**Observations On Data**

**Table 1**

1. Job ID Creation Date are all blank.
2. Job ID Approval Date are all blank.
3. Job ID Release Date have missing values.
4. Only complete observations will be considered as valid data.
5. It is assumed that Job ID is the unique ID for all jobs being posted.
6. Is URL related to Recruiter Code?
7. Recruiter Code have missing values; only complete observations will be considered as valid data.
8. Min & Max CTC have values of 0 and some are ‘3’, ‘4.5’. These will be rejected.

**Table 2**

1. For the same Job ID, there can be multiple Candidate IDs.
2. Rest of the fields have a number of missing values which will not be considered.
3. Data integrity issues – In some cases Offer Acceptance Date is later than Joining Date. For eg. In the 12th row, Acceptance Date is almost 2 years after Date Of Joining. Such data will be rejected.

**Table 3**

1. Data has integrity issues. For eg. Cand Shortlisted has a value of 11 followed by Candidates Screened has a value of 7. However Rejected At Shortlisting Stage has no value. Is it to be assumed that 4 candidates got rejected at shortlisting stage? Likewise for other stages.

**Table 4**

1. Number of missing values in Date Of Cand Record Creation. Null values will be rejected for sake of valid data.

**Table 5**

1. Not understood; has only 1 record.

**Table 6**

1. Employee Code has both numeric and alphanumeric values.
2. Missing Values in Perf Rating.
3. Perf Rating has both numeric values as well as character values like ‘Good’ ‘Excellent’ etc. These need to be converted into one type – either numeric or nominal.

Overall the data has a lot of missing values and the same needs to be provided for sake of completeness. Also, data integrity issues exist in a number of tables. Integrity rules to be defined by EazeHR and violation of these rules will lead to rejection of data.