

# Relatório Técnico: CineDive

## Uma exploração visual dos filmes indicados ao Oscar

Anderson, Gabriel, Ximena

17 de junho de 2025

## Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
<b>2</b>	<b>Dados: Origem, Limpeza e Expansão</b>	<b>2</b>
2.1	Fontes e Estrutura dos Dados . . . . .	2
2.2	Processamento e Limpeza . . . . .	2
2.3	Expansão: Adição de Novos Participantes . . . . .	2
<b>3</b>	<b>Metodologia Visual e Tecnologias</b>	<b>2</b>
<b>4</b>	<b>Big Picture: Visualizações Globais</b>	<b>3</b>
4.1	Mapa Mundial de Filmes . . . . .	3
4.2	Gráfico Radial Multivariado . . . . .	3
<b>5</b>	<b>One vs All Exploration: Grafos e Heatmap</b>	<b>3</b>
5.1	Grafo Filme-Pessoa . . . . .	3
5.2	Grafo Exclusivo de Pessoas . . . . .	3
5.3	Heatmap: Indicações vs. Vitórias . . . . .	3
<b>6</b>	<b>Decisões de Visualização e Justificativas Técnicas</b>	<b>4</b>
<b>7</b>	<b>Conclusão</b>	<b>4</b>
<b>8</b>	<b>Trabalhos Relacionados</b>	<b>4</b>

## 1 Introdução

O projeto **CineDive** tem como objetivo explorar, de forma interativa e visual, o universo dos filmes indicados e premiados pela Academia (Oscar). A proposta é permitir tanto investigações de padrões globais (“big picture”) quanto explorações individuais e comparativas, ampliando a compreensão das conexões, tendências e trajetórias dentro da indústria cinematográfica reconhecida pelo Oscar.

## 2 Dados: Origem, Limpeza e Expansão

### 2.1 Fontes e Estrutura dos Dados

Para a construção do projeto, diferentes fontes de dados foram integradas:

- **IMDb:** Arquivos públicos (`title.basics`, `title.principals`, `name.basics`, etc.), trazendo dados estruturados de filmes (título, ano, duração, gêneros) e profissionais (diretores, atores, roteiristas).
- **Base Oficial do Oscar:** Lista oficial de filmes indicados e vencedores em todas as edições, incluindo ano, categorias, país de origem, número de indicações e prêmios.
- **Web scraping:** Para complementar e validar, utilizei scraping de fontes como Wikipedia e portais de cinema.

### 2.2 Processamento e Limpeza

- **Normalização:** Identificadores (`tconst` para filmes e `nconst` para pessoas) foram padronizados para garantir vinculação correta.
- **Filtragem:** Foram excluídos registros incompletos (sem ano, título ou ligação ao Oscar), duplicatas e obras irrelevantes para o escopo.
- **Agregação:** Os nós dos grafos representam tanto filmes quanto pessoas, e os links codificam os papéis de colaboração (ator, diretor, roteirista, etc).

### 2.3 Expansão: Adição de Novos Participantes

O sistema foi estruturado para ser expansível: novos filmes, pessoas ou premiações podem ser inseridos apenas atualizando as tabelas básicas e reprocessando os arquivos de dados. O pipeline de ETL (Extract, Transform, Load) foi desenhado para automatizar essa integração, permitindo atualizações anuais ou inclusão de eventos como o Oscar internacional.

## 3 Metodologia Visual e Tecnologias

Para implementar as visualizações interativas, utilizei a stack **Svelte** com componentes customizados em **D3.js**. Os gráficos foram divididos em dois grupos principais:

1. **Big Picture:** Visualizações globais que permitem observar padrões, distribuição geográfica e propriedades multidimensionais dos filmes.
2. **One vs All Exploration:** Ferramentas para análise comparativa entre um filme/pessoa selecionado(a) e o restante do universo.

## 4 Big Picture: Visualizações Globais

### 4.1 Mapa Mundial de Filmes

O mapa mundial representa a origem geográfica dos filmes indicados ao Oscar, permitindo filtrar por década, categoria ou país. O canal visual de cor foi escolhido para identificar o volume relativo de indicações/vitórias por país, enquanto o tamanho dos círculos indica o número de filmes ou premiações.

### 4.2 Gráfico Radial Multivariado

O gráfico radial oferece uma comparação intuitiva de múltiplos atributos dos filmes: ano (ângulo), nota média do IMDb, número de votos e duração (codificados em raio, cor ou tamanho). O método de *hexbin* foi utilizado para identificar áreas de alta densidade no espaço de atributos, facilitando a descoberta de padrões (por exemplo, concentração de vencedores do Oscar em certas faixas de avaliação ou duração).

## 5 One vs All Exploration: Grafos e Heatmap

### 5.1 Grafo Filme-Pessoa

O grafo interativo conecta filmes a seus diretores, roteiristas e atores. Para destacar relações, utilizei:

- **Cor da aresta:** Indica o tipo de colaboração (azul para diretores, laranja para atores, verde para roteiristas).
- **Forma do nó:** Estrela para pessoa, círculo para filme.
- **Tamanho:** Proporcional ao rating ou importância do nó.
- **Tooltip:** Mostra detalhes do filme ou pessoa ao passar o mouse.

Os filtros laterais permitem ajustar o período, nota mínima, votos e gêneros, além de mostrar apenas vencedores ou determinados papéis.

### 5.2 Grafo Exclusivo de Pessoas

Ao selecionar um filme, um grafo focado apenas nas pessoas é gerado, destacando conexões indiretas (colaborações comuns) e facilitando a análise de trajetórias. A paleta de cores destaca os selecionados, e é possível visualizar a linha do tempo das participações de cada profissional.

### 5.3 Heatmap: Indicações vs. Vitórias

O heatmap bidimensional exibe a quantidade de filmes com cada combinação de indicações e vitórias ao Oscar.

- **Cor (escala logarítmica):** Mostra a densidade de filmes por célula.

- **Interatividade:** Células podem ser selecionadas para revelar listas de filmes, e modos especiais destacam “campeões absolutos” (maior número de vitórias) ou casos diagonais (todas as indicações resultaram em prêmio).

## 6 Decisões de Visualização e Justificativas Técnicas

- **Cor:** Utilize canais de cor para codificar tipos de conexão (grafo) ou magnitude (heatmap, mapa), evitando sobreposição semântica.
- **Forma e tamanho:** Diferencie pessoas e filmes por símbolos, e ajuste tamanho conforme centralidade ou avaliação.
- **Tooltip e seleção:** As interações via mouse-over e seleção de células facilitam o *drill down* sem sobrecarregar a visualização principal.
- **Escalas logarítmicas:** Adotei escalas logarítmicas em situações de grande variação (ex.: número de filmes por célula no heatmap), para evitar que poucos valores extremos dominem a percepção visual.

## 7 Conclusão

O CineDive demonstra como técnicas modernas de visualização podem revelar padrões e conexões no universo do cinema premiado pelo Oscar. O sistema é flexível, expansível e pode ser adaptado para outras premiações ou para incluir mais dados conforme interesse futuro.

## 8 Trabalhos Relacionados

Diversos projetos prévios exploraram a visualização de dados relacionados ao cinema, cada um com enfoques e soluções técnicas distintas. Entre as iniciativas de maior destaque estão:

- **Tableau Data + Movies:** A iniciativa da Tableau oferece painéis interativos e intuitivos que permitem explorar o desempenho de bilheteria, avaliações de público e tendências de gênero, com interfaces voltadas à clareza e à facilidade de navegação (<https://www.tableau.com/data-plus-movies>).
- **The Movie Database (TMDb):** A plataforma TMDb disponibiliza visualizações dinâmicas que possibilitam a exploração de conexões entre filmes, profissionais e estatísticas agregadas de forma interativa (<https://www.themoviedb.org>).
- **Netflix Research:** O grupo de pesquisa da Netflix desenvolve visualizações e estudos baseados em padrões de consumo dos usuários, analisando preferências e hábitos de audiência em escala global (<https://research.netflix.com>).

Ainda que esses projetos proporcionem interfaces sofisticadas e bancos de dados amplos, o **CineDive** diferencia-se ao focar exclusivamente nas trajetórias e conquistas relacionadas ao Oscar. O sistema enfatiza:

- a construção de redes detalhadas entre filmes e criadores (diretores, atores, roteiristas),
- a exploração relacional e histórica das indicações e vitórias,
- e uma abordagem interativa orientada por storytelling, visando aprofundar o entendimento das conexões e dos padrões de excelência no cinema premiado.

Dessa forma, o CineDive não apenas agrega recursos de explorações presentes em plataformas amplas, mas também oferece uma perspectiva relacional e histórica única sobre a excelência cinematográfica reconhecida pelo Oscar, tornando-se uma ferramenta original para pesquisadores, educadores e cinéfilos.