1. **\*What are the Steps of a Data Science Project?\***
   1. **Define Problem Statement**

Predict if given credit card transaction is **Fraud** or not

* 1. **Get Data**

Acquire Credit card transaction data

* 1. **Clean/preprocess Data**

Clean the data to get it in desired format, delete extra columns or derive new columns

* 1. **Do Exploration of Data (EDA)**

Plot your data in graphs to get better understanding of it and then do further cleaning

* 1. **Train Data Science Model**

Select an Algorithms that suits your problem and train it using your data.

* 1. **Evaluate Model Performance**

Compare the predicted result with actual result to find accuracy, precision, recall etc.

* 1. **Tune Mode to get best result**

Do further tuning of model by changing algorithm parameter, selecting different algorithm or doing data preprocessing etc.

* 1. **Deploy best model.**

Deploy your model to do prediction on real time data.

1. **What is Data Science? What skills it needs to become a data Scientist?**

**Programming** (Python is Best and easy to learn)

**Statistics** (Probability, Statistics, Linear Algebra)

**Domain Knowledge** (Don’t worry there will be Business Analyst to help you)

1. **What are data Science Platforms you are familiar with?**

You should make yourself familiar with following platform. Go to their website and try to see what value they offer.

* Kaggle :- Organizes Data Science Competitions.
* Aanlytics Vidhya:- Data Science Blogs and Competitions
* Towards Data Science: Data Science Knowledge Articles.

1. **What is most and second most important thing in a Data Science Project?**

* Problem Statement is the most Important thing in Data Science
* Data is Second most Important Thing in Data Science

1. **What is difference between Machine Learning, Deep Learning and NLP?**

* ML is about Building systems that can learn from experience.

Example predicting likelihood (probability of happening) of disease.

* Deep Learning is about Building systems that can learn from experience on large data set. Deep Learning is mostly used with Image and Video data.   
  Example is Facial Recognition, Sign Language understanding from video.
* NLP is about Building systems that can understand **language.**

Example: Language translation devices, keypad autocorrect.

1. **What tools you need to install to get started with the Data Science Journey?**

* Python
* Python IDE (Jupyter, Spyder, pycharm etc.)
* Python Libraries (numpy, scipy, pandas, sk-learn)

1. **What are some Data Science Use Cases in real life?**

* Product recommendation (Amazon customer who bought this also bought this)
* Sentiment Analysis (Happy and Sad with your android app based on user comment)
* Anomaly detection
* Fraud detection
* Weather forecasting
* Stock market forecasting
* Object recognition (images & video)
* Speech recognition

1. **What is the use/purpose of following python libraries for data science?**
   1. Numpy
   2. Scipy
   3. Pandas
   4. Sci-kit Learn
   5. NLTK/OpenNLP
   6. OpenCV
   7. TensorFlow
   8. Pytorch
   9. Keras
2. **Which of following data visualization tool you are familiar with?**
   1. Tableau
   2. Power BI
   3. Pyplot
   4. seaborn
3. **What is the full form of following data science terms?**
   1. ML
   2. DL
   3. NLP
   4. EDA (Exploratory Data Analysis)
4. **What is data Pre-processing, why is it necessary?**