**Final Report**

Saint Martin’s University staff will be using test plugins on their Learning Management System, Moodle. Plugins for Moodle are add-ons to the base version of Moodle that add additional features that can be useful for both administrative and instructional purposes. The objective of this project is to analyze the effectiveness of the test plugins to improve the instructional quality of Saint Martin’s University using Moodle.

**Specifications**

This project will use logs generated by Moodle to analyze the effectiveness of plugins that will be tested during the Spring semester. To do this, I will be working with Saint Martin’s Information Technology Services to gather the logs that are stored from the Moodle vendor, eThink. The analysis of the plugins should clearly show the answer to the following:

* How often students are using the plugin?
* Are the students using the plugin more than once?
* Are the students using the plugin over the entire course of the semester?
* What time of day are the students using the plugin?
* Are students using the plugin before a specific activity like quizzes or discussions?
* Do the new plugins change how students interact with other Moodle features?

The answers to these questions will be useful for instructors that are trying to plan effective Moodle courses using these new plugins in the future. To provide a clearer answer, the logs of the test Moodle courses in the Spring will be compared to the logs from Moodle courses in the Fall. The logs from the Fall will be useful in showing if the use of the new plugins affects the usage of other base Moodle features like the viewership of notes and the interactivity of Moodle discussions.

**Design**

The analysis of this project will be done entirely with R. The output of Moodle logs comes in the form of a CSV file which makes it easy to be imported into RStudio. Once the logs have been imported, a significant amount of data wrangling is needed to use the data. The Tidyverse package will provide most of the tools needed for the data cleaning. The data contains a few string multi-valued attributes and redundant log files. For example, the process of turning in an assignment generates four log files for the same action. While useful for error handling, the redundant logs could create misleading data if left alone. All types of redundant files will need to be resolved before analysis is complete to ensure that all data used to examine the usage of a Moodle feature is recorded accurately. To extract useful data from the string multi-valued attributes, regular expressions using the Rebus R package will be used. Finally, the date and time attribute of the log files will be expanded using functions from the Lubridate package. This will make it easy to track the usage of Moodle throughout the duration of the course and during each day.

Once the data is cleaned it will be prepared for visualizations using the Ggplot2 package. These visualizations with commentary will be the main method of presenting the findings of the analysis. They will use the cleaned Moodle logs to answer the questions detailed in the specification section. While it is undetermined at the moment if Saint Martin’s ITS will be able to get access to the logs regularly, if they are able to, the R scripts used for the analysis can be expanded for regular Moodle courses. Should this happen, a separate version of the R script used to create the report for the plugins tested in the Spring will be created to give a more general analysis of a Moodle course and provide visualizations. This result would likely be in the form of an RShiny Dashboard that would allow professors to view how students are interacting with their Moodle course at any time during the semester.

**Schedule**

The currently planned schedule for the project is:

* December – January: Complete and optimize the data cleaning R script, work with ITS and eThink to gain access to logs from the Fall semester Moodle courses
* February: Complete analysis of the features included in the base version of Moodle using logs from the Fall semester
* March: Begin analysis and visualization of the test plugins using logs generated during the first month of the Spring semester
* April: Add logs generated through the end of midterms during the Spring semester to the analysis and visualizations and begin creating a presentation of the findings
* May: Add logs from the entire Spring semester to the analysis and visualizations to complete the dataset and complete presentation of the findings

The process of cleaning the data is mostly completed. The code for that process should work the same for all Moodle course logs, once it is finished the focus will be on manipulating the data for visualizations. I plan to have the data cleaning process finished and optimized by the end of January. When the ITS staff returns from the break after the holiday I will begin working with them on getting access to the logs from the Fall semester. I will also work with ITS to create a plan for getting access to the Spring Moodle logs periodically during the Spring semester to work with it over the semester. I hope to get access to the Fall semester logs by February so I can begin visualizing and analyzing them as soon as the data cleaning script is finished. My hope is also to get access to the Spring logs generated a month into the semester, after mid-terms, and after finals. This will allow me to analyze groups of the data as it is generated during the semester rather than waiting until the Spring semester ends to analyze it all. As of now, it has not been determined how well eThink will be able to handle the requests for data so the data access schedule is partly out of my control.