

### Data Analysis Sequence

#### Question to Answer

Data Acquisition	What ? Where ? How ? Format ?
Data Cleaning	Missing Values? Outliers? String Format? Normalization
Descriptive Analysis	Reporting Aggregate Measures - Scatter plots, Histograms, Statistical Summaries
Exploratory Analysis	Relationships between variables - Inferential, Predictive, Causal, Mechanistic data analysis
Data Modeling	Training and Testing
Quality Assessment of Models	Overfitting ? Underfitting ?
Interpretation of Results	Domain Expertise

### Data Acquisition Pipeline

#### SOURCES      FORMAT      REPRESENTATION

Internet	Plain Text (unstr)	List, Tuple, Set
File / Pickled File	CSV	Array, Matrix
Database	HTML/XML	Frame, Series
	JSON	Dictionary
	Tabular	



By **alexlam**  
[cheatography.com/alexlam/](https://cheatography.com/alexlam/)

Not published yet.  
Last updated 26th February, 2018.  
Page 1 of 1.

Sponsored by **Readable.com**  
Measure your website readability!  
<https://readable.com>