

1. What is the problem you want to solve?
 - Understand reasons for employee turnover based on from information provided about their Salary, performance etc.
2. Who is your client and why do they care about this problem? In other words, what will your client DO or DECIDE based on your analysis that they wouldn't have otherwise?
 - The client is losing employees and facing high turnover. By visualizing the data, client can see the trends and observations. The client can make decisions based on this analysis.
3. What data are you going to use for this? How will you acquire this data?
 - The dataset I am using contains information about each employee. The dataset meets all the criteria to be acceptable as "Tidy data"
 - o Each variable must have its own column.
All metrics associated with an employee are in their own column
 - o Each observation must have its own row.
Each employee has a separate row
 - o Each value must have its own cell.
Each cell in this dataset represent a single value
4. In brief, outline your approach to solving this problem (knowing that you may not know everything in advance and this might change later). This might include:
 - a. Is this a supervised or unsupervised problem?
 - This is a supervised learning problem. We want the model to predict if a present employee is at risk of leaving the company.
 - b. If supervised is it a classification or regression problem?
 - This is a classification problem. Employee needs to be classified as "Could Stay" or "Could Leave"
 - c. What variable is it you are trying to predict?
 - There is a variable called "Left" which contains Boolean values '0' for the employees that did not leave and '1' for employees
 - d. What variables will you use as predictors?
 - Satisfaction_Level
 - Last_Evaluation
 - Number_of_Projects
 - Average_Monthly_Hours
 - Time_Spend_Company
 - Work_Accident
 - Left
 - Promotion_Last_5_Years
 - Department

- Salary

e. What will be your training data?

I will split the same dataset into Training and test sets.

5. What are your deliverables? Typically, this would include code, along with a paper and/or a slide deck.

- Code along with slide deck