**Detaillierte Dokumentation: Univariate Analyse der Einflussfaktoren auf die Heilungsdauer bei Polytrauma-Patienten**

**Projektübersicht**

Dieses Dokument bietet eine umfassende Dokumentation der univariaten Analyse, die durchgeführt wurde, um Faktoren zu identifizieren, die die Heilungsdauer bei Polytrauma-Patienten beeinflussen. Die Analyse ist Teil des umfassenderen Polytrauma-Analyse-Projekts, das darauf abzielt, die Faktoren zu verstehen, die die Genesungszeit und Ergebnisse für Patienten mit mehrfachen traumatischen Verletzungen beeinflussen.

**Analysezweck und -umfang**

Die univariate Analyse wurde konzipiert, um zu untersuchen, wie einzelne Faktoren (ohne Berücksichtigung von Wechselwirkungen) die Heilungsdauer von Polytrauma-Patienten beeinflussen. Insbesondere konzentrierte sich die Analyse auf:

1. Die Auswirkung von Verletzungen bestimmter Körperteile
2. Den Zusammenhang zwischen der Anzahl verletzter Körperteile und der Heilungsdauer
3. Den Einfluss demografischer Faktoren (Alter und Geschlecht) auf die Heilungszeit

Die Analyse dient als Grundlage für nachfolgende multivariate Analysen und potentielle prädiktive Modellierung der Heilungsdauer.

**Datenquelle und -aufbereitung**

**Quelldaten**

* Primärer Datensatz: Polytrauma\_Analysis\_Processed.xlsx (aus Schritt 1 der Verarbeitungspipeline)
* Der Datensatz enthielt 153 Datensätze, die 30 eindeutige Patienten repräsentieren
* Jeder Datensatz stellt einen Patientenbesuch dar, mit mehreren Besuchen pro Patient

**Aufbereitung der Patientendaten**

* Ein Datensatz auf Patientenebene wurde erstellt mit einer Zeile pro Patient (n=30)
* Die Heilungsdauer wurde definiert als Tage vom Unfall bis zum letzten dokumentierten Besuch
* Für kategoriale Felder (z.B. Vorhandensein von Verletzungen) wurde ein Patient als "Ja" markiert, wenn bei einem beliebigen Besuch "Ja" angezeigt wurde
* Für jeden Patienten wurden Informationen zu Verletzungen verschiedener Körperteile extrahiert
* Schlüsselstatistiken für die Heilungsdauer: Min=182 Tage, Max=1233 Tage, Mittelwert=579,4 Tage, Median=540,0 Tage

**Methodischer Ansatz**

**Statistische Methoden**

1. **Normalitätsbewertung**:
   * Shapiro-Wilk-Test zur Überprüfung der Normalverteilung in Untergruppen
   * Verwendet zur Bestimmung geeigneter statistischer Tests (parametrisch vs. nicht-parametrisch)
2. **Körperteilanalyse**:
   * Vergleich der Heilungsdauer zwischen Patienten mit vs. ohne jeden Verletzungstyp
   * T-Tests für normalverteilte Daten
   * Mann-Whitney-U-Tests für nicht normalverteilte Daten oder kleine Stichprobengrößen
   * Berechnung der Effektstärke (Cohen's d) zur Quantifizierung der Größenordnung von Unterschieden
   * Visualisierung mit Violin-Plots mit eingebetteten Box-Plots
3. **Verletzungsanzahlanalyse**:
   * Korrelationsanalyse (Pearson's r) zwischen Verletzungsanzahl und Heilungsdauer
   * Kategorisierung in Schweregrade (1-2, 3-4, 5+ Verletzungen)
   * ANOVA zum Testen auf Unterschiede zwischen Schweregradkategorien
   * Visualisierung mit Streudiagrammen und Box-Plots
4. **Demografische Analyse**:
   * Korrelationsanalyse für Alter vs. Heilungsdauer
   * Altersdekaden-Kategorisierung und ANOVA-Vergleich
   * Geschlechtsbasierter Vergleich (durch Datenverfügbarkeit eingeschränkt)
   * Visualisierung mit Streudiagrammen und Box-Plots

**Signifikanzkriterien**

* Konventionelles statistisches Signifikanzniveau: p < 0,05
* Nahezu signifikant bemerkt bei: 0,05 ≤ p < 0,10
* Interpretation der Effektstärke:
  + Cohen's d < 0,2: vernachlässigbar
  + 0,2 ≤ d < 0,5: klein
  + 0,5 ≤ d < 0,8: mittel
  + d ≥ 0,8: groß

**Ergebnisse und Erkenntnisse**

**Körperteilverletzungen**

**Statistisch signifikante Faktoren**

1. **Abdominale Verletzungen**:
   * 7 Patienten mit Verletzung, 23 ohne
   * Mann-Whitney-U-Test: U=136,00, p=0,0049
   * Effektstärke: Cohen's d = 1,48 (groß)
   * Mittlere Heilungsdauer: Mit Verletzung = 818,3 Tage, Ohne = 479,1 Tage
   * Mittlere Differenz: +339,2 Tage länger mit abdominaler Verletzung
2. **Kopfverletzungen**:
   * 15 Patienten mit Verletzung, 15 ohne
   * T-Test: t=-2,83, p=0,0090
   * Effektstärke: Cohen's d = 1,03 (groß)
   * Mittlere Heilungsdauer: Mit Verletzung = 694,0 Tage, Ohne = 444,9 Tage
   * Mittlere Differenz: +249,1 Tage länger mit Kopfverletzung

**Fast signifikante Faktoren**

1. **Wirbelsäulenverletzungen**:
   * 17 Patienten mit Verletzung, 13 ohne
   * T-Test: t=-1,99, p=0,0591
   * Effektstärke: Cohen's d = 0,76 (mittel)
   * Mittlere Heilungsdauer: Mit Verletzung = 685,6 Tage, Ohne = 492,6 Tage
   * Mittlere Differenz: +193,0 Tage länger mit Wirbelsäulenverletzung
2. **Kombinierte Armverletzungen**:
   * 19 Patienten mit Verletzung, 11 ohne
   * T-Test: t=1,88, p=0,0706
   * Effektstärke: Cohen's d = 0,65 (mittel)
   * Mittlere Heilungsdauer: Mit Verletzung = 625,3 Tage, Ohne = 455,7 Tage
   * Mittlere Differenz: +169,6 Tage länger mit Armverletzung
3. **Rechtsseitige Armverletzungen**:
   * 8 Patienten mit Verletzung, 22 ohne
   * T-Test: t=1,64, p=0,1333
   * Effektstärke: Cohen's d = 0,79 (mittel)
   * Mittlere Heilungsdauer: Mit Verletzung = 722,9 Tage, Ohne = 519,4 Tage
   * Mittlere Differenz: +203,5 Tage länger mit rechtsseitiger Armverletzung

**Nicht-signifikante Faktoren (kleine oder vernachlässigbare Effekte)**

* Linksseitige Armverletzungen: d = 0,33 (klein), p = 0,3721
* Thoraxverletzungen: d = 0,24 (klein), p = 0,5393
* Beckenverletzungen: d = 0,17 (vernachlässigbar), p = 0,4316
* Kombinierte Beinverletzungen: d = 0,15 (vernachlässigbar), p = 0,6909
* Linksseitige Beinverletzungen: d = 0,12 (vernachlässigbar), p = 0,7537
* Rechtsseitige Beinverletzungen: d = 0,01 (vernachlässigbar), p = 0,9828
* Halsverletzungen: Unzureichende Stichprobe (n=1) für statistische Tests

**Verletzungsanzahlanalyse**

* Korrelation zwischen Verletzungsanzahl und Heilungsdauer: r = 0,0473, p = 0,8038 (nicht signifikant)
* Kategorisierte Analyse (ANOVA): F = 2,79, p = 0,0793 (nähert sich der Signifikanz)
* Mittlere Heilungsdauer nach Kategorie:
  + 1-2 Verletzungen (n=9): 464,0 Tage
  + 3-4 Verletzungen (n=10): 728,0 Tage
  + 5+ Verletzungen (n=11): 538,7 Tage
* Dies deutet auf eine nicht-lineare Beziehung hin, bei der moderates Polytrauma längere Heilungszeiten erfordern kann als entweder geringfügige oder sehr schwere Fälle

**Altersanalyse**

* Gesamtkorrelation: r = -0,2829, p = 0,1298 (nicht signifikant)
* Altersdekaden ANOVA: F = 3,74, p = 0,0102 (signifikant)
* Mittlere Heilungsdauer nach Dekade:
  + 20er (n=6): 513,5 Tage
  + 30er (n=3): 690,7 Tage
  + 40er (n=5): 943,6 Tage (höchste)
  + 50er (n=7): 450,9 Tage
  + 60er (n=4): 582,8 Tage
  + 70er (n=2): 399,0 Tage
  + 80er (n=2): 304,5 Tage (niedrigste)
* Die Ergebnisse zeigen ein nicht-lineares Muster, wobei Patienten mittleren Alters (40er) die längsten Heilungsdauern haben

**Geschlechtsanalyse**

* Die Protokolldatei vermerkte unzureichende Geschlechtsdaten für eine aussagekräftige Analyse: "WARNING - Nicht genügend Geschlechtsdaten für Analyse vorhanden"

**Technische Implementierungsdetails**

**Software und Bibliotheken**

* Python für Datenverarbeitung und statistische Analyse
* Hauptbibliotheken:
  + pandas: Datenmanipulation und -analyse
  + numpy: Numerisches Computing
  + scipy.stats: Statistische Tests (t-Test, Mann-Whitney-U, ANOVA)
  + matplotlib und seaborn: Datenvisualisierung
  + logging: Umfassendes Protokollieren von Prozess und Ergebnissen

**Implementierungsstruktur**

Die Analyse wurde im Skript univariate\_analysis.py mit den folgenden Hauptfunktionen implementiert:

1. **setup\_logging()**: Konfigurierte Protokollierung für das Skript
2. **load\_dataset()**: Lud und validierte den Datensatz
3. **prepare\_patient\_level\_data()**: Erstellte einen Datensatz auf Patientenebene für die Analyse
4. **test\_normality()**: Testete, ob Daten einer Normalverteilung folgen
5. **analyze\_body\_part\_impact()**: Analysierte die Auswirkung von Körperteilverletzungen auf die Heilungsdauer
6. **analyze\_injury\_count\_impact()**: Analysierte den Zusammenhang zwischen Verletzungsanzahl und Heilungsdauer
7. **analyze\_demographic\_impact()**: Analysierte die Auswirkung demografischer Faktoren auf die Heilungsdauer
8. **export\_results\_to\_excel()**: Exportierte Ergebnisse in Excel-Format
9. **generate\_summary\_report()**: Generierte einen umfassenden Markdown-Bericht der Ergebnisse

**Ausgabedateien und -orte**

1. **Datendateien**:
   * Patientendaten: data/output/step4/univariate\_analysis/patient\_level\_data.xlsx
   * Analyseergebnisse: data/output/step4/univariate\_analysis/univariate\_analysis\_results.xlsx
2. **Visualisierungsdateien** (alle gespeichert in plots/step4/univariate\_analysis/):
   * Körperteilvisualisierungen (z.B. Kopf\_Heilungsdauer.png)
   * Verletzungsanzahlvisualisierungen:
     + Verletzungsanzahl\_Heilungsdauer\_Scatter.png
     + Verletzungsanzahl\_Heilungsdauer\_Box.png
   * Altersvisualisierungen:
     + Alter\_Heilungsdauer.png
     + Altersdekade\_Heilungsdauer.png
3. **Dokumentation**:
   * Analysebericht: data/output/step4/univariate\_analysis/univariate\_analysis\_report.md
   * Detailliertes Protokoll: logs/step4/univariate\_analysis.log

**Einschränkungen und Überlegungen**

**Statistische Einschränkungen**

1. **Stichprobengröße**: Die Stichprobe von 30 Patienten begrenzt die statistische Power, besonders für Subgruppenanalysen
2. **Multiples Testen**: Tests über mehrere Körperteile hinweg erhöhen das Risiko von Typ-I-Fehlern (falsch positive Ergebnisse)
3. **Ungleiche Untergruppen**: Einige Verletzungstypen hatten wenige Fälle (z.B. Halsverletzung n=1)
4. **Fehlende Daten**: Geschlechtsanalyse war aufgrund unzureichender Daten nicht möglich

**Methodische Überlegungen**

1. **Definition der Heilungsdauer**: Zeit bis zum letzten Besuch entspricht möglicherweise nicht perfekt der vollständigen physiologischen Erholung
2. **Univariater Ansatz**: Die Analyse berücksichtigt keine Wechselwirkungen zwischen Faktoren
3. **Patientenauswahl**: Ergebnisse könnten durch die spezifische Patientenkohorte beeinflusst sein und nicht auf alle Polytrauma-Fälle verallgemeinert werden können
4. **Zeitliche Aspekte**: Die Analyse berücksichtigt keine möglichen Änderungen in Behandlungsansätzen über die Zeit

**Empfohlene nächste Schritte**

Basierend auf den Ergebnissen dieser univariaten Analyse werden die folgenden nächsten Schritte empfohlen:

1. **Multivariate Analyse**:
   * Multiple Regressionsmodellierung unter Verwendung signifikanter Faktoren (Bauchverletzungen, Kopfverletzungen, Wirbelsäulenverletzungen, Alter)
   * Berücksichtigung von Interaktionstermen, insbesondere zwischen Verletzungstypen
2. **Überlebensanalyse**:
   * Implementierung einer zeitabhängigen Analyse, um den zeitlichen Charakter der Heilung besser zu erfassen
   * Erwägung verschiedener Definitionen von "Erholungsereignissen" über den letzten Besuch hinaus
3. **Merkmalsgewichtung**:
   * Fokus auf die signifikantesten Faktoren (Bauch- und Kopfverletzungen) in nachfolgenden Analysen
   * Eingehendere Untersuchung der Altersmuster, um die nicht-lineare Beziehung zur Heilungsdauer zu verstehen
4. **Klinische Korrelation**:
   * Überprüfung der Ergebnisse mit medizinischen Spezialisten, um die Befunde anhand klinischer Erfahrung zu validieren
   * Erwägung der Einbeziehung von Lebensqualitäts- oder funktionellen Outcome-Maßnahmen neben der Heilungsdauer
5. **Datenverbesserung**:
   * Berücksichtigung zusätzlicher Variablen, die die kontraintuitiven Befunde erklären könnten (z.B. Behandlungsintensität nach Altersgruppe)
   * Adressierung fehlender Datenfelder, insbesondere Geschlechtsinformationen

**Schlussfolgerung**

Die univariate Analyse hat erfolgreich mehrere Faktoren identifiziert, die die Heilungsdauer bei Polytrauma-Patienten signifikant beeinflussen. Bauch- und Kopfverletzungen erwiesen sich als die signifikantesten Faktoren mit großen Effektstärken und statistischer Signifikanz. Die beobachteten nicht-linearen Beziehungen sowohl mit der Verletzungsanzahl als auch mit dem Alter deuten auf komplexe Muster hin, die weitere Untersuchungen durch multivariate Ansätze rechtfertigen. Trotz der begrenzten Stichprobengröße bietet diese Analyse wertvolle Einblicke zur Lenkung nachfolgender analytischer Schritte und klinischer Entscheidungsfindung für Polytrauma-Patienten.