The first challenge I have faced in this project is to parse the unstructured address. We have to find out the city and state name for each individual post. But the address format is not structured. For example, some may have “Oxford, MS” type address, some may have “Oxford MS” and other may have some garbage values like “Call XXXXXXX”. Another challenge was the parsing the car attributes for each individual car. The HTML of that attributes have same class, so there is no easy way to identify each individual attribute. For example, when I call BeautifulSoup function to find the “span” which has the class “attribute”, it returns a list of attributes. I need to parse the text “condition: good” to get the information for the condition parameter. As each car doesn’t have same number of attributes, I need to match the parameter name to ensure which parameter now I am working with.