TDS3751 – Social Media Computing
Session 2 2022/2023
Assignment 2

OBJECTIVES

This assignment worksheet details the SECOND assignment component for the subject TDS3751 in Trimester 2320 with the following points:

- The assignment is to be done in the same groups from assignment 1
- The final submission will be in the form of report (softcopy) alongside functional code.

ASSIGNMENT TASKS

Due to the predicted but unfortunately disastrous changes to how Twitter developer accounts can be used, the second part of the assignment for TDS3751 has to be radically changed.

Your second part is to investigate alternate methods to scrape Twitter for data – some examples have been shown in earlier tutorials and those are permitted to be used as reference but you need to expand on how you are implementing it in your assignment

The details for this part of the assignment is divided into the following parts
PART 1

- Research into *at least* three (3) different approaches to scraping data from Twitter – these could include manual methods (i.e. search, copy, and paste), scripting (e.g. BeautifulSoup, Scrapy, etc), alternate access APIs (e.g. TWINT, snscrape, scweet, etc), paid methods and so on
- Come up with a conclusion table to the advantage/disadvantage of the three methods you have chosen to research into. You may choose the comparison points between the methods – for example price, speed, ease of use, flexibility, accuracy, etc
- Once you have the conclusion table it should point towards the method(s) you are going to use to collect enough data to complete Part 1 of your assignment – you may use multiple methods if necessary
- Collect only the minimum amount of data necessary to make calculations statistically valid – typically 30-50 individual items. No marks are allotted for excessive data collection

PART 2

- Show how you use the methods chosen in part 1 to collect the data – you may include any talking points including issues, problems, requirements, etc
- Show example code or screenshots of the method in execution to collect your data
- Include your data collected (i.e. dataset(s)) as appendix into your submission report later

DELIVERABLES

1) Code for data collection (if available) else explain in report if using a paid/free trial method with its own interface/GUI
2) Code for individual students processing/cleaning of the brand's Twitter channel (including comments and citations to sources used if any)

3) Final dataset(s) (all content collected/student) – these could be different dataset from Part 1 of the assignment, so label them correctly or place them in different subfolders
4) Report and explanation of dashboard content (i.e. explanations of peaks, spikes, drops, why one brand has more X compared to Y etc)
    a. You may have separate dashboards per brand or one dashboard that consolidates all the brands' data together.
    b. The dashboard does NOT need to be interactive (i.e. no requirement for animated graphs, clickable hotspots, etc) but the content displayed must be dynamic and generated onload and not a static pre-drawn image loaded when displayed

RUBRICS

Changes to mark distribution are being made here to affect both part 1 and part 2 of the assignment

This assignment part is allocated 10% for the subject assignment component. The breakdown for the tasks is as follows (subject to changes and normalization where and when necessary)

- Method research 5%
- Method implementation (collection and cleaning) 5%

The rubrics from part 1 is revised and mapped directly to 40% to take into account the unforeseen changes to the social media platform planned earlier. Thus the revised mark distribution for part 1 is as follows (all rubrics remain the same but scaled accordingly)

- Domain analysis 10%
- Data analysis (metric calculations) 10%
- Brand analysis and comparisons 10% (5 metrics ×4%)
- Dashboarding and reporting 10%

The tasks will be graded using the following rubrics based on the individual questions asked.

Incomplete 0%	Below average 25%	Partial 50%	Above average 75%	Complete 100%
No implementation and no output from execution	Incomplete solution with partial execution output	Implementation completed but execution does not correspond to requirements (i.e. wrong output, no output etc)	Completed implementation but with partially correct execution (i.e. errors generated, exceptions etc)	Completed implementation with proper execution as required

* where execution refers to either program running capability, processing output accuracy, justification, explanation etc.