# Report:

- Dirty Data:-

  - **Job**: column in the data set there is unnecessary information such as the yearly package, technology, company name, and work type (remote), etc added in the title.Consistency
  - **no_of_application**: this column has a value that should not be present ('hours', 'minutes'). Consistency
  - **posted_day_ago**: In this column, time is added in the form of a string for Exampal (9 hours, 8 minutes). Validity
  - **Alumni**: with the alumni count the string is present with the count value ('company alumni'). Validity
  - **company_id**: This column is completely blank. Completeness
  - **Hiring_person**: In this column, there are some nicknames added within parentheses Consistency
  - **linkedin_followers**: In the linkedin_followers column the 'linkedin_followers' string is present with numbers and numbers are separated with (' , ') Validity

- Messy Data:-

  - **Full_time_remote**: In this column involvement and level are both combined in one column and separated by ('·'). Validity
  - **No_of_employ**: In this column employees_count and industry are both combined in one column and

the 'employees' string is added extra separated ('·'). Validity

- o **Location**: In the location column country, state and city are added in one column and separated by (',') Validity
- o **Job**: In the Job column some are starting with the capital letter and some are in small letter Consistency

  -**company_name**: in the company_name also there is inconsistency in the name some of the letters are starting with a small letter and some of them are stares with the capital letter. Consistency

- From correlation matrix alumni and linkedin_followers columns have correlation of -0.50 so they are dependent on each other.
- Alumni are between 0.0-0.1 and 1.1-1.3 no vales are in between showing some error.
- LinkedIn followers have alumni near 0 and 1.2e6
- Most employees are full-time employees.
- Internship employees have more number of followers and mid-senior level employees have least number.
- There are less number of applications for volunteer and more for full-time.
- Lead Java Software engineer job is the job with highest job postings.
- The job-postings belong to IT industry.
- Most job postings  are in Bangalore.