

**Projektbericht zum Modul Information Retrieval und
Visualisierung Sommersemester 2021**

Titel des Dokuments

Richard Brennecke

Matrikelnummer:

Inhaltsverzeichnis

1.	Inhaltsverzeichnis	2
2.	Einleitung	3
2.1	Anwendungshintergrund	3
2.2	Zielgruppen	3
2.3	Überblick und Beiträge	4
3.	Daten	4
3.1	Technische Breitstellung der Daten	4
3.2	Datenvorverarbeitung	4
4.	Visualisierung	4
4.1	Analyse der Anwendungsfälle	5
4.2	Anforderungen an die Visualisierungen	5
4.3	Präsentation der Visualisierung	5
4.3.1	Visualisierung Eins	6
4.3.2	Visualisierung Zwei	6
4.3.3	Visualisierung Drei	6
4.4	Interaktion	6
5.	Implementierung	6
6.	Anwendungsfälle	6
6.1	Anwendung Visualisierung Eins	7
6.2	Anwendung Visualisierung Zwei	7
6.3	Anwendung Visualisierung Drei	7
7.	Verwandte Arbeiten	7
8.	Zusammenfassung und Ausblick	7
9.	Anhang	8

Seitenabstand Ränder alle 2

Schriftart: Times new roman

Zeilenabstand: 1,5

Zitierung: mit eckige Klammer mit Nummer dahinter

1. Einleitung

- Zielproblem:
 - Analyse von verschiedenen Wein Daten
 - Interessante Zusammenhänge herausfinden
 - Erkenntnisgewinn aus diesen Daten
- Fragen welche Beantwortet werden sollen
 - Gibt es Zusammenhänge zwischen dem Körper/Body/Süße/Alk. und dem Preis eines Weines?
 - Scatterplot
 - Hängen Daten über mehrere Dimensionen zusammen?
 - Parallele Koordinaten
 - Wo kommen die meisten Weine her?
 - Baumhierarchie

1.1 Anwendungshintergrund

- Erklärung der Informationsvisualisierungen
 - Was ist ein Scatterplott/ parallele Koordinaten/ Baumhierarchie?
- Hintergrund Daten bereitstellen?
 - Informationen zu verschiedenen Daten
 - Wie kann der Körper/ Süße/ Säure/ Gerbstoffe bestimmt werden?
 - Was haben die Jahre für Einfluss auf die Weine?
 - Was bedeuten die Felder Verwendung und Type (Sorte, Lokal, Produzent)

1.2 Zielgruppen

- Weininteressierte
 - Vorwissen
 - Kaum bis gar nicht
 - Erkenntnisse:
 - Zusammenhang von verschiedenen Kriterien beim Wein
 - Zusammenhänge kurz erklären
 - Informationsgewinnung
 - Kennenlernen von Weinen
 - Entdeckung neuer Weine welche sie trinken möchten
- Weineinkäufer
 - Vorwissen
 - Vorhanden bis Exzellent
 - Erkenntnisse:
 - Entdecken von neuen Sorten die sein Sortiment ergänzen
 - Entdeckung von neuen Sorten die ggf. Außergewöhnlich sind
 - Beratung der Kunden die gewisse Vorlieben haben
- Weinexperte
 - Vorwissen
 - Gut bis ausgeprägt
 - Erkenntnisse:
 - Entdecken von neuen Sorten die seinem Geschmack entsprechen
 - Bessere Einschätzung seiner bisherigen Weine

1.3 Überblick und Beiträge

- Erklären welche Daten verwendet wurden
 - o Oberkategorien kurz erklären
- Visualisierungstechniken erklären
- Beiträge
 - o Mehrwert der Techniken für die Darstellung der Daten
 - Scatterplot -> Gegenüberstellung von Dimensionen

2. Daten

- Beschreibung der gegebenen Daten
- Eignung der Daten für die Zielgruppen
 - o Weintressierte
 - Gut
 - o Weinexperte/ Weinverkäufer
 - Teilweise
 - Daten können unvollständig sein -> und haben zu wenig Aussagekraft mit Body, Süße usw.
- Fragestellungen
 - o Gut da Dimensionen erkannt werden können
 - o Herkunft der Daten erkennbar -> Teilweise aber unvollständig
- Daten Ergänzung
 - o Mussten mit Geo Daten ergänzt werden für Baumhierarchie da sonst kein Ursprungsknoten
 - o Und wo kommen die her?

2.1 Technische Breitstellung der Daten

- Daten Zugänglich?
 - o Sind über GitHub für die verschiedenen Darstellungen erreichbar
- Formate
 - o CSV -> Für alle Daten die nicht ergänzt wurden
 - o JSON -> Für alle Geo Daten
- Besonderheiten
 - o CSV
 - 0 oder nichts bedeutet dort ist nichts vorhanden
 - Trennung durch normales komma
 - o JSON
 - Nur Name und Beziehung (Eltern – Kind) in der Datei vorhanden
 - Länder welche keine Weine Produzieren wurden außen vor gelassen

2.2 Datenvorverarbeitung

- Datenverarbeitungsschritte
 - o Bekanntmachen mit den Daten (Umwandeln in besser Lesbares Format -> Excel)
 - o Bearbeitung der Daten
 - Namen überarbeiten (aus den Zahlen entfernen)
 - Zahlen bearbeiten

- Umrechnung von WON in Euro
 - Durchschnitte Bilden
 - Namen werden Überarbeitet (Umlaute und Apostrophe nicht richtig konvertiert)
 - Übersetzten der Spaltenüberschriften
 - Bereitstellung der Daten für JSON
 - Herauslösen der Datensätze, welche nicht komplett sind
 - Überführung der Daten
- Daten weggelassen
 - Daten nicht mehr lesbar (eine Japanischer Wein wo nicht mehr Rückschlüsse gezogen werden könnten)
 - Nichts Werte
 - Durschnitte
 - Gebildet über die Trinktemperatur, Alkoholgehalt
 - Daten konnten ansonsten nicht eingelesen werden
 - Außerdem lag der unterschied durchschnittlich nicht bei mehr als 2 Gard
 - Aussagekräftiger?
 - So hat man noch Toleranz beim der Temperatur und Alkohol ohne dass sich die anderen Werte ändern müssten (Body, Süße usw.)

3. Visualisierung

- Analyse kann erst gemacht werden, wenn Visualisierungen fertig sind

3.1 Analyse der Anwendungsaufgaben

- Analyse kann erst gemacht werden, wenn Visualisierungen fertig sind
- Anwendungsaufgaben
 - Wie helfen die Darstellungen die genannten Problemstellungen zu beantworten?
 - Hauptziel möglich viele Einblicke in das Thema der Weine zu erhalten
 - Zusammenhänge zwischen den einzelnen Eigenschaften
 - Zusammenhang von Produktionsmenge eines Landes und der Anzahl der Weine
 - Hervorstechen von Datensätze aus Allgemeinheit, Zusammenhänge oder Trends von Eigenschaften
- Mentale Modelle
 - Welche Visualisieren eignen sich um das alles zu kombinieren von Wissen und den Modellen (intuitiv erkennbar)

3.2 Anforderungen an die Visualisierungen

- Analyse kann erst gemacht werden, wenn Visualisierungen fertig sind

3.3 Präsentation der Visualisierung

- Analyse kann erst gemacht werden, wenn Visualisierungen fertig sind
- Vorstellen, Interaktivität, Designentscheidungen begründen Diskutieren wieso nicht anderen Techniken verwenden worden sind

3.3.1 Visualisierung Eins

- Wird ein Scatterplot

3.3.2 Visualisierung Zwei

- Wird Parallele Koordinaten

3.3.3 Visualisierung Drei

- Wird eine Baumhierarchie

3.4 Interaktion

- Scatterplot und Parallele Koordinaten
 - o Buttons zum verändern/ verschieben der Dimensionen
- Baumhierarchie
 - o Keine nur anschauen
- Zweck der Interaktion
- Warum wurden andere Interaktionen umgesetzt und nicht andere?
- Begründung Interaktion zwischen denen nicht mit dabei

4. Implementierung

- Kann erst eingeschätzt werden, nachdem es fertig gestellt worden ist
 - o Aktuell hoher Aufwand und nur Baumhierarchie konnte sehr einfach aus Übung übernommen werden
- Gliederung des ELM Codes
- Übungsadaption
- Datenstruktur Modells bei den verschiedenen Interaktionen
- Bei uns in einem Record gespeichert im Main und dann im Update wird auf einen record zugegriffen

5. Anwendungsfälle

- Erst nach Fertigstellung der Visualisierungen möglich
- Spezifischer Anwendungsfall -> wo Muster da sind oder nicht was es zu was besonderen macht
- Relevanz für die Zielgruppe

- Möglichkeit Umsetzung mit anderen Personen

5.1 Anwendung Visualisierung Eins

- Anwendungsfall für Scatterplot

5.2 Anwendung Visualisierung Zwei

- Anwendungsfall für Parallele Koordinaten

5.3 Anwendung Visualisierung Drei

- Anwendungsfall für Baumhierarchie

6. Verwandte Arbeiten

- Aktuell noch nicht recherchiert
- Zwei Artikel diskutieren
 - Gemeinsamkeiten und Unterschiede dabei herausstellen

7. Zusammenfassung und Ausblick

- Ausblick er bei fertigem Projekt möglich
- Zusammenfassung der Beiträge
- Mehrwert für Zielgruppe und Personen
- Erweiterungen für Ebene und Datenebene

Anhang