MSDS 6306: Doing Data Science

Case Study 02

Due: One hour before Live Session 15

Description: DDSAnalytics is an analytics company that specializes in talent management solutions for Fortune 1000 companies. Talent management is defined as the iterative process of developing and retaining employees. It may include workforce planning, employee training programs, identifying high-potential employees and reducing/preventing voluntary employee turnover (attrition). To gain a competitive edge over its competition, DDSAnalytics is planning to leverage data science for talent management. The executive leadership has identified predicting employee turnover as its first application of data science for talent management. Before the business green lights the project, they have tasked your data science team to conduct an analysis of existing employee data.

Your team has been given a dataset (CaseStudy2Data.zip) to conduct exploratory data analysis (EDA) to determine factors that lead to attrition. You should identify (at least) the top three factors that contribute to turnover. There may or may not be a need to create derived attributes/variables. The business is also interested in learning about any job role specific trends that may exist in the data set (e.g., "Data Scientists have the highest job satisfaction"). You can also provide any other interesting trends and observations from your analysis. The analysis should be backed up by robust experimentation and (where applicable) appropriate visualization. Experiments and analysis must be conducted in R. If needed, consult with the business owner to get clarifications on the dataset or business goals.

Deliverables:

This is a group project so it's the responsibility of the group members to collaborate accordingly. Team work is important.

The due date for submission is 1 hour before live session 15. During live session 15, the group will have up to 20 minutes to present their case study report. The goal is to communicate the findings of the project in a clear, concise and scientific manner.

Create a GitHub repository with an RMarkdown file containing an executive summary, introduction to the project, all supporting code and analysis, and the group presentation. Submit a link to the GitHub repository via the space provided for the Case Study 02 page in 2DS.