

## Question 1:

**Intent :** Learn the basics of how web applications are built and what technologies could be used for web applications.

Suppose you were given to build a web application for <https://www.boats.com/> from scratch as a startup . Imagine you would be solely responsible to build it's frontend, backend, design database, setup media storage, deployment, ci/cd automation and everything else. Thus, do some research on your part and write down your response to the following questions.

- a. Which backend framework would you prefer to use and why?
- b. Which frontend framework would you prefer to use and why?
- c. Which database would you prefer to use and why?
- d. What would you use for version control of the codebase ?
- e. Which platform would you prefer for media storages ( eg. storage of product images) ?
- f. Where would you deploy your application and which web server would you use and why?

## Question 2:

**Intent:**

We would like you to learn the basics of python and data science to load a dataset, read it and perform some operations to find multiple mathematical metrics such as average, maximum, minimum and such.

Here is a dataset for autos.

<https://drive.google.com/file/d/1QP21K5tiJAjt5NA7W2FxSe9Wam9-tlcQ/view?usp=sharing>

Flow:

1. Download this dataset.
2. Write basic python script to load csv and read it as dataframe
3. Use the dataframe to perform following:
  - a. Find Average price of autos ( using **price** column of dataset)
  - b. Print the list of different possible types of **VehicleType** found in dataset
  - c. Calculate and print lowest **yearOfRegistration** and highest **yearOfRegistration**
  - d. Find and print standard deviation of column **kilometer**
  - e. Draw a bar graph to represent count of different type of column **brand**
  - f. Find out which **VehicleType** is sold minimum and maximum
  - g. Create a pie chart to represent different types of **gearbox** count

**Question 3:**

**Intent:** Check your problem solving approach on machine learning

Consider the dataset on Question 2. Now,

A client has solicited your services to develop a machine-learning model that can forecast the approximate value of their customers' used cars. The objective is to provide accurate quotations to customers on the price to offer for the purchase of their used cars. You have been furnished with a dataset of used cars, and your task is to:

1. conduct exploratory data analysis to identify crucial features that will be utilized in the model.
2. Please justify the selection of these features and aim to incorporate as many as possible.
3. kindly identify any potential challenges or limitations you anticipate/encounter during the feature selection process. (if any)
4. (Bonus) Try to propose a good model you feel would be able to best fit the features you have selected to make predictions.

**Question 4:**

Given a string, find the length of the longest repeating subsequence, such that the two subsequences don't have the same string character at the same position, i.e. any  $i$ th character in the two subsequences shouldn't have the same index in the original string.

**Examples:**

**Input:** `str = "abc"`

**Output:** 0

There is no repeating subsequence

**Input:** `str = "aab"`

**Output:** 1

The two subsequence are 'a'(first) and 'a'(second).

Note that 'b' cannot be considered as part of subsequence as it would be at same index in both.

**Input:** `str = "aabb"`

**Output:** 2

**Input:** `str = "axxy"`

**Output:** 2

