

Assignment -1

Deadline : Oct 4th, 5 pm (1 hour before class)

Total points : **100**

Question 1 - 15 Points

- Create a list of numbers which are multiple of 3 and 7 between 1 to 500 (5 points).
- For the list, find what is the mean. Print what is 25th quartile, 50th quartile and 75th quartile. (10 points)

Question 2 - 15 Points

- Read the file 'ques2.txt'. Print the number of words. (5 points)
- Find out what is the average, median and mode for word length (character length) (10 points).

Question 3 - 35 Points

- There are 1500 json files in 'Lecture03/question03'. Find out how many of the files contain the string, 'User not found' (case insensitive). (15 points)
- ```
{"code":"USER-100","name":"User Not Found","message":"Unknown user Zyee1Ale1Bx8N15"}
```
- File name is user\_id.json
- Create a dictionary, containing (user\_id,present). present is 0 if 'user not found' else 1. Print random 5 entries (20 points).

# Question 4 - 35 Points

- For the same Json files, there is a field, **`segmentUid`** which can either start with 0xxxxx(lets call them 1st party data) or 1xxxxx (lets call them 3rd party data).
- For each user, count the number of first party segments and third party segments. Points (35 points)
- Hint : {userID: {first:20,third:30}}

- {"dxId":"ly16MJAb1BAmyG5","optOut":false,"profile":{"attributes":null,"siloeData":null},"flightFrequencyCaps": [],"flightImpressionCounts": [],**"retargetingSegments":**[{"**"segmentUid":"1SMkVD","timestamps":****[1472730860],"activityCount":1**}, {"segmentUid":"1SU7KQ","timestamps":[1471921416],"activityCount":1}, {"segmentUid":"1SZD81","timestamps":[1471921416],"activityCount":1}, {"segmentUid":"1SRIZO","timestamps": [1471921416],"activityCount":1}, {"segmentUid":"1SaReX","timestamps":[1474775638],"activityCount":1}, {"segmentUid":"1S4SfB","timestamps":[1472547850],"activityCount":1}, {"segmentUid":"0SFXYQJYy","timestamps":[1474344699],"activityCount":1}, {"segmentUid":"1SUTxw","timestamps":[1471921416],"activityCount":1}, {"segmentUid":"1SbPVN","timestamps": [1471834934],"activityCount":1}, {"segmentUid":"1SeNbC","timestamps":[1471921416],"activityCount":1}, {"segmentUid":"0SF5glSMor","timestamps":[1474344661,1474877858],"activityCount":3}, {"segmentUid":"0Sd4aFi7Hr","timestamps":[1473442782],"activityCount":1}, {"segmentUid":"1SQiWu","timestamps":[1472730860],"activityCount":1}, {"segmentUid":"1SPw0E","timestamps": [1472547850],"activityCount":1}, {"segmentUid":"1SNkNm","timestamps":[1472730860],"activityCount":1}, {"segmentUid":"0SJPJUMn4p","timestamps":[1472584744],"activityCount":1}, {"segmentUid":"1SHcuo","timestamps":[1472547850],"activityCount":1}, {"segmentUid":"0SuVvVNte2","timestamps":[1471834818,1474458240],"activityCount":2}, {"segmentUid":"1SaALG","timestamps":[1472547850],"activityCount":1}, {"segmentUid":"0SZnH7iq1B","timestamps":[1471921368],"activityCount":1}, {"segmentUid":"1SR9Z9","timestamps":[1472730860],"activityCount":1}, {"segmentUid":"1SxK2W","timestamps": [1472730860],"activityCount":1}, {"segmentUid":"0Svn4L7KgH","timestamps": [1471921368,1473605157],"activityCount":2}, {"segmentUid":"1SXW3S","timestamps": [1471921416],"activityCount":1}, {"segmentUid":"0SyYlclL2x","timestamps":[1474458238],"activityCount":1}, {"segmentUid":"1SlwqD","timestamps":[1472547850],"activityCount":1}, **{"segmentUid":"0SrkdCp0dZ","timestamps":****[1474877834],"activityCount":1**}, {"segmentUid":"1SIPOb","timestamps":[1472547850],"activityCount":1}, {"segmentUid":"1Ssnug","timestamps": [1472547850],"activityCount":1}, {"segmentUid":"0S1Rflclwz","timestamps":[1473183994],"activityCount":1}, {"segmentUid":"1S6Gkx","timestamps":[1471921416],"activityCount":1}, {"segmentUid":"1Sc2bL","timestamps": [1472730860],"activityCount":1}, {"segmentUid":"1S4ByK","timestamps":[1472023437],"activityCount":1}, {"segmentUid":"0S3n7PJTms","timestamps":[1474877810],"activityCount":1}], "messages":null,"dataSource":"Cache"}