Individual Final Project

|  |
| --- |
| **HR Analytics: Job Change of Data Scientists: This dataset contains information about data scientists who are looking**  **For a job change. The dependent variable is "target" with two factors: "0" indicating that the employee is not**  **looking for a job change, and "1" indicating that the employee is looking for a job change. The independent**  **variables include both qualitative and quantitative variables such as "city", "education\_level","company\_size", and**  **"last\_new\_job".** |



|  |  |
| --- | --- |
| **enrollee\_id** | Their unique ID |
| **city** | The city they are working |
| **city\_development\_index** | The city’s development index |
| **gender** | The gender of that person |
| **relevent\_experience** | Their relevant work experience |
| **enrolled\_university** | Their enrolled university |
| **education\_level** | Their education level if is masters or bachelors |
| **major\_discipline** | Their major in university |
| **experience** | Total work experience |
| **company\_size** | The company size they are working |
| **company\_type** | Which type is the company they are working in |
| **last\_new\_job** | Their last new job |
| **training\_hours** | The hours they are trained |
| **target** | If they want to change the company.1 means they want a job change whereas 0 means they are not looking for a job change |

Questions:

1.Find if there is a false positive score using target score

2.Do plot if there is any intersection between the genders

3.Do Wilk's Lambda and F test for the dataset

4.Check if you can predict the number of your employees are leaving the company for sure

5.Do PCA on the data and find if the people are leaving the company or not