

Project 3 Proposal: IBM Employee Attrition Rate

Question/need:

- The purpose of the model/system I plan to build is to predict employee churn, what factors influence churn and to predict if a specific employee will churn.
- Who benefits from exploring this question or building this model/system? The organization benefits from knowing churn rate because they can then take actions to prevent attrition. Additionally, employees themselves benefit from exploration of this question because they may be compensated appropriately by the organization ahead of time as an effort to retain the employee.

Data Description:

- What dataset(s) do you plan to use, and how will you obtain the data? The dataset I will be using is a fictional IBM dataset created for IBM data scientists, uploaded to [Kaggle](#):
- What is an individual sample/unit of analysis in this project? What characteristics/features do you expect to work with? For this model an individual unit of analysis is whether or not the employee will churn considering features such as monthly, daily, and hourly rate, age, distance from home, environment satisfaction, job satisfaction, years in current role, years at company, years with current manager, years since last promotion, work life balance etc.
- If modeling, what will you predict as your target? The target for this model is whether or not the employee will churn.

Tools:

- How do you intend to meet the tools requirement of the project? Tools needed include excel/google sheets, tableau, python, matplotlib, seaborn

MVP Goal:

- What would a [minimum viable product \(MVP\)](#) look like for this project? A preliminary model including some of the features and their impact on whether or not an employee will churn.