Please follow this

I've included the information below for you to look over.

- Coding Assignment pointers

Based on the Scenario given candidate has to consider the below pointers in coding assignment.

* Clean up the data and transform it into structured formats suitable for analysis

* Create the master table & perform aggregations from the given dataset sources and storing it in a single format.

* Processing the data pipelines with a Test driven development approach (TDD) like formation of a number of things into a cluster.

* Code structure should be good & Solutions provided should be clear.

* Write Different Pass/Fail test cases for multiple scenarios.

* Good documentation should be done like provide justification with good comments with Flow diagrams.

* Mention what are the things the candidate has considered.

* What points should be added for future enhancements?

* Can create a read-me file and send a GITHUB link or ZIP file for the review