Continuous Data:

<https://towardsdatascience.com/understanding-feature-engineering-part-1-continuous-numeric-data-da4e47099a7b>

Categorical Data:

<https://towardsdatascience.com/understanding-feature-engineering-part-2-categorical-data-f54324193e63>

Feature Engineering:

<https://towardsdatascience.com/exploratory-data-analysis-feature-engineering-and-modelling-using-supermarket-sales-data-part-1-228140f89298>

<https://towardsdatascience.com/feature-engineering-and-data-preparation-using-supermarket-sales-data-part-2-171b7a7a7eb7>

**Creating Machine Learning Models**

<https://towardsdatascience.com/creating-machine-learning-models-b48bb72a791f>

Data Cleaning:

<https://towardsdatascience.com/data-cleaning-with-python-and-pandas-detecting-missing-values-3e9c6ebcf78b>

Interview Question Links:

<https://github.com/iNeuronai/interview-question-data-science->

<https://github.com/iNeuronai>

Dependent and Independaent Variables

<https://medium.com/machine-learner/independent-and-dependent-variables-1-10d8553ad616>

EDA:

<https://towardsdatascience.com/exploratory-data-analysis-eda-techniques-for-kaggle-competition-beginners-be4237c3c3a9>

**Statistics**:

Sampling Distribution: <https://www.youtube.com/watch?v=hqiMcHqlZ4s>

**EDA**It is most important Part of ML, taking 60% of time. You cant start building model without exploring and cleaning it.  
  
Here is what you mostly should do:  
  
1)Data Profiling Start with exploring all variables  
Check data types of all variables  
Check mean mode median for continuous variables  
Find null values in each column and replace it with  
mean or median for continuous variables or mode for  
categorical variables.  
2)Outliers - Check Kurtosis , which will give outliers. Boxing can also be used to find out Outliers. Treatment of outliers is very important  
3)Normalization - In many cases we have to Normalize the dataset, to use a common scale  
4)Data Visualization - Visualizing all variables by plotting them, often used plot are given below:  
Box plot for categorical variables  
Scatter plot  
Histogram plot  
Heatmap plot  
Bar plot  
Factor plot  
5)Converting categorical variables to numeric - Dummy values