



Universität Ulm,
Institut für Epidemiologie und Medizinische Biometrie,
D-89070 Ulm

**Institut für Epidemiologie und
Medizinische Biometrie**

Leiter: Prof. Dr. D. Rothenbacher
Schwabstr. 13, 89075 Ulm
Tel. +49 731 / 5026901

Übung 1 im Fach "Biometrie / Q1"

Aufgabe 1:

Um bei Ihren Untersuchungen die adäquaten statistischen Größen zu berechnen, müssen Sie eine Typisierung Ihrer Merkmale vornehmen.

Im Folgenden sind Variablen des Übungsdatensatzes mit ihren Ausprägungen aufgeführt. Geben Sie zu jeder Variablen den Merkmalstyp und die entsprechende Skala an. [s. Merkblatt "Merkmalstypen und Skalen"]

LÖSUNG:

Variable	Ausprägungen	Skala	Merkmalstyp
BMI (Body Mass Index)	Zahlen mit Nachkommastellen Einheit: kg/m ²	Intervallskala (Verhältnisskala)	Stetiges Merkmal
GESCHL (Geschlecht)	1 = männlich 2 = weiblich	Nominalskala	Alternativmerkmal
PUBSTKL (Klassierung nach pubertärer Stufe)	1 = prae 2 = intra 3 = post	Ordinalskala	Rating-Merkmal

Aufgabe 2:

a)

Der BMI beschreibt das Verhältnis von Körpergewicht zur Körpergröße und korreliert eng mit der Menge körperlichen Fettgewebes. Er gilt weltweit als Orientierung für das Sollgewicht eines Erwachsenen, hat sich eingeschränkt auch zur Bestimmung von Übergewicht im Kindes- und Jugendalter international durchgesetzt.

Erstellen Sie eine Variable mit dem Namen "BMI_NEU", die den BMI-Wert für jeden Patienten enthält.

Berechnen Sie den Wert nach folgender Formel:

$$\text{BMI} = \frac{\text{Körpergewicht in kg}}{(\text{Körpergröße in m})^2}$$

Da die Körpergröße im Datensatz in cm vorliegt, muss sie vor dem Quadrieren durch 100 dividiert werden.

LÖSUNG: *Transformieren → Variable berechnen*

$$\text{BMI_NEU} = \text{KGWKG} / ((\text{KGRCM} / 100) ** 2)$$

Aufgabe 3:

Um Ihr Studienkollektiv beschreiben zu können, werten Sie unter anderem das stetige Merkmal "BMI" deskriptiv aus.

Berechnen Sie dazu folgende statistische Kenngrößen. Tragen Sie Ihre Ergebnisse auf zwei Nachkommastellen gerundet in die Tabelle ein:

LÖSUNG: *Analysieren → Deskriptive Statistiken → Häufigkeiten*

Statistiken		
BMI		
N	Gültig	219
	Fehlend	0
Mittelwert		32,47139223
Median		31,85185185
Standardabweichung		5,542684702
Minimum		23,25501863
Maximum		50,76934242
Perzentile	25	28,14760026
	50	31,85185185
	75	35,73145329

Ergebnisse auf zwei Nachkommastellen gerundet:

Mittelwert	Std.Abw.	Minimum	25 % Q	Median	75 % Q	Maximum
32.47	5.54	23.26	28.15	31.85	35.73	50.77

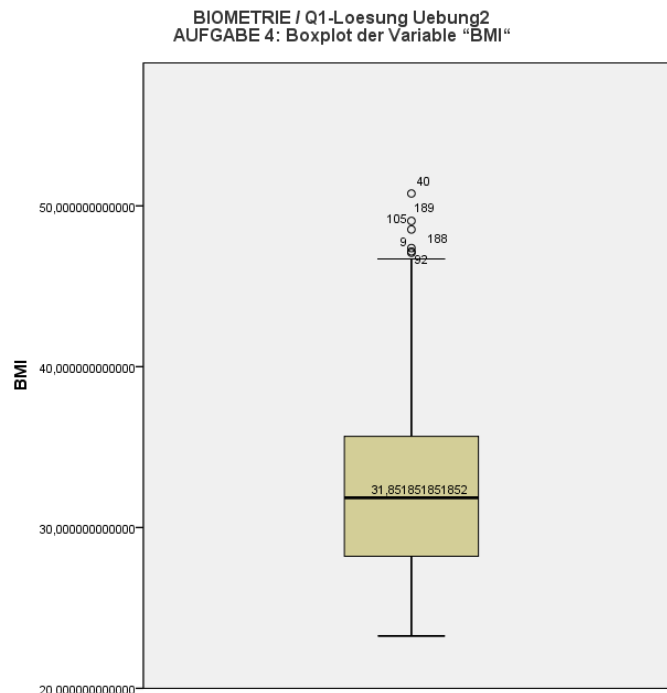
Die Einheit der Werte ist "kg/m²".

Aufgabe 4:

Zeichnen Sie mit den in Aufgabe 3 ermittelten Werten einen horizontalen Box-&whisker-Plot, um die Verteilung der Ausprägungen in der Stichprobe abzubilden!

Erstellen Sie mit SPSS einen Box-&whisker-Plot zur Kontrolle Ihrer Grafik. Wählen Sie **Diagramme → Diagrammerstellung** und dann *Boxplot → Einfacher Boxplot*.

LÖSUNG:

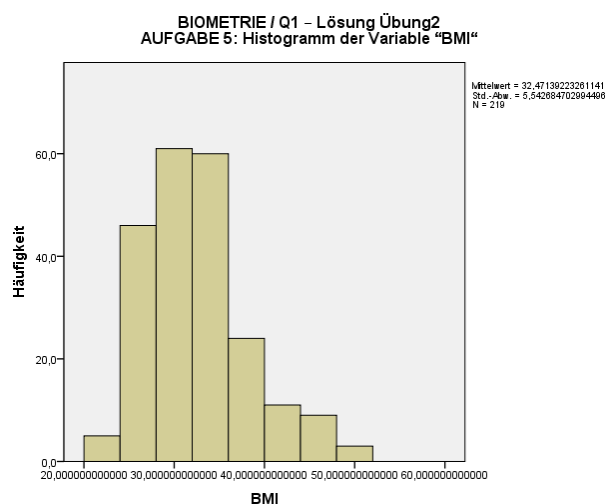


Aufgabe 5:

Erstellen Sie mit Hilfe von SPSS ein Histogramm für "BMI". Gehen Sie dabei nicht über die deskriptive Statistik, sondern über **Diagramme → Diagrammerstellung**.

Drücken Sie den Button **Elementeigenschaften**. Danach unter *Statistiken* den Button **Parameter festlegen**, um in die Einstellungen für die Anzeige des Histogramms zu gelangen. Tragen Sie bei *Klassengrößen* die **Intervallbreite 4** ein, um die Einteilung auf der x-Achse zu steuern.

LÖSUNG:



Interpretation:

Der BMI ist in unserer Stichprobe **eingipflig** und **rechtsschief** verteilt. Wegen der Eingipfligkeit ist ein **Maß der zentralen Tendenz** aussagekräftig. Da die Verteilung aber **nicht symmetrisch** ist, fällt die Wahl auf den **Median**.

Aufgabe 6:

Die Einteilung der Patienten wurde nach den Richtlinien der "American Diabetes Association" durchgeführt. Patienten deren Blutzuckerwert nüchtern zwischen 100 mg/dl und 140 mg/dl lag, wurden in die Kategorie "Impared Fasting Glucose" (IFG) eingestuft.

Werten Sie die qualitative Variable "IFG" deskriptiv aus. Tragen sie die absoluten und die relativen Häufigkeiten in die Tabelle ein. Runden sie die relativen Häufigkeiten auf zwei Nachkommastellen.

Um zwei Nachkommastellen in der Tabelle anzuzeigen: **Rechtsklick** auf Tabelle → **Inhalt bearbeiten** → **In separatem Fenster**. In dem neu geöffneten Fenster die Werte markieren, die geändert werden sollen. Dann wieder **Rechtsklick** → **Zelleneigenschaften** → **Reiter Formatwert** und anschließend Dezimalstellen auf 2 schalten.

LÖSUNG: *Analysieren → Deskriptive Statistiken → Häufigkeiten*

IFG					
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	nein	207,00	94,52	94,52	94,52
	ja	12,00	5,48	5,48	100,00
Gesamt		219,00	100,00	100,00	

Interpretation:

5% unseres Patientenkollektives wurden nach ihren Blutzuckerwerten in die Kategorien "Impared Fasting Glucose" eingestuft.

Aufgabe 7:

Leptin ist ein Hormon, das nicht nur appetitzügelnd wirkt, es kann auch Fettzellen dazu bringen, ihre Fettreserven zu verbrennen und es bremst den Aufbau von Fett im Körper.

Patienten deren Blutzuckerwert nüchtern über 100 mg/dl und zwei Stunden nach dem Oraln Glukosetoleranztest über 140 mg/dl lag, wurden in die Kategorie "Blutzucker pathologisch" eingestuft. Untersuchen Sie, ob sich in unserer Stichprobe eine Tendenz zwischen den Patientengruppen mit normalen und pathologischen Blutzuckerwerten (BZPATH: 0 = normal, 1 = pathologisch) bezüglich des Leptins abzeichnet.

Erstellen Sie eine gruppenspezifische Auswertung mit SPSS. Führen Sie dazu eine deskriptive Auswertung der Variablen "LEPTIN", gruppiert nach der Variablen „BZPATH“ durch und lassen Sie sich die Ergebnisse graphisch mit Box-&-Whisker-Plots darstellen. Notieren Sie folgende statistischen Kenngrößen.

LÖSUNG:**Analysieren → Deskriptive Statistiken → Explorative Datenanalyse**

Verarbeitete Fälle

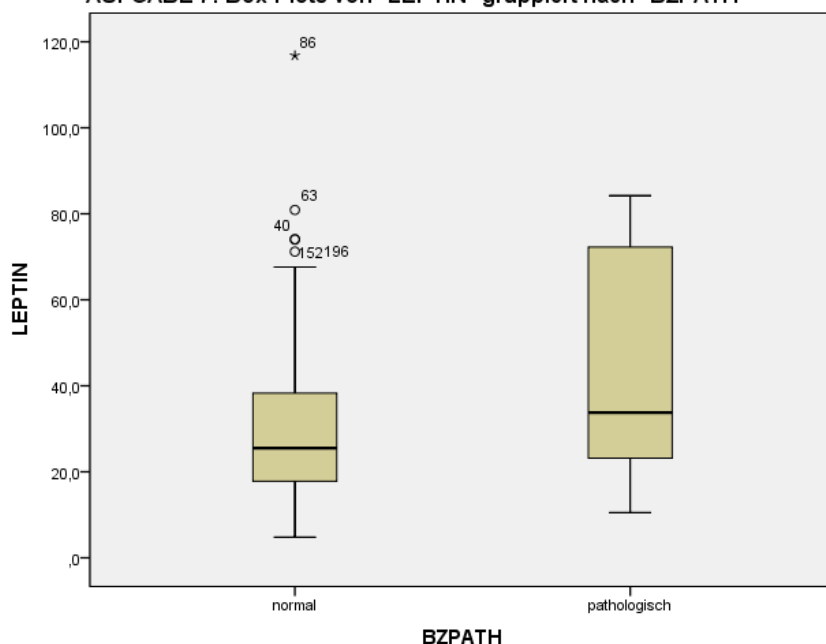
BZPATH	Fälle					
	Gültig		Fehlend		Gesamt	
	N	Prozent	N	Prozent	N	Prozent
LEPTIN normal	185	93,4%	13	6,6%	198	100,0%
pathologisch	20	95,2%	1	4,8%	21	100,0%

Deskriptive Statistik

BZPATH				Statistik	Standardfehler
LEPTIN normal	Mittelwert			29,998	1,2794
	95% Konfidenzintervall des Mittelwerts	Untergrenze		27,474	
		Obergrenze		32,523	
	5% getrimmtes Mittel			28,773	
	Median			25,500	
	Varianz			302,823	
	Standardabweichung			17,4018	
	Minimum			4,8	
	Maximum			116,8	
	Spannweite			112,0	
	Interquartilbereich			20,7	
	Schiefe			1,336	,179
	Kurtosis			2,846	,355
pathologisch	Mittelwert			44,505	5,7339
	95% Konfidenzintervall des Mittelwerts	Untergrenze		32,504	
		Obergrenze		56,506	
	5% getrimmtes Mittel			44,189	
	Median			33,750	
	Varianz			657,543	
	Standardabweichung			25,6426	
	Minimum			10,5	
	Maximum			84,2	
	Spannweite			73,7	
	Interquartilbereich			53,0	
	Schiefe			,343	,512
	Kurtosis			-1,451	,992

Betrachten Sie zuzüglich die Box-&-Whisker-Plots. Können Sie eine Tendenz erkennen und wenn ja, welche?

BIOMETRIE / Q1 – Lösung Übung2
AUFGABE 7: Box-Plots von "LEPTIN" gruppiert nach "BZPATH"



Interpretation:

Die mediane Leptin-Konzentration ist bei Probanden mit pathologischen Blutzuckerwerten höher als bei denen mit normalem Blutzuckerwert.

Aufgabe 8:

Die BMI-Werte der untersuchten Kinder und Jugendlichen lagen zwischen 23 und 51. Nehmen Sie eine Einteilung der Variable "BMI" entsprechend der unten abgebildeten Tabelle vor. Die neue Klassenvariable soll den Namen "BMI_Klassen" erhalten.

Setzen Sie unter der Variablenansicht Wertelabels, um den Klassen sprechende Namen geben zu können (s. Tabelle unten).

Stellen Sie anschließend die absoluten Häufigkeiten der klassierten Daten in Abhängigkeit davon dar, ob die Blutzuckerwerte als normal oder als pathologisch eingestuft wurden (Variable: "BZPATH"). Verwenden Sie dazu **Analysieren → Deskriptive Statistiken → Kreuztabellen**. Wählen Sie "BMI_Klassen" als **Zeile(n)** und "BZPATH" als **Spalten** aus. Über den Button **Zellen** unter **Prozentwerte** haben Sie die Möglichkeit sich verschiedene Prozentangaben mit ausgeben zu lassen.

LÖSUNG:

BIOMETRIE/Q1 - Loesung Uebung 2 / Gruppe B
Aufgabe 8: Kreuztabelle "BMI_Klassen" und "BZPATH"

Anzahl

		BZPATH		Gesamt
		normal	pathologisch	
BMI_Klassen	20 - 30	80	5	85
	30 - 40	102	9	111
	40 - 50	15	7	22
	50 - 60	1	0	1
Gesamt		198	21	219

**Hinweis:**

Falls die Gesamtzahl kleiner ist als der Stichprobenumfang liegt das daran, dass in der Kreuztabelle keine Missings enthalten sind.

Um Missings zu berücksichtigen, gehen Sie wie folgt vor:

Kodieren Sie die fehlenden Werte in eine Zahl um und definieren diese unter der **Variablenansicht** bei **Wertelabels** als **MISSINGS**.