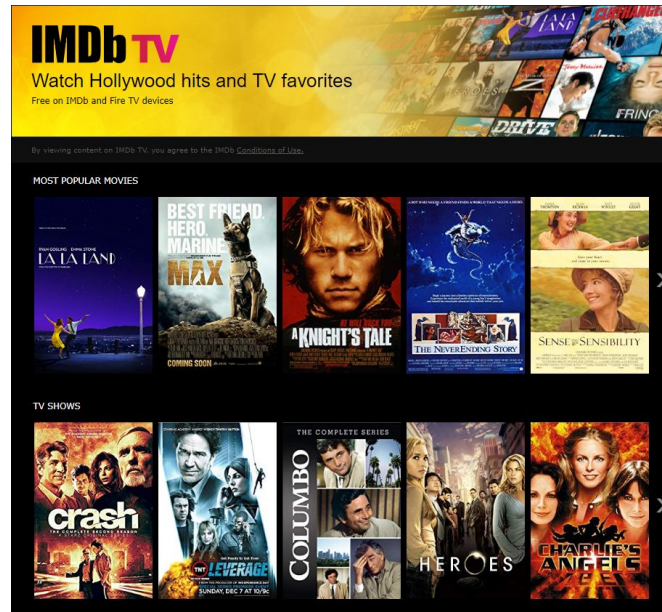


Visualizing the Internet Movie Database (IMDb)

By: Dillon Gorlesky & Brett Dale

What is IMDb?

- An online database of information related to films, television programs, home videos, video games, and streaming content online – including cast, plot summaries, ratings, fan and critical reviews, and more.
- IMDb is the most authoritative source of entertainment information, with features designed to help fans explore the world of movies and shows and decide what to watch.



What Problem We are Currently Solving

- Interpreting data can be confusing and difficult, and the difficulty increases as the size of the data set grows.
- Having a tool that can simplify and condense the information into a visualization is vital in a variety of situations
- Our objective is to take the IMDb dataset of the Top 1000 films from 2006 through 2016, and visualize its contents to offer a better and more interactive way to interpret the data.

	A	B	C	D	E	F	G	H	I	J	K	L
1	Rank	Title	Genre	Director	Actors	Year	Runtime	Rating	Votes	Revenue	Metascore	
2	1	Guardians of the Galaxy	Action, Adventure	James Gunn	Chris Pratt, Zoe Lister-Jones, Dave Bautista, Michael Rooker, Karen Gillan	2014	121	8.1	757074	333.13		76
3	2	Prometheus	Adventure, Sci-Fi	Ridley Scott	Noomi Rapace, Michael Fassbender, Idris Elba, Logan Lerman, Michael Sheen	2012	124	7	485820	126.46		65
4	3	Split	Horror, Thriller	James McGuinness	M. Night Shyamalan, James Mcguinness, Haley Joel Osment, Anthony Mackie, Bruce Campbell	2016	117	7.3	157606	138.12		62
5	4	Sing	Animation, Comedy	Chris Buck	Matthew McConaughey, Reese Witherspoon, Robert Dyer, John Goodman, Angelina Jolie	2016	108	7.2	60545	270.32		59
6	5	Suicide Squad	Action, Adventure	David Ayer	Will Smith, Jared Leto, Diezel Yarbrough, Jai Rodriguez, Adewale Akinnuoyefea	2016	123	6.2	393727	325.02		40
7	6	The Great Wall	Action, Adventure	Yimou Zhang	Matt Damon, Pedro Pascal, Wen Jiang, Wen Jiang, Wen Jiang	2016	103	6.1	56036	45.13		42
8	7	La La Land	Comedy, Drama, Romance	Damien Chazelle	Ryan Gosling, Emma Stone, Rosemarie DeWitt, Faye Dunaway, Corey Stoll	2016	128	8.3	258682	151.06		93
9	8	Mindhorn	Comedy	A has-bee Sean Foley	Essie Davis, Sean Foley, Sean Foley, Sean Foley, Sean Foley	2016	89	6.4	2490			71
10	9	The Lost City	Action, Adventure	A true-life James Graham	Charlie Hulse, Charlie Hulse, Charlie Hulse, Charlie Hulse, Charlie Hulse	2016	141	7.1	7188	8.01		78
11	10	Passenger 57	Adventure, Sci-Fi	A space-cr Morten Tyldum	Jennifer Lawrence, Ryan Reynolds, Michael Fassbender, Michael Fassbender, Michael Fassbender	2016	116	7	192177	100.01		41
12	11	Fantastic Beasts and Where to Find Them	Adventure, Fantasy	The wizard David Yates	Eddie Redmayne, Katherine Waterston, Ezra Miller, Ezra Miller, Ezra Miller	2016	133	7.5	232072	234.02		66
13	12	Hidden Figures	Biography, Drama	The story Theodore Melfi	Taraji P. Henson, Octavia Spencer, Kirsten Dunst, Kirsten Dunst, Kirsten Dunst	2016	127	7.8	93103	169.27		74
14	13	Rogue One: A Star Wars Story	Action, Adventure	The Rebel Gareth Edwards	Felicity Jones, Diego Luna, Alan Tudyk, Alan Tudyk, Alan Tudyk	2016	133	7.9	323118	532.17		65
15	14	Moana	Animation, Adventure	In Ancient Hawaii Ron Clements	Auli'i Cravalho, Dwayne Johnson, Rachel Zegler, Rachel Zegler, Rachel Zegler	2016	107	7.7	118151	248.75		81
16	15	Colossal	Action, Comedy	Gloria is a Nacho Vigalanti	Anne Hathaway, Anne Hathaway, Anne Hathaway, Anne Hathaway, Anne Hathaway	2016	109	6.4	8612	2.87		70
17	16	The Secret Life of Walter Mitty	Adventure, Drama	The quiet Chris Kenner	René Louis C.K., René Louis C.K., René Louis C.K., René Louis C.K., René Louis C.K.	2016	87	6.6	120259	368.31		61
18	17	Hacksaw Ridge	Biography, War	WWII American Mel Gibson	Andrew Garfield, Andrew Garfield, Andrew Garfield, Andrew Garfield, Andrew Garfield	2016	139	8.2	211760	67.12		71
19	18	Jason Bourne	Action, Thriller	The CIA's iPaul Greengrass	Matt Damon, Matt Damon, Matt Damon, Matt Damon, Matt Damon	2016	123	6.7	150823	162.16		58
20	19	Lion	Biography	A five-year Garth Davis	Dev Patel, Dev Patel, Dev Patel, Dev Patel, Dev Patel	2016	118	8.1	102061	51.69		69
21	20	Arrival	Drama, Mystery	When two Denis Villeneuve	Amy Adams, Amy Adams, Amy Adams, Amy Adams, Amy Adams	2016	116	8	340798	100.5		81
22	21	Gold	Adventure, Drama	Kenny Weir	Stephen Graham, Stephen Graham, Stephen Graham, Stephen Graham, Stephen Graham	2016	120	6.7	19053	7.22		49

Dataset Explanation

- Our dataset has 12 features, with movies being ranked on popularity
- IMDb has software titled IMDbPro that uses 'proprietary algorithms' to take into account several measures, with the primary measure being who and what films people are searching for on IMDb.com

	A	B	C	D	E	F	G	H	I	J	K	L
1	Rank	Title	Genre	Descriptive	Director	Actors	Year	Runtime	(Rating)	Votes	Revenue	(Metascore)
2	1	Guardians of the Galaxy	Action, Adventure	A group of intergalactic	James Gunn	Chris Pratt, Zoe Lister-Jones, Dave Bautista, Karen Gillan, Michael Rooker, Bradley Pitt	2014	121	8.1	757074	333.13	76
3	2	Prometheus	Adventure, Sci-Fi	Following Ridley Scott	Nolan	Noomi Rapace, Michael Fassbender, Iden	2012	124	7	485820	126.46	65
4	3	Split	Horror, Thriller	Three girls	M. Night Shyamalan	James McAvoy, Haley Joel Osment, Anya Taylor-Joy	2016	117	7.3	157606	138.12	62
5	4	Sing	Animation	In a city of animals	Christoph Ludwig	Matthew McConaughey, Reese Witherspoon, Robert Delpy	2016	108	7.2	60545	270.32	59
6	5	Suicide Squad	Action, Adventure	A secret government	David Ayer	Will Smith, Jared Leto, Diezel Yaxley	2016	123	6.2	393727	325.02	40
7	6	The Great Wall	Action, Adventure	European	Yimou Zhang	Matt Damon, Wen Jiang, Pedro Pascal	2016	103	6.1	56036	45.13	42
8	7	La La Land	Comedy, Drama	A jazz pianist	Damien Chazelle	Ryan Gosling, Emma Stone	2016	128	8.3	258682	151.06	93
9	8	Mindhorn	Comedy	A has-been	Sean Foley	Essie Davis, David Morris	2016	89	6.4	2490		71
10	9	The Lost City	Action, Adventure	A true-life	James Gray	Charlie Hunnam, Jennifer Lawrence	2016	141	7.1	7188	8.01	78
11	10	Passenger	Adventure, Sci-Fi	A spacecraft	Morten Tyldum	Jennifer Lawrence, Chris Pratt	2016	116	7	192177	100.01	41
12	11	Fantastic Beasts and Where to Find Them	Adventure, Fantasy	The adventures	David Yates	Eddie Redmayne, Katherine Waterston, Dan Fogarty	2016	133	7.5	232072	234.02	66
13	12	Hidden Figures	Biography	The story of	Theodore Melfi	Taraji P. Henson, Octavia Spencer, Kirsten Dunst	2016	127	7.8	93103	169.27	74
14	13	Rogue One: A Star Wars Story	Action, Adventure	The Rebel Alliance	Gareth Edwards	Felicity Jones, Diego Luna, Alan Tudyk	2016	133	7.9	323118	532.17	65
15	14	Moana	Animation	In Ancient Hawaii	Ron Clement	Auli'i Cravalho, Dwayne Johnson	2016	107	7.7	118151	248.75	81
16	15	Colossal	Action, Comedy	Gloria is a	Nacho Vigalondo	Anne Hathaway, Jesse Plemons	2016	109	6.4	8612	2.87	70
17	16	The Secret Language of Flowers	Animation	The quiet	Chris Renaud	Louis C.K., Amy Poehler	2016	87	6.6	120259	368.31	61
18	17	Hacksaw Ridge	Biography, War	WWII American	Mel Gibson	Andrew Garfield, Sam Worthington, Luke Bracey	2016	139	8.2	211760	67.12	71
19	18	Jason Bourne	Action, Thriller	The CIA's	Paul Greengrass	Matt Damon, Taylor Kitsch, Michael Keaton	2016	123	6.7	150823	162.16	58
20	19	Lion	Biography	A five-year	Garth Davis	Dev Patel, Nicole Kidman, Russell Crowe	2016	118	8.1	102061	51.69	69
21	20	Arrival	Drama, Mystery	When two	Denis Villeneuve	Amy Adams, Jeremy Renner, Forest Whitaker	2016	116	8	340798	100.5	81
22	21	Gold	Adventure	Kenny Williams	Stephen C. Miller	Matthew McConaughey, Eddie Redmayne	2016	120	6.7	19053	7.22	49

Our Goals of this Project

1. Our primary goal is building an interactive graph visualization that is easy and informative to use.
2. Shows all the nodes and links connecting each other.
3. Allow the user to see which node was clicked
4. Being able to search for Actors and Movies
5. When a node or link is hovered over, it shows all the connections related to it.

MoSCoW Analysis

Must Have	Should Have	Could Have	Won't Have
Read in dataset stored in CSV	A way to look at one specific piece of information (i.e. movie, actor, director, etc.) and all connections	A formatted table near the graph to show in input data	Adding new names/actors after creation
Connections between movie and top four actors that played in said movie	Be able to zoom in and out of the graph	Implement a database to push and pull data from (ArangoDB)	Adding new edges after creation
Allow the user to view the graph visualization on our website	A way to search by actor name or movie name	Nodes/Edges storing more information (i.e. Rank, Metascore, Revenue, etc.)	Allow for an edge to be deleted
Create a graph visualization using D3	When a node is clicked it will change colors and the name of the node will be displayed	Allow the user to choose what information the graph is created around (directors, actors, etc.)	Allow the user to upload a new CSV file while one is already loaded in

The Tools We Are Using

React:

A JavaScript library for building user interfaces or UI components.

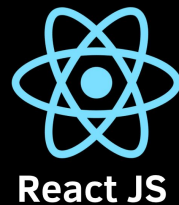
D3 Graph:

A JavaScript library for producing dynamic, interactive data visualizations in web browsers

ArangoDB:

A native multi-model database system that supports three data models with one database core and a unified query language AQL

How We're Using React



- React essentially allows us to have a website interface ran on a localhost to have a visualization of our data.
- Allows us to implement a variety of packages and tools, such as D3, search functionality, and interactive components to make our website unique and conformed to our purposes.
- React handles any sort of data passing and handling to make sure things are running correctly together.
- We also use Provider classes that act as Event Listeners that allow for easy communication of information being passed between classes.

```
import React from "react";

const Hello = ({ name }) => (
  <>
    <h1>Hello, {name}!</h1>
    <p>
      Welcome to React. A UI component framework
      for deterministic view renders.
    </p>
  </>
);
```




How We're Using D3

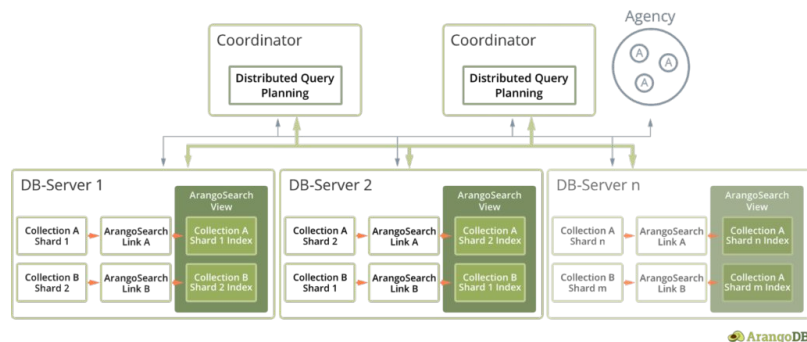
- The React-D3 package allows us to essentially build our graph from the nodes and links created from the IMDb dataset.
- When implemented into React, it allows for an interactive graph users can use to see a visualization of the data.
- D3 also allows for users to change different aspects of the graph, such as deleting nodes, changing node sizes and colors, zooming in and out, and even dragging and dropping nodes.
- Currently, we're using a community based react-d3 package called Graph for our visualization.

```
import { Graph } from "react-d3-graph";
const data = {
  nodes: [{ id: "Harry" }, { id: "Sally" }, { id: "Alice" }],
  links: [{ source: "Harry", target: "Sally" },
    { source: "Harry", target: "Alice" },
  ],
};
const myConfig = {
  nodeHighlightBehavior: true,
  node: {
    color: "lightgreen",
    size: 120,
    highlightStrokeColor: "blue",
  },
  link: {
    highlightColor: "lightblue",
  },
};
return(
  <Graph
    id="graph-id" // id is mandatory
    data={data}
    config={myConfig}
  />
);
```

How We Plan to Use ArangoDB



- Currently our project is loading the dataset locally
- Our plan is to eventually have an ArangoDB database that will allow us to pull the IMDb dataset contents.
- This may also be a solution to one of our past problems of not being able to select a specific file from a File Browser button for upload.



Problems We've Ran Into Thus Far

1. Before this project we had never heard or worked with any of our tools, let alone how to connect them all together, and we had zero experience with writing HTML code and mixing it with JavaScript.
2. Allowing a user to upload a file from file explorer wasn't possible because of browser security.
3. There was a lot of stop and go, and stop and back peddling because we'd find a solution that worked for one problem, but problems down the line wouldn't work because of them.

Live Demo

Our Future Work

1. Add more ways to filter the data, such as just movie or actor connections.
2. Implementing ArangoDB into our program
3. Other graph visualizations to represent our data (i.e. Bar, line, or pie graphs)
4. The possibility of deleting nodes from the graph
5. Allowing the usage of other datasets
6. A load button so users can upload their own datasets
7. Allow the user to search for actors or movies and only show direct connections

Visualizing the Internet Movie Database (IMDb)

Listed: Chris
Pratt

