|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Univariate Statistik - Grundlagen** | | | | | |
| **Nominalskala** (keine Reihenfolge) ledig, verheiratet, verwitwet, geschieden | | **Ordinalskala** (klare Reihenfolge) sehr teuer, teuer, günstig, sehr günstig | | **Intervallskala** (Masseinheit, kein Nullp.) Zeit, Temperatur | |
| **Verhältnisskala** (Intervallsk. + Nullpunkt) | | Kosten, Anzahl Mitarbeiter, Alter, Eink. | | Prozentuale Veränderung! | |
| **Qualitativ** = Nicht metrisch (Nominal + Ordinal) | | | **Quantitativ** = metrisch (Intervall + Verhältnis) | | |
| **Diskret** = bestimmte Werte auf der Skala (1 oder 2 Kinder, aber nicht 1.5), **stetig** = jeder Wert ist definiert (BIP in CHF) | | | | | |
| **Relative Häufigkeit** (in %) | |  | | \* 100 | |
| **Dichte** (Histogramm) (Klassenbreite in 10000? 2,3,5,10…) | |  | | Klasse mit der höchsten Dichte = **Modalklasse** | |
| **Univariate Statistik - Mittelwerte** | | | | | |
| **Arithmetisches Mittel** (Durchschnitt) z.B Durchschnittlich verlangter Preis | |  | | Daten in L1 => [2nd data] =>  1-VarStats: L1, ONE => x̅ | |
| **Gewichteter Durchschnitt** (Klassenmitte + abs. Häufigkeit als Gew.)z.B Durchschnittlich bezahlter Preis | |  | | L1: Klassenmitte, L2: abs. Häufigkeit => [2nd data] =>  1-VarStats: L1, FRQ: L2 => x̅ | |
| **Median** (interpolieren) Perzentile (100), Dezile (10), Quart. (4) | |  | | L1: Klassenoberg., L2: rel. H. kumuliert,  => [2nd data] => LinReg: L2,L1,Y => f(50) | |
| rechtsschief: Modus < Median < arithmetisches Mittel / linksschief: arithmetisches Mittel < Median < Modus | | | | | |
| **Geometrisches Mittel** (zeitliche Abfolge) z.B Wachstum, Umsatzentw., Renditen z.B bei 5 Faktoren 5. Wurzel! | |  | | L1: Faktoren => L2 = ln(L1) =>  1-VarStats: L2 (allenf. Gewicht in L3) => Resultat = ex̅ (StatVars: x̅) | |
| **Durchschnittliche Wachstumsrate** | |  | |  | |
| **Univariate Statistik - Streuungsmasse** | | | | | |
| **Spannweite** | |  | | Höchster Wert – kleinster Wert | |
| **Quartilsabstand** (Spannweite der mittleren 50% der Messwerte) | |  | | Kumulierte Häufigkeit: 75% - 25% | |
| **Empirische Varianz** und **Empirische Standardabweichung** | |  | |  | |
| **Empirische Varianz gew.** | |  | |  | |
| **Empirische Standardabweichung gew.** | |  | | L1: Daten, L2: Gewicht => [2nd data] =>  1-VarStats: L1, FRQ: L2 =>Sx | |
| **Variationskoeffizient** Streuung im Verhältnis zum Mittelwert | |  | | In % angeben | |
| **Univariate Statistik – Indexzahlen** (z.B Durchschnittliche Veränderung über 10 Jahre = 10. Wurzel von Index2000/Index1990) | | | | | |
| **Zusammengesetzter Index** (Alles Index oder Preis!) | | | **Umbasierung** (gleiches Jahr 100) | | **Verkettung** |
|  | | |  | |  |
| **Mit Preisen** Gewicht = P0 \* M0 |  | |
| **Mit Index** Index t0 = 100 |  | |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Bivariate Statistik – Korrelation** (funktionaler Zusammenhang zwischen x und y / **positiv** = +/+ und -/-, **negativ** +/- und -/+) | | | | |
| **Kovarianz** (Daten zentrieren) und **empirische Kovarianz** | |  |  | |
| **Korrelationskoeffizient** (Daten Standardisieren) | |  | L1: x, L2: y => LinReg: L1, L2, YES => r | |
| 0 - ± 0.25 => kein funktionaler Zusammenh. / ± 0.25 - ± 0.5 => schwach / ± 0.5 - ± 0.75 => mittel ± / ± 0.75 - ± 0.1 => stark ± | | | | |
| **Bivariate Statistik – Regression** | | | | |
| **Kleinste Quadrate** (Gerade möglichst nahe bei allen Punkten) | |  | 2 Gleichungen => Sys-Solv => y = ax+b | |
| **Lineares Modell** | |  | LinReg ax+b | y\*x+z (andere Var) |
| **Polynom Modell** | |  | QuadraticReg  CubicReg | |
| **Logarithmisches Modell** | |  | LnReg a+blnx | |
| **Potenz Modell** | |  | PwrReg ax^b | |
| **Exponentielles Modell** | |  | ExpReg ab^x | |
| **Bestimmtheitsmass** **R2 oder r2** (in %) LinReg => r2 (wie stark wird y von x beeinflusst?) | |  | | |
| **Zeitreihenanalyse** | | | | |
| **Gleitender Durchschnitt**  (Ordnung entsprechend der Anzahl Perioden) | |  | | |
| **Trendwert, Saisonkomponente additiv**  LinReg: L1, L2, YES => L3 = f(L1) | | => arith. Mittel pro Saison | | |
|  | Prognose = Trendwert f(L1) + Saisonkomp. |
| **Saisonkomponente multiplikativ** | | => geom. Mittel pro Saison (Wurzel!) | | |
|  | Prognose = Trendwert f(L1) \* Saisonfaktor |
| **Saisonbereinigung additiv** | |  |  | |
| **Saisonbereinigung multiplikativ** | |  |  | |
| **Veränderung in Prozent** | | L1: erste Zahl löschen, L2: Letzte Zahl löschen => L3 = ((L1/L2) -1) \*100 | | |