

**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	3
3.	Daten	4
3.1	Technische Breitstellung der Daten	4
3.2	Datenvorverarbeitung	4
4.	Visualisierung	5
4.1	Analyse der Anwendungsfälle	5
4.2	Anforderungen an die Visualisierungen	5
4.3	Präsentation der Visualisierung	5
4.3.1	Visualisierung Eins	5
4.3.2	Visualisierung Zwei	5
4.3.3	Visualisierung Drei	5
4.4	Interaktion	5
5.	Implementierung	5
6.	Anwendungsfälle	6
6.1	Anwendung Visualisierung Eins	6
6.2	Anwendung Visualisierung Zwei	6
6.3	Anwendung Visualisierung Drei	6
7.	Verwandte Arbeiten	6
8.	Zusammenfassung und Ausblick	6
9.	Anhang	7

1. Einleitung

- Problem:
 - Analyse von verschiedenen Wein Daten -> Interessante Zusammenhänge herausfinden
- Fragen welche Beantwortet werden sollen
 - Gibt es zusammenhänge zwischen den verschiedenen Daten?
 - Scatterplot
 - Hängen Daten über mehrere Dimensionen zusammen?
 - Parallele Koordinaten
 - Wo kommen die meisten Weine her?
 - Baumhierarchie

1.1 Anwendungshintergrund

- Erklärung der Informationsvisualisierungen
- Kleine Beantwortung der Fragen mithilfe der Visualisierung

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
 - Kurz auf Aufbereitung eingehen
- Visualisierungstechniken erklären
- Beiträge???

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

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 Sichten der Daten (Umwandeln in besser Lesbares Format -> Excel)
 - o Bearbeiten der Daten
 - Durchschnitte Bilden
 - Namen überarbeiten
 - Übersetzen
 - o Rücküberführung zur CSV
- Daten weggelassen
 - o Daten nicht mehr lesbar (eine Japanischer Wein wo nicht mehr Rückschlüsse gezogen werden könnten)
- Durschnitte
 - o Gebildet über die Trinktemperatur, Alkoholgehalt
 - o Daten konnten ansonsten nicht eingelesen werden
 - o Außerdem lag der unterschied durchschnittlich nicht bei mehr als 2 Gard
- Aussagekräftiger?

- So hat man noch Toleranz beim der Temperatur und Alkohol ohne das sich die anderen Werte ändern müssten (Body, Süße usw.)

3. Visualisierung

- Analyse kann erst gemacht werden, wenn Visualisierungen fertig sind

3.1 Analyse der Anwendungsfälle

- Analyse kann erst gemacht werden, wenn Visualisierungen fertig sind

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

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
 - Buttons zum verändern/ verschieben der Dimensionen
- Baumhierarchie
 - Keine nur anschauen

4. Implementierung

- Kann erst eingeschätzt werden, nachdem es fertig gestellt worden ist

- Aktuell hoher Aufwand und nur Baumhierarchie konnte sehr einfach aus Übung übernommen werden

5. Anwendungsfälle

- Erst nach Fertigstellung der Visualisierungen möglich

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

7. Zusammenfassung und Ausblick

- Ausblick er bei fertigem Projekt möglich

Anhang