

## Data used

We analyze the following page

<http://www.healthysf.org/bdi/outcomes/zipmap.htm>,

in order to obtain the data that is in the table of postal codes and to transform the data into a pandas data frame.

```
: import requests # library to handle requests
from bs4 import BeautifulSoup
import pandas as pd

: response = requests.get("http://www.healthysf.org/bdi/outcomes/zipmap.htm")
soup = BeautifulSoup(response.text, "lxml")
table = soup.find_all("table")
df = pd.read_html(str(table))
df = pd.DataFrame(df[4])

: df.columns = df.iloc[0]
df = df.iloc[1:-1, :-1]
sf_data = df
sf_data.head()
```

	Zip Code	Neighborhood
1	94102	Hayes Valley/Tenderloin/North of Market
2	94103	South of Market
3	94107	Potrero Hill
4	94108	Chinatown
5	94109	Polk/Russian Hill (Nob Hill)