

OUT PUT:

```
<class 'pandas.core.frame.DataFrame'>
```

```
RangeIndex: 5175 entries, 0 to 5174
```

```
Data columns (total 4 columns):
```

| # | Column     | Non-Null Count | Dtype  |
|---|------------|----------------|--------|
| 0 | Unnamed: 0 | 5175 non-null  | int64  |
| 1 | label      | 5175 non-null  | object |
| 2 | text       | 5175 non-null  | object |
| 3 | label_num  | 5175 non-null  | int64  |

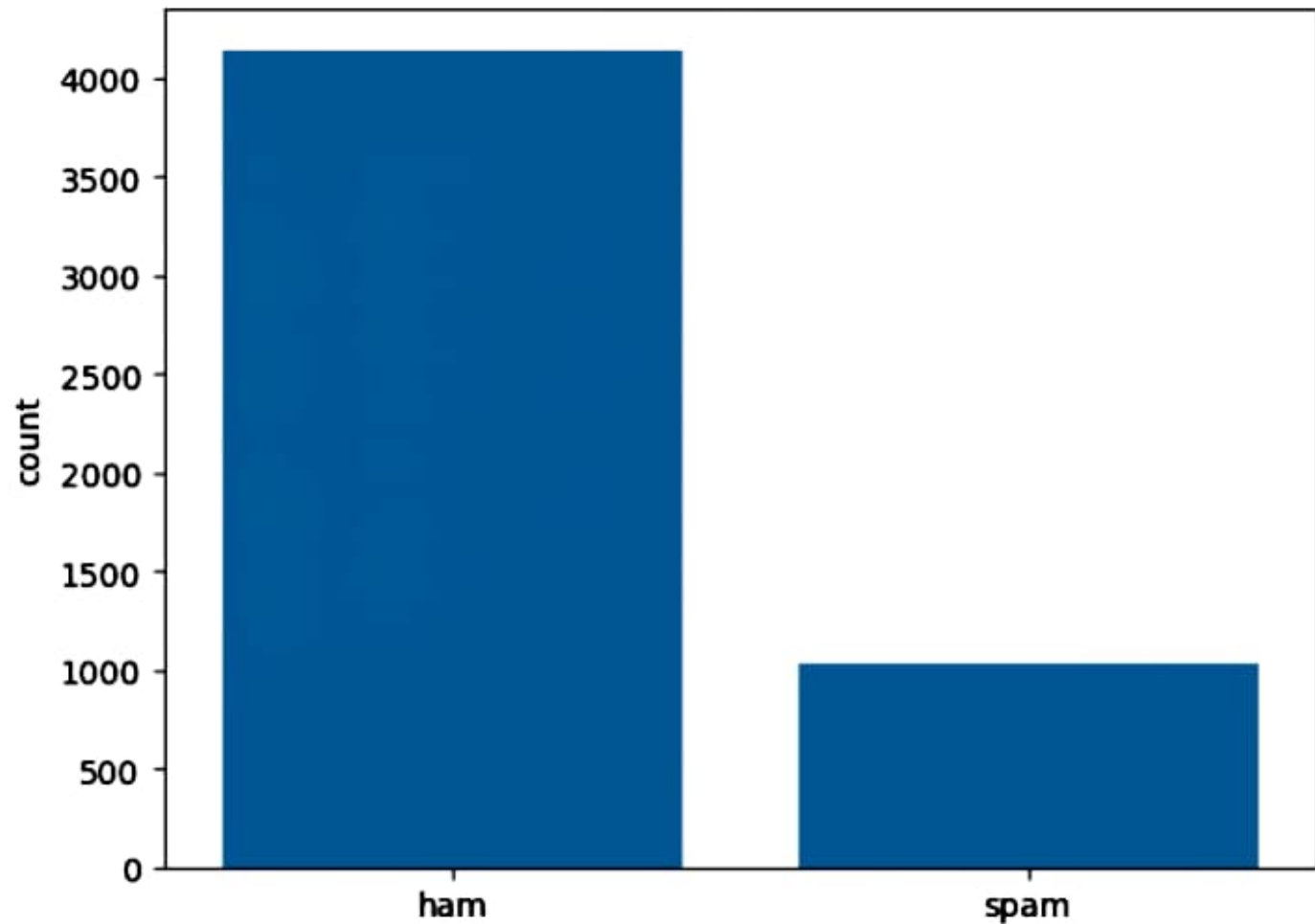
```
dtypes: int64(2), object(2)
```

```
memory usage: 161.8+ KB
```

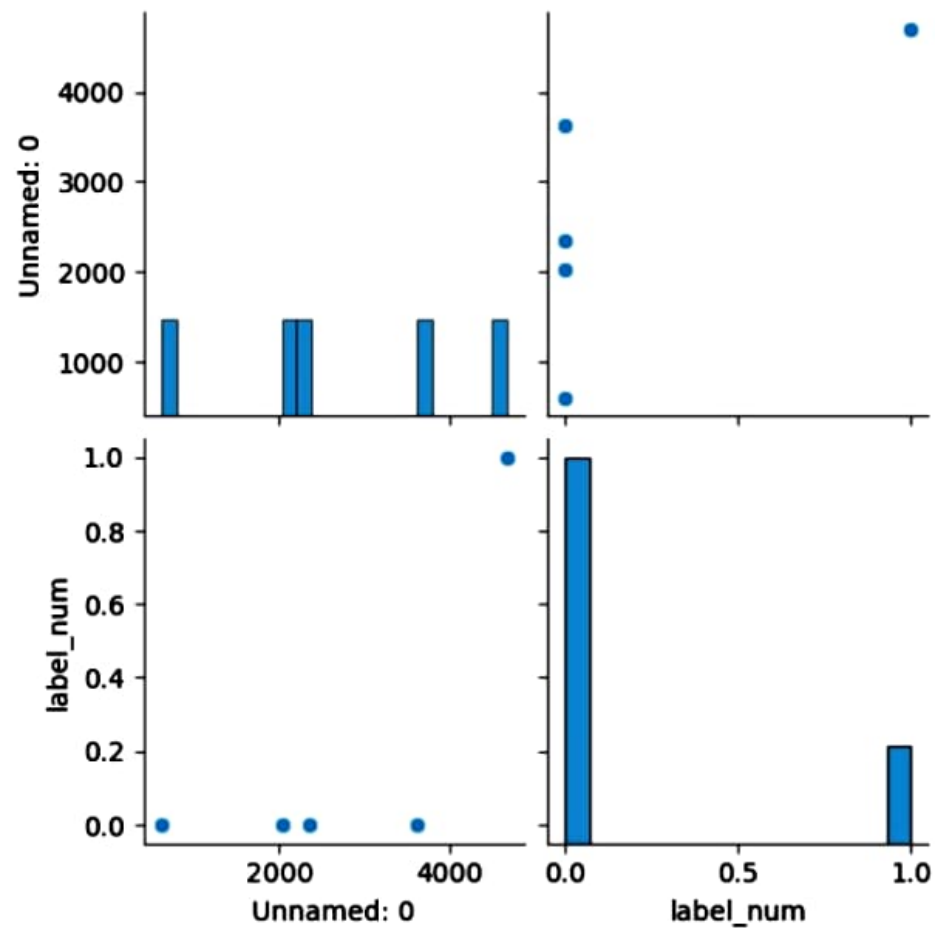
```
None
```

|   | Unnamed: 0 | label | text \  |
|---|------------|-------|---|
| 0 | 605        | ham   | Subject: enron methanol ; meter # : 98829         |
| 1 | 2349       | ham   | Subject: hpl nom for january 9 , 2001             |
| 2 | 3624       | ham   | Subject: neon retreat who ho ho , we're ar...     |
| 3 | 4685       | spam  | Subject: photoshop , windows , office . cheap ... |
| 4 | 2030       | ham   | Subject: re : indian springs this deal is t...    |

|       | label_num   |             |
|-------|-------------|-------------|
| 0     | 0           |             |
| 1     | 0           |             |
| 2     | 0           |             |
| 3     | 1           |             |
| 4     | 0           |             |
|       | Unnamed: 0  | label_num   |
| count | 5175.000000 | 5175.000000 |
| mean  | 2658.600000 | 0.200000    |
| std   | 1396.312355 | 0.400039    |
| min   | 605.000000  | 0.000000    |
| 25%   | 2030.000000 | 0.000000    |
| 50%   | 2349.000000 | 0.000000    |
| 75%   | 3624.000000 | 0.000000    |
| max   | 4685.000000 | 1.000000    |



**label**



By: Karthika. M