

Cheatography

data scientist statistics Cheat Sheet

Simple statistic

moyenne	faire la moyenne d'une liste de valeur
somme	faire la somme
eccart type	faire l'eccart type

Probabilité

Complement	$P(A) + P(A') = 1$
Intersection	$P(A \cap B) = P(A)P(B)$
Union	$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Regression

linéaire	linéaire
logistique	logistique
logarithmique	logarithmique

Variability

Percentiles	A measure that indicates the value below which a given percentage of observations in a group of observations falls.
Quantiles	Values that divide the number of data points into four more or less equal parts, or quarters.
Interquartile Range (IQR)	A measure of statistical dispersion and variability based on dividing a data set into quartiles. $IQR = Q3 - Q1$
Variance	The average squared difference of the values from the mean to measure how spread out a set of data is relative to mean.
Standard Deviation	The standard difference between each data point and the mean and the square root of variance.
Standard Error	An estimate of the standard deviation of the sampling distribution.