INNOMATICS®
RESEARCH LABS

INNOVATION. AUTOMATION. ANALYTICS

PROJECT ON

Exploratory Data Analysis
(AMCAT Dataset)

Submitted By,
Atluri Vigneshwar Rao
Intern ID: IN9249218

About me

- Hi there!! I am Atluri Vigneshwar Rao, a data science enthusiast, currently learning various things to crack an opportunity to go further.
- Apart from this, I possess the problem solving ability and I am good at learning new things that makes me an ideal candidate to follow my dreams.
- I have previously worked as a ML enginner intern at Data point solutions company and right now I am doing my internship at innomatics research labs, I am doing the internship to refine my skills at their highest level.
- Feel free to reach out! You can do so by following the below links:
  - https://github.com/vigneshkewk
  - https://www.linkedin.com/in/vignesh-atluri-9615b32b7/

INNOMATICS RESEARCH LABS

OBJECTIVE OF THE PROBLEM

- This exploratory data analysis of “AMCAT DATASET” focuses on understanding various factors that might influence the level of salaries indicated in the dataset. We consider education and experience, gender, specialization, and job roles and observe how they are related in order to understand a factor that influences higher or lower levels of salaries. The critical steps which indicate the analysis involved creating a mental image of the data, establishing trends and patterns, testing many hypotheses post observations to finally build insightful results which could be used as guidelines for any decision making process that could further calibrate salary prediction models.

SUMMARY OF THE DATASET

- There are ***38 columns*** in total that are used to find the individual impacts on salary.
- Out of 38 columns, there are ***29 numerical columns*** and ***9 categorical columns***.
- With ***3998 Datapoints*** that make our analysis to the optimal insights with all the necessary information.

DATA CHECKS TO PERFORM

- Check missing values, duplicated values and various different columns.
- Check the datatypes and also look at the unique number of columns.
- Check statistics of data set
- Check various categories present in the different categorical column.
- Drop unnecessary columns

Conclusion

The analysis of the AMCAT dataset provides insightful conclusions regarding salary trends, specialization, and skill sets of fresh graduates in different roles. Here are some key takeaway:

- The average salaries for roles like Programming Analyst, Software Engineer, Hardware Engineer, and Associate Engineer align with industry standards as reported in the Times of India.
- Graduates with Computer Science and IT-related specializations tend to command higher salaries, reflecting the strong demand for these skills in the tech industry.
- There is an uneven distribution of male and female graduates across different job roles, indicating potential gender biases or disparities in certain specializations and job roles.
- Technical skills like programming, computer science, and other related fields are strongly correlated with higher salaries, emphasizing their significance in securing well-paying jobs.

THANK YOU

INNOMATICS RESEARCH LABS