**Hints on Solving the First Problem in ML ( Data Collection) in Python:**

**Web Scrapping to collect hotel data from website tripadvisor.in**

1. **Libraries you would need:** 
   1. **Numpy : to handle arrays**
   2. **Pandas: library to perform data manipulation, you will be using to update and transform your data to store it in your result.**
   3. **Urllib: library to perform url actions, access website etc.**
   4. **Beautifulsoup: parsing library to extract data from html tags**
2. **Select your data collection website**
3. **Go though its pages, data you would collect**
4. **Analyze APIs you will be hitting, html tags that contains your data. MAKE A LIST OF IT.**
5. **Collect data in a tabular form (in csv or any text file). If you are storing images, collect it in a folder. For example: product details from flipkart.com store product images as file name: dove\_shampoo\_1 that is <product\_name>\_<product\_category>\_<img\_number>**
6. **Try adding label to the data from collected information so that we have annotated data of our own.**
7. **Play around with it as much as you can and share information retrieved.**

**Note: Keep in mind data is to be collected which will help implement some algorithm to analyze and predict further information. For example, for hotel data: collect hotel name, price, discount, location, description and images (if possible)**

**Some other links can be:**

* **Goibibo.com/robots.txt**
* **Traveltriangle.com**
* **Or any other of your choice with justification**