or tea with saccharine or Equal.] (Note: Variable PHQ010 was not released as part of this data set because of disclosure issues)

[68]:
```
fig, ax = plt.subplots(1,3,figsize=(10,5))

total_df[total_df['year']==2020]['PHQ030'].hist(alpha=0.5,ax=ax[0])
total_df[total_df['year']==2000]['PHQ030'].hist(alpha=0.5,ax=ax[0])

total_df[total_df['PHQ030']==1].groupby(['year']).count()['PHQ030'].
 ↪plot(ax=ax[1],marker='o',mfc = 'r')
total_df[total_df['PHQ030']==2].groupby(['year']).count()['PHQ030'].
 ↪plot(ax=ax[2],marker='o',mfc = 'r')
plt.tight_layout()
ax[0].legend(['1999','2017'])
ax[0].set_title('1 = Yes, 2 = No')
ax[1].set_title('Yes')
ax[2].set_title('No')
plt.show()
```



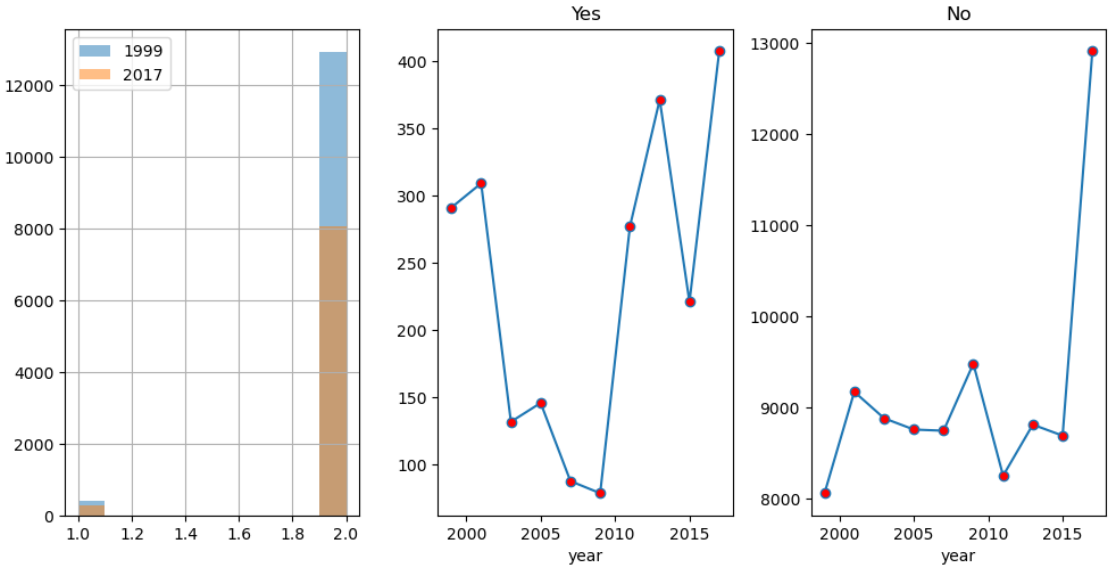
4 Gum, breath mints, lozenges or cough drops, or other cough or cold remedies?

4.0.1 Have you had any of the following since the fasting time in variable PHQ010. PHQ010: When was the last time you ate or drank anything other than plain water? [Do not include diet soda, black coffee or tea with saccharine or Equal.] (Note: Variable PHQ010 was not released as part of this data set because of disclosure issues)

[13]:
```
fig, ax = plt.subplots(1,3,figsize=(10,5))

total_df[total_df['year']==2017]['PHQ040'].hist(alpha=0.5,ax=ax[0])
total_df[total_df['year']==1999]['PHQ040'].hist(alpha=0.5,ax=ax[0])

total_df[total_df['PHQ040']==1].groupby(['year']).count()['PHQ040'].
 ↪plot(ax=ax[1],marker='o',mfc = 'r')
total_df[total_df['PHQ040']==2].groupby(['year']).count()['PHQ040'].
 ↪plot(ax=ax[2],marker='o',mfc = 'r')
plt.tight_layout()
ax[0].legend(['1999','2017'])
ax[1].set_title('Yes')
ax[2].set_title('No')
plt.show()
```



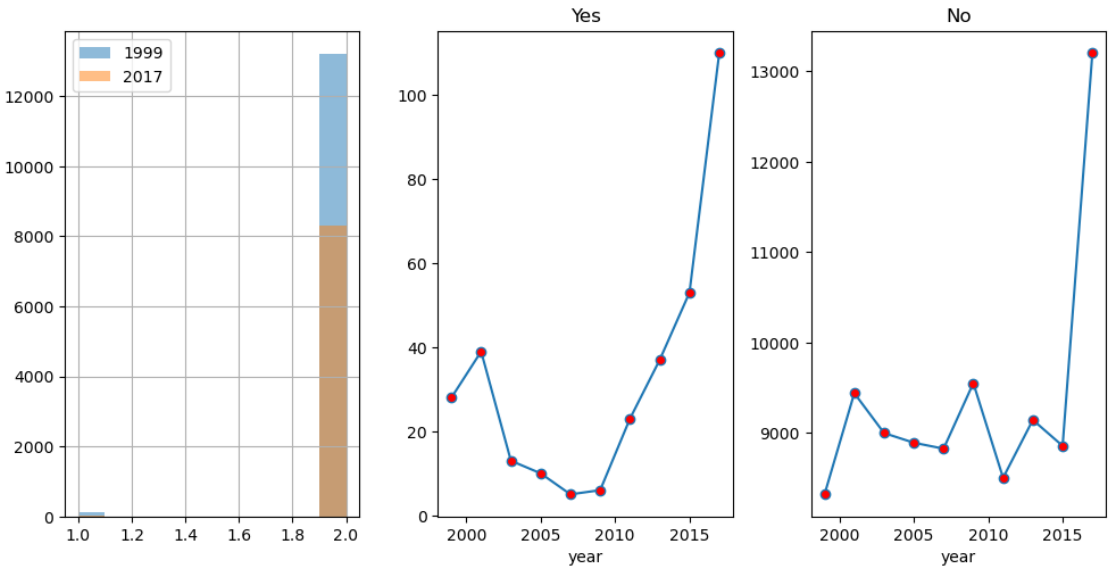
5 Antacids, laxatives, or anti-diarrheals?

5.0.1 Have you had any of the following since the fasting time in variable PHQ010. PHQ010: When was the last time you ate or drank anything other than plain water? [Do not include diet soda, black coffee or tea with saccharine or Equal.] (Note: Variable PHQ010 was not released as part of this data set because of disclosure issues)

[14]:
```
fig, ax = plt.subplots(1,3,figsize=(10,5))

total_df[total_df['year']==2017]['PHQ050'].hist(alpha=0.5,ax=ax[0])
total_df[total_df['year']==1999]['PHQ050'].hist(alpha=0.5,ax=ax[0])

total_df[total_df['PHQ050']==1].groupby(['year']).count()['PHQ050'].
 ↪plot(ax=ax[1],marker='o',mfc = 'r')
total_df[total_df['PHQ050']==2].groupby(['year']).count()['PHQ050'].
 ↪plot(ax=ax[2],marker='o',mfc = 'r')
plt.tight_layout()
ax[0].legend(['1999','2017'])
ax[1].set_title('Yes')
ax[2].set_title('No')
plt.show()
```



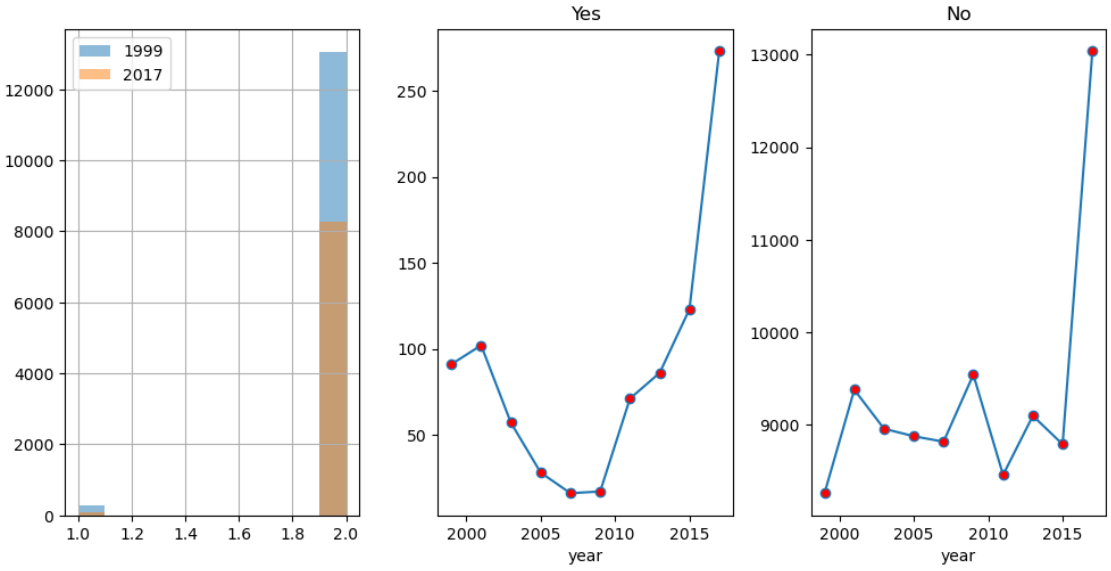
6 Dietary supplements such as vitamins and minerals? [Include multivitamins and single nutrient supplements.]

6.0.1 Have you had any of the following since the fasting time in variable PHQ010. PHQ010: When was the last time you ate or drank anything other than plain water? [Do not include diet soda, black coffee or tea with saccharine or Equal.] (Note: Variable PHQ010 was not released as part of this data set because of disclosure issues)

[15]:
```
fig, ax = plt.subplots(1,3,figsize=(10,5))

total_df[total_df['year']==2017]['PHQ060'].hist(alpha=0.5,ax=ax[0])
total_df[total_df['year']==1999]['PHQ060'].hist(alpha=0.5,ax=ax[0])

total_df[total_df['PHQ060']==1].groupby(['year']).count()['PHQ060'].
 ↪plot(ax=ax[1],marker='o',mfc = 'r')
total_df[total_df['PHQ060']==2].groupby(['year']).count()['PHQ060'].
 ↪plot(ax=ax[2],marker='o',mfc = 'r')
plt.tight_layout()
ax[0].legend(['1999','2017'])
ax[1].set_title('Yes')
ax[2].set_title('No')
plt.show()
```



7 Conclusion: The dietary habits and the eating habits are the same as a decade before. It did not change much of fasting behavior.