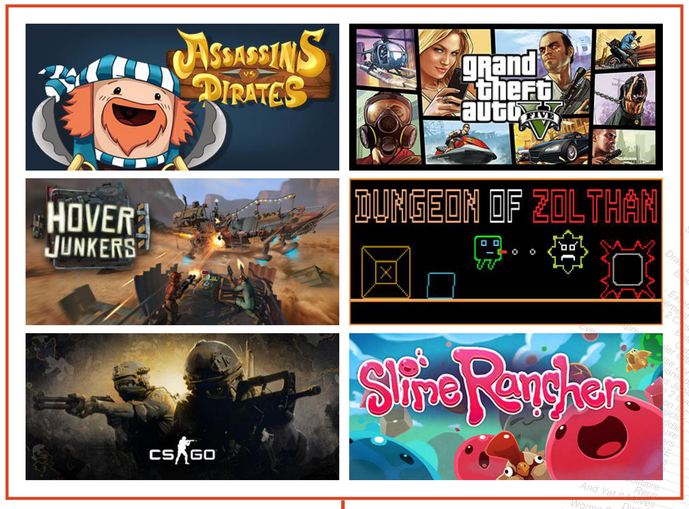
Steam™ Achievement Classification & Analytics Project

Problem



"Action Games"

Video games are notoriously hard to categorize. Unlike traditional media, video games have an interactive element that is difficult to sum up in a single, over-arching category.

The digital distribution platform Steam™ groups games by genre, but these genres are often so broad as to be meaningless; all of the above games fall into the "Action" genre, even though they have nothing in common with each other.

Users can add custom tags, but they can also be misleading or irrelevant to gameplay. For example, some of the more popular tags are "great soundtrack", "funny", and "colorful".

We wanted to find a better way to classify games, in particular one that takes advantage of that interactive aspect.

Solution

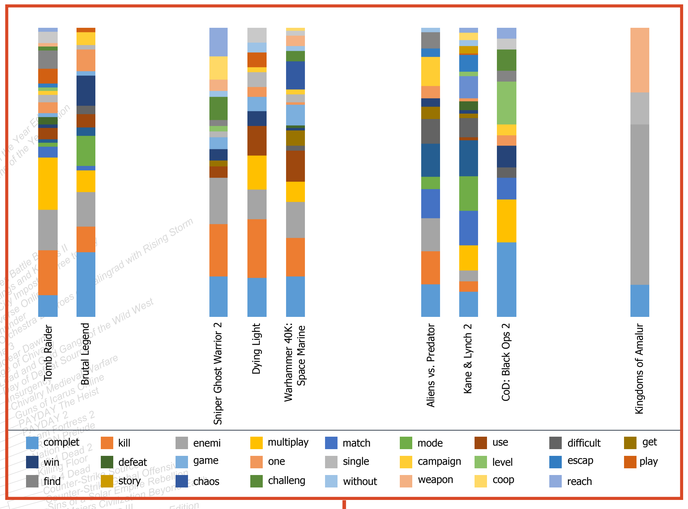
We decided to use video game achievements.

Achievements are rewards for completing game objectives or performing pre-determined actions in a game. They often explicitly describe these actions as well as the things players interact with (e.g. "Go here, do this, get that"). We apply data science methods to compare these descriptions:

- Extract individual terms to create a corpus
- Remove "stop" words from corpus
  (e.g. ~~and~~, ~~or~~, ~~there~~, ~~can~~, ~~the~~)
- Stem remaining terms to create unique unigrams
  (e.g. complete~~d~~, kill~~s~~, construction~~.~~)
- Produce document-term matrix for analysis

Similar Games,
Similar Achievements

Result



The DNA of Achievements

These stemmed unigrams can be seen as the metaphorical "genes" that determine the achievement make-up of each game. Based on the relative frequency of the unigrams, we used a hierarchical clustering algorithm to sequentially pair games whose compiled "genome" most closely resemble one another. The end result is a dedrogram; a tree-like diagram displaying the "family" that each game belongs to.

In contrast to the thematic and meta-element nature of genres and tags, the strength of achievement-based analysis lies in its insight into the gameplay and structural elements within each game. Unigrams such as "discov" distinguish games that encourage open-world gameplay, while achievements frequently using the term "campaign" flag games utilizing long-form strategic gameplay.

Information School
UNIVERSITY of WASHINGTON

Adam Lathrop
adl5@uw.edu

Λ3

Daniel Wilson
drwilson@uw.edu

Project Sponsor
Jin Ha Lee

GAMER
GAME RESEARCH GROUP