**About Dataset**

**users.csv**  
User\_id: Unique identifier of user  
Country\_code: Country code where the user registered  
**assets.csv**  
Show\_type: Type of content, whether the asset is a movie or an episode of a TV series  
Genre: Genre of content  
Running\_miutes: Runtime of content (Playable number of minutes)  
Source\_language: Production language of content  
Asset\_id: Unique identifier of video content at the most granular level (a movie or an episode of a TV series)  
Season\_id: Unique identifier of content at season level. This is only applicable to TV series  
Series\_id: Unique identifier of content at series level. This is only applicable to TV series  
Studio\_id: Unique identifier of production studio for the content  
**plays.csv**  
Platform: Platform of consumption  
Minutes\_viewed : Total number of minutes viewed, rounded to the nearest integer (0 means less than 30 seconds)  
**Demographics.csv Psychographics.csv**  
The dataset identifies psychographic and demographic tags about some iflix users. Each user-tag pair has an associated confidence score (1 is the highest, and 0 is the lowest confidence). Each trait can have up to 3 levels, depending on its granularity. Some traits can be identified by only considering the first two levels. At the same time, there are others that make more sense when all the three levels are considered, e.g., ‘iflix Viewing Behaviour’ is a level 2 psychographic trait that only makes sense when it is looked at in combination with the level 3 traits corresponding to it (‘casual,’ ‘player’ and ‘addict’). These traits represent different levels of viewing behavior of iflix users. Casual users have less than five viewing days in a month, player users have 5 to 12 viewing days in a month, and people with an addiction have more than 12 viewing days in a month. Traits are available corresponding to a user\_id in the dataset only if we have certain confidence that the user belongs to the trait.  
Column and Description  
Level\_1: Identifies the first level of the trait (psychologic or demographic)  
Level\_2: Identifies the second level of the trait (e.g., Music Lovers, Movies Lovers)  
Level\_3 : Identifies the third level of the trait, if available/relevant (e.g. Malay Movies Lovers, Indonesian TV Fans)  
Confidence\_score: Confidence in associating the said trait (level\_1, level\_2, level\_3) with the user