SUBMISSION:

1. You will create a folder in timberlake named lab1. (Timberlake is a cse server).
2. Every notebook should have your name only at the top of the notebook and your team member's name in the second line.
3. Store or transfer all the notebooks to lab1 folder on timberlake: yourLastNameLab1Part1.ipynb, yourLastNamePart2.ipynb, youLastNamePart3.ipynb, **all the data used including curated tweet**s; we need the **data** to run your notebooks to make sure we **can reproduce your results**.
4. On timberlake tar the lab1 files into yourLastNameLab1.tar
5. Documentation: Use Jupyter markdown to document you analysis.
6. Submit using submit-cse487 filename.tar or submit-cse587 filename.tar

DUE DATE: 3/8/2017 BY MIDNIGHT. ONLINE SUBMISSION ON TIMBERLAKE.

HOW CAN DO WELL IN THIS LAB?

- Start working on it today. For example, visit the fluview page and download all the data and the corresponding graphs (this is in a single powerpoint file). Next week they may be gone or different.
- You can work in parallel on the Part2 and Part 3. Set up the Oauth code [8] and start collecting the tweets for Part 3. You may not get the data you want in the last minute. You cannot copy data from others.
- Plan, design, prototype, test and iterate through these steps.
- Choose a partner so that you can complement each other in skills and learn from each other.
- Attend TA office hours and recitations every week. Attend any number of office hours by any TA until your questions are answered.
- Enroll in Piazza (CSS4/587) and ask questions. Don't post code. Be civil. This is a public forum.
- Login into timberlake.cse.buffalo.edu and make sure you have an account on cse servers. If not send mail to cse-consult@cse.buffalo.edu to get an account.
- Create a lab1 folder with dummy files for 3 python notebooks, tar/zip the file, submit the zip file and check it out it goes without any problem.
- Finally, no cheating. Do not copy or get the code from somebody. By this you are building a disadvantage. You are missing a golden opportunity to learn. The lab, the languages and tools may be hard for non-programmers, but that is no substitute for hard work. Of course, we will make sure people who cheat are appropriately penalized.

REFERENCES:

1. Jupyter. http://jupyter.org/, last viewed 2019.
2. The R Language. https://cran.r-project.org/, last viewed 2019.
3. R-Studio. https://www.rstudio.com/, last viewed 2019.
4. Lecture handout on Jupyter, R and RStudio, Lecture on Feb 2019.
5. Lecture demo on R-language demo, Lecture on Feb, 2019.
6. Twitter API. Twitter Developer https://dev.twitter.com/, https://developer.twitter.com/en/docs last viewed 2019.
7. TwitteR package. https://cran.r-project.org/web/packages/twitteR/twitteR.pdf, last viewed 2019.

8. D. Kahle and H. Wickham. ggmap: Spatial Visualization with ggplot2. The R Journal Vol. 5/1, June 2013, https://journal.r-project.org/archive/2013-1/kahle-wickham.pdf.
9. OAuth2.0. OAuth2.0: https://oauth.net/2/, last viewed 2019.
10. https://www.cdc.gov/flu/weekly , interactive data analysis of various flu parameters, last viewed 2019.
11. https://www.cdc.gov/flu/, CDC Weekly report on Flu Activity, last viewed 2019.
12. J. Gentry. TwitteR Vignette: A Twitter Client for R. http://geoffjentry.hexdump.org/twitteR.pdf, last viewed 2017.
13. Fluview data download https://gis.cdc.gov/grasp/fluview/fluportaldashboard.html. Last viewed 2019.
14. Twitter geolocations. https://developer.twitter.com/en/docs/geo/places-near-location/api-reference/get-geo-search.html, last viewed 2019.