Report

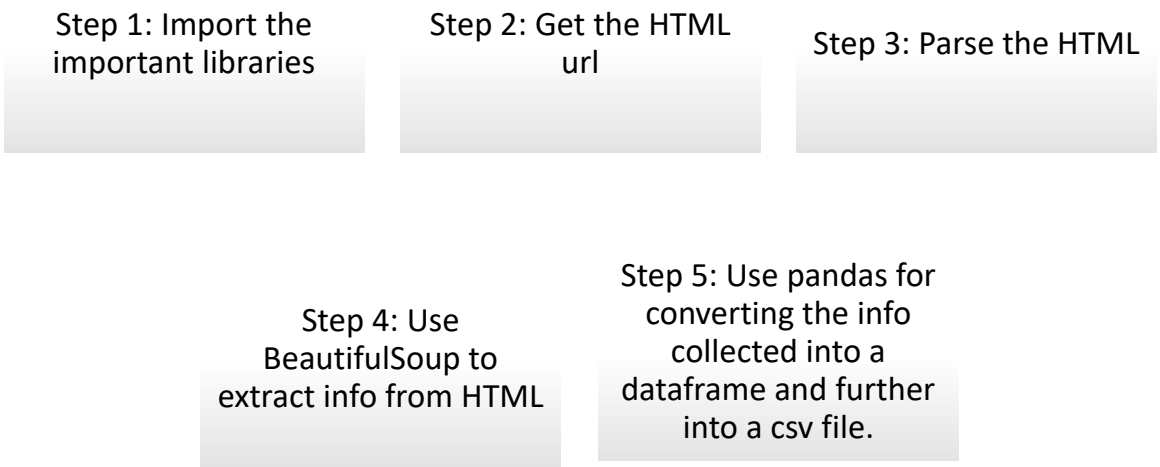
Problem Statement:
Projects and Tenders: California Procurement

Solution:

1. **Libraries Imported:**
    1. Request: Allows to send HTTP request using python. This HTTP request returns a response object with all the response data(content, encoding, status etc.).
    2. BeautifulSoup: Used for pulling data out of HTML file.
    3. Pandas: Used to convert the collected data into a dataframe and further into a csv file.

2. **HTML url:** https://dot.ca.gov/programs/procurement-and-contracts/contracts-out-for-bid

3. **Steps involved:**



4. **Structure:**
   Through exploration of the whole website, the structure of the tabular data could be reduced to:

```
<tbody>
        <tr>
              <td>
              <td>
              <td>
        </tr>
         <tr>
              <td>
              <td>
              <td>
        </tr>…….(lets say some 'n' number of times)
<tbody>
```

Here, <td>= table class, defines the structure, beginning and ending of table.
<tr> = defines the rows of the table.
<td>= defines the elements (data in the rows).

There are 3 important data. The first contains the **EVENT ID** which contains link(anchor tag). The second contains **EVENT NAME** and the third contains **END DATE**.