

# Spotify Challenge

---

Isa Kilic, Pablo Schneider, Malte Schwenker

# Roadmap

- Ansatz:
  - Assoziationsanalyse (Apriori Algorithmus)
- Aktueller Stand
  - Wo sind wir derzeit?
  - auf welche Probleme sind wir gestoßen?
- Erste Ergebnisse
- Ausblick
  - wie wollen wir weiter machen?

# Ansatz: Assoziationsanalyse

- Assoziationsregeln definieren:
  - Wenn Track X und Track Y sich in Playlist befinden, befindet sich wahrscheinlich auch Track Z in Playlist.
- Eigenschaften:
  - Support: Wie häufig Track X und Y zusammen vorkommen:
  - Konfidenz: Wenn man Track X und Y in einer Playlist hat, wie wahrscheinlich wäre dann Z.

# Strategie

- 1 Million Playlists aufteilen/sortieren für kleinere Datenmengen
- häufige Itemsets finden
  - Optimaler min\_support muss weise gewählt werden.
- mit den gefilterten Itemsets kann man Konfidenz berechnen.
- Evaluieren

# Aktueller Stand

- Umsetzen der Assoziationsanalyse
  - mithilfe von Apriori
  - justieren der Werte wie min\_support etc. für bessere Ergebnisse
- Verbindung zu Spotify mit Spotipy
  - weitere Infos aus Tracks ziehen
  - für weiteres Vorgehen filtern über bspw. Genres

# Erste Ergebnisse...

min\_supp = 0.01, Itemsets: max\_len = 3, Playlists: 3000

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
752	(spotify:track:6F609ICg9Spjrw1epsAnpa, spotify:track:0VgkVdmE4gld66l8iy6jgx, spotify:track:7yyRTcZmCiyzzJlNz6C90L, spotify:track:0v9Wz8o0BT8DU38R4ddjeH, spotify:track:0SGkqnVQo9KPytSri1H6cF, spotify:track:7KXjTSCq5nL1LoYtL7XAWS)	(spotify:track:1xzn6GDRh1oQq0xbwXa3)	0.010521	0.044589	0.010020	0.952381	21.359016	0.009551	20.063627
704	(spotify:track:0VgkVdmE4gld66l8iy6jgx, spotify:track:7yyRTcZmCiyzzJlNz6C90L, spotify:track:0v9Wz8o0BT8DU38R4ddjeH, spotify:track:0SGkqnVQo9KPytSri1H6cF, spotify:track:7KXjTSCq5nL1LoYtL7XAWS)	(spotify:track:7KXjTSCq5nL1LoYtL7XAWS)	0.013026	0.048096	0.012024	0.923077	19.192308	0.011398	12.374749
614	(spotify:track:7yyRTcZmCiyzzJlNz6C90L, spotify:track:0v9Wz8o0BT8DU38R4ddjeH, spotify:track:0SGkqnVQo9KPytSri1H6cF, spotify:track:7KXjTSCq5nL1LoYtL7XAWS)	(spotify:track:5hTpBe8h35rJ67eAWHQsJx)	0.012525	0.039579	0.011022	0.880000	22.233924	0.010526	8.003507
744	(spotify:track:0v9Wz8o0BT8DU38R4ddjeH, spotify:track:0SGkqnVQo9KPytSri1H6cF, spotify:track:7KXjTSCq5nL1LoYtL7XAWS)	(spotify:track:7yyRTcZmCiyzzJlNz6C90L)	0.013026	0.043086	0.011022	0.846154	19.638640	0.010461	6.219940
553	(spotify:track:0SGkqnVQo9KPytSri1H6cF, spotify:track:7KXjTSCq5nL1LoYtL7XAWS)	(spotify:track:7KXjTSCq5nL1LoYtL7XAWS)	0.013026	0.048096	0.011022	0.846154	17.592949	0.010396	6.187375
..	...	...	...	...	...	...	...	...	...
825	(spotify:track:7KXjTSCq5nL1LoYtL7XAWS)	(spotify:track:4Km5HrUVYTasUfiSGPJeqR, spotify:track:2771LMNxxwF62FTAdpJMQfM)	0.048096	0.018537	0.010020	0.208333	11.238739	0.009128	1.239743
247	(spotify:track:7KXjTSCq5nL1LoYtL7XAWS)	(spotify:track:2ANLarE8yHVsLWW21nj79M)	0.048096	0.021042	0.010020	0.208333	9.900794	0.009008	1.236578
268	(spotify:track:7KXjTSCq5nL1LoYtL7XAWS)	(spotify:track:6fwdbPMwP1zVStm8FybmK0)	0.048096	0.021543	0.010020	0.208333	9.670543	0.008984	1.235946
506	(spotify:track:7KXjTSCq5nL1LoYtL7XAWS)	(spotify:track:3kxfsdsCpFgN412fPnW85Y)	0.048096	0.025050	0.010020	0.208333	8.316667	0.008815	1.231516
390	(spotify:track:7KXjTSCq5nL1LoYtL7XAWS)		0.048096	0.030561	0.010020	0.208333	6.816940	0.008550	1.224554

[1014 rows x 9 columns]

To Do: Regeln mit geeigneten confidence auf die Playlist anwenden.

# Ausblick

- weiter mit den Parametern Ergebnisse evaluieren
- Fertigstellung vom Aufsplitten der Daten
  - relevante Daten aus Spotify Tracks herausfiltern
  - damit weitere Parameter erstellen
- ggbs. weitere Algorithmen gegentesten