Data Science/Analysis : Remote Work Assessment

Congratulations on making it to the remote work assessment stage for the **Data Science** Curriculum **Designer** and **Facilitator** position at Moringa School.

This document will be your guide in helping you design a deliverable for the **Product Team** at Moringa School thus enabling you to experience a day in the life of a **Content Designer** at Moringa School.

Our work assignments are inspired from real world scenarios that graduates from Moringa School have experienced and given feedback on. This allows us to give you a chance to sample the expectations levied on the quality of your work as a curriculum designer from a students perspective and from the Outcomes point of view.

The idea is for you to experience the type of work your role will entail, empowering you with the type of problem solving mentality you will have to embody at Moringa School. At the same time, it allows us to understand how you would approach the role and how best to support you in the role.

We wish you the best and hope you enjoy the experience as you step into the shoes of a Data science **Curriculum Developer** at Moringa School.

Assignment

Moringa is looking to improve it's module-to-module retention rates for its Software Development programme. Students enrolled in the programme must complete 5 modules in this order: Intro to Software Development, Angular, Java/Python, Android/Django and Professional Development.

To improve retention, we would need to find out why students may be dropping out between modules even when they pass their prerequisite modules. And to help us do this, we have collected performance and engagement metrics for a randomly selected Software Development class. For each student in each module, we have the following metrics:

Metric	Description	Type
Attendance	Percentage of times present for each attendance taken	Percentage
IP Grade Average	The average of the combined set of Independent project grades a student has obtained	Percentage
Total Activity Time	Total amount of time in seconds a student spends engaging with courseware.	Integer
Participation Level	Number between 0 and 3 that represents relative participation in learning activities.	Integer
Page View Level	Number between 0 and 3 that represents amount of content viewed relative to others in the same class	Integer

The data for each student for each of the modules is included in the attached spreadsheet. In addition to the metrics listed above it also includes the obfuscated IDs for the students enrolled in each module.

For each of the questions below, make sure to include the process & methods used to come up with your answers.

Question 1.

Clean up the provided data with the following assumptions:

- Students that have none of the above metrics stated is presumed to have dropped out of the module and should not count towards the final enrolment number for that module
- Students that scored below 50% in a module and are not shown to be enrolled in one of the next modules are presumed to have failed and should be excluded when calculating module-to-module retention (See Q2.)
- Students that pass Angular can enrol in either Java or Python but not both.
- Students that pass Java can proceed to take Android but not Django.
- Students that pass Python can proceed to take Django but not Android.
- Students that pass either Android or Django can proceed to take Professional Development.

MORINGA
Discover · Grow · Transform

Question 2.

Determine the module-to-module retention rates for each of the modules.
(e.g. Intro to Software Development -> Angular : x%).

Question 3.

Which of the metrics above is most correlated to the module-to-module retention rate? List them in order of most to least correlated.
(Include any charts and tables necessary to explain your answer.)

Question 4.

If Moringa had to pick 2 metrics out of the above listed ones to help improve module-to-module retention, which should we pick and why?
(Include any charts and tables necessary to explain your answer.)

Question 5.

What other metrics, in your opinion, do you think Moringa should be measuring to better predict our module-to-module retention and why?
(Note that no data has been provided for this question, so your response is expected to be speculative, but backed with well-reasoned arguments)