Deskriptive Datenanalyse

Variable ist kategorial, sie hat (qualitativ) verschiedene Ausprägungen; Messung auf Nominalskalenniveau			
Kennwerte	Bedeutung	Formel	Befehl Excel
Häufigkeit	Häufigkeit eines Ereignisses bzw. Häufigkeit einer Variablenausprägung	f *frequency*	=zählenwenn(Bereich; „AusprägungVariable“)
Häufigkeits-verteilung	Häufigkeiten aller Ausprägungen der Variablen		=HÄUFIGKEIT(Daten; Klassenobergrenzen) Matrixbefehl mit Strg+Shift+Enter bestätigen
Relative Häufigkeit	Gibt die Häufigkeit einer Variablenausprägung relativ zu allen Ereignissen an	$\frac{f(Ausprägung\ Variable)}{f(alle\ Beobachtungen)}$	
Wahrscheinlichkeit	Grenzwert der relativen Häufigkeit für unendlich viele Ereignisse	$p = \frac{f(Ereignis)}{f(alle\ Ereignise)}$	
Modus/Modalwert	Die Ausprägung der Variablen, die am häufigsten ist		=MODALWERT(Bereich)

Variable hat verschiedene Ausprägungen, die eine Reihenfolge bilden; Messung auf Ordinalskalenniveau			
Kennwert	Bedeutung	Formel	Befehl Excel
Median	Mittlerer Wert in einer Rangreihe der Messwerte (wenn es keinen einzelnen mittleren Wert gibt, Mittel der beiden angrenzenden Werte	Median	=MEDIAN(Bereich)
Quartile	Wert, der die unteren 25%/75% vom Rest der Werte trennt	25% Quartil 50% Quartil (Median) 75% Quartil	=QUARTILE(Bereich;1) =QUARTILE(Bereich;2) =QUARTILE(Bereich;3)
Quartilsabstand	Maß für die Streuung der Messwerte	QA = Quartil75 – Quartil25	= QUARTILE(Bereich;3) - QUARTILE(Bereich;1)

Variable ist kontinuierlich, sie hat quantitativ verschiedene Ausprägungen; Messung auf Intervallskalenniveau			
Kennwert	Bedeutung	Formel	Befehl Excel
Mittelwert (arithmetisches Mittel)	Wert, von dem die Messwerte insgesamt am wenigsten abweichen	$M_X = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$	=MITTELWERT(Bereich)
Varianz	Mittlere quadrierte Abweichung der Messwerte vom Mittelwert	$S^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - M_X)^2$ (Groß S)	=VAR.P(Bereich)
Streuung/Standard-abweichung	Mittlere Abweichung der Messwerte vom Mittelwert	$S = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - M_X)^2}$	=STABW.N(Bereich)
z-Standardisierung	Transformation der Werte auf eine Skala mit M=0 und S=1	$z = \frac{(x_i - \bar{x})}{S_X}$	=STANDARDISIERUNG(x; Mittelwert();STABW.N())

Kennwerte/Parameter der zugehörigen theoretischen Verteilungen: 1. Normalverteilung		
Mittelwert in der Population	μ	
Schätzung mittels Stichprobe	$\mu = M_X$	
Varianz in der Population	σ^2	
Schätzung Varianz mittels Stichprobe	Bei einer Stichprobe $\hat{\sigma}^2 = s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - M_X)^2$ Bei zwei Stichproben $\hat{\sigma}^2 = {s_{iB}}^2 = \frac{(n_1 - 1) * {s_1}^2 + (n_2 - 1) * {s_2}^2}{n_1 + n_2 - 2}$ Vereinfacht bei n1 = n2 $\hat{\sigma}^2 = {s_{iB}}^2 = \frac{{s_1}^2 + {s_2}^2}{2}$	=VAR.S(Bereich) =eigene Formel
Streuung in der Population	σ	
Schätzung Streuung mittels Stichprobe Bei zwei Stichproben: Streuung innerhalb der Bedingungen	Bei einer Stichprobe $\hat{\sigma} = s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - M_X)^2}$ Bei zwei Stichproben $\hat{\sigma} = s_{iB} = \sqrt{\frac{(n_1 - 1) * {s_1}^2 + (n_2 - 1) * {s_2}^2}{n_1 + n_2 - 2}}$	=STABW.S(Bereich) =eigene Formel
2. t-Verteilung (Verteilung der standardisierten Mittelwerte)		
Mittelwert der Mittelwerte	$\mu = M_{Mx}$	
Geschätzt mittels Stichprobe	$M_{Mx} = M_X$	
Streuung der Mittelwerte = Standardfehler des Mittelwerte	σ_M	
Geschätzt mittels Stichprobe	$SE_M = \frac{s}{\sqrt{n}}$	=STABW.S(Bereich)/wurzel(n)
Konfidenzintervall (t-Verteilung)		=t.inv(0,05;df) * SE_M =konfidenz.t(0,05; stabw.s(Bereich);n)
Konfidenzintervall (z-Verteilung = Normalverteilung)		=konfidenz.norm(0,05; stabw.s(Bereich);n)
Freiheitsgrade	df	bei 1 Mittelwert df = n-1 bei 2 Mittelwerten df = $n_1 + n_2 - 2$