**Client Requirement Details**

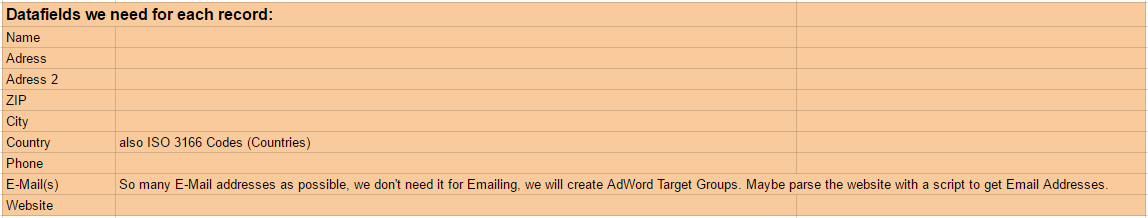
Hi,

We need the contact details of all US boat dealers.

Here is the Job description:

Important:

* you have to search in ALL states in the USA: <http://state.1keydata.com/state-abbreviations.php>
* Extract all required data fields
* Search on URL for more email addresses (we need them for Ad Words and Social Media Targeting not for E-Mailing)
* Clean duplicates and merge data. Identify duplicates by phone number or website...
* Data to be scrapped from URL: http://www.yellowpages.com/search?search\_terms=Boat+Dealers&geo\_location\_terms



**Work Procedure:**

**Tech Stack:**

Python, Python Selenium and Python Pandas

**Procedure:**

1. Using Selenium open the requested URL to scrap

2. Analysis the CSS Selectors or Xpaths to get the urls

* Listed all States in list[] and using for loop appended it to URL
* Page numbers (page numbers are calculated by taking results of the page and by reviewing the page, I have come to know that each page consists of 30 records so did some mathematical calculation to bring the page number count) are also appended to URL.
* Listed the result in List (and where I used DICT inside it for representing the name and variable name as key and value respectively) and used pandas.

3, Scrap the data and store it in csv using Pandas

**Output**

****