Zeitpunkte (Datumswerte) wären super für Änderungen

Stopwörter 🡪 Liste anhängen, z.B. laden aus: <https://gist.github.com/sebleier/554280>

Actual ↓ Predicted →

# Explorativer Teil (Intro-Part, nur auf Rohtexte)

Rohtexte laden

Auszählung und Timeline nach Subreddits:

* Cluster-Balkendiagramm für Verteilung
* Small-Multiple Line für Subreddits – zeitliche Entwicklung
  + Format 3 Spalten, 2 Zeilen 🡪 gutes Steigungsverhältnis
  + Logarithmische Skala für faire Vergleiche bei unterschiedlichen Niveaus
  + Title für X- und Y-Achse aus
  + Title umbenennen
* Filterinteraktion zwischen Balkendiagramm und SMM aufheben

Word-Cloud (von Microsoft) 🡪 Alternative ist „Tag Cloud“ von xViz 🡪 mächtig, aber teilweise kostenpflichtig

* Stop-Wörter manuell
* Default-Stop-Wörter kann man testen (in Englisch gut, in Deutsch eher nicht wirklich.. (Test mit DataGroup-Daten gemacht)
* Eigene Stopwörter, die ggf. auch bestimmte Standardbegriffe aus dem Umfeld sind 🡪 auch für Endanwender editierbar über CSV/XLS machen 🡪 via eigene Tabelle und Join

Stopwörter dranhängen? 🡪 ggf. erst für Lemmas wichtig 🡪 hilft dann besser 🡪 natürlich auch in Rohtexten schon denkbar

* Stopwörter laden
* Transpose-Transformation 🡪 Merged Columns mit Seperator „Space“ 🡪 damit als ein Feld über Relation an NLP-Eergebnisse nutzbar 🡪 von WordCloud-Visual nutzbar
* Column mit ID 1 für Merge, gleiche Pseudocolumn an NLP-Ergebnisse hängen 🡪 dann Relation herstellen im Model-View
* In Visual unter „Excludes“ einfügen, manuelle Stopwörter zusätzlich oder Liste (XLS/CSV) anpassen und neu laden (auch mal zeigen, aber Achtung: Mehr Wörter führen am Ende zu neuen Spalten 🡪 MergeColumns muss neu gemacht oder erweitert werden, auch zeigen)

# Auswertung Prediction (späterer Part)

Card Browser (Microsoft) nutzen (Variante: Strippet-Browser) für Detail-Drill-Down bis auf Text

* Index für Texte nötig 🡪 wenn nicht in Rohdaten, dann Index-Column ergänzen

Confusion-Matrix zeigen

* die False-Negatives drill-Down

Textfilter

* sehr restriktiv, aber hilft für Fokussierung
* ansonsten über Filtereinstellungen der Visuals

Boxplots zur Anzeige der Verteilung der Kommentare

* Boxplot-Visual aus App-Source laden
* Formateinstellungen feinjustieren: Categories auf 45°, Data Colors nicht „show all“, Titel vergeben

URL anzeigen

* URL formatieren in Daten: Column Tool 🡪 Data Category 🡪 Web URL
* In Tabelle reinziehen, unter Format „Link-Symbol“ aktivieren = nur noch Symbol angezeigt

# Sonstiges

Sonstige Visuals evtl. interessant:

* Text Highlighter (zeigt Sentiments in einem Text, benötigt aber Stichwörter dazu, eher nicht skalierbar).
* Smart narrative feature in Power BI:
* WebPage in PowerBI dynamisch anzeigen:
  + https://p3adaptive.com/2018/11/dynamically-embedding-web-pages-in-power-bi/

Ausblick:

* Netzwerkdarstellungen für Linked-Entities und Q&A-Ergebnisse, auch Topics etc (Network Navigator Chart)
* Time-Intelligence für Veränderungsanalysen
* NER im Strippet-Browser