***Team 1 Python Write-Up***

***3. Instead of extracting the visible reviews in the first page, can you improve the code and extract all***

***the reviews posted on the restaurant page? (Code how much you can and write down in words any***

***other ideas)***

**Steps:**

1. **Get the number of pages for review**
   1. **Use a for loop on the soup.find\_all( ) on < div class\_= Page Number”>**
   2. **Search for class\_ = “pageNum last”to get num\_pages\_review= value of final page number**
   3. **Run a for loop with iterator on range(0,num\_pages\_review)**
   4. **Define a variable url**
   5. **For trip\_advisor, the pattern for each page changes with** 
      1. [**https://www.tripadvisor.com/Restaurant\_Review-g187147-d1751525-Reviews-Cafe\_Le\_Dome-Paris\_Ile\_de\_France.html**](https://www.tripadvisor.com/Restaurant_Review-g187147-d1751525-Reviews-Cafe_Le_Dome-Paris_Ile_de_France.html)
      2. [**https://www.tripadvisor.com/Restaurant\_Review-g187147-d1751525-Reviews-or10-Cafe\_Le\_Dome-Paris\_Ile\_de\_France.html**](https://www.tripadvisor.com/Restaurant_Review-g187147-d1751525-Reviews-or10-Cafe_Le_Dome-Paris_Ile_de_France.html)
      3. [**https://www.tripadvisor.com/Restaurant\_Review-g187147-d1751525-Reviews-or20-Cafe\_Le\_Dome-Paris\_Ile\_de\_France.html**](https://www.tripadvisor.com/Restaurant_Review-g187147-d1751525-Reviews-or20-Cafe_Le_Dome-Paris_Ile_de_France.html)

**so we need to add “or”+ page\_number**

**url = “https://www.tripadvisor.com/Restaurant\_Review-g187147-d1751525-Reviews-“+ “or”+ page\_number**

* **Run a for loop on each url**
* **Get a list[] containing reviews for each of these pages**

***4. Can you improve the sentiment analysis algorithm in question 1 further? Can you figure out the***

***approximate sentiment of a word that does not exist in the “word\_sentiment.csv” file? (Code how***

***much you can and write down in words any other ideas)***

**get approximate sentiment of word that does not exist:**

1. **Determine sentiments of all other words present in that review**
2. **If some of them present, allot the average sentiment of those words to other words not present**
3. **So the overall average sentiment of that review remains the same**
4. **Create a dictionary of common prepositions eg – is, this, was, it,etc. and do not allot any sentiment to these neutral words**
   1. **These will have value =0**