Perform **Exploratory Data Analysis (EDA)** on the data-set given below. Consider **Salary** as a target variable. You should submit a Jupyter Notebook(i.e. .ipynb file) along with a 5 Page report in [**Innomatics Report Template**](https://docs.google.com/presentation/d/1se6cqgykNSVhM9Sa0vOld76Sh_8ih8FX/edit#slide=id.p1). Zip both the files together and upload on LMS.

Download the dataset from this link. [**CLICK HERE TO DOWNLOAD DATASET**](https://docs.google.com/spreadsheets/d/1yKll2Vg_mRRrNP_yrUaGr53y1VZMygQo/edit?usp=sharing&ouid=101183190533718048323&rtpof=true&sd=true)

The detailed description about the dataset can be found here. [**Dataset Description**](https://docs.google.com/document/d/1xvo6kN5Q4QG-ezZKrNwg2YMD6S_fwlmz/edit?usp=sharing&ouid=101183190533718048323&rtpof=true&sd=true)

**If you are facing any difficulty in performing EDA, follow the steps mentioned below:**

**Step - 1 -** Introduction -> Give a detailed data description and objective

**Step - 2 -** Import the data and display the head, shape and description of the data.

**Step - 3 -** Univariate Analysis -> PDF, Histograms, Boxplots, Countplots, etc..

* Find the outliers in each numerical column
* Understand the probability and frequency distribution of each numerical column
* Understand the frequency distribution of each categorical Variable/Column
* Mention **observations** after each plot.

**Step - 4 -** Bivariate Analysis

* Discover the relationships between numerical columns using Scatter plots, hexbin plots, pair plots, etc..
* Identify the patterns between categorical and numerical columns using swarmplot, boxplot, barplot, etc..
* Identify relationships between categorical and categorical columns using stacked bar plots.
* Mention **observations** after each plot.

**Step - 5 - Research Questions**

* Times of India article dated Jan 18, 2019 states that “*After doing your Computer Science Engineering if you take up jobs as a Programming Analyst, Software Engineer, Hardware Engineer and Associate Engineer you can earn up to 2.5-3 lakhs as a fresh graduate.*” Test this claim with the data given to you.
* Is there a relationship between gender and specialization? (i.e. Does the preference of Specialisation depend on the Gender?)

**Step - 6 -** Conclusion

**Step - 7 -** (Bonus) Come up with some interesting conclusions or research questions (such as step-5).

**NOTE:** Mention **observations** after each plot.

**-------------------------------------------------- END OF TASK---------------------------------------------------**

***Disclaimer: You can share this project on your LinkedIn to showcase your internship work. You can also add it to your portfolio via GitHub.***

***Submission: You should submit a Jupyter Notebook(i.e. .ipynb file) along with a 5 Page report (PPT or Word Document). Zip both the files together and upload on LMS.***