**Please download and open in Excel:**

* *Retention.txt*
* *Training.txt*

Trodecco is a company operating in LA. Over time, the company has found that after recruiting and training employees, many of their brightest and most experienced employees leave the company for other opportunities. Trodecco wants to understand what causes employees to leave and how to retain their employees. They compiled data for the past 3 years in the file: Retention.txt

These data include (for each employee):

* **Satisfaction:** (0.0- 1.0) Employee’s self-reported satisfaction level, collected through a recent survey at the company (1.0 is highest)
* **Last\_evaluation** (0.0 – 1.0): The employee’s latest annual performance evaluation (1.0 is highest)
* **Project**: (No. of Projects) Average number of projects per month the employee was assigned
* **Monthly\_hours**: (hours) Average number of hours the employee worked each month
* **Time**: (years) Number of years the employee has been with the company
* **Left**: (1=yes, 0 = no) Whether the employee quit
* **Promotion**: Whether the employee was promoted in the last 5 years (1=yes, 0 =no)
* **Department**: Department the employee worked in.
* **Salary**: (low, medium, high). These scales are relative to the particular job role and department.

The second file “Training.txt” gives the average cost of training a new employee for each department by salary level.

Use these data to better learn why employees are leaving and make recommendations to prevent turnover below. Where appropriate, always provide brief justifications of your response and thinking.

* 1. (4 pts) Compare retention rates, i.e., the % of employees that have NOT left, for all departments. Which department has the lowest retention rate? How much is it?
  2. (6 pts) Use the training cost table in the “Training\_Cost.txt” to find the training cost for each employee.
  3. (3 pts) What is the total training cost spent for all the employees?
  4. (3 pts) Out of this total, what is the percentage spent on those who has left the company?
  5. (2 pts) There is a belief that the best employees , i.e., the ones with the highest evaluation scores, are the ones that are leaving the company. Do you agree? Please quantify your answer with an appropriate p-value.   
     ***Hint:*** *You want to compare the evaluation scores between people who stayed with the company and people who left and see if there is a statistically significant difference. You can do this with an A/B test where “Variant 1”correspond to the people who stayed with the company and “Variant 2” corresponds to the people who left.*
  6. (4 pts) A smart manager points out that the problem might be that the best employees might be leaving because they feel underpaid. If that were true, you might expect a (statistically significant) difference in evaluation scores between employees who left and employees who stayed, particularly at the “high” salary level. Do you agree? Please quantify your result with difference, p-value and the significant level of 5%.
  7. (4 pts) Many of those who left complained that they had long hours at work. If this is the case, the average number of monthly hours of those who left would be more than the average number of monthly hours of those who haven’t left. Is the statement true for each department? Briefly explain your findings.

* 1. (2 pts) Consider the following metric: Average satisfaction level. Based on our discussion in class regarding good KPI’s, do you think that this is a good KPI? Briefly explain.