1. Describe the main principles of statistical learning in short and concise words using each of these terms.

Solution.
Statistical Learning refers to problem solving framework involving vast amount real-world data and statistical theory principles. Some of the modern day examples include email spam classification and stock market prediction. In short, we can call it as learning from data.

Broadly, these learning problems can be classified as either **supervised learning** – learning process guided by input and **outcome variables** or **unsupervised learning** – learning methodology where we observe only **input features** and without any measurements of responses. Based on the input data, the supervised learning methods can be of two types: **classification** - problem of identifying to which of a set of a category a new data belongs or **regression** – estimating a function from the inputs and numeric/ continuous output variables. In the context of classification tasks, we usually think of target variables as **qualitative** (Dogs vs. Cats classification) usually categorical labels and in regression analysis, outputs are usually **quantitative variables** i.e., numerical values.

The general learning framework for statistical models involve creating data in the form of **training data**, which has both input features and outcomes and **test data**, which has access only the inputs variables. As the name suggests, training data is used to train a model to learn the underlying data distribution. The trained model is usually used for either **prediction** tasks that involve predicting outcomes for new data points (such as test data) or for **inference** which primarily involves understanding how a stochastic model generates its estimates.

The process of building statistical models usually involve making different assumptions to approximate the underlying function. Usually, **parametric models** assume that a finite set parameters define/control the data distribution whereas **non-parametric models** assume that the underlying data cannot assumed with a finite set of parameters. Good examples for parametric and non-parametric models include linear regression models and k-nearest neighbour method respectively. ■