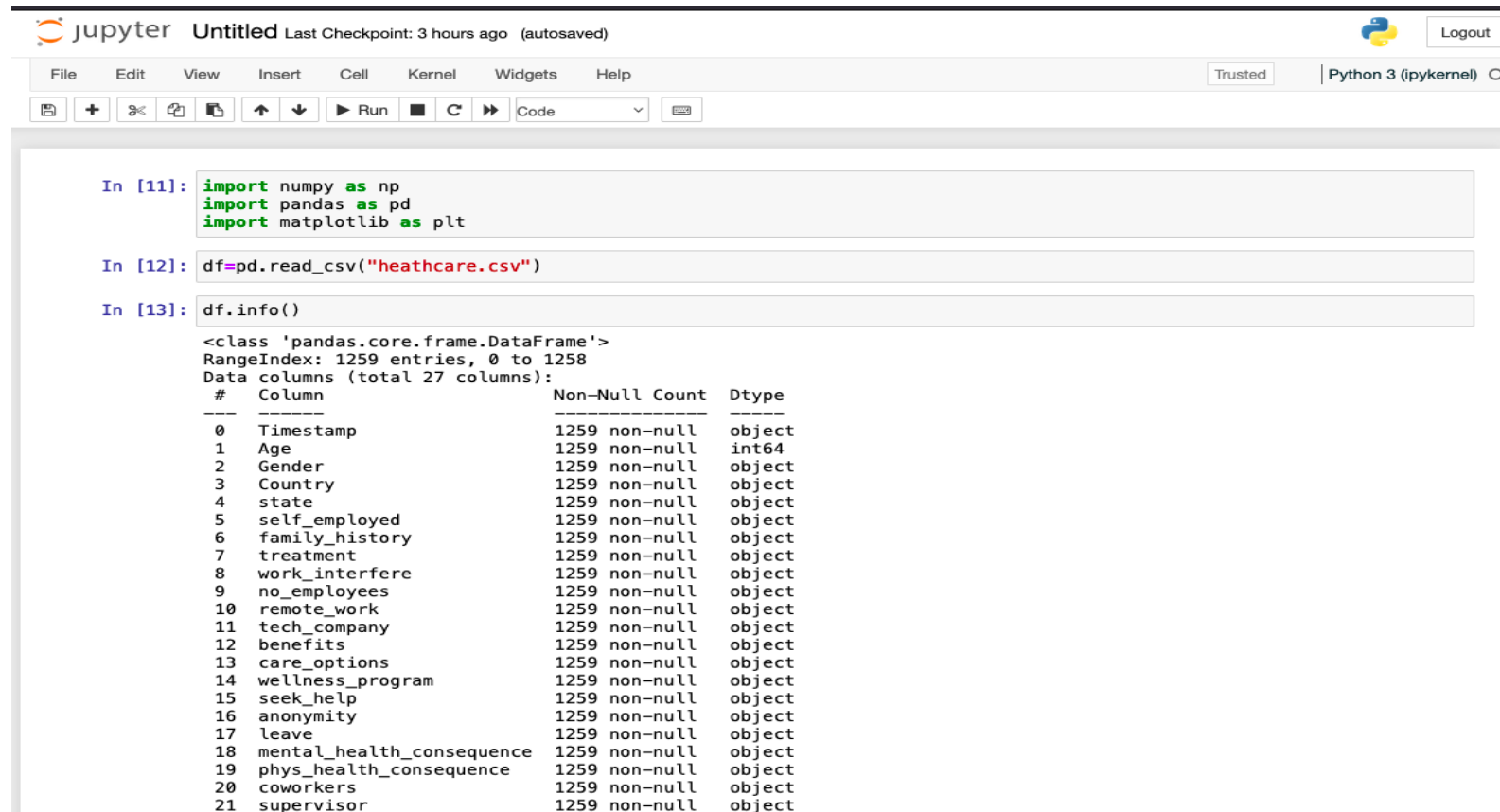




Public Health Awareness Campaign Analysis

Importing the python library's in jupyter (numpy,pandas)



The screenshot shows a Jupyter Notebook window titled "Untitled" with a last checkpoint 3 hours ago. The interface includes a menu bar (File, Edit, View, Insert, Cell, Kernel, Widgets, Help) and a toolbar with icons for file operations, running, and code execution. The notebook contains three input cells:

```
In [11]: import numpy as np
import pandas as pd
import matplotlib as plt
```

```
In [12]: df=pd.read_csv("heathcare.csv")
```

```
In [13]: df.info()
```

The output of the third cell shows the DataFrame's structure:

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1259 entries, 0 to 1258
Data columns (total 27 columns):
#   Column                                     Non-Null Count  Dtype
---  ---
0   Timestamp                                1259 non-null   object
1   Age                                       1259 non-null   int64
2   Gender                                   1259 non-null   object
3   Country                                  1259 non-null   object
4   state                                    1259 non-null   object
5   self_employed                           1259 non-null   object
6   family_history                           1259 non-null   object
7   treatment                                1259 non-null   object
8   work_interfere                           1259 non-null   object
9   no_employees                             1259 non-null   object
10  remote_work                              1259 non-null   object
11  tech_company                             1259 non-null   object
12  benefits                                  1259 non-null   object
13  care_options                             1259 non-null   object
14  wellness_program                         1259 non-null   object
15  seek_help                                1259 non-null   object
16  anonymity                                 1259 non-null   object
17  leave                                    1259 non-null   object
18  mental_health_consequence               1259 non-null   object
19  phys_health_consequence                  1259 non-null   object
20  coworkers                                1259 non-null   object
21  supervisor                               1259 non-null   object
```

Verifying the head of the given data set

```
jupyter Untitled Last Checkpoint: 3 hours ago (autosaved) Python 3 (ipykernel) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted

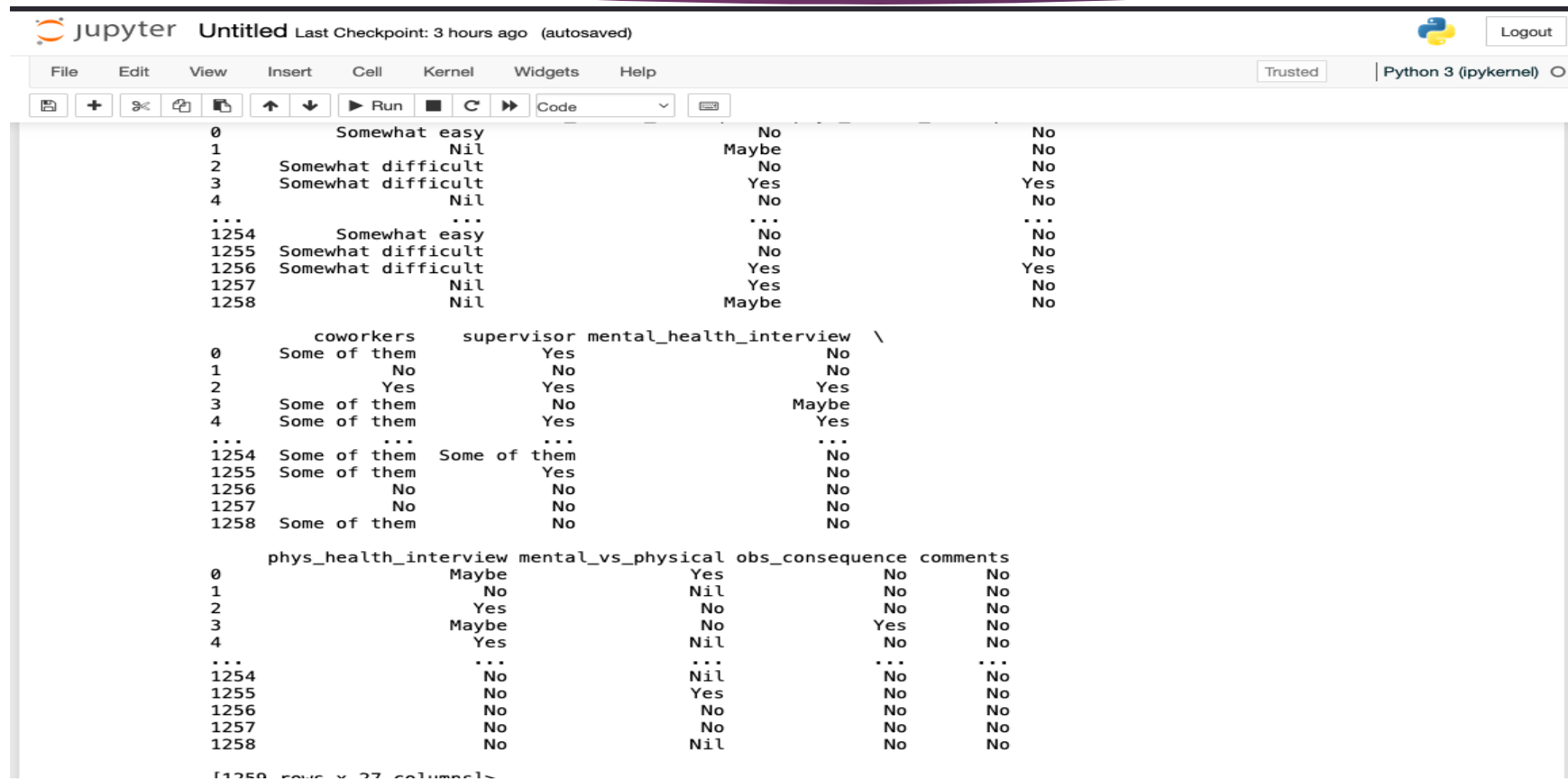
In [14]: df.head

Out[14]: <bound method NDFrame.head of
0      27-08-2014 11:29 37 Female United States IL No
1      27-08-2014 11:29 44 M United States IN No
2      27-08-2014 11:29 32 Male CaNoda No No
3      27-08-2014 11:29 31 Male United Kingdom No No
4      27-08-2014 11:30 31 Male United States TX No
...
1254 12-09-2015 11:17 26 male United Kingdom No No
1255 26-09-2015 01:07 32 Male United States IL No
1256 07-11-2015 12:36 34 male United States CA No
1257 30-11-2015 21:25 46 f United States NC No
1258 01-02-2016 23:04 25 Male United States IL No

family_history treatment work_interfere no_employees ... \
0 No Yes Often Jun-25 ...
1 No No Rarely More than 1000 ...
2 No No Rarely Jun-25 ...
3 Yes Yes Often 26-100 ...
4 No No Never 100-500 ...
...
1254 No Yes No 26-100 ...
1255 Yes Yes Often 26-100 ...
1256 Yes Yes Sometimes More than 1000 ...
1257 No No No 100-500 ...
1258 Yes Yes Sometimes 26-100 ...

leave mental_health_consequence phys_health_consequence \
0 Somewhat easy No No
1 Nil Maybe No
2 Somewhat difficult No No
3 Somewhat difficult Yes Yes
4 Nil No No
...
1254 Somewhat easy No No
1255 Somewhat difficult No No
```

continues



The image shows a Jupyter Notebook interface with a code cell containing a large text-based dataset. The dataset is organized into three sections, each with a header row and a series of rows indexed from 0 to 1258. The first section has three columns, the second has four, and the third has five.

```
0      Somewhat easy      No      No
1      Nil                Maybe   No
2      Somewhat difficult  No      No
3      Somewhat difficult  Yes     Yes
4      Nil                No      No
...
1254   Somewhat easy      No      No
1255   Somewhat difficult  No      No
1256   Somewhat difficult  Yes     Yes
1257   Nil                Yes     No
1258   Nil                Maybe   No

coworkers supervisor mental_health_interview \
0      Some of them      Yes     No
1      No                No      No
2      Yes                Yes     Yes
3      Some of them      No      Maybe
4      Some of them      Yes     Yes
...
1254   Some of them      Some of them      No
1255   Some of them      Yes     No
1256   No                No      No
1257   No                No      No
1258   Some of them      No      No

phys_health_interview mental_vs_physical obs_consequence comments
0      Maybe            Yes     No      No
1      No                Nil     No      No
2      Yes                No     No      No
3      Maybe            No     Yes     No
4      Yes                Nil     No      No
...
1254   No                Nil     No      No
1255   No                Yes     No      No
1256   No                No     No      No
1257   No                No     No      No
1258   No                Nil     No      No
```

[1259 rows x 27 columns]

Verifying the head & tail of data frame

jupyter Untitled Last Checkpoint: 3 hours ago (autosaved) Python 3 (ipykernel) Logout

File Edit View Insert Cell Kernel Widgets Help

Trusted

In [15]: `df.shape`

Out[15]: (1259, 27)

In [16]: `df.head()`

Out[16]:

	Timestamp	Age	Gender	Country	state	self_employed	family_history	treatment	work_interfere	no_employees	...	leave	mental_health_consequenc
0	27-08-2014 11:29	37	Female	United States	IL	No	No	Yes	Often	Jun-25	...	Somewhat easy	N
1	27-08-2014 11:29	44	M	United States	IN	No	No	No	Rarely	More than 1000	...	Nil	Mayt
2	27-08-2014 11:29	32	Male	CaNoda	No	No	No	No	Rarely	Jun-25	...	Somewhat difficult	N
3	27-08-2014 11:29	31	Male	United Kingdom	No	No	Yes	Yes	Often	26-100	...	Somewhat difficult	Ye
4	27-08-2014 11:30	31	Male	United States	TX	No	No	No	Never	100-500	...	Nil	N

5 rows x 27 columns

In [17]: `df.tail()`

Out[17]:

	Timestamp	Age	Gender	Country	state	self_employed	family_history	treatment	work_interfere	no_employees	...	leave	mental_health_consequ
1254	12-09-2015 11:17	26	male	United Kingdom	No	No	No	Yes	No	26-100	...	Somewhat easy	
1255	26-09-2015 01:07	32	Male	United States	IL	No	Yes	Yes	Often	26-100	...	Somewhat difficult	
1256	07-11-2015 12:36	34	male	United States	CA	No	Yes	Yes	Sometimes	More than 1000	...	Somewhat difficult	

Preprocessing the data set & cleaning the excess objective ,rows & etc

jupyter Untitled Last Checkpoint: 3 hours ago (autosaved) Python 3 (ipykernel) Logout

File Edit View Insert Cell Kernel Widgets Help

4 27-08-2014 11:30 31 Male United States TX No No No Never 100-500 ... Nil

5 rows x 27 columns

```
In [17]: df.tail()
```

```
Out [17]:
```

	Timestamp	Age	Gender	Country	state	self_employed	family_history	treatment	work_interfere	no_employees	...	leave	mental_health_consequ
1254	12-09-2015 11:17	26	male	United Kingdom	No	No	No	Yes	No	26-100	...	Somewhat easy	
1255	26-09-2015 01:07	32	Male	United States	IL	No	Yes	Yes	Often	26-100	...	Somewhat difficult	
1256	07-11-2015 12:36	34	male	United States	CA	No	Yes	Yes	Sometimes	More than 1000	...	Somewhat difficult	
1257	30-11-2015 21:25	46	f	United States	NC	No	No	No	No	100-500	...	Nil	
1258	01-02-2016 23:04	25	Male	United States	IL	No	Yes	Yes	Sometimes	26-100	...	Nil	

5 rows x 27 columns

```
In [18]: df=df.dropna()  
df=df.drop_duplicates()
```

```
In [19]: df.to_csv('cleaned_data.csv',index=False)
```

```
In [ ]:
```



THANK YOU

1. Varun .M
2. Sriram .V
3. Amarnath
4. Venmathiraj