SMM635 — Mid-Term Project
Mid-Term Project Package

Submission deadline	November 14th at 4:00 PM
Template	Mandatory, available as .docx and .tex format
Wordl imit	True, see template documents

1 Project Description

Consider the Human Resources Dataset.[1] Then, make the following assumptions. First, you are a junior business analyst working for a global consultancy company. Second, the data regards the company AlfaBetaGamma, your employer's client. Third, you are supposed to contribute to creating a deck to use for a meeting with the client. Specifically, you are supposed to create a set of visualizations that help to address the following questions raised by the client:

1. How do employees' performance scores change within and across managers?

2. What is the overall diversity profile of the organization?

3. What are our best recruiting sources if we want to ensure a diverse organization?

4. What are the most frequent factors for job termination?

5. Are there areas of the company where pay is not equitable?

2 Deliverables

By November 14th, the students must submit:

- One Matplotlib visualization for each of the five questions above

- The Python code necessary to replicate the Matplotlib visualizations

- An executive summary covering the following aspect

  - The design choices behind the proposed visualizations
  - The rationale behind the individual design choices
  - A set of actionable business analytics recommendations that are grounded in the main results of the project

The students are required to use the provided template — available in .docx and .tex format — and respect the word limits applying to the various boxes/sections of the template. The total length of the document is 2,500 words.

---

[1]https://www.kaggle.com/datasets/rhuebner/human-resources-data-set

3   Key Features of the Dataset

The dataset contains 311 observations (aka, employees) and 36 variables, whose meaning is described in Table 1.

4   References and Contextual Information

The students are warmly encouraged to liaise with the module leader about possible readings and contextual information on the role of data visualization in human resources analytics.

Table 1: Codebook of the Variables

Variable	Description
Employee Name	Employee's full name
EmpID	Employee ID is unique to each employee
MarriedID	Is the person married (1 or 0 for yes or no)
MaritalStatusID	Marital status code that matches the text field MaritalDesc
EmpStatusID	Employment status code that matches text field EmploymentStatus
DeptID	Department ID code that matches the department the employee works in
PerfScoreID	Performance Score code that matches the employee's most recent performance score
FromDiversityJobFairID	Was the employee sourced from the Diversity job fair? 1 or 0 for yes or no
Salary	The person's yearly salary. $ U.S. Dollars
Termd	Has this employee been terminated - 1 or 0
PositionID	An integer indicating the person's position
Position	The text name/title of the position the person has
State	The state that the person lives in
Zip	The zip code for the employee
DOB	Date of Birth for the employee
Sex	Sex - M or F
MaritalDesc	The marital status of the person (divorced, single, widowed, separated, etc)
CitizenDesc	Label for whether the person is a Citizen or Eligible NonCitizen
HispanicLatino	Yes or No field for whether the employee is Hispanic/Latino
RaceDesc	Description/text of the race the person identifies with
DateofHire	Date the person was hired
DateofTermination	Date the person was terminated, only populated if, in fact, Termd = 1
TermReason	A text reason / description for why the person was terminated
EmploymentStatus	A description/category of the person's employment status. Anyone currently working full time = Active
Department	Name of the department that the person works in
ManagerName	The name of the person's immediate manager
ManagerID	A unique identifier for each manager.
RecruitmentSource	The name of the recruitment source where the employee was recruited from
PerformanceScore	Performance Score text/category (Fully Meets, Partially Meets, PIP, Exceeds)
EngagementSurvey	Results from the last engagement survey, managed by our external partner
EmpSatisfaction	A basic satisfaction score between 1 and 5, as reported on a recent employee satisfaction survey
SpecialProjectsCount	The number of special projects that the employee worked on during the last 6 months
LastPerformanceReviewDate	The most recent date of the person's last performance review.
DaysLateLast30	The number of times that the employee was late to work during the last 30 days
Absences	The number of times the employee was absent from work.