

Feedback Questionnaire Form
CA675 – Cloud Technologies

Your group ID: O

Group ID of the report being reviewed: Q

Section 1 – General Idea

What I find bad/can be improved in this idea is We find that the presentation of the idea is vague, no details on the demographics dataset have been included.

What I find good/has potential in this idea is Exploring the various factors influencing an election is an interesting analysis. The insights gained from this analysis can prove to be a game-changer in the election process.

One problem/challenge you might encounter with this idea is The primary dataset being considered for the realisation of the idea has very small number of records.

Section 2 – Technology

Technology 1

One key technology used in this project is Spark (using PySpark)

Does this technology fit the general idea?

☐ Yes ☒ No

If Yes, one way the app can be expanded/improved using such technology is

If No, why?

Having reviewed the primary dataset being used, a Spark runtime (with Hadoop) would not be necessary for a 17mb dataset. (We are assuming that the volume of a limited set of records would not considerably increase post the merger with the undefined demographics dataset)

Technology 2

Another key technology used in this project is Tableau

Does this technology fit the general idea?

☒ Yes ☐ No

If Yes, one way the app can be expanded/improved using such technology is We find that Tableau can be used to build a live dashboard by integrating it with Hadoop. We believe this integration is key for the solution to have an end to end functionality vis-à-vis data ingest to data visualisation.

If No, why?

Section 3 – Task allocation and user roles

Is the task allocation fair/reasonable?

☐ Yes ☒ No

If Yes, who do you think is a good leader candidate for the group and why?

If No, say why and indicate at least one advice for improving fairness in task allocation.

There is no ownership of a specific functionality across the proposed solution. The general observation is that all the members are performing all the tasks.

An approach that can be considered is - Task distributions based on specific functionality. Eg. Data cleaning Function1 created by Member1 can be called by Data analysis Function2 in Member2's code.

We also observe that one member has been entasked with the responsibility of the report creation, however, we believe each member has to document his/her contribution.

