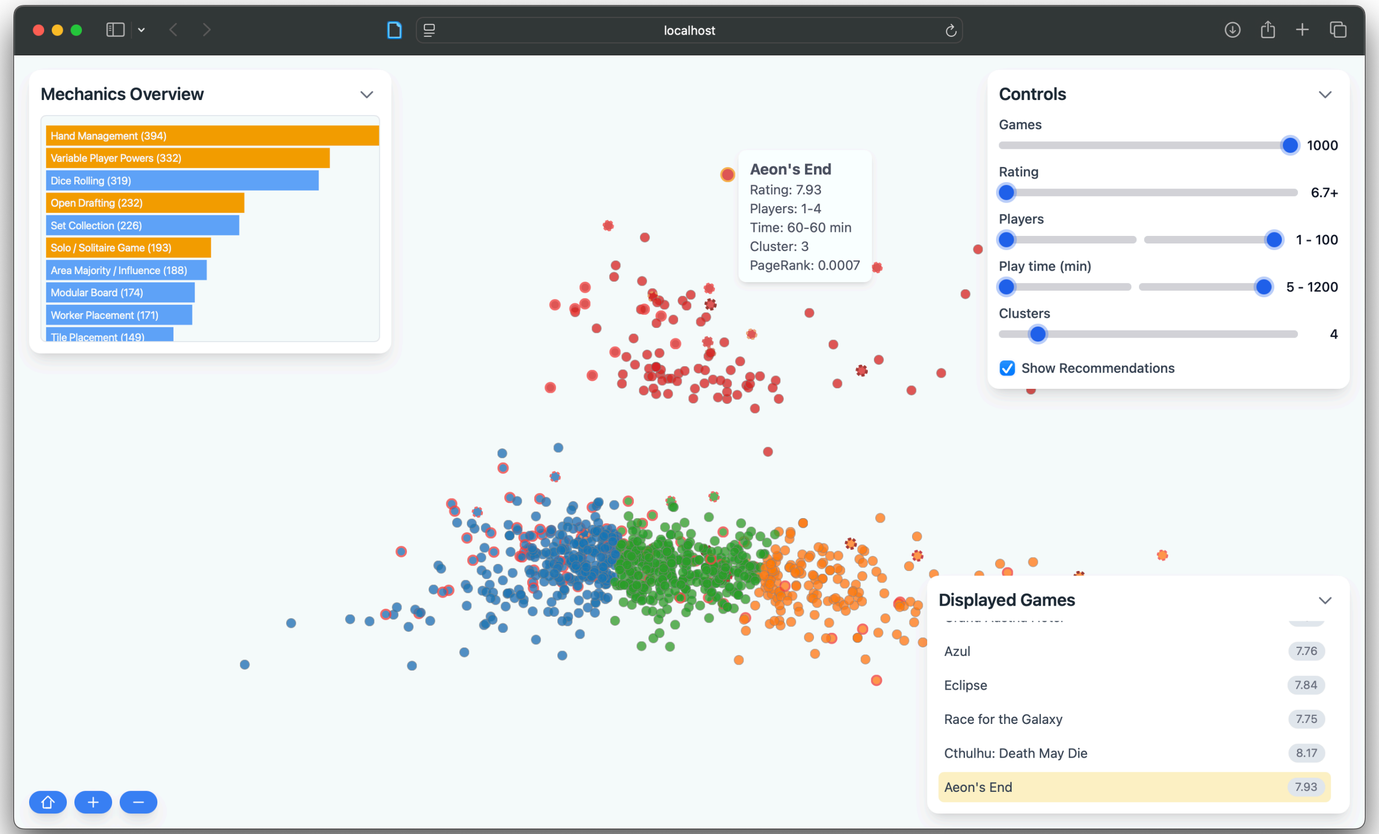
Visual Analytics Dashboard for Boardgames

Task 2: Group games that follow similar trends

Task 3: Investigate impact of mechanics on board game recommendations



Girish Mohan: Design (0.50 hrs), D3 visualization (3 hrs), Interactions (1 hr), Presentation (0.5 hr)
Shashwata Sourav Roy Samya: Design (0.50 hrs), Preprocessing (3 hrs), Presentation (0.5 hr)

The Data

- Two files: ***bgg_Gameitems.csv*** & ***recommendations-2021-12-31.csv***

- ***bgg_Gameitems.csv*** shape:
  113,904 rows, 37 columns

  recommendations-2021-12-31.csv shape:
  1000 rows, 37 columns

- ***bgg_Gameitems.csv*** key observations:
  Several ID-columns (**bga_id**, **dbpedia_id**, etc.) were 100% null.
  Metadata flags like **cooperative**, **implementation**, **integration** also >90% null

- ***recommendations-2021-12-31.csv*** key observations:
  28 recommendation columns; after **6**, null rate jumps to ~33%.
  Other fields (**Name**, **URL**, **Thumbnail**) are complete.

The Data Cleaning & Processing

- For **bgg_Gameitems.csv**:

  Dropped extremely sparse columns **(>90% null)**
  Reduced from **37 → 25 columns**

  Removed redundant player-count variants
  Originally three sets of min/max player counts: **raw**, **recommended** and **best.**
  For most analyses, the raw counts suffice. Columns pruning leaves only **min_players** & **max_players.**

  Dropped the original **rank** column, **re-ranked** games by **Bayesian estimate**

  Parsing list-style fields
  mechanic, **family**, **categories** now a true Python **list[str]**

  Downcasting & **type refinement**
  year, ***play times***, ***ages***, ***ratings***, ***votes***: casted from **float64** to **float32**
  name, ***publisher***, ***game_type***, ***artist***: converted to pandas **category** dtype

The Data Cleaning & Processing

- For **recommendations-2021-12-31.csv**:

  Lowercased and **renamed** columns for consistency
  (e.g. **ID** → **bgg_id**, **Average** → **avg_rating**, etc.).

  Dropped sparse recommendation columns (>33% null): **recommendation7** to **recommendation28**
  Collapsed the remaining recommendation columns into a single list-column **fans_liked**

  Dropped the original **rank** column, **re-ranked** games by **Bayesian estimate**

  Parsing list-style fields
  mechanic, **family**, **categories** now a true Python **list[str]**

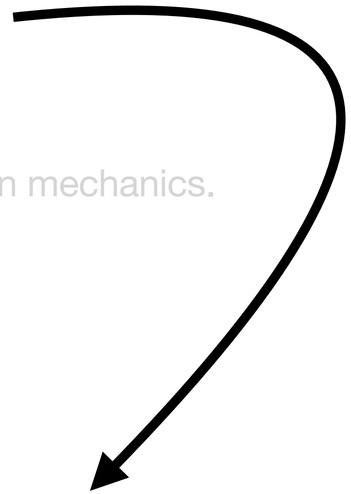
  Downcasting & **type refinement**
  Year, ***Users rated***, ***Rank*** to **int32**; **ratings/votes** to **float32**
  Name, ***URL***, ***Thumbnail*** to pandas **category**

The Data Cleaning & Processing

- **Merged** both dataframes which carries over **full metadata plus top-6 recommendations**.
- Ensured ***mechanic***, ***family***, ***categories***, and ***fans_liked*** are all Python **list[str].**
- **Sorted** the merged dataframe **descending** by **bayes_rating**
- Assigned a new integer **rank** on the **index** column
- **Constructed** JSON records and **exported** it to further visualization

Why these tasks?

- Task 2:
  “Which games have similar trends?"
  📌 An analyst can explore trends within clusters and understand what defines a cluster.
- Task 3:
  “Why a board game is recommended from another?"
  📌 An analyst can compare the similarity of recommended games and correlate based on mechanics.



Goal: *Explore*

Means: *Identify*

Target: *Trends*

Attribute: Mechanics

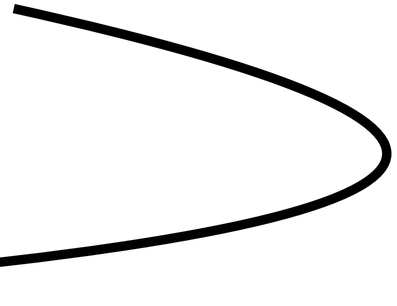
Cardinality: *All*

The user wants to (explore) and (identify) (trends) in board game (mechanics) for (all) games.

Why these tasks?

- Task 2:
  “Which games have similar trends?"
  📌 An analyst can explore trends within clusters and understand what defines a cluster.

- Task 3:
  “Why a board game is recommended from another?"
  📌 An analyst can compare the similarity of recommended games and correlate based on mechanics.



Goal: *Describe*

Means: *Locate*

Target: Correlate

Attribute: Score + Recommendation + Features

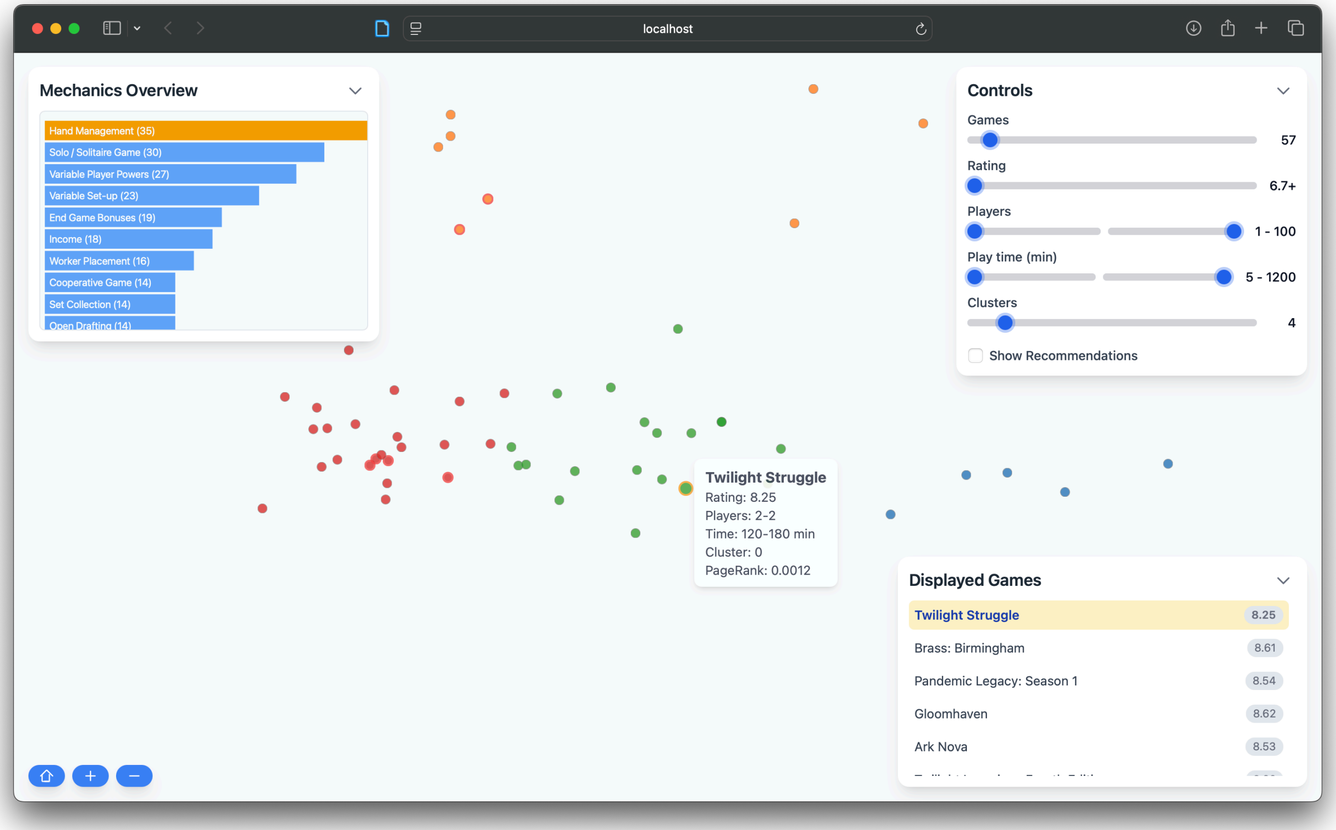
Cardinality: All

The user wants to (describe) and (locate) (correlations) that impact boardgames recommendations for (all) games.

How are these helpful?

- Task 2:
  “Which games have similar trends?"
  📌 An analyst can explore trends within clusters and understand what defines a cluster.

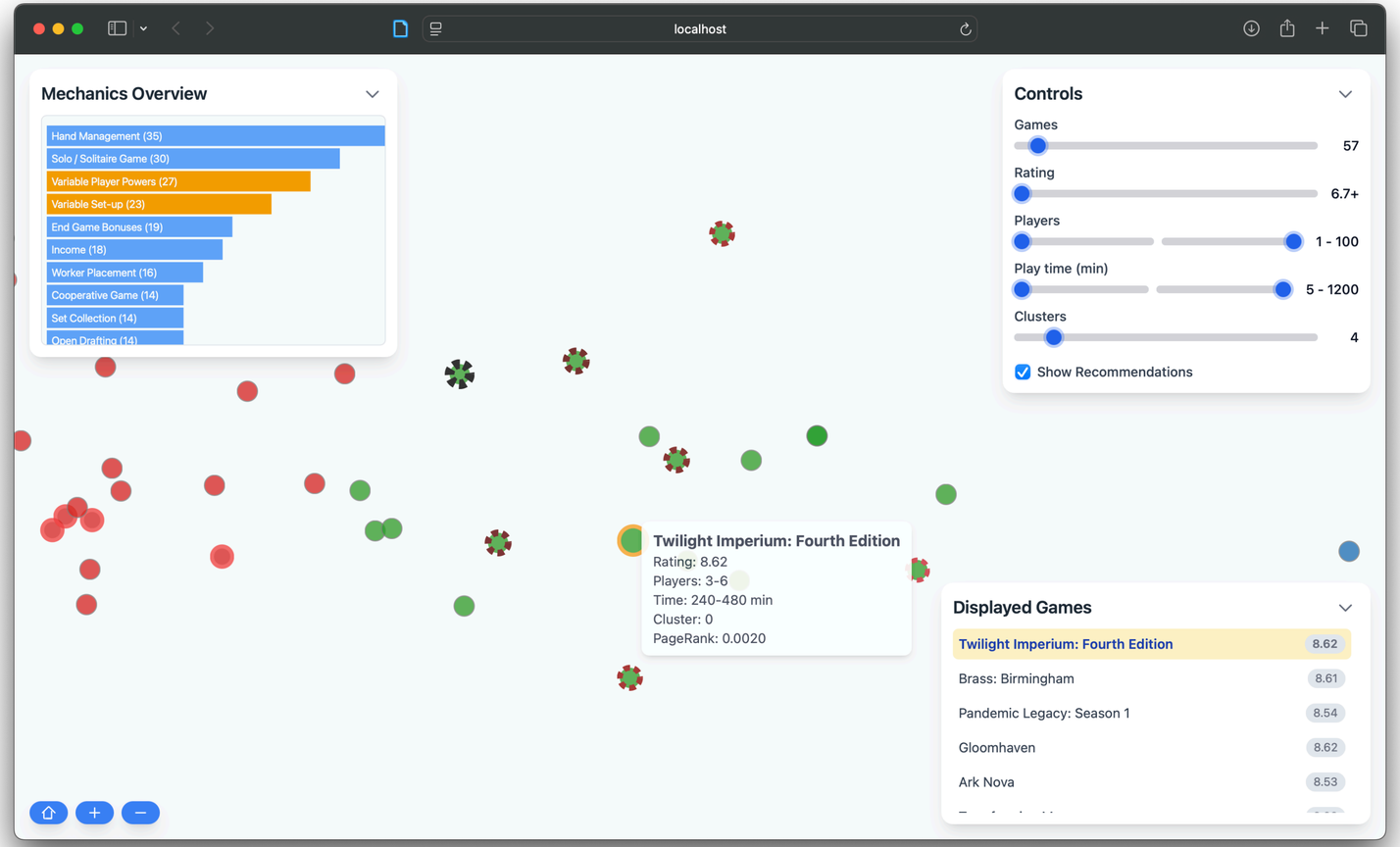
- Task 3:
  “Why a board game is recommended from another?"
  📌 An analyst can compare the similarity of recommended games and correlate based on mechanics.



How are these helpful?

- Task 2:
  “Which games have similar trends?"
  📌 An analyst can explore trends within clusters and understand what defines a cluster.

- Task 3:
  “Why a board game is recommended from another?"
  📌 An analyst can compare the similarity of recommended games and correlate based on mechanics.



Demo

http://localhost:3000

What could be improved?

- Projection based on more features
- Panels clutter scatterplot
- More interactions