Mini Project

Sub-Topics	Max Marks
1.1 Data manipulation using Python (30 marks) 1.2 Statistical Analysis using Python (30 Marks) 1.3 Exploratory Data Analysis using Python (40 marks)	**100**

Domain:

HR Analytics

About:

HRWorks Pvt Ltd is a Bangalore based start-up that commenced its operations in the summer of 2010. HRWorks was conceived by a team of HR practitioners.

HRWorks sees itself as the first true end-to-end Talent Acquisition Solutions organization which has the passion to bring together decades of experience in Technology Consulting and Talent Acquisition areas to usher in a paradigm shift in the way Talent Acquisition is practiced in today's ultra-demanding business environment. HRWorks not only advises its customers on where their Talent Acquisition practices are, but also recommends and implements individually tailored, viable solutions using analytics.

Business process re-engineering with its three tenets − People Capability, Process Maturity and Technology Adoption − form the core ability of the company to provide customers with an enterprise-class customized solution to address their Talent Acquisition challenges. They bring in deep domain knowledge of how Talent Acquisition happens in corporates and provide viable recommendations to their customers.

Challenges:

Client service is all about the quality of the people involved in delivering business. However, one of the major challenges for HRWorks and its clients revolved around managing a quality workforce. Organizations spend tremendous amount of time and energy to create a homogenous environment where people thrive and succeed. Despite all the effort to keep an environment that is conducive, people leave organizations in search of better opportunities. In order to fill the vacuum, HR is bound to recruit new talent, thus forming a vicious circle in between attrition and recruitment; and to mitigate this, organizations keep trying to bridge the gap by strengthening their recruitment processes and creating a culture of inclusivity.

HRWorks wanted to find a unique solution which goes beyond the process aspect of human resource management. At first, HRWorks identified and prioritized the renege problem and put forward in a subtle way:
“A significant proportion of the candidates do not join the company that has made an offer. If we can identify them in advance, then companies don’t have to waste their resources.”

What is Expected?

HRWorks supports several information technology (IT) companies in India with their talent acquisition. One of the challenges they face is that about 30% of the candidates who accept the job offer do not join the company. This leads to huge loss of revenue and time as the companies initiate the recruitment process again to fill the workforce demand.

Being a data analyst, you must come up with a first step document that lists output of your exploratory analysis, any issues or problems you may see with data that need follow up, and some basic descriptive analysis that you think highlights important outcomes/findings from the data. Based on your findings, the next level of analysis will be charted out.

.

Data Dictionary:

- Sno: Sl number auto increment
- Candidate_Ref: Candidate reference number
- DOJ_Extended: Date of Joining of extended
- Duration_to_accept offer: Duration to accept the offer by candidate
- Notice_Period: Notice period of previous employer
- Offered_Band: E1 < E2 < E3 and so on
- Percent_hike_expected_in_CTC: expected hike by candidate
- Percent_hike_offered_in_CTC: hike offered by joining organisation
- Percent_difference_CTC: difference between expected and offered
- Joining_Bonus: any joining bonus offered.
- Candidate_relocate_actual: relocating required or not
- Gender: Gender of candidate
- Candidate_Source: How candidate applied or reached
- Rex_in_Yrs: years of exp
- Postal_Code : Postal Code of the area of organisation
- Domicile: Name of Home state
- Location: Name of the location
- Region_Name: Name of the Region
- Age: Age of candidate
- Date_of_Interview : Date on which Interview was conducted.
- LOB: Line of business

Target variable:

- Status: joined or not.

Task 1.1(Data Manipulation using Python)

Here are some indicative types of analysis you can perform. Please note that this is not an exhaustive list, you may add more

- Come up with appropriate results for the following:
    - Analysis of percentage joined of offer released.
    - What are the key drivers that influence the candidate joining/not joining a company?
    - Are there specific locations where candidates are not joining?
    - Joining status depends on the duration to accept an offer?
    - Hike offered has an impact on joining status.?

Task 1.2 (Statistical Analysis using Python)

- Descriptive statistics for both numerical and categorical and draw few insights from them.
- Perform relevant hypothesis testing (t, chi-Square, Anova tests)

TASK 1.3 (Exploratory Data Analysis)

Data Preparation/Analysis tasks including (but not limited to) the following.

- Univariate, Bi- Variate Analysis and Multi- Variate Analysis
- Missing values identification and treatment
- Outlier analysis and treatment
- Data scaling using min-max and/or Z-score normalisation
- Data transformation
- Feature Engineering

Deliverables/Submission guidelines

1. You have to prepare a power point presentation with screenshots of outputs (10 -15 slides)
2. Mention Problem Statement and Your approach to the problems
3. You need to submit all the code files .
4. All comments/inferences/insights/reasons for doing a particular tasks etc should be written as a ‘markdown text’, but **NOT** using a comment lines with # or ‘’’.
5. Submit the code file as HTML file format (Also, submit Jupyter Notebook).
6. Upload all the deliverables in the LMS