Steps for data science

1. Formulate the business problem and identify the need and expectation for data science
2. Identify data sources and collect information
3. Import the data and prepare it to desired format
4. Glimpse at the dataset
   1. Missing percentage
   2. Variable type
   3. Unique levels
   4. Eyeball some sample records
5. Univariate analysis
   1. Tables (frequency, percentage)
   2. Charts (histograms, pie charts)
6. Bivariate analysis
   1. Statistics (Pearson correlation, chi-squares, information value, etc.)
   2. Tables (two-way crosstabs)
   3. Charts (box plot, scatter plot)
7. Clustering analysis
8. Pre-processing and ETL