

# Assignment 6

## Dataset

Use data\_dictionary.xlsx and data.csv and metadata.json

## Tasks to be performed

Focus on applying the learnt data analytics concepts and try to share your findings on following aspects:

- What are the highest paid Skills in the US market?
- What are the job categories, which involve above mentioned niche skills?
- Applying clustering concepts, please depict visually what are the different salary ranges based on job category and years of experience.

1. Data Preparation [10 marks]
  - a. Schema must be normalized; field types must be appropriate as per fields available. Proper data model
2. Identification of Variables [10 marks]
3. Variable Selection [10 marks]
4. Feature Engineering [10 marks]
5. Missing value or data [10 marks]
6. Analysis [50 marks]