**MDS 560 Week 7 Hands-On Accelerator**

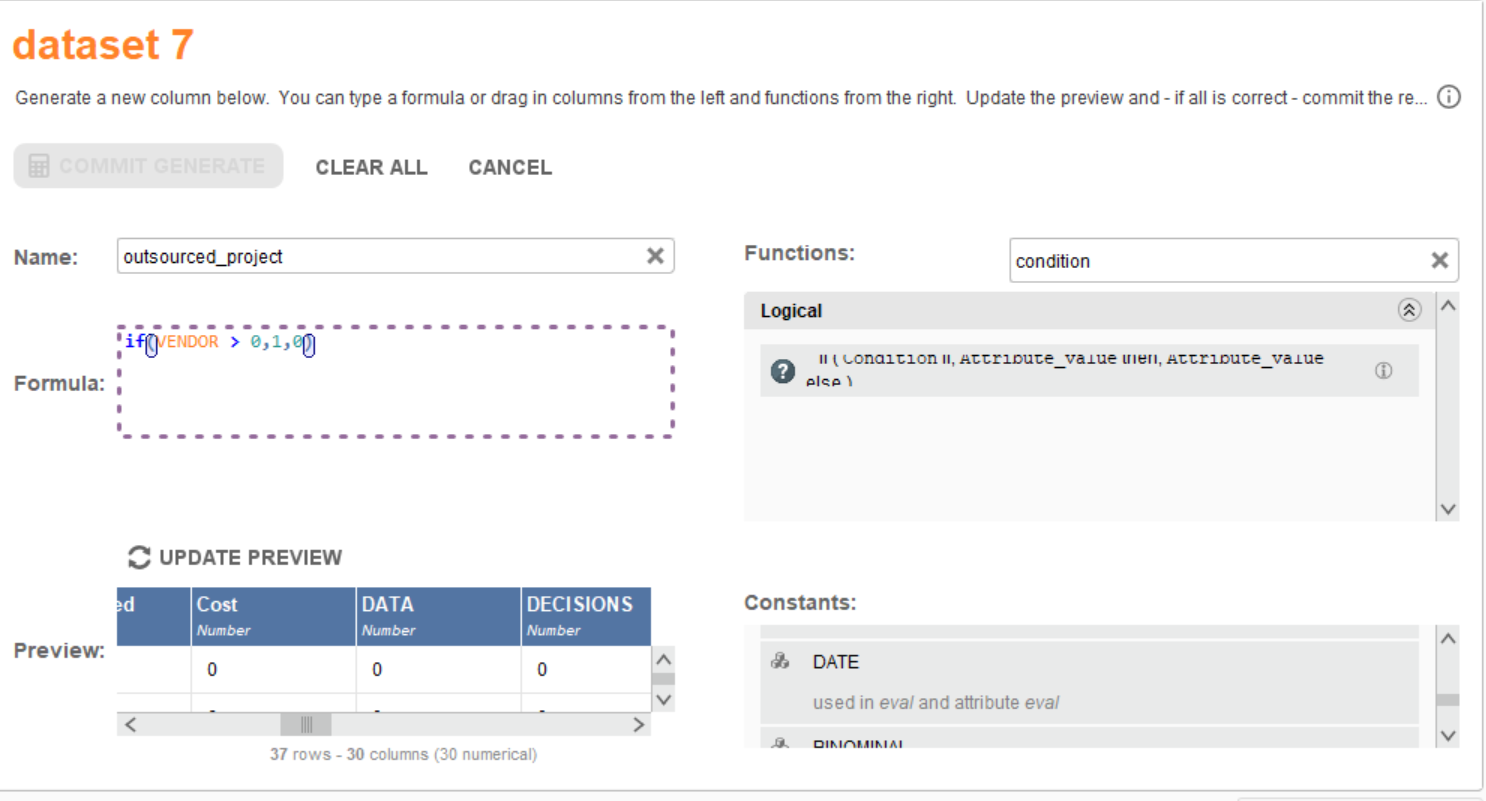
Your deliverables and hands-on activities for this week are:

1. Make substantial progress on your course project assignment.

Results: I made significant progress on my course project this week.

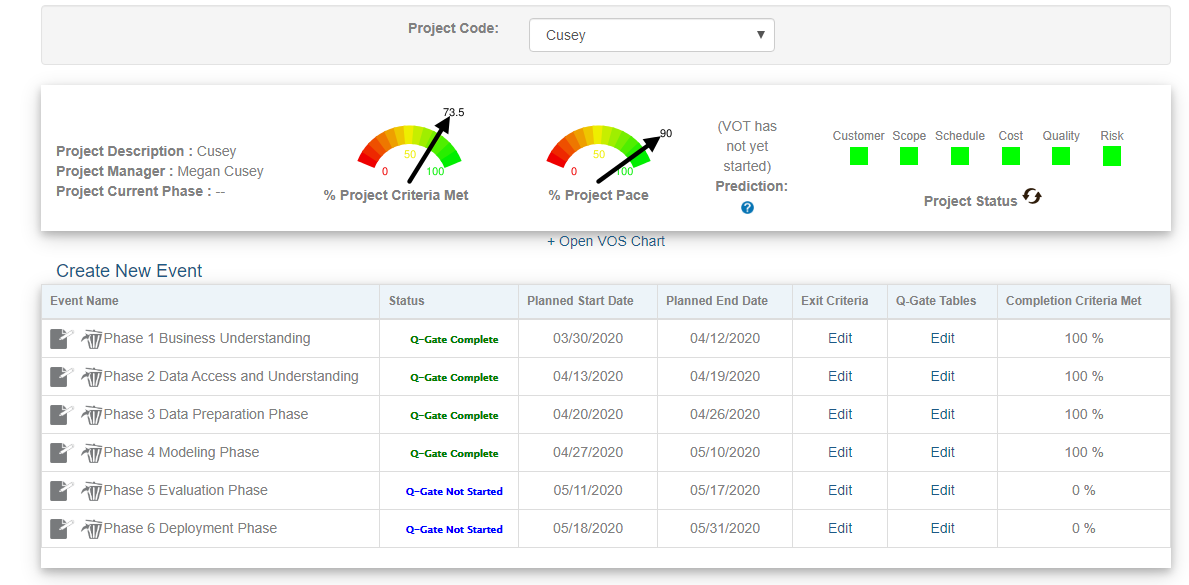
2. Build a simple example of feature engineering using the Generate function of Turbo Prep in Rapidminer using dataset 7.

Results: Below I am creating “outsourced\_project” variable. I am looking to see if there is any mention of the word “Vendor”. If so, I am flagging the project as an outsourced project which indicates that at least one component of the project is outsourced. This may be a stronger predictor variable then counting the number of times vendor is used.



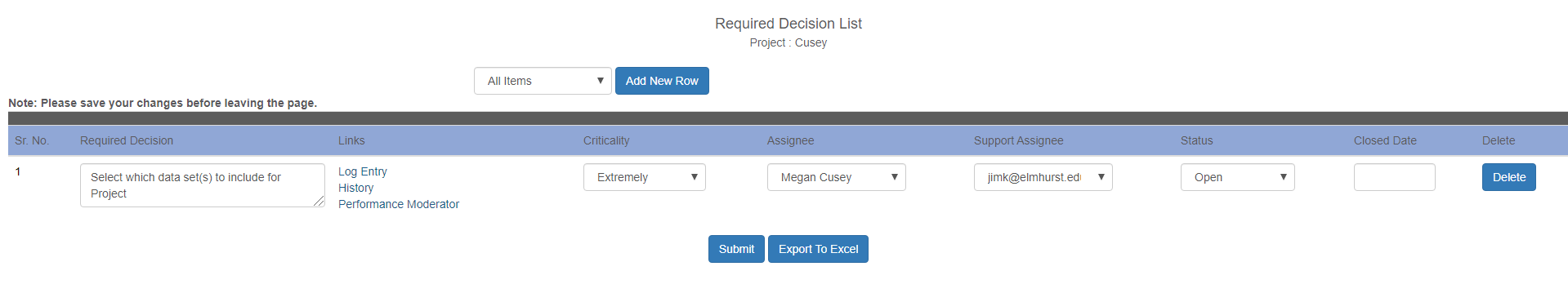
3. Go to your project in PRIMMS and begin closing out activities and quality criteria for the “Data Prep” and “Modeling” phases of the CRISP-DM life cycle.

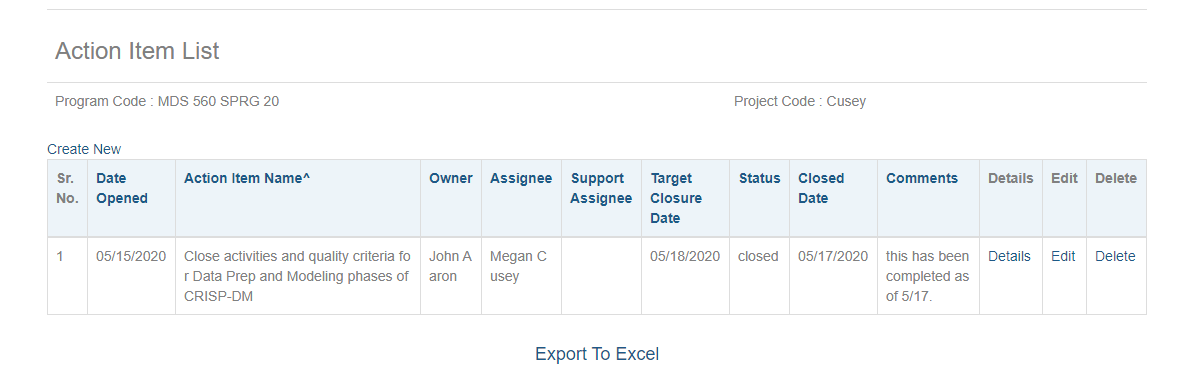
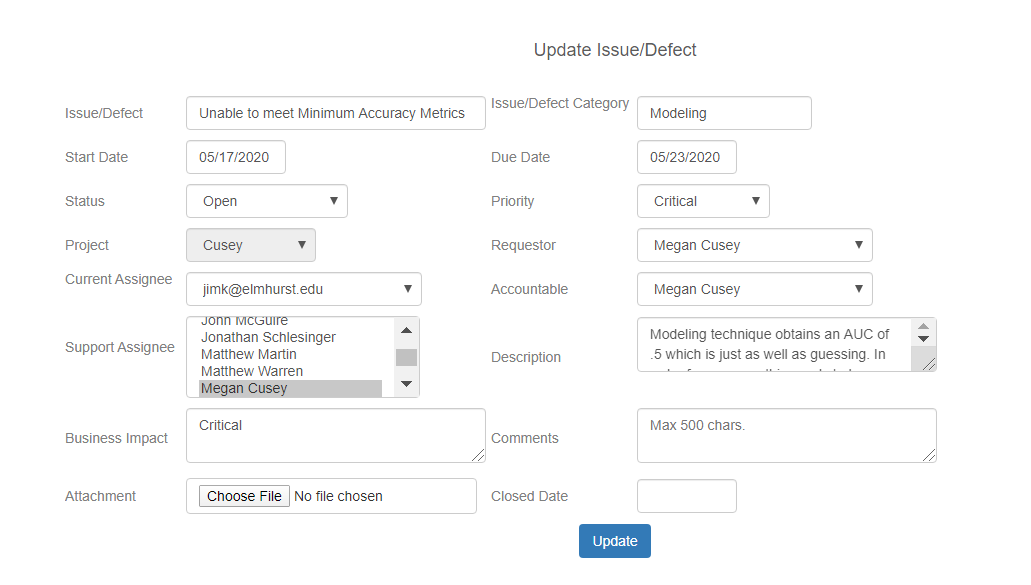
Results: I successfully updated PRIMMS to close out all activities and quality criteria for the Data Prep and Modeling Phases of the CRISP-DM life cycle. My project metrics updated as a result.



1. Within your PRIMMS project add the following*:*
   * An open decision
   * An action item
   * An issue/defect

Results: I successfully added an open decision, action item, and issue/defect in PRIMMS. Please see screenshots below:





1. In the PRIMMS Resource Center download the project management training documents of interest. Reflect upon the foundational chapter and consider why project management skills are important for the data scientist. It is also recommended that read the 1993 Quality Gate paper.

Results: Project Management helps monitor the quality, scope, time and cost of each project. It also helps to identify requirements, set achievable goals, and creates a flexible roadmap for project execution. All of these concepts that are applied to many different project structures (IT, construction, marketing, etc.) also apply to data science endeavors. For example, project management helps set clear expectations of what the overall goal of a data mining project needs to achieve and the quality requirements that accompany it. If you were building a machine learning decision making tool to flag fraud, applying project management will assist you in knowing that the scope of the project includes a classification problem with a single binary output (0 or 1) for fraud. Perhaps the project objective is a little less generic and they only want to flag fraud on commercial accounts with more than 500,000 in the bank account or more than 500,000 credit limit, etc. It may also state that it is more important to focus on identifying positive fraud cases rather then missing a true fraudulent case. It needs to have at least 75% of recall. Project management also tracks cost and time it takes to work on a project which is typical in deciding if the project should be prioritized ahead of others or is feasible to complete.