# DAT 650 Use Case Descriptions Document

# Employee Attrition

The human resources department within GE has recently become aware that many high-potential employees have left the company to pursue other opportunities. This awareness was raised by many middle managers and supported by the recent increase in job postings. Given the need to remain competitive and the total cost and time required to train new employees, a need is present to identify how talent may be retained within GE. In the current environment the average cost of attrition for an individual is 80% of annual their salary.

The current data environment is an HR web based desktop system that contains information about all employees, current and past, including their Attrition status of YES – they have left or NO – they have not left. This envornment includes meta data about each employee. It is maintained by HR staff via a web based java client server application. The data is stored in an Oracle database in transactional form. The IT department has built a data warehouse that is updated each evening with the current day data. The HR team uses this Datawarehouse data as the source to their reports and have the ability to have an adhoc extract to select data into excel for their unique research needs. The HR extracts are limited in rows as well as fields that have been corporately pre approved for extraction.

GE has compiled a file for the purpose of this pilot project using the extract tool. The dataset provided,Employee\_Attrition.csv, includes information on employees which the GE HR team believes to be relavent to analyzing this problem.

The HR team would like to determine if this data can be used to identify Attrition of employees that may leave. It is important to be able to understand attrition drivers for meta data like high performers, role types and other pertinent groupings which comes from the analysis.

The pilot will need to only show basis for this data to be able to describe and generally identify employees that may leave. The management team expects to make a GO or NO GO business decision based on the pilot recommendation. If there is a GO, then GE will allocate new project dollars to arrange for GE resources to develop a full enterprise deployed predictivie analytic model. Note that the results of this pilot will be used as a basis for that next project.