

Mesures statistiques : mesures de tendance centrale, de dispersion et de caractérisation d'une distribution

Mesures	Définition	Caractéristiques de la mesure	Excel/Calc (français, anglais)
MESURES DE TENDANCE CENTRALE : résumer en une seule valeur la grandeur typique, le milieu ou le centre d'un ensemble de données			
Moyenne	Somme des valeurs d'une série de données divisée par le nombre de données dans la série <i>(point d'équilibre de la série)</i>	<ul style="list-style-type: none"> La plus familière et la plus couramment utilisée Influencée par la valeur de toutes les observations Valeur peut être biaisée par quelques valeurs extrêmes La plus efficace quand il s'agit de faire certaines inférences sur la population à partir de données d'échantillonnage Moins représentative pour distributions asymétriques 	=MOYENNE(plage de données) =AVERAGE(plage de données)
Médiane	Valeur qui occupe la place du milieu dans une série de données ordonnée	<ul style="list-style-type: none"> Facile à calculer et à comprendre Affectée par le nombre d'observations mais pas par la valeur de celles-ci Généralement moins efficace que la moyenne lorsqu'elle est utilisée à des fins d'inférence statistique Plus représentative que la moyenne pour distributions asymétriques 	=MEDIANE(plage de données) =MEDIAN(plage de données)
Mode	Valeur la plus fréquente dans une série de données	<ul style="list-style-type: none"> Généralement moins utilisée Peut ne pas exister dans certaines séries d'observations OU il peut y en avoir plus d'un dans d'autres distributions N'est pas influencée par les valeurs extrêmes d'une distribution 	=MODE(plage de données) =MODE(plage de données)
MESURES DE DISPERSION : évaluer la dispersion à l'intérieur d'une distribution : éparpillement, étalement, dispersion ou variation			
Étendue	Différence entre la plus grande et la plus petite valeur d'une série	<ul style="list-style-type: none"> La plus simple à calculer N'implique que les valeurs extrêmes 	=MAX(plage de données) - MIN(plage de données) =MAX(plage de données) - MIN(plage de données)
Écart type	Dispersion des valeurs d'une série autour de la moyenne	<ul style="list-style-type: none"> La plus fréquemment utilisée Valeur de chacune des observations affecte la valeur de l'écart type Valeur peut être fortement influencée par quelques valeurs extrêmes 	=ECARTYPE(plage de données) =STDEV(plage de données)
Variance	Moyenne des carrés des écarts entre chaque valeur d'une série et la moyenne des valeurs de la série	<ul style="list-style-type: none"> Carré de l'écart type 	=VAR(plage de données) =VAR(plage de données)
Coefficient de variation	Rapport de l'écart type à la moyenne, en pourcentage	<ul style="list-style-type: none"> Permet de comparer la dispersion entre deux séries de données Donne une idée de l'homogénéité ou l'hétérogénéité d'une distribution : En dessous de 15%, bonne homogénéité des données 	=ECARTYPE(plage de données) / MOYENNE(plage de données) * 100 =STDEV(plage de données) / AVERAGE(plage de données) * 100
MESURES DE CARACTÉRISATION D'UNE DISTRIBUTION : décrire la forme d'une distribution			
Coefficient d'asymétrie	Sens et force de l'asymétrie d'une distribution	<ul style="list-style-type: none"> Valeur nulle = distribution symétrique Valeur positive = distribution asymétrique à droite (c'est-à-dire les données se regroupent principalement dans les basses valeurs) Valeur négative = distribution asymétrique à gauche (c'est-à-dire les données se regroupent principalement dans les hautes valeurs) Plus le coefficient en valeur absolue est grand, plus l'asymétrie est marquée 	=COEFFICIENT.ASYMETRIE(plage de données) =SKEW(plage de données)
Coefficient d'aplatissement (Kurtosis)	Type et force de l'aplatissement d'une distribution	<ul style="list-style-type: none"> Valeur nulle = distribution ayant un aplatissement normal Valeur positive = distribution moins aplatie qu'une distribution normale (c'est-à-dire les données sont plus fortement regroupées autour de la moyenne que la normale) Valeur négative = distribution plus aplatie qu'une distribution normale (c'est-à-dire les données sont moins fortement regroupées autour de la moyenne que la normale) Plus le coefficient en valeur absolue est grand, plus l'aplatissement (ou le non aplatissement) est marqué 	=KURTOSIS(plage de données) =KURT(plage de données)