<https://www.analyticsvidhya.com/blog/2016/09/40-interview-questions-asked-at-startups-in-machine-learning-data-science/>

**Differences between supervised and unsupervised learning.**

### ****Important skills to have in Python with regard to data analysis?****

### ****What is Selection Bias?****

### ****What do you understand by the term Normal Distribution?****

### ****What do you understand by statistical power of sensitivity & specificity and how do you calculate it?****

### ****What are the differences between overfitting and under fitting and how will you avoid it?****

### ****How data cleaning plays a vital role in analysis?****

### ****Differentiate between univariate, bivariate and multivariate analysis.****

### ****What is Cluster Sampling?****

### ****Can you cite some examples where a false positive is important than a false negative?****

### ****Can you cite some examples where a false negative important than a false positive?****

### ****Can you cite some examples where both false positive and false negatives are equally important?****

### ****Explain cross-validation.****

### ****What are the various classification algorithms?****

### ****How can outlier values be treated?****

### ****What are the various steps involved in an analytics project?****

### ****During analysis, how do you treat missing values?****

### ****How will you define the number of clusters in a clustering algorithm?****

### ****In any 15-minute interval, there is a 20% probability that you will see at least one shooting star. What is the proba­bility that you see at least one shooting star in the period of an hour?****

### ****A certain couple tells you that they have two children, at least one of which is a girl. What is the probability that they have two girls?****

### ****What are Eigenvectors and Eigenvalues?****