

Übungen – Einheit 2:

Aufgabe 2.1:

Unser Datensatz beinhaltet Informationen zum Alter in Form des Geburtsjahres (ybirth). Berechnen Sie für jede Erhebungseinheit (Person) das Alter. Bedenken Sie, dass die Umfrage im Jahr 2009 durchgeführt wurde. Erstellen Sie eine Häufigkeitsverteilung! Fassen Sie dabei das Merkmal Alter in Intervalle mit einer Intervallbreite von 5 Jahren zusammen (wobei die Kategorie über 15 bis 20 Jahren die 1. Kategorie darstellt). Erstellen Sie ein Histogramm und formatieren Sie das Histogramm ansprechend. Wie ist das Merkmal „Alter“ skaliert? Handelt es sich dabei um ein stetiges oder ein diskretes Merkmal? Unterscheidet sich die Skalierung von Alter und Geburtsjahr?

Berechnungen: siehe Excel.

Das Merkmal „Alter“ ist metrisch und verhältnisskaliert. Weiters ist es stetig (man kann auch 25,43 Jahre alt sein), liegt aber nur in diskretisierter Form vor).

Das Geburtsjahr ist ebenfalls metrisch, allerdings "nur" intervallskaliert. Es handelt es sich dabei um ein diskretes Merkmal, da es nur ganzzahlige Werte annehmen kann.

Aufgabe 2.2:

Berechnen Sie für das Merkmal „Alter“ die Lage- und Streuungsmaße, die wir in der Präsenzeinheit behandelte haben (nämlich Mittelwert, Median, Quartile, Spannweite, Varianz, Standardabweichung und Variationskoeffizient). Interpretieren Sie Mittelwert, Median, Quartile und Spannweite sehr kurz (max. 1 Satz). Berechnen Sie die Lage- und Streuungsmaße einzeln (mit den entsprechenden EXCEL-Befehlen), sowie mit der Analysefunktion „Populationskenngrößen“.

Berechnungen: siehe Excel.

Interpretation:

Die Personen der Stichprobe sind durchschnittlich 49.5 Jahre alt (Mittelwert).

Die durchschnittliche Person ist 49 Jahre alt (Median). Das bedeutet, dass zumindest 50% der Personen 49 Jahre oder jünger sind, und zumindest 50% der Personen 49 Jahre oder älter sind.

Das 1. Quartil beträgt 36 Jahre. Das bedeutet, dass zumindest 25% der Personen 36 Jahre oder jünger sind, und zumindest 75% der Personen 36 Jahre oder älter sind.

Das 3. Quartil beträgt 63 Jahre. Das bedeutet, dass zumindest 75% der Personen 63 Jahre oder jünger sind, und zumindest 25% der Personen 63 Jahre oder älter sind.

Die Spannweite beträgt 83 Jahre. Das bedeutet, dass der Unterschied zwischen der ältesten und der jüngsten Person in der Stichprobe 83 Jahre beträgt.

Aufgabe 2.3:

Die Variable „seval“ gibt die Einschätzung der Personen über die „Angemessenheit des Wohnraums in der Wohneinheit“ an. Es gibt 5 mögliche Ausprägungen (Much too small; A bit too small; Just right; A bit too large; Much too large). Starten Sie mit dem Tabellenblatt „Variablen (Namen)“. Kodieren Sie

die Ausprägungen des Merkmals mit den Zahlen 1 (für Much too small) bis 5 (für Much too large). Verwenden Sie dafür die EXCEL-Befehl „WENN“. Wie ist die Variable skaliert?

Berechnungen: siehe Excel.

Das Merkmal ist ordinal skaliert, da eine Reihenfolge (eine Ordnungsbeziehung von „much too small“ bis „much too large“) hergestellt werden kann. Der Abstand zwischen zwei Merkmalsausprägungen ist aber nicht klar definiert und nicht interpretierbar.

Aufgabe 2.4:

Erstellen Sie ein Stabdiagramm der Variable „seval“! Achten Sie darauf, dass das Diagramm ausreichend viel Informationen aufweist, sodass die Abbildung verständlich ist (Titel, Fußnote, Achsenbeschriftung). (Abbildungen sollten auch lesbar sein, wenn man den Fließtext zur Abbildung nicht liest!) Berechnen Sie jene Lage- und Streuungsmaße, die für diese Variable sinnvoll sind, und interpretieren Sie diese!

Berechnungen: siehe Excel.

Interpretation:

Sinnvolle Lagemaße sind der Median und die Quartile:

Das 1. Quartil beträgt „just right“. Das bedeutet, dass zumindest 25% der Personen ihren Wohnraum als „just right“ oder kleiner, und zumindest 75% der Personen ihren Wohnraum als „just right“ oder größer bezeichnen würden.

Das 2. Quartil (der Median) beträgt „just right“. Das bedeutet, dass zumindest 50% der Personen ihren Wohnraum als „just right“ oder kleiner, und zumindest 50% der Personen ihren Wohnraum als „just right“ oder größer bezeichnen würden.

Das 3. Quartil beträgt „just right“. Das bedeutet, dass zumindest 75% der Personen ihren Wohnraum als „just right“ oder kleiner, und zumindest 25% der Personen ihren Wohnraum als „just right“ oder größer bezeichnen würden.

Aufgabe 2.5:

Kodieren Sie die Variable „sex“ so, dass der Mittelwert (der kodierten Variable) dem Anteil der Männer entspricht!

Berechnungen: siehe Excel.