

Projektbericht zum Modul Information Retrieval und
Visualisierung Sommersemester 2023

Visualisierungen zur Analyse von Einflussfaktoren auf die Schlaff Effizienz

Mick Stewart Wörner

28.12.2023

Link zum GitHub-Repository (URL des Projekt):

<https://github.com/EinKnopfzu/Sleep-Efficiency-Dataset/tree/main>

Letzter Commit: 977f7c9

Datum Commit: 28.12.2023 - 02:00

1 Einleitung

Im folgenden Paper werden drei Visualisierungstechniken zur Analyse von Schlafqualität dargestellt. Dabei wird ein Fokus auf die Analyse der Verteilung, Einflüsse und Korrelation mit Verhaltensweisen gelegt. Die vorgestellten Visualisierungen ermöglichen es dem Anwender ein tiefes Verständnis über die Aussagefähigkeit der Daten und Zusammenhänge zu erschaffen.

1.1 Anwendungshintergrund

Hintergrund dieser Anwendung ist die Darstellung von im Internet bereitgestellten Daten. Diese beziehen stellen unterschiedliche Attribute zur Beschreibung der Schlafqualität dar und erhebt dazu Daten zu verschiedenen Verhaltensweisen. Die Daten können durch die Visualisierung beurteilt, in Kontext gesetzt und mit Verhaltensweisen korreliert werden.

1.2 Zielgruppen

Die Schlafqualität spielt eine entscheidende Rolle für die Lebensqualität, deswegen richtet sich die Anwendung an eine Zielgruppe, die ein Interesse daran hat, die Zusammenhänge und Einflüsse auf die Schlafqualität zu verstehen. Diese Zielgruppe könnte aus Forschern und Privatpersonen bestehen, die daran interessiert sind, wie verschiedene Lebensstile die Schlafqualität beeinflussen können und die Schlafzyklen sich zueinander verhalten. Ein tiefgehendes Wissen über die Schlafphasen ist nicht erforderlich, aber ein generelles Interesse an der Analyse des Zusammenhangs zwischen Verhalten und Schlafqualität charakterisiert den Benutzer. Die Benutzer sollten sich darüber im Klaren sein, dass die Visualisierungen Hinweise liefern können, jedoch nicht notwendigerweise kausale Zusammenhänge zeigen.

Die Visualisierungen, die in diesem Projekt präsentiert werden, sollen dabei helfen, Einflüsse und Zusammenhänge bei der Schlafqualität zu identifizieren. Dabei werden verschiedene Attribute wie Koffein-, Alkohol- und Tabakkonsum, Sport, sowie Alter und Geschlecht der Personen untersucht. Die Zielgruppe sollte ein Grundverständnis für die Bedeutung dieser Attribute in Bezug auf die Schlafqualität haben.

Die Benutzer sollten sich bewusst sein, dass der Datensatz, der in dieser Anwendung verwendet wird, nur 452 Personen umfasst und durch die Datenverarbeitung weiter reduziert wird. Daher ist es wichtig, die statistische Validität der Ergebnisse zu berücksichtigen, insbesondere bei der Anwendung von Filtern. Statistische Aussagekraft ist bei $n < 30$ nicht mehr gegeben, und daher müssen die Benutzer dies bei der Benutzung berücksichtigen.

1.3 Überblick und Beiträge

Im diesem Abschnitt wird ein Überblick auf die Daten und verwendeten Visualisierungstechniken gegeben.

2 Daten

Der verwendete Datensatz bietet eine umfangreiche Sammlung von 15 Attributen, die Informationen über die Schlafqualität und Lebensgewohnheiten von 452 Personen liefern. Auf der Kaggle Seite wird nach Angaben des Authors behauptet, dass der Datensatz im Kontext einer Studie von der ENSIAS, Marroco gesammelt wurde. Innerhalb einer Literatur Recherche konnten weder auf der Webseite der ENSIAS noch in weitergehender Literaturrecherche eine Veröffentlichung zu diesem Datensatz identifizieren. Daher sollten die Daten und daraus entwickelten Ergebnisse, nicht unreflektiert übernommen werden.

Der Datensatz hat drei Nominale (Id, Gender und Raucher) und zwölf Quantitative (Age, Bedtime, Wakeup Time, Sleep Duration, Sleep Efficiency, REM Sleep percentage, Deep Sleep Percentage, Light Sleep Percentage, Awakenings, Caffeine Intake, Alkohol Intake, Tobako Intake, Exercise frequency) Aufgrund der unklaren Herrkunft des Datensatzes muss werden die Attribute im folgenden näher beschrieben.

- ID: Die ID identifiziert eine Person eindeutig. Da keine ID mehrfach aufgeführt wurde, ist anzunehmen, dass jede Person nur einmal an der Studie teilgenommen hat. Die ID wird als Integer bereitgestellt, und der Datenbereich reicht von 1 bis 452.
- Age: Gibt das Alter an, das die Person zum Zeitpunkt der Erfassung hatte. Das Alter wird als Integer angegeben und ist daher diskret, zum Beispiel 43 Jahre. Die Verteilung des Datenbereichs wird im Folgenden in diesem Format angegeben: (Quantile [Min, 25, 50, 75, Max]), Quantile [9, 29, 40, 52, 69].
- Gender: Das Geschlecht wird als String abgespeichert und nimmt nur zwei Werte an: "Male" oder "Female". Dabei gibt es einen Anteil von 50 Prozent Männern und 50 Prozent Frauen.
- Bedtime: Gibt die Uhrzeit an, zu der die Person ins Bett gegangen ist. Hierbei ist nicht klar, ob damit der Zeitpunkt gemeint ist, zu dem die Person eingeschlafen ist oder zu dem die Person sich ins Bett gelegt hat. Die Information werden als DateTime angegeben. Die Daten steigen in 30 Minuten schritten und ist trotz DateTime Formati damit Diskret.
- WakeUp time: Gibt das Datum und die Uhrzeit an, zu dem die Person erwacht. Steigt analog zu der Bedtime in halben Stunden Schritten an. Das Datum ist bei beiden DateTime-Formaten nicht von weiterem Interesse, da es keine zeitliche Entwicklung der erfassten Personen gibt.
- Sleep Duration: Die Schlafdauer ist wie die Bedtime unklar in ihrer Interpretation, da sich der Wert immer aus der Differenz zwischen Bedtime und WakeupTime berechnet. Daher ist unklar, ob es sich um die geschlafene oder um die im Bett verbrachte Zeit handelt.

Die Spalte wird als Float angegeben und steigt aufgrund der halbstündlichen Sprünge der BedTime und WakeupTime auch in 0.5 Schritten. Die Daten haben Quantile von [5.0, 7.0, 7.5, 8.0, 10.0] Stunden. Die Interpretation der Sleep Duration wird weiter dadurch erschwert, dass im weiteren Datensatz die Anzahl angegeben wird, wie oft eine Person in der Nacht wach wird, ohne Angabe, wie lange diese SSchlafpausen spezifisch sind.

- Sleep efficiency: Gibt den prozentualen Anteil an, den eine Person schlafend im Bett verbracht hat. Die Daten werden als Float mit zwei Nachkommastellen angegeben. Die Daten haben die Quantile: [0.5, 0.7, 0.82, 0.9, 0.99]. Eine Person, die 5 Stunden im Bett verbracht hat und davon eine Stunde wach war, hat also eine Schlafeffizienz von 80 Prozent. Da hier wieder eine Interpretationsproblematik besteht, wird im weiteren davon ausgegangen, dass die Schlafeffizienz angibt, welchen Anteil die Person nach erstmaligem Einschlafen, schlafend verbringt. Weiter wird diese Kennzahl in der Datenverarbeitung auf einen absoluten Stundenwert hochgerechnet.
- REM Sleep percentage: Die REM steht für Rapid Eye Movement Schlaf, dies ist einer der drei Schlafzyklen, die ein Mensch im Schlaf durchführt. Die CDC empfiehlt einen Anteil von 25 Prozent Healthline . Der REM Percentage gibt den prozentualen Anteil an, den die schlafende Person im REM verbracht hat, also den Anteil REM an Schlaf Effizienz. Die Daten werden als Integer abgespeichert und haben Quantile von [15, 20, 22, 25, 30]. Weiter wird diese Kennzahl in der Datenverarbeitung auf einen absoluten Stundenwert hochgerechnet.
- Deep sleep percentage: Der Tiefschlaf-Prozentsatz gibt den Anteil am Schlaf an, der im Tiefschlaf verbracht wurde. Die Daten werden als Integer angegeben und haben Quantile von [18, 51, 58, 63, 75]. Weiter wird diese Kennzahl in der Datenverarbeitung auf einen absoluten Stundenwert hochgerechnet.
- Light sleep percentage: Gibt den prozentualen Anteil am Schlaf an, der im Leichtschlaf verbracht wurde. Die Daten werden als Integer angegeben und haben Quantile von [7, 15, 18, 40, 63]. Weiter wird diese Kennzahl in der Datenverarbeitung auf einen absoluten Stundenwert hochgerechnet.
- Awakenings: Gibt die absolute Anzahl an, wie oft eine Person aufgewacht ist. Die Daten werden im Datensatz als Float abgespeichert. 0.0 bedeutet, dass eine Person durchgeschlafen und nur einmal final am Morgen aufgewacht ist. Die Daten reichen von [0.0, 1.0, 1.0, 3.0, 4.0].
- Caffeine intake: Gibt an, wie viel Koffein die Person in den letzten 24 Stunden zu sich genommen hat. Die Maßeinheit hierbei beträgt mg. Die Daten werden als Float abgespeichert und haben Quantile von [0.0, 0.0, 25.0, 50.0, 200.0].

- Alcohol intake: Gibt an, wie viel Alkohol die Personen in den letzten 24 Stunden zu sich genommen haben in Oz. Die Daten haben Quantile von [0.0, 0.0, 0.0, 2.0, 5.0].
- Tobacco intake: Gibt an, ob die Person raucht. Die Daten sind als String abgespeichert: "Yes" für Raucher und "No" für Nichtraucher. 154 Personen geben an zu rauchen und 298 geben an Nichtraucher zu sein.
- Exercise frequency: Gibt an, wie viele Einheiten Sport die Person in der Woche macht. Dabei ist nicht angegeben, welche Maßeinheit diese Einheiten Sport haben. Die Daten haben Quantile von [0, 0, 2, 3, 5], es gibt 6 fehlende Werte.

Unter der Annahme, dass die Daten valide sind und die angegebene Interpretation der Attribute korrekt ist, gewährt dieser Datensatz detaillierte Einblicke in die Schlafqualität, insbesondere hinsichtlich der Schlafphasen und ihrer Dauer. Allerdings weist die Erfassung der Kaffee- und Alkoholeinnahme Schwächen auf, da diese Substanzen im menschlichen Körper eine geringe Halbwertszeit haben, und somit der zeitliche Kontext der Einnahme relevant wäre. Weitere Einflussfaktoren auf die Schlafqualität werden nicht erfasst, ebenso wenig wird die Länge individueller Schlafunterbrechungen differenziert oder den drei Schlafphasen zugeordnet. Daher könnte es für den Anwender von Interesse sein zu erfahren, welche der Schlafphasen durch bestimmte Verhaltensweisen gestört werden. Zudem lässt sich argumentieren, dass die Verteilung der Verhaltensweisen sehr linksseitig ist, dies ist möglicherweise auf kulturelle Unterschiede zurückzuführen. Daher sollte die Aussagekraft des Datensatzes auf den marokkanischen Lebensstil eingeschränkt werden.

3 Die Daten

Der Datensatz wurde auf Kaggle veröffentlicht und ist die dritte Version. Die Daten sind als CSV mit einer Grösse von 9 kB zugänglich. Sie sind im GitHub-Repository abgelegt, was eine zentrale Aktualisierung der Daten ermöglicht, falls dies notwendig sein sollte.

Das Programm ruft die CSV-Datei mittels eines HTTP-GET-Requests auf und speichert sie als einen String ab. Falls der Request erfolgreich ist, wird die Nachricht 'Got Text Ok fulltext' an die Funktion 'update' gesendet. Der String wird dabei durch 'fullText' repräsentiert. Daraufhin wird im 'Model' das 'datenladen' auf 'Success' gesetzt. Auf den String wird die Funktion 'stringtoUnverarbeitete' angewandt, deren Ziel es ist, den String in eine 'Liste (UnverarbeiteteDaten)' zu transformieren. Die Funktion verwendet das Paket 'BrianHicks/elm-csv' (im Code als Decode bezeichnet), um die Namen der Felder aus der ersten Zeile im String zu extrahieren. Die Funktion 'decode' decodiert die Eingabedaten, wobei zu beachten ist, dass auch leere Felder im String vorkommen dürfen. Die Funktion 'Decode.blank' gibt ein 'Nothing' zurück, falls das Feld leer ist. Wenn es dennoch zu einem Fehler kommt, gibt die Funktion 'stringtoUnverarbeitete' eine leere Liste an das Modell zurück. Wenn die Daten nun in Form von 'UnverarbeiteteDaten' vorliegen, wird die Funktion 'sleep2Point' angewandt. Diese entfernt ein Tupel, wenn eines seiner Felder ein

‘Nothing’ enthält. Weiter werden die Werte für REM, Tiefschlaf und Leichtschlaf in tatsächlich verbrachte Stunden transformiert, indem sie mit 0.1 in Prozent umgerechnet und dann mit der Schlafeffizienz und Schlafdauer multipliziert werden. Dies verbessert die Interpretierbarkeit und stellt die Vergleichbarkeit her. Gender und Raucher werden von einem String zu einem Float mittels Case-Handling konvertiert (Funktionen ‘genderToFloat’ und ‘raucherToFloat’). Dies ermöglicht das Rechnen mit den Daten und erleichtert die weitere Visualisierung. Die Funktion ‘sleep2Point’ hat einen Output vom Typ Alias ‘AussortierteDaten’. Weiter werden die Attribute Bedtime und Wakeup Time in den aktuellen Visualisierungen nicht verwendet, könnten jedoch bei der Weiterentwicklung interessant werden. Daher wurden sie im Datenkonstrukt belassen, aber nicht weiter behandelt man könnte von einer sanften Projektion sprechen. Innerhalb der Einstellungen kann der Anwender eine Selektion durchführen und Einschränkungen auf das zu untersuchende Attribut anwenden. So werden nur Datentupel an die Visualisierungen übergeben, die dem Kriterium entsprechen.

4 Visualisierungen

4.1 Analyse der Anwendungsaufgaben

Die konkreten Anwendungsaufgaben, die von den Anwendern bearbeitet werden müssen, zielen darauf ab, die Daten kritisch zu analysieren, Zusammenhänge zwischen verschiedenen Attributen explorativ zu erkennen und den Einfluss der Verhaltensweisen zu erkennen.

1. Datenlage analysieren:

- Die Nutzer müssen zunächst die Datenlage kritisch analysieren und die Güte der einzelnen Attribute bewerten. Hierbei ist ein Verständnis für die statistische Interpretation und die Relevanz von Normalverteilungen zur Aussagefähigkeit der Visualisierungen wichtig.

2. Vergleich von Attributen:

- Nach der Bewertung der Güte der Attribute, soll ein Nutzer den Zusammenhang dieser mit anderen Attributen verstehen. Es soll hierbei auch auf versteckte Zusammenhänge stoßen können und die indirekte Beziehungen erkennen.

3. Analyse der Einflüsse des Verhaltens:

- Die Einflüsse des Verhaltens auf die Attribute müssen analysiert werden. Hierbei ist es notwendig, den Einfluss von multiplen Verhaltensindikatoren auf die Ausprägung eines Merkmals zu schätzen.

Bei diesen Aufgabe sind die folgenden mentalen Modelle von Bedeutung:

1. Statistische Interpretation:

- Nutzer sollten Grundlagen der statistischen Interpretation und die Relevanz von Normalverteilungen verstehen. Visualisierungen wie QQ-Plots können dabei helfen, Abweichungen von der Normalverteilung zu identifizieren.

2. Korrelation interpretieren:

- Ein Verständnis für die Interpretation von Korrelationen ist wichtig. Visualisierungen können durch Farbgebung (z. B. Rot für negative und Grün für positive Korrelationen) unterstützen, auch wenn dies nicht zwingend notwendig ist.

3. Parallele Koordinaten Interpretieren:

- Benutzer sollten verstehen, dass jede Achse ein Attribut repräsentiert, und Linien zwischen den Achsen die Beziehung zwischen Werten in verschiedenen Dimensionen anzeigen. Die Interpretation von Parallelenkoordinaten erfordert die Fähigkeit, Muster in den Linienverläufen zu erkennen und so Zusammenhänge zwischen den Attributen zu verstehen.

4.2 Anforderungen an die Visualisierungen

Die Visualisierungen müssen die Anwenderinnen bei der explorativen Analyse unterstützen. Somit müssen die Daten expressiv und effektiv dargestellt werden. Die Visualisierung muss die Daten korrekt wieder geben, dabei den Anwender unterstützen, Muster und Trends in den Daten zu identifizieren. Weiter müssen die Attribute sinnhaftig auf Ihre Beziehungen untereinander dargestellt werden und die Anwenderin dabei unterstützen auch indirekte Beziehungen zu erkennen. Ausserdem müssen die Visualisierungen die Einflüsse der Verhaltensweisen auf die Schlafqualität leicht interpretierbar darstellen. Die Visualisierungen sollten barrierefrei gestaltet sein, um sicherzustellen, dass Benutzer mit unterschiedlichen visuellen Fähigkeiten die Informationen analysieren können.

4.3 Präsentation der Visualisierungen

Im folgenden Abschnitt werden die drei verwendeten Visualisierungstechniken vorgestellt, deren Nutzen für den Anwender dargelegt und alternative Visualisierungen besprochen.

4.3.1 Visualisierung Eins: QQ-Plot Datenlage Normalverteilt?

Die erste Visualisierung ist die Darstellung der Daten in einem zweidimensionalen Diagramm, spezifisch eines Normalverteilungs-QQ-Plots.

Ein QQ-Plot (Quantile-Quantile-Plot) ist ein nützliches grafisches Werkzeug in der statistischen Datenanalyse, insbesondere wenn es darum geht, die Verteilung von Daten zu überprüfen und diese mit einer theoretischen Verteilung zu vergleichen. In der Anwendung wird aufgrund des Implementierungsaufwandes nur der Vergleich mit der Normalverteilung vorgenommen. Hierzu werden die Daten berechnet, die Verteilung der Daten berechnet, die das Attribut hätte, falls es einer Normalverteilung folgen würde, und in einem Scatterplot dargestellt.

Die typischen Aufgabenfelder eines QQ-Plots sind die Normalitätsüberprüfung, die Identifikation von Abweichungen, die Bestätigung der Normalverteilungshypothese und Identifikation von Ausreißern.

Die Visualisierung zeigt den Wert auf der Y-Achse an, den ein Datenpunkt hätte, falls der Datensatz einer Normalverteilung folgen würde. Die X-Achse zeigt den Wert des Datenpunktes an. Dann werden die Datenpunkte basierend auf ihrer Reihenfolge im Scatterplot gezeichnet. Ein perfekt normalverteiltes Attribut würde somit eine gerade Linie im 45-Grad-Winkel ergeben.

Dies ermöglicht dem Anwender eine Normalitätsüberprüfung durchzuführen. Somit kann dieser sich ein eigenes Bild darüber machen, ob das Attribut normalverteilt ist und seinen Anforderungen entspricht. Bei einem perfekt normalverteilten Attribut würden im QQ-Plot alle Punkte auf einer $f(x) = x$ -Linie liegen. Abweichungen von dieser Linie deuten auf Abweichungen von der Normalverteilung hin; Krümmungen und Biegungen weisen auf Verschiebungen hin. Hier ist es an dem Anwender, diese zu interpretieren. Die hypothetische $f(x) = x$ - Linie ist zur Orientierung auch eingezeichnet. Weiter ist eine rote Linie vom 25. Quantil bis zum 75.

Quantil eingezeichnet. Dies erleichtert dem Benutzer den Datensatz zu überprüfen und einer Verunsicherung durch die Abweichungen an den Extrema entgegenzuwirken.

Die Identifikation von Ausreißern kann im Normal QQ-Plot an den visuellen Enden identifiziert werden. Starke Abweichungen sind durch eine große Distanz zur schwarzen Linie erkennbar. Diese können durch das Filtern des Datenbereichs entfernt werden. Weiter ist es möglich, einen Vergleich verschiedener Verteilungen anzustellen. Dies ist aufgrund des Umfangs des Projektes nicht in der Anwendung implementiert, könnte dem Anwender aber einen Mehrwert bieten. Somit sollte der Anwender durch den Scatterplot in der Lage sein, die Datenverteilung und Interpretierbarkeit einzuschätzen und diese in geringem Maße mit den Filtern zu selektieren.

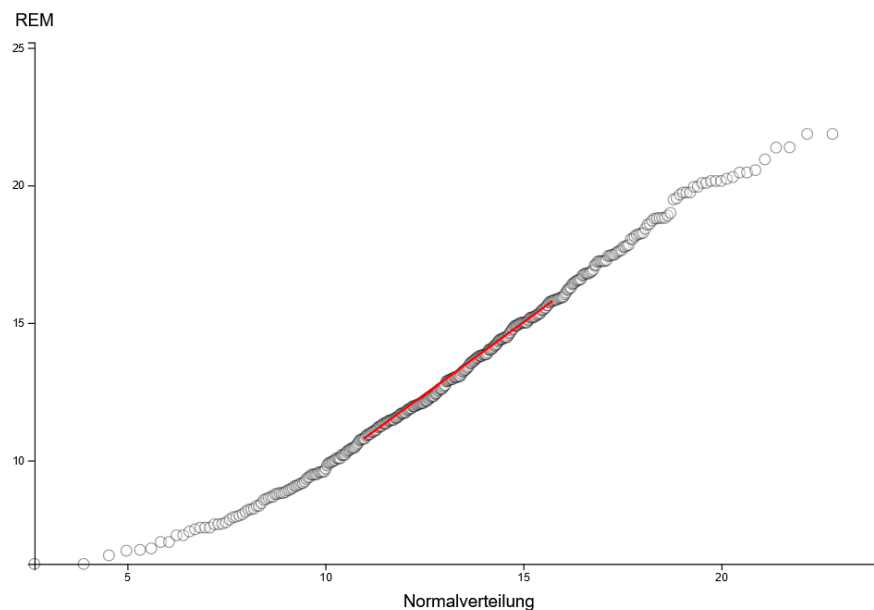


Abbildung 1: Visualisierung Eins Norm QQ-Plot REM

Aufgrund des Anspruchs, die Interaktion von Daten analysieren zu können, sind zwei weitere Scatterplots unter dem des Hauptattributes dargestellt. Deren Funktion ist ebenfalls die Interpretation der Güte.

Das oben erwähnte Auswählen und das Filtern geschieht in der Seitenleiste. Hier kann der Benutzer das Attribut auswählen, das er untersuchen möchte, und direkt darunter dieses in seinem Wertebereich filtern. Die Seitenleiste wird im Abschnitt ‘Interaktion’ genauer erklärt. Der eingeschränkte Datenbereich wird dann in den anderen Visualisierungen verwendet. Der Benutzer kann in der zweiten Visualisierung zwei weitere Attribute zum Untersuchen auswählen. Diese werden daraufhin unter dem Haupt-Scatterplot als kleinere QQ-Plots dargestellt. Dies ermöglicht auch für sie eine schnelle Einschätzung der Güte.

Die Visualisierung mittels Scatterplot ist im Kontext der Datenbewertung effektiv, da sie die zugrunde liegenden Fragestellungen visuell gut unterstützt und dem Anwender bei multiplen

Fragestellungen Antworten liefern kann. Die Expressivität ist limitiert, da die Interpretation nicht ohne Kenntnis der Interpretation von QQ-Plots möglich ist und möglicherweise die Attribute, welche in den kleinen Scatterplots eine Unterordnung zum Haupt-Scatterplot impliziert. Alternativ hätte auch ein Histogramm verwendet werden können. Bei diesem hätte man die Hürde zur Analyse reduziert, aber die Identifikation von Ausreißern erschwert.

4.3.2 Visualisierung Zwei: Parallele Koordinaten

Präsentieren der Visualisierung und Interaktionsmöglichkeiten.

In der Visualisierung werden drei Dimensionen abgebildet: das im QQ-Plot ausgewählte Hauptattribut sowie zwei weitere Attribute. Die Achsen sind parallel zueinander angeordnet und gleich groß. Um die Ausprägungen eines Datentupels darzustellen, wird die Ausprägung des Tupels am jeweiligen Attribut auf die Y-Achse skaliert. Anschließend wird eine gerade Linie zwischen den Positionen auf den benachbarten Achsen gezogen, wobei die Endpunkte der Liniensegmente mit den Attributwerten korrespondieren.

Das Hauptattribut, das in Visualisierung 1 ausgewählt wurde, befindet sich in der Mitte, das Y-Attribut auf der linken und das Z-Attribut auf der rechten Seite. Es wurde bewusst darauf verzichtet, weitere Attributsdimensionen hinzuzufügen. Dies dient einerseits dazu, dass der Anwender seinen Fokus auf die Analyse des Hauptattributes legt, und andererseits könnten bei mehr als 3 Attributen suggerierte Zusammenhänge entstehen, die nicht zwangsläufig vorhanden sein müssen. Eine Clusterung zwischen Dimension 1 und 2 sowie zwischen 2 und 3 bedeutet nicht zwangsläufig eine Clusterung zwischen 1 und 3. Dies wäre der Effektivität abträglich, da so Zusammenhänge suggeriert werden könnten, die nicht beabsichtigt sind.

Parallele Koordinaten eignen sich besonders gut für mehrdimensionale, quantitative und endliche Daten. Diese sind in unserem Datensatz gegeben. Die Visualisierung ermöglicht es dem Anwender, Muster und Trends in den Daten zu identifizieren, die auf Beziehungen zwischen den Attributen hinweisen. Zum Beispiel können parallele Linien auf eine starke Korrelation zwischen den entsprechenden Variablen hinweisen. Die Visualisierung ermöglicht es dem Anwender, Rückschlüsse auf den Zusammenhang von Attributen zu ziehen und diese mit Visualisierung 1 zu vergleichen. Die Untersuchung der Attribute und das Visualisieren kann mittels eines interaktiven Anpassens der Pixelgröße und RGBA-Werte erfolgen, dies ist nützlich, da viele der Inputdaten wenige unterschiedliche Wertausprägungen haben, was die Clusterung der Inputdaten erschwert. Um dennoch interessante Muster erkennen kann. So kann der Benutzer basierend auf seinen eigenen visuellen Bedürfnissen Anpassungen vornehmen. In Bild 3 wird dies vorgestellt. Dazu müssen dem Anwender Konzepte wie Clustering, Farbwahrnehmung und Abstrahierung bekannt sein.

Alternativ hätte eine Radviz Visualisierung angewandt werden können, diese hätte allerdings den Fokus weg vom zu untersuchendem Hauptattribut gelenkt. Daher ist sie keine alternative gewesen obwohl sie die gleichen Daten darstellt und zu ähnlichen mentalen Modellen führt basiert. Die Visualisierung ist weniger effektiv, da der Anwender verschiedene Kombinationen

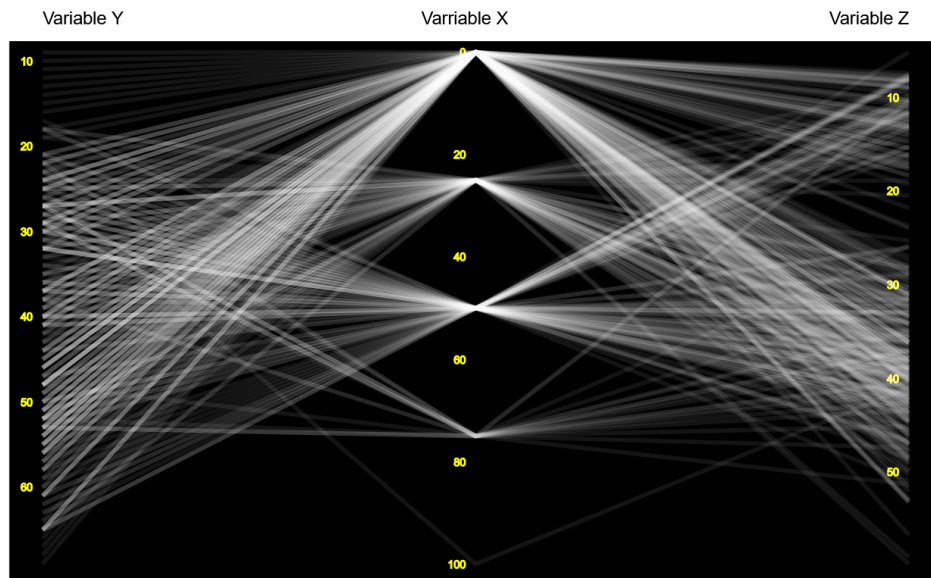


Abbildung 2: Visualisierung Zwei: Parallele Koordinaten REM

von Attributen und Pixel ausprobieren muss um sich ein Bild zu machen. Dies ist allerdings auch der Stärke der Visualisierung, da der Anwender sich die Visualisierung so anpassen kann, dass Sie seinen Bedürfnissen entspricht. Hinsichtlich der Expressivität, kann es dazu kommen, dass der Anwender durch Farbgebungen z.b. Rote Linien. Eine nicht intendierte Interpretation erzeugt.

4.3.3 Visualisierung Drei: Weihnachts Stern

Die dritte Visualisierung lässt sich wohl den pixelorientierten Techniken zuordnen. Dabei werden die im Abschnitt Daten erwähnten Verhaltensweisen und Attribute wie Alter, Geschlecht, Koffeinkonsum, Alkoholkonsum, Tabakkonsum und Sporteinheiten kreisförmig um das Hauptattribut angeordnet. Basierend auf ihrer Korrelation zu diesem Hauptattribut werden sie sowohl farblich als auch in ihrer Grösse modifiziert. Die in Visualisierung 2 hinzugefügten Attribute können ebenfalls integriert werden. Somit wird die Korrelation doppelt dargestellt. Die Farbgebung erfolgt durch die Skalierung der Korrelation auf ein Farbschema: negativ wird durch Rot dargestellt, positiv durch Grün. Die Intensität der Farbe wird positiv durch die Korrelation beeinflusst, während der Betrag der Korrelation die Grösse des Kreises beeinflusst. Dadurch werden Daten mit hoher Korrelation besonders hervorgehoben. Um die Expressivität beizubehalten, sind die Merkmale und ihre Korrelation als Schriftzug eingefügt. Es sollte jedoch beachtet werden, dass eine qualitative Interpretation suggeriert werden könnte. In der Abbildung ist beispielsweise die Anzahl der Erwachungen dargestellt.

Obwohl das Erwachen in der Nacht keine Steigerung der Lebensqualität ist, könnte angenommen werden, dass die geröteten Attribute grundsätzlich Negativ sind.

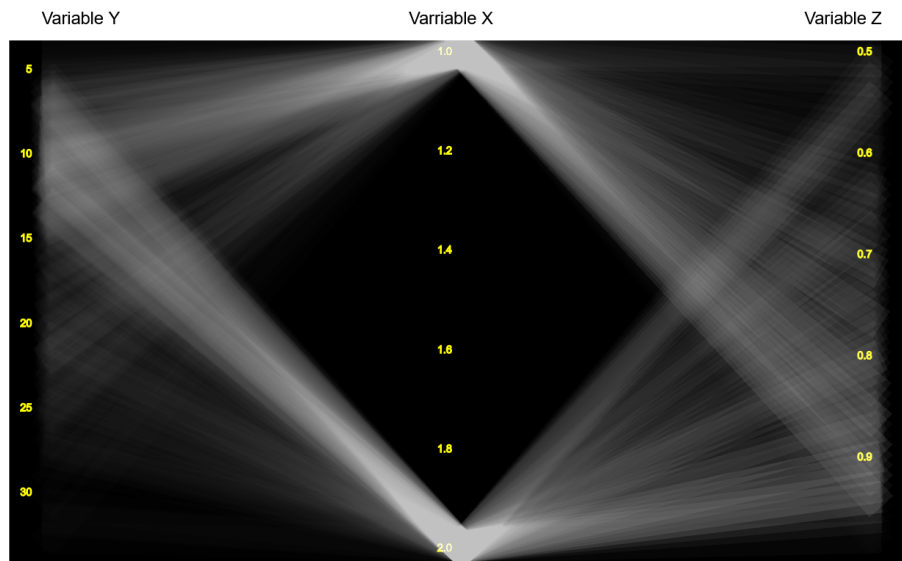


Abbildung 3: Visualisierung Zwei: Parallele Koordinaten:Pixel 30, RGBA (255, 255, 255, 0.006)

Der Anwender kann somit mithilfe dieser Visualisierung schnell Rückschlüsse auf den Einfluss des Verhaltens auf das untersuchte Hauptattribut ziehen. Wichtig dabei ist zu verstehen, dass Korrelation nicht gleichbedeutend mit Kausalität ist. Alternativ hätte eine Korrelationsmatrix genutzt werden können um die Korrelation darzustellen. Dies wäre auch zu einem ähnlichen Mentalem Modell geführt.

4.4 Interaktion

Die Visualisierungen sind durch Interaktion miteinander verknüpft. Da es sich um eine Onepage Application handelt ist die Interaktion über die Einstellungen umgesetzt um so unnötiges klicken zu vermeiden.

In dem Scatterplot kann das Hauptattribut ausgewählt werden und der Datenbereich eingeschränkt werden. Diese setzt also den Fokus für die weiteren Visualisierungen.

Die Einschränkung wird auf die verwendeten Daten in den anderen Visualisierungen angewendet so dass nur relevante Daten dargestellt werden.

Das in 1. gewählte Attribut wird Hauptattribut in dem Parallelen koordinatem System auf die X Position angeordnet. In Visualisierung 2 können zwei weitere Attribute hinzugefügt werden. Diese werden in Vis 1. als Junior QQ Plotts dargestellt.

Die dritte Visualisierung wird durch die beiden anderen beeinflusst. So ist das in Vis 1 gewählte Attribut das Hauptattribut auf das korreliert wird und die in Vis 2 gewählten Attribute werden als zusätzliche Zapfen dargestellt.

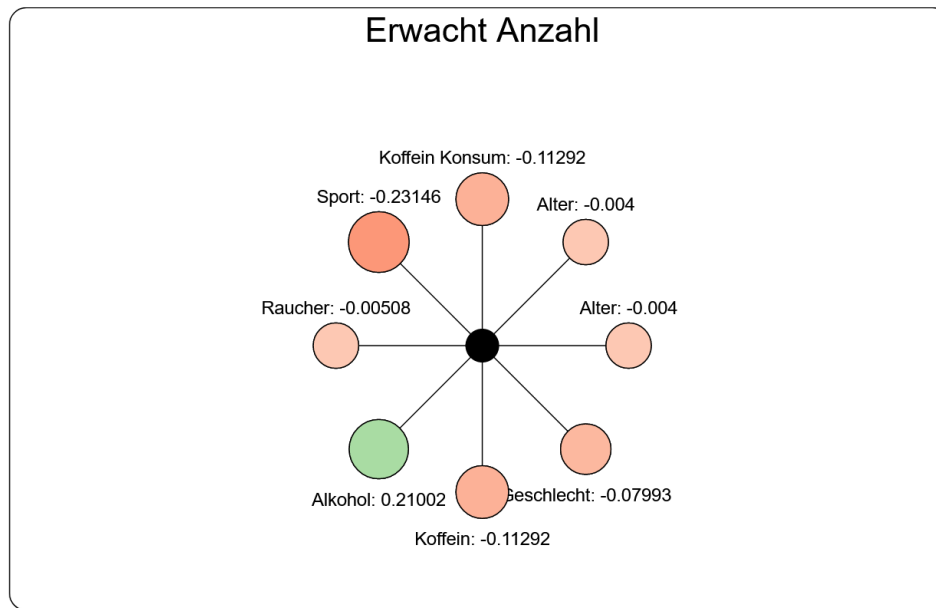


Abbildung 4: Visualisierung Drei: WeihnachtsStern: Anzahl Erwachungen

5 Implementierung

Die Visualisierung ist als One-Page-Application konzipiert, was aufgrund der benötigten Interaktion mit den Filtern und der häufigen Auswahl der Attribute nützlich ist. Das 'Main.elm' ruft zu Beginn die 'Init'-Funktion auf, die den Type Alias 'Modell' zunächst auf den Standardwert setzt und die Daten einliest. Die Kommunikation mit dem View erfolgt über den Typ 'Msg', der verschiedene Ausprägungen über Änderungen am Modell an die 'Update'-Funktion weitergibt. Die 'Update'-Funktion gibt die angehängte Information an das Modell zurück. Besonderer Erklärung bedarf hier die Abspeicherung des Datensatzes, dieser wird im Modell direkt als 'List (Aussortierte_Daten)' abgespeichert. Der Typ 'Aussortierte_Daten' ist ein Tupel, das die Merkmalsausprägungen abspeichert. Hierbei werden nur die ID, Schlafenszeit und die Aufwachzeit als String abgelegt. Die anderen Attribute sind alle in Floats transformiert. Da die Interaktion vor allem durch Projektion und Selektion umgesetzt ist, erfolgt dies in der Anwendung durch die Filter und die Auswahl der Attribute, die im Modell gespeichert werden.

Die im 'Main' wichtigste Funktion ist das 'View'. Hier werden verschiedene 'div' erzeugt, in denen die Interaktionen und Visualisierungen erstellt sind. Die Auswahl der Attribute erfolgt über ein Dropdown. Dieses übergibt einen String an das 'Main' und die Variable bezieht mithilfe eines 'case'-Handlings den gewünschten Wert. Der in Visualisierung 1 gewählte Filterbereich wird als 'daten' an die anderen Funktionen übergeben. Die so gefilterten Daten werden in einzelne Listen aufgebrochen und die benötigten Attribute wieder zusammengesetzt (Funktionen wie 'combined...'). Die Daten werden dann im 'div' an die in den anderen Modulen definierten Funktionen zum Zeichnen der Visualisierungen übergeben. Optional werden auch Zustände aus

dem Modell mit übergeben.

5.0.1 QQ Plott

Der QQ-Plott (`scatterplott.elm`) definiert zwei wichtige Funktionen die erzeugung von Normalverteilten Daten und das zeichnen des Plotts.

Die Funktion ‘`addNV`’ funktioniert wie folgt:

1. Die Eingabedaten (`data`) werden nach den y-Werten sortiert.
2. Es werden verschiedene Berechnungen durchgeführt, um die Normalverteilung zu approximieren und Quantile der Daten zu erhalten.
3. Eine Transformation (`transformierer`) wird definiert, um Werte von der Normalverteilung auf die ursprüngliche Verteilung zurückzuübertragen.
4. Die Funktion `invNormalCdf` berechnet die inverse Normalverteilungskumulative Funktion.
5. Quantile werden mit Schätzungen verglichen, um Werte auf der Normalverteilung zu erhalten.
6. Die Daten werden entsprechend transformiert, um die Normalverteilung darzustellen.
7. Ein QQ-Plot wird erstellt, indem die transformierten Werte gegen die ursprünglichen Datenpunkte aufgetragen werden.

Die Funktion `drawScatterplot` skaliert die Datenpunkte auf ihre relative XY-Position und zeichnet diese in einem SVG. Eine rote Linie wird zwischen dem 25. und 75. Quantilpunkt gezogen, während die schwarze Linie die Funktion $f(x)=x$ indirekt abbildet. Da diese Funktion aus der Übung bekannt ist, wird sie hier nicht weiter erläutert.

5.0.2 Parallele Koordinaten

Die Funktion ‘`blackbox`’ zeichnet die Parallele Koordinaten Visualisierung.

1. Die Funktion nimmt eine Liste von mehrdimensionalen Punkten (`model`), Pixelgrösse (`pixel`), sowie RGB- und Opazitätswerte für die Farbe entgegen.
2. Die Min Max werden berechnet um eine Skalierung zu ermöglichen.
3. Die Daten werden auf Ihre basierend auf der Dimension skaliert und auf die entsprechenden Achsen projiziert.
4. Die Verbindungslinien werden zwischen den Punkten an Ihren relativen Positionen gemalt.
5. Die Farbgebung und Opazität der Linien können durch die Parameter `rgb1`, `rgb2`, `rgb3` und `opacity` aus dem `Main.Model` gesteuert werden.

Die Funktionen zur Berechnung der Quantile, hat ungenutzte Funktionen um die Quantile in in einer anderen Art zu berechnen. Diese wurden im Programm belassen, um diese Funktion dem Anwender später zu ermöglichen.

5.0.3 Weihnachtsstern

Die Funktion ‘graph’ zeichnet die Weihnachtsstern Visualisierung.

Hierbei werden die Informationen (Hauptattribut, Verhaltensweisen und ausgewählten Attribute) als Datentyp ImpactGraph übergeben. Dieser Speichert den Namen des Attributes und eine Liste (Name, List (Float))ab.

```
exposed | 2 references
type alias ImpactGraphData =
    { xdescriptor : ImpactData
    , attribute : List ImpactData
    }

exposed | 8 references
type alias ImpactData =
    { name : String
    , data : List (Float)
    }
```

Abbildung 5: type Alias ImpactGraphData und ImpactData

Hier hervorzuheben ist die reaktive Positionierung der Kreise um das ‘Hauptattribut’ herum. Die Funktion `winkelEinteilung` berechnet gleichverteilte Winkelschritte mit $2 * \pi / (\text{length data} + 1)$ Die Funktion `Kreis` übernimmt diese und weitere relevante Informationen über die Grösse und Positionierung der Kreise. Hervorstellen sind hierbei die Funktionen `xPosition` und `yPosition` welche basierend auf dem Index einer Liste eine relative Positionierung um den Kreise berechnen.

```

6 references
xPosition : Float
xPosition = xa + (cos (toFloat index * winkel) * radiusUmkreis)

7 references
yPosition : Float
yPosition = ya + (sin (toFloat index * winkel) * radiusUmkreis)

```

Abbildung 6: Relative XY Positionierung der Kreise.

Da Elm keine Rückwärts kompatibilität erzwingt, ist es manchmal nicht möglich die neuste version eines Paketes zu verwenden. Dies ist allerdings etwas frustrierend, wenn man ein Nützliches paket gefunden hat. Es ist allerdings möglich den Quellcode auf Github, lokal abzuspeichern und als anderes Paket zu importieren, solange die Abhängigkeiten erfüllt sind. Das Modul JXStat ist das Paket Stat aus dem Elm-Stat 6.0.2.

Für die Zeichnung des Scatterplots und die Implementierung der Parallelen Koordinaten waren die Übungen sehr hilfreich. Die Implementierung der Datenübergabe, Manipulation und auch Visuell ansprechende Darstellung war auf persönlicher Ebene am schwierigsten.

6 Anwendungsfälle

Im folgendem wird die Anwendung der Visualisierungen anhand von Beispielen erläutert.

6.1 Anwendung Visualisierung Eins

Zu Beginn wird das zu untersuchende Attribut ausgewählt. In diesem Fall ist es die Schlaf Effizienz. Diese wurde gefiltert auf einen Datenbereich von [0.55; 0.85] Hier lässt sich eine gute Annäherung an die Normalverteilung erkennen. Das REM Attribut ist auch annähernd Normalverteilt und hat nur schwache Ausreißer. Der Leichtschlaf Anteils hingegen weist eine schlängelung auf. Dies ist relevant für die spätere Analyse des Parallelen koordinaten Systems.

6.2 Anwendung Visualisierung Zwei

Im zweiten Schritt der Visualisierung wurden die Attribute REM und Leichtschlaf Anteil untersucht auf Interaktion aufeinander untersucht. Man kann gut erkennen, dass es einen Zusammenhang zwischen niedrigem Leichtschlaf und hoher Schlafeffizienz gibt. Dies ist allerdings durch die rechts Verschiebung der Daten zu erklären. Bei hohem leichtschlaf Anteil geht ein weiteres Cluster zu einer reduzierten Schlafeffizienz. Der Zusammenhang zwischen REM und Schlafeffizienz, besteht aus parallel Absenkenden Strichen, dies deutet darauf hin, dass es eine direkte Positive Korrelation zwischen dem REM Schlafanteils die Schlaff Effizienz gibt.

So lässt sich nur aus der Visualisierung ohne Vorwissen die Hypothese entwickeln, dass mit

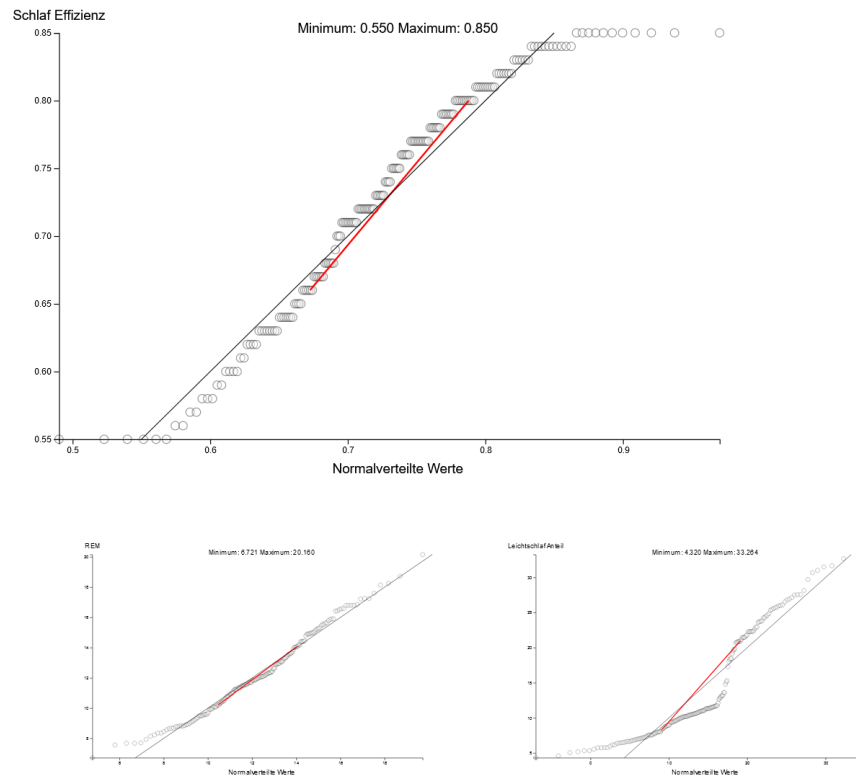


Abbildung 7: Visualisierung Eins: Anwendungsbeispiel

geringerer Schlafeffizienz die REM Schlafanteile sinken und der Anteil des leichten Schlafes erhöht wird.

6.3 Anwendung Visualisierung Drei

Die durch Visualisierung 2 entwickelte Hypothese, kann durch Visualisierung 3 unterstützt werden. Die Repräsentation der Korrelierung durch die Kugeln in gröSse und Farbe zeigen stützen, dass die Korrelation von REM auf die Schlafeffizienz positiv ist und die korrelation des Leichtschlafanteils negativ. Weiter kann man erkennen, dass die Standardmässige angezeigten Verhaltensweisen Attribute mit der Schlafeffizienz korrelieren. Obwohl die Cluster des REM und Leichtschlafes verschieden sind haben Sie dennoch beide einen starken Einfluss auf die Schlafeffizienz. Verhaltensweisen wie Alkohol konsum und auch das Rauchen scheinen die Schlafeffizienz zu reduzieren.

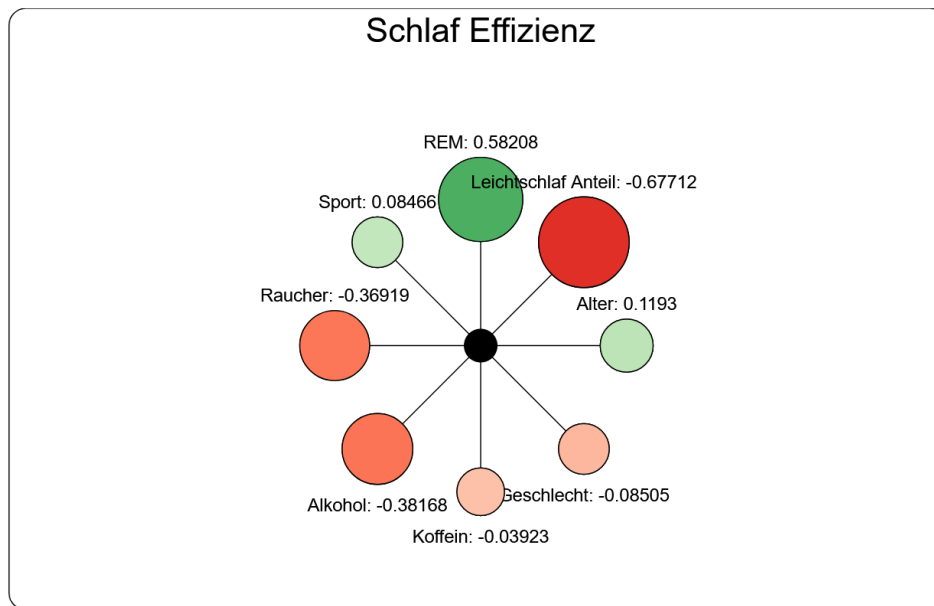


Abbildung 8: Visualisierung Eins: Anwendungsbeispiel

7 Verwandte Arbeiten

Um eine kurze Literaturrecherche durchzuführen, wurde die Datenbank des ScienceDirect¹ mit dem folgenden Suchstring durchsucht:

(„informationvisualization“ ∨ „visual analytics“) ∧ („sleep“ ∨ „sleep data“ ∨ „sleep efficiency“)

Um die gefundenen Ergebnisse weiter einzugrenzen, wurden lediglich Artikel aus dem aktuellen Jahr 2023 betrachtet, welche über das für diese Anwendung passende Journal „Sleep Health“ veröffentlicht wurden. Im Folgenden werden die beiden Artikel „Race and sex differences in the longitudinal changes in multidimensional self-reported sleep health characteristics in aging older adults“ von Tapia et al.[2] und „The thief of (bed)time: Examination of the daily associations between bedtime procrastination and multidimensional sleep health“ von Carlson et al.[1] betrachtet.

Zwei unabhängige Studien zum Thema Schlafgesundheit wurden analysiert, wobei Gemeinsamkeiten und Unterschiede hervorgehoben wurden. Die erste Studie von Tapia et al. umfasste eine umfangreiche Stichprobe von über 3000 Teilnehmern und legte besonderen Fokus auf demografische Merkmale wie Geschlecht und Rasse. Deskriptive Statistiken wurden präsentiert, wobei Teilnehmermerkmale wie Bildungsniveau, Alkoholkonsum und der frühe sozioökonomische Status berücksichtigt wurden. Die Studie analysierte Schlafgesundheitsdimensionen wie Schlafzufriedenheit, Tagesmüdigkeit, Probleme beim Ein- und Durchschlafen sowie die Schlafdauer.

¹sciencedirect.com

Interessanterweise wurden signifikante Wechselwirkungen zwischen Alter, Geschlecht und Rasse identifiziert, die Auswirkungen auf verschiedene Schlafcharakteristika hatten.

Die zweite Studie von Carlson et al. konzentrierte sich auf eine kleinere Stichprobe und untersuchte spezifisch Schlafprokrastination in Verbindung mit Chronotyp und anderen SchlafgesundheitsmaSSen. Durchschnittliche tägliche Schlafmessungen wurden präsentiert, und bivariate Korrelationsanalysen zeigten Assoziationen zwischen Schlafprokrastination, Chronotyp und verschiedenen Schlafindikatoren. Multilevel-Analysen verdeutlichten, dass höhere Schlafprokrastinationswerte mit schlechterer Schlafgesundheit auf täglicher Ebene verbunden waren. Interessanterweise unterschieden sich die beiden Studien auch hinsichtlich ihrer Fokussierung auf zusätzliche Variablen; die erste Studie berücksichtigte sozioökonomische und demografische Aspekte, während die zweite Studie tiefer in die Untersuchung von Schlafprokrastination eintauchte.

Zusammenfassend reflektieren diese Studien die Vielfalt der Fragestellungen in der Schlafforschung. Während die erste einen breiten Überblick über die Schlafgesundheit unter Berücksichtigung verschiedener Einflussfaktoren bot, konzentrierte sich die zweite auf die detaillierte Analyse von Schlafprokrastination und deren Beziehung zu anderen Schlafcharakteristika. Beide trugen zum Verständnis der komplexen Zusammenhänge bei, betonten jedoch unterschiedliche Schwerpunkte in Bezug auf StichprobengröSse, Variablenauswahl und Analysemethoden.

8 Zusammenfassung und Ausblick

Zusammenfassend lässt sich sagen, dass die Visualisierungen dem Anwender einen hohen Nutzen bei der explorativen Analyse der Daten bieten. Die Analyse kann auf die Schlafenszeit sowie die Aufwachzeit erweitert werden. Weiter wäre die umsetzung einer log-Transmutation bei Analyse der Daten nützlich um sie der Normalverteilung anzunähern und somit bessere Konklusionen ziehen zu können.

Anhang: Git-Historie

* 977f7c9 (HEAD -> main, origin/main, origin/HEAD) (Rechtscheib fehler korrigiert und Zusammenfassung Anteilung entfernt., 2023-12-28)
12956a1 (Die PDF compiliert nicht richtig, 2023-12-28)
07b894b (gitHistorie angefügt, 2023-12-28)
934584f (Rechtschreib fehler korrigiert, 2023-12-28)
852732c (Rechtschreib Fehler korrigiert, 2023-12-28)
a34b7e0 (Kleine Bugfixes Im latech, 2023-12-28)
dee1d42 (Unnötige latex dokumente entfernt, 2023-12-28)
41e4d35 (Anwendungshintergrund hinzugefügt und Github Link angefügt, 2023-12-28)
0d0dcc2 (Nicht benötigte Pakete entfernt, 2023-12-28)
64ce9d0 (Bild umbenannt, 2023-12-28)
085a2b7 (Bild umbenannt und gepusht, 2023-12-28)
64da624 (Anwendung Visualisierung ausgeführt, 2023-12-28)
cc6f029 (Bilder eingefügt, 2023-12-28)
a9767f8 (Implementierung beschrieben, 2023-12-28)
38fd1d5 (NV erklärt, 2023-12-28)
4f88c36 (Implementierung: main beschrieben, 2023-12-28)
16b39a7 (Interaktion beschrieben Bild fehlt, 2023-12-27)
5a6dfb5 (Bilder umbenannt und versuch am Bug Fix, 2023-12-27)
be79e77 (Visualisierung 3 beschrieben., 2023-12-27)
0a6a26c (Weihnacht Stern beschrieben, 2023-12-27)
b92a381 (Anwendung 2 weiter ausgeschrieben, 2023-12-27)
2ffd502 (Layout und rechtschreib kontrolle, 2023-12-27)
d3ddf38 (Literatur Recherche eingefügt, 2023-12-27)
4ec1fa7 (Literatur in die bib eingefügt, 2023-12-27)
c53b5a8 (Analyse der Anwendungsaufgabe formatiert und strukturiert., 2023-12-27)
4ca80e3 (Rechtschreib Kontrolle, 2023-12-27)
5009a50 (Konkulation der Daten korrigiert., 2023-12-27)
5b16f18 (Rechtschreib korrektur restliche Datenattribute, 2023-12-27)
6b7092c (ID Agegender und Bedtime rechtscheib kontrolle, 2023-12-27)
ece3051 (zielgruppe ordentlich formuliert., 2023-12-27)
962a4ec (Beschreibung impactGraph Anwendungsbeispiel eingefügt mit Bild, 2023-12-27)
0abc8d3 (Bild eingefügt, (Davor Bbliotheken auf ImapctGraph entfernt, wurde nicht korrekt

gepusht.), 2023-12-27)

ac29634 (Bsp Bild QQ plot eingefügt, 2023-12-27)

63505f4 (Überschüssige Pakete entfernt aus ImpactGraph entfernt, 2023-12-27)

5309d1f (Skalierung Bug bei negativen Zahlen entfernt, 2023-12-27)

10080f9 (Wetere beschreibung eingefügt, 2023-12-27)

cc856dd (Beschreibung und erklärng der Interaktion, 2023-12-27)

5a9def8 (Main geordnet, 2023-12-27)

911be71 (Interaktion beschrieben und Main reload entfernt, 2023-12-27)

4ab6dc4 (Änderungen an der Ansicht der Einstellungen, 2023-12-27)

c2bc8ee (Interaktion grob beschrieben., 2023-12-27)

34ffdd7 (Beispiel Bild und Erklärung eingefügt, 2023-12-27)

f482e74 (Default strich farbe auf 255 255 255 gesetzt, 2023-12-27)

5656186 (RGBA und pixel können übergeben werden., 2023-12-27)

a2019b5 (Model angepasst, so dass RBA Opacity und Pixel abgespeichert werden können., 2023-12-27)

9eae24b (Parallele koordinaten erklärt, 2023-12-27)

8c30edc (breite Scatterplott angepasst, 2023-12-27)

047c423 (Min Max Anzeige Eingefügt, 2023-12-27)

eb7e031 (NV Gerade eingefügt, 2023-12-27)

b1c4312 (Aufgabe der Visualisierung1 beschrieen., 2023-12-27)

e0eab45 (Seitenumbrüche eingefügt, 2023-12-27)

66a7038 (Pakete neu geladen und Datei umbenannt, 2023-12-27)

f7a75ed (Bild hinzugefügt und neue Latex pakete hinzugefügt (csquotes und bookmark), 2023-12-27)

92cf9e3 (Bug fix in graph, droppdown3 wurde bei leerem droppdown2 in die [] eingegeben und dann gepusht, 2023-12-27)

4834c42 (Problem bei colorTon gelöst. redsInterpolar braucht werte zwischen [0-1] und kann nicht mit negativen zahlen arbeiten, 2023-12-27)

330b7be (Beschriftungen im Main eingefügt und weitere Veränderungen an der Aufgabenstellung im Bericht, 2023-12-27)

397c9c3 (Analyse der Anwendungsaufgaben, 2023-12-27)

a9c1114 (Koordinaten Ansicht repariert, 2023-12-27)

20fda00 (Beschriftung Scatterlpott entfernt, 2023-12-27)

63ac669 (Programm so umgebaut, dass die Ausgewählten Attribute für die koordinaten auch eingefügt werden, 2023-12-27)

818def5 (Ungenutze Pakete entfernt im Main, 2023-12-27)

b915681 (View angepasst und kleine Scatterplots eingefügt, diese werden sowohl im QQ Plot als auch in der visualisierung force graph hinzugefügt, 2023-12-27)

c3b4341 (Kleine Änderungen an dem Anforderung an das Design.ünd Blueprint für Interatktion

von 2 an 3, 2023-12-26)
 1590189 (In Nominale und Quantitative unterteilt, 2023-12-26)
 7bae36a (Daten: Aufarbeitung weiter ergänzt und umgeschrieben, 2023-12-26)
 aa97925 (Datenvorverarbeitung erklärt, 2023-12-26)
 795189f (sleep2point erklärt, 2023-12-26)
 a17fcdd (sleep2Point angepasst, so dass die drei Schlafphasen als absolute Daten angegeben werden., 2023-12-26)
 74fb29e (Beschreiben der Datenvorverarbeitung., 2023-12-26)
 b60137b (Bereitstellung der Daten über Repository, 2023-12-26)
 5436a86 (Kritische Analyse der Anwendungsfähigkeit des Datensatzes, 2023-12-26)
 084e660 (Attribute beschrieben., 2023-12-26)
 5cb06e0 (Datenpunkte: Age, gender, Bedtime, Wakuptime Sleep Duration, Schlaf Effizienz und REM erklärt., 2023-12-26)
 75482b1 (Abschnitt Daten: Daten beschrieben und Zielgruppe umformuliert, 2023-12-26)
 0344d6b (Layout angepasst, 2023-12-25)
 38887c2 (Versuch beim Hovern den Korrelationswert anzeigen zu lassen, 2023-12-25)
 1fa6d27 (Dropdown Fehler in my Leichtschlaf. Default MaxFilter von 1 auf 10000 geändert., 2023-12-25)
 47d6f08 (Color Scale rot und grün hinzugefügt und statt rgba eingefügt, 2023-12-25)
 aa14bc1 (Layout Änderungen am div, 2023-12-25)
 3189e51 (Farbe angepasst rot -> gelb besser sichtbar, 2023-12-25)
 25d6cb5 (Ansicht wieder normalisiert und basierend auf vh eingeführt, 2023-12-25)
 694dcac (Kleine Änderungen an den Div Attributen. (Nächster Commit Änderungen an der Darstellung der Kreise in ImpactGraph, 2023-12-25)
 9e710b1 (Farbskalierung nicht notwendig, da rgba hier mit Werten zwischen 0-1 arbeitet., 2023-12-25)
 303134a (Linie eingefügt um die "Impactors" mit dem Attribut zu verbinden. (Anpassung der Größe und Farbe linear., 2023-12-25)
 e0102a6 (Fehler in my filteralkohol behoben: Falsches Attribut, 2023-12-24)
 3dee53d (Input für Max hinzugefügt und Layout etwas aufgeräumt, 2023-12-24)
 c3fc2fd (Model hat minWert und maxWert bekommen. Input Feld implementiert um den hinterlegten Wert zu aktualisieren, 2023-12-24)
 ebda702 (Raucher und gender werden jetzt direkt in sleep2Point in Floats umgewandelt., 2023-12-24)
 0836332 (Anpassungen an der Größe und Position der HTML Elemente. Button entfernt. Nächster Schritt = Einfügen einer Funktion den Wertebereich einzugrenzen., 2023-12-24)
 7e13571 (Beschreibung für ImpactGraph hinzugefügt, 2023-12-24)
 a351036 (Platzhalter Headertexte eingefügt., 2023-12-24)
 830dde7 (Beschriftungen für Röntgen angepasst. "Reaktive Benennung" wird eine Anpassung

auf Datenvorverarbeitung benötigen., 2023-12-24)

2457645 (Grafiken angepasst, so dass diese in einer Reihe sind., 2023-12-24)

56a2824 (Quader eingefügt im Impact graph, "Rahmen", 2023-12-23)

f846937 (Wenn Korrelation positiv ist -> Grün, ansonsten rot. basierend auf der Stärke der Korrelation wird die Opacity eingesetzt -> Je stärker desto stärker, Kreise sind noch zu klein und nicht umrandet, 2023-12-23)

293dd12 (Name werden angezeigt. Und Funktion Runden eingefügt, die eine Zahl auf die 4 Nachkommastelle hochrundet., 2023-12-23)

c2baf5 (Fehler Gefunden, in der Festlegung der Variablen wurde nicht der richtige List (Float) ausgewählt, 2023-12-23)

65c82c4 (ImpactGraph wird jetzt gezeichnet. Beschriftung und Linien müssen noch eingefügt werden. die etwas Stimmt nicht bei der Daten aufarbeitung hinsichtlich Koffein, 2023-12-23)

3a48f7b (indiziereImpactData angepasst, so dass die List (IndexedImpactData entspricht, weiter sollte Kreis jetzt Kreise um einen Kreis zeichnen. (Farbskalierung kommt noch), 2023-12-23)

c735f2d (winkleEinteilung hinzugefügt, -> Gibt uns an an wie groß die Schritte in dem Kreis sind., 2023-12-23)

58ac811 (Indizierende funktion und aufruf der Funktion eingefügt. Sorgt dafür, dass wir den listen den index anfügen., 2023-12-23)

a0c6c1f (Anzahl Punkte berechnet-> Gibt an, wie viele Kreise um "X"gemalt werden. basierend auf der länge der ImpactData.data, 2023-12-23)

4ccd099 (Änderungen an ImpactgraphData Aufbau so angepasst, dass spätere erweiterung möglich ist., 2023-12-23)

06b3ad2 (Weitere Kreise eingefügt. Radius ist aktueller Indikator, Relative x und y position muss noch berechnet werden., 2023-12-23)

fb8b1f1 (Bessere bezeichner für die Filter eingefügt, 2023-12-22)

c8da39e (Droppdown Menü aufgestockt, 2023-12-22)

4df7a63 (ImpactGraphData hinzugefügt, neue Droppdowns anlegt. Graph (noch die Version die nur einen Kreis zeichnet eingepflegt, 2023-12-22)

f1ce2f7 (Impact Graph: Soll erstmal einen Kreis in der Mitte des Viewbox zeichnen. Rechtschreibung angepasst., 2023-12-22)

b33d5eb (Syntax Fehler entfernt List.. -> list, 2023-12-22)

b011911 (Zusammenfügen der Listen eingefügt, 2023-12-22)

7359ee2 (data change, 2023-12-22)

d4f6a3e (Funktion eingefügt combineList die zwei Listen (Float) als List (Float, Float), 2023-12-22)

8b29bb3 (rValue = Korrelationskoeffizient von X auf Y. Wichtig X muss die Werte des Attributs beinhalten, 2023-12-21)

9090239 (Data angelegt = Liste der X und Y Daten., 2023-12-21)

7f8c299 (Svg struktur in ImpactGraph eingeführt und kleine Änderungen an viewbox -> Die

höhe wird von oben und unten reduziert, wobei wir die GrösSe der Grafik beibehalten., 2023-12-21)

833a69c (dat change. egal, 2023-12-21)

861902c (import entfernt, Problem bei der Umbenennung, 2023-12-21)

cd6f099 (Kopie von ELM-Stat.Stat 6.0.2, 2023-12-21)

fb64a1a (Summe hinzugwfügt um die Summe einer Liste zu berechnen., 2023-12-21)

b358792 (ImpactGraph angelegt., 2023-12-21)

684911f (Boolean werte können nicht NV sein., 2023-12-21)

c52ce41 (Raucher und Gender im Dropdown eingefügt, 2023-12-21)

3e25d26 (Helferfunktionen für Raucher und Gender eingebaut, 2023-12-21)

1392f1b (Geschlecht wird jetzt umgewandelt zu 1 für NO und 2 Für Yes, 2023-12-21)

c54f79c (Doppelte Variable entfernt, 2023-12-21)

03aae77 (Update d.dat file, 2023-12-21)

664711b (Verschieben der Definitionen für die GrösSe des SVG in Scatterplott, da es ansonsten zu Zirkelabhängigkeiten kommt., 2023-12-21)

ff27c95 (Variablen die die gröSse der Grafiken bestimmen werden jetzt im Main festgelegt und an die anderen Funktionen übergeben, 2023-12-21)

8791093 (Weitere Beschreibung der Daten und anpassung, des Namens der Im NV plott angezeigt wird. Datein und ein paar Funktionen umbenannst, so dass diese mehr Sinn ergeben, 2023-12-21)

0d9326f (Aufgaben und Zielstellung der Visualisierung. Oberflächliche Beschreibung und Logik., 2023-12-21)

330f687 (Plan und erklärung für die Manipulation der Daten, 2023-12-21)

4cb2a6c (Daten erklärung im Bericht eingefügt, 2023-12-21)

6d5214b (Rechtschreib Fehler korrigiert: Schlaf Effizienz (Effizient), 2023-12-21)

6f37e7a (Guter Zustand. Morgen anlegen von entweder Pixelorientierter Technik oder Vektororientierter Technik., 2023-12-20)

2d6678d (Anpassung der viewbox, 2023-12-20)

bb936d9 (Line in NV Funktioniert jetzt: Das x1,y1,x2,y2 ist jetzt als: TypedSvg.Attributes.In-Px... angelegt, 2023-12-20)

264e984 (NV funktioniert jetzt, Noch falsche anzeigeposition, 2023-12-20)

80fcb15 (berechnung der NV eingeführt+ die funktionen die in der HA bereits gebaut worden sind, 2023-12-20)

ee4af34 (Boxplott war nicht im scope der let function, daher ist das Problem gelöst. Weiter: Planänderung mit dem Scatterplott. Soll die X-Variable anpassen., 2023-12-20)

13e67d2 (Anpassungen an dem Boxplott, unnötiges entfernt und Funktion agelegt zum Überge-

ben der Losten an die Boxplott funktion, 2023-12-20)

b64c9ab (Ein paar packete importiert, 2023-12-20)

1e71f5c (Exposed Boxplott in Main, 2023-12-20)

a3d6072 (Boxplott angelegt, 2023-12-20)

c421ec0 (Code für den Scatterplott übertragen., 2023-12-20)

d579e78 (Buttons eingefügt. Scatterplott berechnungen noch verändern, 2023-12-20)

b221249 (Model Msg und update angepasst um einen wechsen auf verscheidene Seiten möglich zu machen, 2023-12-20)

5092532 (Footer in die View eingefügt., 2023-12-20)

73c9e09 (Neue Filter angelegt., 2023-12-20)

b55068c (Model angepasst und neuen Typ Grafik angelegt, 2023-12-20)

e87fc5f (Einladen gelöst. Laag am leerzeichen, 2023-12-20)

1576ed4 (Bezeichner des Dropdown angepasst und Problemmit dem hüpfenden Scatterplott gelöst. (Lag an dem String: ,Html.text (toString combinedList))) Alter wird nciht richtig eingelesen- > Liste bleibt leer... [], 2023-12-20)

6f5db64 (Angepasste position der Axen bezeichner. Problem, wenn gewisse werte ausgewählt werden Hüpf die grafik nach unten, 2023-12-20)

a9e5bb7 (Gebastelt an der SVG viewBox., 2023-12-20)

c7654bb (div so angepasst, dass die grafik immer neben der Sidebar angezeigt wird und nicht-mehr unter dem Header, 2023-12-20)

60d4207 (Änderungen am view scuess um absätze und anpassung der schriftart, 2023-12-20)

dbe6d92 (div angepasst für den scatterplott, 2023-12-20)

3b34974 (Veränderung an der gröSSe der Scale. Problem ist, dass das div sich dynamisch anpasst., 2023-12-19)

6803be1 (Scatterplott sollte jetzt getrackt werden, 2023-12-19)

53ed700 (Positionierung geändert anhand vom div., 2023-12-19)

8329061 (Kleine Änderung an der Positionierung, 2023-12-19)

46b575e (Scatterplott wird gezeichnet aber an der falschen Position., 2023-12-19)

5b4a4c7 (Änderungen an den paketen, Problem identifiziert: Falsched wideextent:, 2023-12-19)

42c8963 (Änderung des datentyps von daten zu List(angepasste_Daten) durch nutzen von fil-

terMap, 2023-12-19)
 f290e19 (Geschlecht entfernt und die Daten aufarbeitung verändert, 2023-12-19)
 6e0aeac (Funktioniert: test daten zurückgesetzt, 2023-12-19)
 827c7e0 (Überprüfen ob das aussortieren auch beim Rauchrn funtkioniert, 2023-12-19)
 efdc5c7 (Änderungen an der Dartsellung der Filter:, 2023-12-19)
 e49b0e9 (Filter im dropp down eingefügt., 2023-12-19)
 f0404b6 (Änderungen an sleep2Point zurückgedreht, 2023-12-19)
 e4af450 (Update und view angepasst die neuen Model.Typen zu akzeptieren, 2023-12-19)
 22478e6 (model erweitert, 2023-12-19)
 ac8337f (Model anders definier, so dass wir zustände specihern können., 2023-12-19)
 16f1259 (ich muss die Model Struktur umbauen., 2023-12-19)
 ead37ab (Ein paar Variablen angelegt für die positionierung der Grafiken, 2023-12-19)
 a9117be (Überprüfen ob die sleeptopoint funktion funktioniert, 2023-12-19)
 89f02b9 (einsetzen der stringtounverarbeite, 2023-12-19)
 c3ed763 (Kleine anpassung am decoder und anpassung am ablauf, 2023-12-19)
 334edf9 (Fix an der Modul Bezeichnung, 2023-12-19)
 4aabb2 (Änderung am header und Footer, 2023-12-19)
 ae6c248 (Bug fixes am sleep2Point, 2023-12-19)
 3499dce (Kleiner Change am Dat, 2023-12-19)
 2eb674f (Anpassungen an der transformation der daten, 2023-12-19)
 e17f7d9 (Aussortierte Daten Typ angelegt, 2023-12-18)
 f74cf01 (Überprüfung des umgangs mit Nothing, 2023-12-18)
 7bf283a (Neuen Decoder implementiert. Emöglicht Decoding pipeline wie in json, 2023-12-18)
 2fbbe15 (Verschieben der Strukturtypen in das Main zum vorbereiten der Daten, 2023-12-18)
 3aad4f6 (Scatterplott typ erklärungen, 2023-12-18)
 0c9b1d2 (Daten strukturen angelegt für den Scatterplott, 2023-12-18)
 1d40326 (Pakete installiert und Scatterplott angelegt, 2023-12-18)
 d378c59 (view angepasst-> head, inhalt und footer wird angezeigt, 2023-12-18)
 6abe3fc (footer hinzugefügt, 2023-12-18)
 0367826 (inhalt hinzugefügt, 2023-12-18)
 9da026d (Header angelegt, 2023-12-18)
 60e0279 (Link zum einladen der daten angepasst., 2023-12-18)
 443a483 (Dateiname geändert, 2023-12-18)
 dbb0768 (Änderungen am Bericht, 2023-12-18)
 806cdf7 (Änderungen am Bericht, 2023-12-18)
 5608717 (Plan für das Paper und To Do lits eeingefügt, 2023-12-18)
 f62dea3 (Zielgruppe und Visualisierungs Techniken bearbeitet, 2023-12-17)
 b19aecd (Bricht Änderung, 2023-12-17)
 8368dc0 (Anpassung an den Types und decoder, 2023-12-17)
 831dae (Link für csv daten satz, 2023-12-17)
 04518d6 (typen aktualisiert, 2023-12-17)
 fecd3b7 (Pakete Csv, Color und http installiert, 2023-12-17)
 fd80fc4 (csv datei hinzugefügt, 2023-12-17)
 981756e (Bericht vorbereitet, 2023-12-17)
 7475d73 (Initialisierung, 2023-12-17)
 9ea934c (Elm Init und Main kopiert, 2023-12-17)

Literatur

- [1] Steven E. Carlson u.a. „The thief of (bed)time: Examination of the daily associations between bedtime procrastination and multidimensional sleep health“. In: *Sleep Health* 9.6 (2023), S. 903–909. ISSN: 2352-7218. DOI: <https://doi.org/10.1016/j.sleh.2023.08.002>. URL: <https://www.sciencedirect.com/science/article/pii/S2352721823001717>.
- [2] Amanda L. Tapia u.a. „Race and sex differences in the longitudinal changes in multidimensional self-reported sleep health characteristics in aging older adults“. In: *Sleep Health* 9.6 (2023), S. 947–958. ISSN: 2352-7218. DOI: <https://doi.org/10.1016/j.sleh.2023.08.008>. URL: <https://www.sciencedirect.com/science/article/pii/S235272182300178X>.