## BACKLOG

### VORBEREITUNG:

* GITHUB einrichten
* Mögliche Themen besprechen
* Bereiche aufteilen
* Rollen klären

### DATEN:

* Recherchen
* Festlegen
* Download
* Bereinigen
* Für ML vorbereiten
* Kategorische Variable in nummerische
* …….
* ……

### MACHINE LEARNING:

* ML Modelle testen
* ML Model festlegen
* Zusätzliche Parameter (Schwellenwert)
* ……
* …..
* ….

### STREAMLIT

* 3 Bereiche festlegen

# Mögliche Analyseaspekte für das Buch-Projekt

**Grundlegende Buchmerkmale**

| **Merkmal** | **Beispiele für Fragen / Metriken** |
| --- | --- |
| Genre | Welche Genres sind weltweit/am beliebtesten? |
| Sprache / Herkunftsland | Welche Sprachen dominieren den Buchmarkt? |
| Erscheinungsjahr | Wie verändert sich der Trend bestimmter Genres über die Jahre? |
| Seitenzahl | Haben moderne Bestseller mehr oder weniger Seiten als früher? |
| Verlag | Welche Verlage veröffentlichen am meisten? |

# Ökonomische Aspekte

| **Thema** | **Beispiele** |
| --- | --- |
| Verkaufszahlen | Welche Bücher verkaufen sich weltweit am besten? |
| Bestseller-Dynamik | Wie lange bleibt ein Buch Bestseller? |
| Preisentwicklung | Wie verändern sich Buchpreise über die Jahre / nach Genre? |
| Einnahmen nach Region | Welche Länder erwirtschaften am meisten Umsatz mit Büchern? |

# Verbreitung & Einfluss

| **Thema** | **Beispiel-Fragen** |
| --- | --- |
| Übersetzungen | In wie viele Sprachen wird ein Bestseller übersetzt? |
| Ländervergleich | Welche Länder lesen welche Genres? |
| E-Book vs. Print | Wie stark wächst der E-Book-Anteil pro Genre? |
| Alter der Zielgruppe | Welche Altersgruppen lesen welche Bücher? |

# Illegale Verbreitung / Schwarzmarkt (optional)

| **Thema** | **Hinweise / Tools** |
| --- | --- |
| Piraterie-Statistiken | Wie häufig werden Bestseller illegal heruntergeladen? |
| Downloads vs. Verkäufe | Verhältnis geschätzter Downloads zu offiziellen Verkäufen |
| Quellen: Reddit, Piratebay-Tracker | Crawling oder Scraping von Foren, wenn erlaubt/ethisch |
| Ländervergleich Schwarzmarkt | In welchen Ländern ist die Piraterie am stärksten vertreten? |

# Mögliche Visualisierungen

| **Visualisierung** | **Datenbeispiel** |
| --- | --- |
| Balkendiagramm | Verkäufe pro Genre / Sprache |
| Liniendiagramm | Entwicklung der Verkaufszahlen über Jahre |
| Heatmap | Genre vs. Sprache oder Genre vs. Land |
| Weltkarte | Verkaufszahlen / Piraterie nach Land |
| Donut / Pie-Chart | Anteil Print / E-Book / Hörbuch |
| Scatterplot | Seitenzahl vs. Verkaufsrang |

# Mögliche Datenquellen

* 📚 Open Library API
* 💲 NPD BookScan (teilweise kostenpflichtig)
* 🌐 Google Books API
* 📊 Kaggle Datasets: z. B. "Top 1000 Bestsellers", "Books.csv", "Goodreads"
* 🧠 Wikipedia / Wikidata für freie Metadaten zu Büchern