

## Case Studies – Fortnightly Task 3

*Vishal Patnaik Damodarapatruni - s3811521*

*21/09/2020*

#####

### INTERVIEWEE DETAILS

- Name: UDAY SAI TYADA
- Company: YourOwnRoom Pvt Ltd.
- ROLE: Machine Learning Engineer and Software Developer
- Linked in Profile: <https://www.linkedin.com/in/uday-sai-tyada-31601a12b/>

### INTERVIEW DURATION

12:50 AM to 1:51 AM – 1 hour and 1 minute.  
20/09/2020

### INTERVIEW QUESTIONS & ANSWERS

- INTERVIEWER:  
What biggest data set you have ever processed, and how, what outcomes did you obtain?  
INTERVIEWEE:  
“One of the biggest data sets that I dealt with is all the real estate data, and all the amenities that are provided in almost every area of Bengaluru. The data is initially cleaned by making essential conversions and handling typos and missing values. Then the outliers are handled. The primary outcome of the exercise was to build a recommendation system based on people's search history and making the each real-estate property in terms of vectors.”
- INTERVIEWER:  
What is mostly needed by a data scientist?  
INTERVIEWEE:  
“A data scientist must have an in-depth understanding in knowledge related to a business. One of the key prerequisites for a data scientist is to have an analytical mindset with a solid statistical background and great information on data structures and machine learning algorithms. They should be solid in Python or R and ought to be agreeable in dealing with huge data sets.”
- INTERVIEWER:  
What are the hours like?

INTERVIEWEE:

“10:00 AM to 7:00 PM”

➤ INTERVIEWER:

How do you handle missing data? What imputation techniques do you recommend?

INTERVIEWEE:

“The method or technique varies from data set to data set or from analysis to analysis. I recommend, Imputation Using (Mean/Median) Values, Imputation Using (Most Frequent) or (Zero/Constant) Values as some of the best ways to tackle missing data problems”

➤ INTERVIEWER:

What is the curse of big data?

INTERVIEWEE:

“Can't see results fast.”

➤ INTERVIEWER:

What solution do you propose for identifying plagiarism?

INTERVIEWEE:

“Word vectorization and tf-idf are some of the great ways to find the plagiarism.”

➤ INTERVIEWER:

Most admired person? That person must belong to data science community.

INTERVIEWEE:

“Andrew NG.”

➤ INTERVIEWER:

Most admired company?

INTERVIEWEE:

“Google, because of its RND programs.”

➤ INTERVIEWER:

What biggest areas of questions or opportunities that you like to tackle with?

INTERVIEWEE:

“Real-estate, medicine, angel investment.”

➤ INTERVIEWER:

Most proud moment in your career?

INTERVIEWEE:

“Being a data scientist, I drew a risk analysis for my company, which saves nearly 3 Lakh Indian rupees a month.”

## ADDITIONAL DISCUSSION

- He told me about internships.
- What kind of internships or internship providers we should apply, how to choose them, and how to apply them?
- How the internship environment would be?
- He also told something about the role of a data science team and its key requirements.