```python
import pandas as pd
```

```python
data = pd.read_csv("https://raw.githubusercontent.com/YBI-Foundation/Dataset/main/YouTube%20Spam.csv")
```

```python
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1956 entries, 0 to 1955
Data columns (total 5 columns):
 #   Column      Non-Null Count  Dtype
---  ------      --------------  -----
 0   COMMENT_ID  1956 non-null   object
 1   AUTHOR      1956 non-null   object
 2   DATE        1711 non-null   object
 3   CONTENT     1954 non-null   object
 4   CLASS       1956 non-null   int64
dtypes: int64(1), object(4)
memory usage: 76.5+ KB
```

```python
data.isnull().sum()
```

```
COMMENT_ID     0
AUTHOR         0
DATE         245
CONTENT        2
CLASS          0
dtype: int64
```

```python
data.dropna(axis = 0, how = 'any')
```

	COMMENT_ID	AUTHOR	DATE	
0	LZQPQhLyRh80UYxNuaDWhIGQYNQ96IuCg-AYWqNPjpU	Julius NM	2013-11-07T06:20:48	Hu che
1	LZQPQhLyRh_C2cTtd9MvFRJedxydaVW-2sNg5Diuo4A	adam riyati	2013-11-07T12:37:15	Hey g o chann
2	LZQPQhLyRh9MSZYnf8djyk0gEF9BHDPYrrK-qCczIY8	Evgeny Murashkin	2013-11-08T17:34:21	ju h m
3	z13jhp0bxqncu512g22wvzkasxmvvzjaz04	ElNino Melendez	2013-11-09T08:28:43	me s sexy cha
4	z13fwbwp1oujthgqj04chlngpvzmtt3r3dw	GsMega	2013-11-10T16:05:38	v=vtaR Chec

```python
data.isnull().sum()
```

```
COMMENT_ID     0
AUTHOR         0
DATE         245
CONTENT        2
CLASS          0
dtype: int64
```

```python
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1956 entries, 0 to 1955
Data columns (total 5 columns):
 #   Column      Non-Null Count  Dtype
---  ------      --------------  -----
 0   COMMENT_ID  1956 non-null   object
 1   AUTHOR      1956 non-null   object
 2   DATE        1711 non-null   object
 3   CONTENT     1954 non-null   object
 4   CLASS       1956 non-null   int64
```

```
    dtypes: int64(1), object(4)
    memory usage: 76.5+ KB
```

```
data.dropna(how='any')
```

	COMMENT_ID	AUTHOR	DATE	
0	LZQPQhLyRh80UYxNuaDWhIGQYNQ96IuCg-AYWqNPjpU	Julius NM	2013-11-07T06:20:48	Hu che
1	LZQPQhLyRh_C2cTtd9MvFRJedxydaVW-2sNg5Diuo4A	adam riyati	2013-11-07T12:37:15	Hey g c chann
2	LZQPQhLyRh9MSZYnf8djyk0gEF9BHDPYrrK-qCczIY8	Evgeny Murashkin	2013-11-08T17:34:21	ju h m
3	z13jhp0bxqncu512g22wvzkasxmvvzjaz04	ElNino Melendez	2013-11-09T08:28:43	me s sexy cha
4	z13fwbwp1oujthgqj04chlngpvzmtt3r3dw	GsMega	2013-11-10T16:05:38	v=vtaR Chec

```
df_records_dropped = data.dropna(axis=0, how='any')
df_records_dropped.info()
```

```
    <class 'pandas.core.frame.DataFrame'>
    Int64Index: 1709 entries, 0 to 1955
    Data columns (total 5 columns):
     #   Column      Non-Null Count  Dtype
    ---  ------      --------------  -----
     0   COMMENT_ID  1709 non-null   object
     1   AUTHOR      1709 non-null   object
     2   DATE        1709 non-null   object
     3   CONTENT     1709 non-null   object
     4   CLASS       1709 non-null   int64
    dtypes: int64(1), object(4)
    memory usage: 80.1+ KB
```

```
data.isnull().sum()
```

```
    COMMENT_ID      0
    AUTHOR          0
    DATE          245
    CONTENT         2
    CLASS           0
    dtype: int64
```

```
df_records_dropped = data.dropna(axis=0, how='all')
df_records_dropped.info()
```

```
    <class 'pandas.core.frame.DataFrame'>
    Int64Index: 1956 entries, 0 to 1955
    Data columns (total 5 columns):
     #   Column      Non-Null Count  Dtype
    ---  ------      --------------  -----
     0   COMMENT_ID  1956 non-null   object
     1   AUTHOR      1956 non-null   object
     2   DATE        1711 non-null   object
     3   CONTENT     1954 non-null   object
     4   CLASS       1956 non-null   int64
    dtypes: int64(1), object(4)
    memory usage: 91.7+ KB
```

```
data.mean()
```

```
    <ipython-input-28-abc01cf6c622>:1: FutureWarning: Dropping of nuisance columns in DataFrame reductions (with 'numeric_only=None') is dep
      data.mean()
    CLASS    0.513804
    dtype: float64
```

```python
data.fillna(data.mean().head(136))
```

```
<ipython-input-31-694300ae6e5c>:1: FutureWarning: Dropping of nuisance columns in
data.fillna(data.mean().head(136))
```

	COMMENT_ID	AUTHOR	DATE
0	LZQPQhLyRh80UYxNuaDWhIGQYNQ96IuCg-AYWqNPjpU	Julius NM	2013-11-07T06:20:48
1	LZQPQhLyRh_C2cTtd9MvFRJedxydaVW-2sNg5Diuo4A	adam riyati	2013-11-07T12:37:15
2	LZQPQhLyRh9MSZYnf8djyk0gEF9BHDPYrrK-qCczIY8	Evgeny Murashkin	2013-11-08T17:34:21
3	z13jhp0bxqncu512g22wvzkasxmvvzjaz04	ElNino Melendez	2013-11-09T08:28:43
4	z13fwbwp1oujthgqj04chlngpvzmtt3r3dw	GsMega	2013-11-10T16:05:38

```python
print(data)
```

```
                                            COMMENT_ID                 AUTHOR  \
0       LZQPQhLyRh80UYxNuaDWhIGQYNQ96IuCg-AYWqNPjpU             Julius NM
1       LZQPQhLyRh_C2cTtd9MvFRJedxydaVW-2sNg5Diuo4A            adam riyati
2       LZQPQhLyRh9MSZYnf8djyk0gEF9BHDPYrrK-qCczIY8       Evgeny Murashkin
3               z13jhp0bxqncu512g22wvzkasxmvvzjaz04        ElNino Melendez
4               z13fwbwp1oujthgqj04chlngpvzmtt3r3dw                 GsMega
...                                             ...                    ...
1951    _2viQ_Qnc6-bMSjqyL1NKj57ROicCSJV5SwTrw-RFFA           Katie Mettam
1952    _2viQ_Qnc6-pY-1yR6K2FhmC5i48-WuNx5CumlHLDAI  Sabina Pearson-Smith
1953    _2viQ_Qnc6_k_n_Bse9zVhJP8tJReZpo8uM2uZfnzDs          jeffrey jules
1954    _2viQ_Qnc6_yBt8UGMWyg3vh0PulTqcqyQtdE7d4Fl0        Aishlin Maciel
1955    _2viQ_Qnc685RPw1aSa1tfrIuHXRvAQ2rPT9R06KTqA            Latin Bosch

                         DATE  \
0         2013-11-07T06:20:48
1         2013-11-07T12:37:15
2         2013-11-08T17:34:21
3         2013-11-09T08:28:43
4         2013-11-10T16:05:38
...                       ...
1951  2013-07-13T13:27:39.441000
1952  2013-07-13T13:14:30.021000
1953  2013-07-13T12:09:31.188000
1954  2013-07-13T11:17:52.308000
1955  2013-07-12T22:33:27.916000

                                                CONTENT  CLASS
0     Huh, anyway check out this you[tube] channel: ...      1
1     Hey guys check out my new channel and our firs...      1
2                just for test I have to say murdev.com      1
3        me shaking my sexy ass on my channel enjoy ^_^      1
4            watch?v=vtaRGgvGtWQ   Check this out .          1
...                                                 ...    ...
1951  I love this song because we sing it at Camp al...      0
1952  I love this song for two reasons: 1.it is abou...      0
1953                                                wow      0
1954                          Shakira u are so wiredo        0
1955                      Shakira is the best dancer        0

[1956 rows x 5 columns]
```
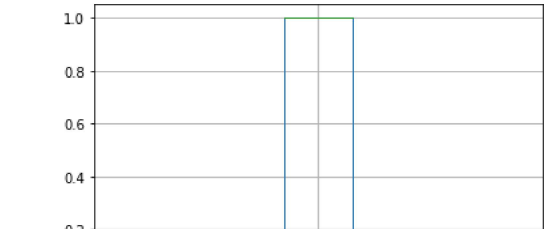
```python
import matplotlib.pyplot as plt
```

```python
boxplot = data.boxplot(column=['CLASS'])
```

```
data.dtypes
```

```
COMMENT_ID    object
AUTHOR        object
DATE          object
CONTENT       object
CLASS          int64
dtype: object
```