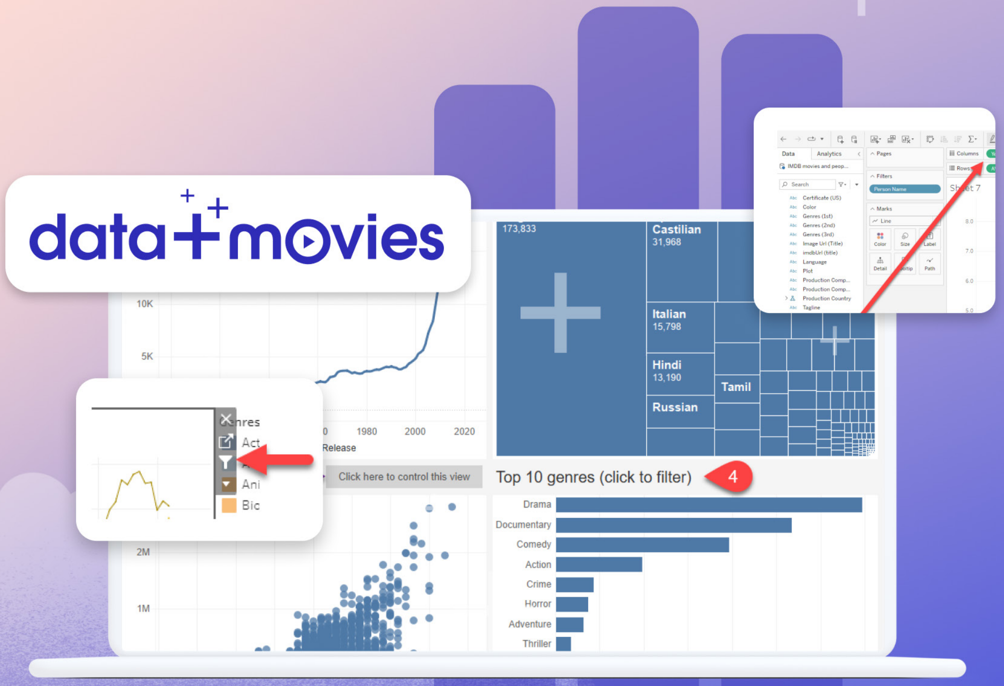
tableau from Salesforce

Tableau Data+Movies スターターキット

シニアデータエバンジェリスト、Andy Cotgreave



Powered by IMDb

目次

はじめに 1
準備 – Tableau Publicアカウントを作成する 3
スターターダッシュボードの構成を理解する 4
スターターダッシュボードのコピーを作成する 5
グラフを作成する – 好きな俳優の出演作品を見つける 6
グラフを作成する – ジャンルの変遷を探ってみる 10
グラフを作成する – 投票数の多い映画を確認する 11
ダッシュボードを作成する – 2つのグラフをダッシュボードにまとめる 12
書式を設定する 14
ワークブックを保存し、表示する 15
ワークブックを共有する 16
Vizにハッシュタグ「#DataPlusMoviesJP」を追加する 16

付録 – データフィールドリスト 17
Movie 17
Cast and Crew 19
Person’ s Role in the Movie 20
Award Nominations 21
Age Rating 22

はじめに

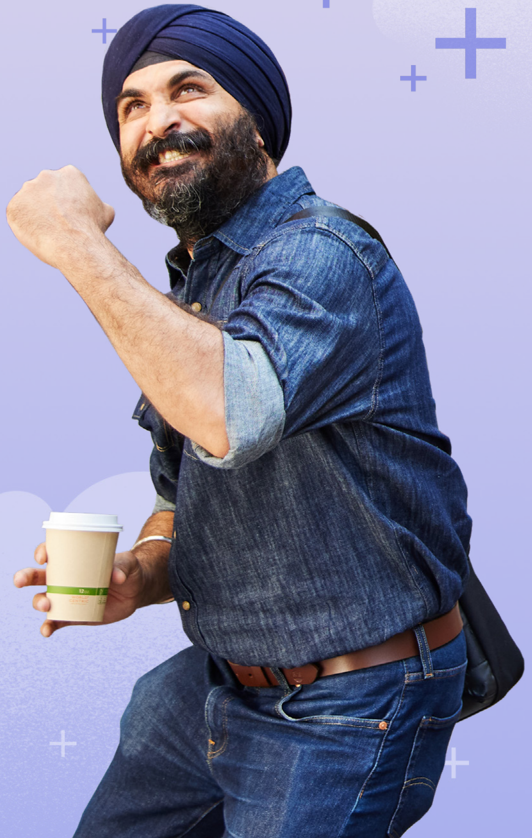
映画ファンの皆さま、こんにちは。

『Tableau Data + Moviesスターターキット』へようこそ！このたびIMDbとの提携により、映画に関するデータをTableauで活用できるようになりました。IMDbは、世界的に知られる、権威ある映像関連のデータベースです。映画やテレビ番組、俳優、監督に関する情報を網羅しています。ぜひ、あなたの好きな映画に関するデータからストーリーを紡ぎ、共有してください。映画のデータから、どのようなことがわかるでしょうか。気になる俳優の全出演作品？マーベル映画とDC映画はどちらが人気？あるいは、世界ではどんな言語の映画が製作されているのかを調べてみるのも面白そう。

すぐに始められるように、[スターターダッシュボード]とデータを用意しました。Tableauを使ったことがなくても大丈夫。このガイドで、映画をテーマにしたビジュアライゼーション（Viz）の作成方法を順を追って説明します。すでにTableauを使っている方は、最初の手順はスキップして、さっそくVizを作ってみてください。

テーマ別のスターターダッシュボードも用意しました！

SF・スターターダッシュボード（英語）
ホラー&ハロウィーン・スターターダッシュボード（英語）
ホリデー・スターターダッシュボード（英語）



このガイドの内容

- Tableau Publicを使ってみる
- スターターダッシュボードの構成を理解する
- 4つの例で学ぶデータ探索とVizの作成
- 書式設定のアイデア

初めて使う方のために、Tableauを詳しく学べるさまざまなリソースを用意しています。

Vizが完成したら、ぜひ共有してください。Vizの共有方法については、後ほど詳しく説明します。

データを提供し、Data + Moviesチャレンジにご協力いただいたIMDbに、心から感謝いたします。

TableauのシニアデータエバンジェリストであるAndy Cotgreaveが、このガイドで皆さまをお手伝いします。では、さっそく始めましょう！

Andy

お願い

商標および著作権の制限により、作成したビジュアライゼーションをTableauで共有する場合は、商標または著作権を侵害するコンテンツを使用しないでください。IMDbのWebサイト上にある素材は、すべてこの制限の対象になります。Tableauアンバサダーは、Vizを作成するための画像の使用許可を得ています。使用制限に関する規則の詳細については、こちらをお読みください（英語）。

Andy Cotgreaveは『The Big Book of Dashboards』の共著者であり、Tableauのシニアデータエバンジェリストです。Tableauの動画シリーズ『If Data Could Talk』や『Chart Chat』に出演し、『Information Age』ではコラムを執筆。オックスフォード大学にアナリストとして勤務後、データビジュアライゼーションとBI分野で15年以上の経験を積む。あらゆる業界でお客様やメディア、アナリストがデータを視覚化し、理解できるように支援しています。ビジュアルアナリティクスでトレンドを発見し、データ探索スキルを向上するうえで役立つ、技術的なアドバイスやアイデアを多くの人々に提供、指導しています。

準備

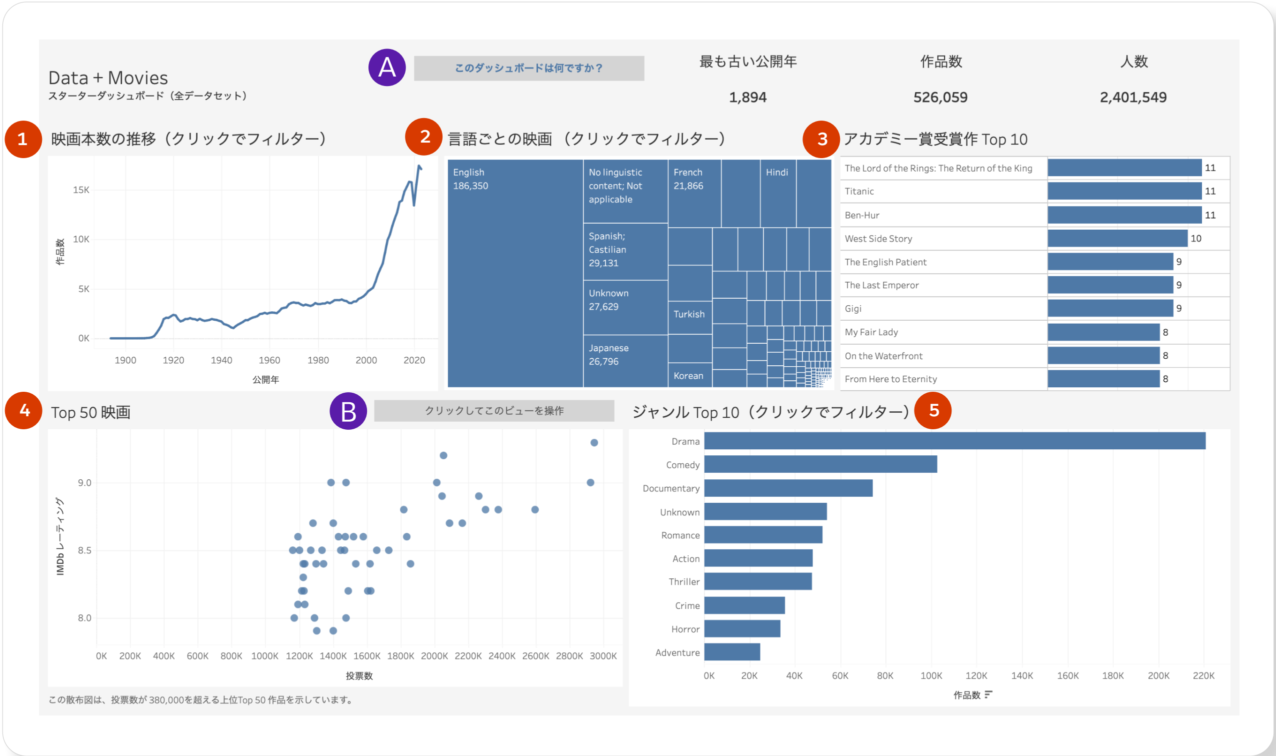
Tableau Publicアカウントを作成する

Tableau Publicアカウントを持っていない場合は、まずアカウントを作成してください。Tableau Publicのサインインページで、[アカウントの作成] をクリックします。

手順に従ってアカウントを作成し、メールアドレスの確認が済んだら、準備完了です。

スターターダッシュボードの構成を理解する

まずは、スターターダッシュボードがどのようなものか、見てみましょう。



すぐに使える、構築済みの4つのビュー

1. 折れ線グラフ – 現在までの公開作品数を年ごとに表示します。
2. ツリーマップ – 言語ごとの作品数を表示します。
3. 棒グラフ – アカデミー賞受賞数の多い上位10作品を表示します。ツールチップには、受賞部門が表示されます。
4. 散布図 – IMDbの評価と投票数にもとづいて、評価の高い映画を表示します。ツールチップには、主な出演者が5名程度表示されます。
5. 棒グラフ – 作品数の多い上位10ジャンルを示します。

折れ線グラフ、ツリーマップ、棒グラフ（左の画像の1、2、5）をクリックすると、ダッシュボードをフィルタリングできます。

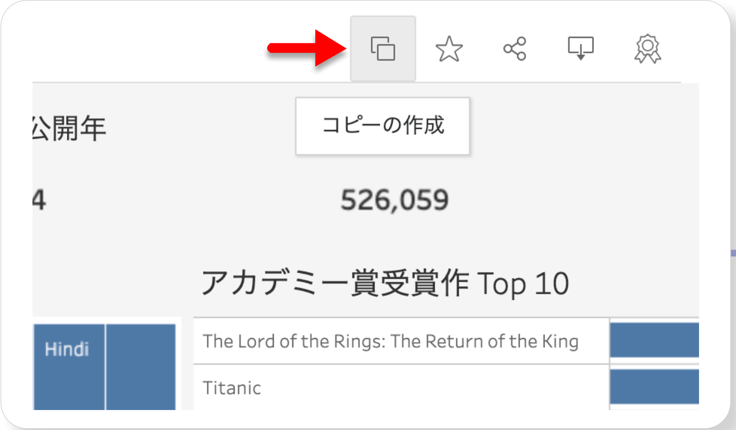
散布図には、投票数50万票以上を集めている上位2,500本の映画作品がデフォルトで表示されます。このしきい値は、［クリックしてこのビューを操作］ボタン（左の画像のBマーク）をクリックして変更できます。

このダッシュボードの詳細を確認するには、Aボタンをクリックしてください。

スターターダッシュボードのコピーを作成する

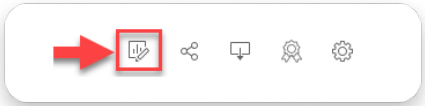
さっそく、映画スターターダッシュボードのコピーを作ってみましょう。Tableauを初めて使う場合は、ブラウザーでの操作をおすすめします。手順は以下のとおりです。

1. スターターダッシュボードにアクセスします。
2. ページの右上にある［コピーの作成］ボタンをクリックします。



3. Tableau PublicでVizとダッシュボードを編集するウィンドウが開きます。ダッシュボードのコピーを作成したら、データ探索の準備は完了です。

さっそくデータを探索して、ストーリーを作ってみましょう。次ページ以降の各セクションで、データ探索のアイデアをいくつか紹介します。実行する際には、編集モードになっていることを確認してください。下図のボタンをクリックすると、編集モードになります。



ヒント

ローカルで作業したい場合は、ワークブックをダウンロードできます。［ダウンロード］ボタンをクリックしてください。このワークブックのデータセットは容量が非常に大きいため、コンピューターのパフォーマンスに影響する可能性があります。また、2000年から2023年の映画のデータのみで構成されるワークブックもダウンロード可能です。

ローカルで作業するには、コンピューターにTableauのインストールが必要です。

Tableauの無料トライアル版は、こちらから。

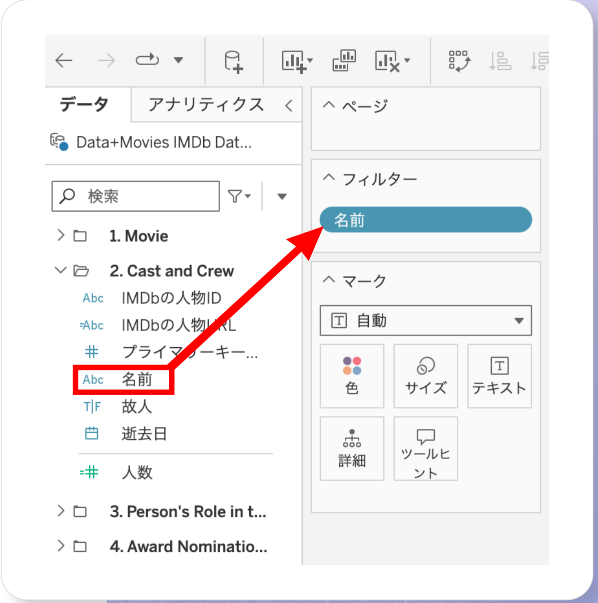
グラフを作成する – 好きな俳優の出演作品を見つける

3人の俳優モーガン・フリーマン（Morgan Freeman）、メリル・ストリープ（Meryl Streep）、ニコラス・ケイジ（Nicolas Cage）についてのインサイトを発見してみましょう。出演作品の評価が、時代とともにどのように変化しているのかを調べてみることにします。

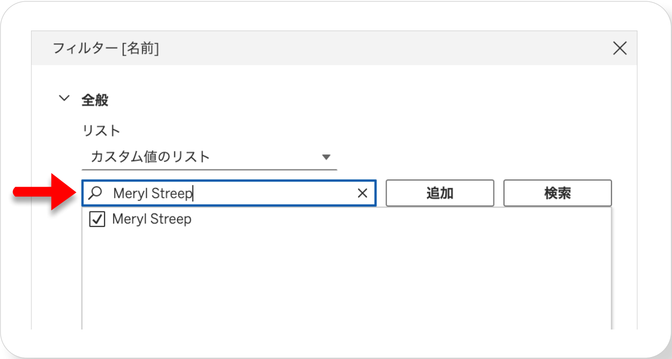
［新しいワークシート］ボタンをクリックします。

探索したい俳優だけを表示するフィルターを作成します。ここでフィルターを作っておかないと、映画関係者150万人ものデータを使ったグラフになってしまいます。

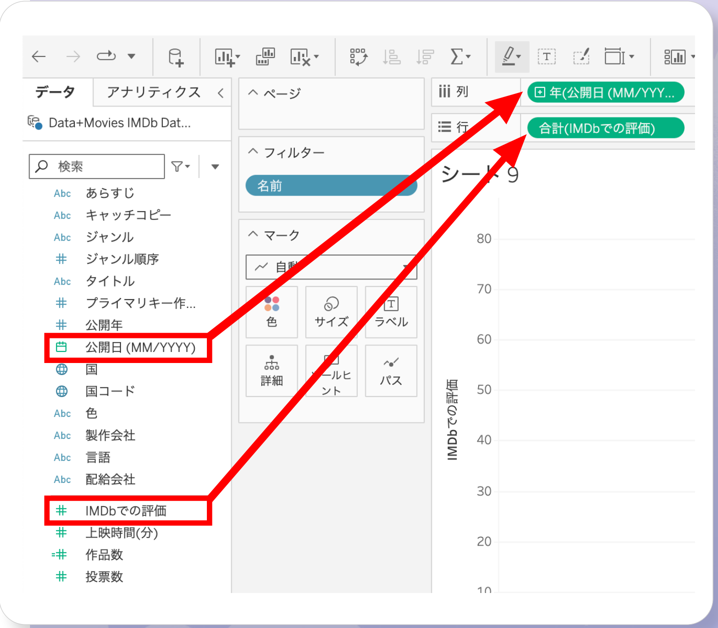
1. 左側の［データ］ウィンドウから［名前］フィールドを［フィルター］シェルフにドラッグします。



2. 探索したい俳優の名前をテキストボックスに入力します。* その際、［カスタム値のリスト］が選択されていることを確認してください。俳優の名前を入力して、［追加］ボタンをクリックします。
* 俳優の名前は、英語での入力を推奨。

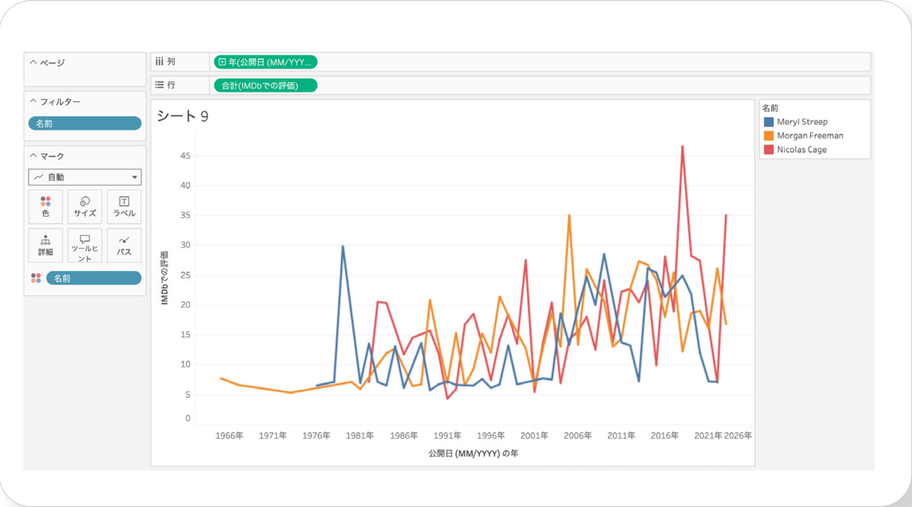


3. フィールドをビューにドラッグします。［公開日 (MM/YYYY)］を［列］シェルフに、［IMDbでの評価］を［行］シェルフにドラッグします。



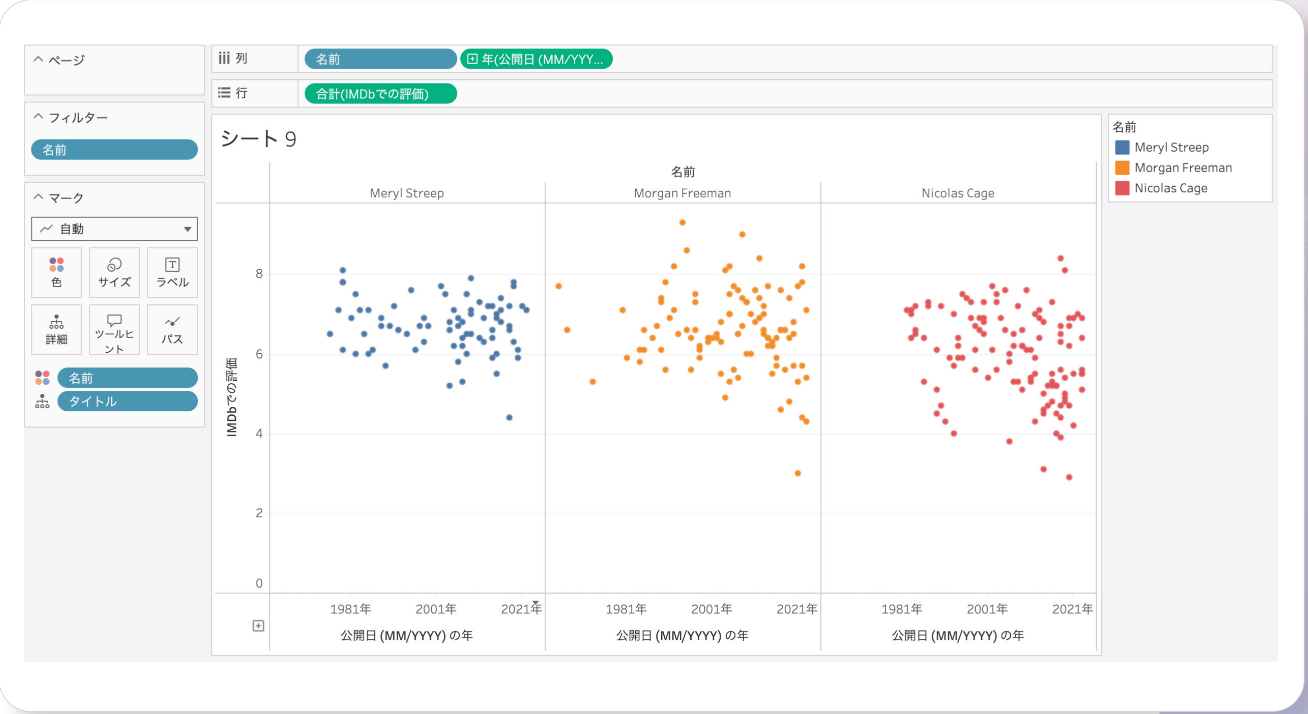
折れ線グラフが表示されました。このグラフは、3人の俳優のいずれかが出演した映画の平均評価の変化を示しています。しかし、確認したいのは各俳優の出演作の評価です。

[名前] を [データ] ウィンドウから [色] シェルフにドラッグ&ドロップすると、俳優ごとの折れ線グラフが作成されます。これで最初のビューが完成しました。以下のように表示されています。



このグラフからでもインサイトを得られますが、もう少しわかりやすくしてみましょう。

それぞれの出演作を確認できるようにしてみます。[データ] ウィンドウから [タイトル] を [詳細] シェルフにドラッグします。その後、新たに [名前] を [列] シェルフにドラッグします ([公開日 (MM/YYYY)] の左側にドラッグ)。次ページの画像のようなビューが表示されます。点にカーソルを合わせると、詳細を確認できます。



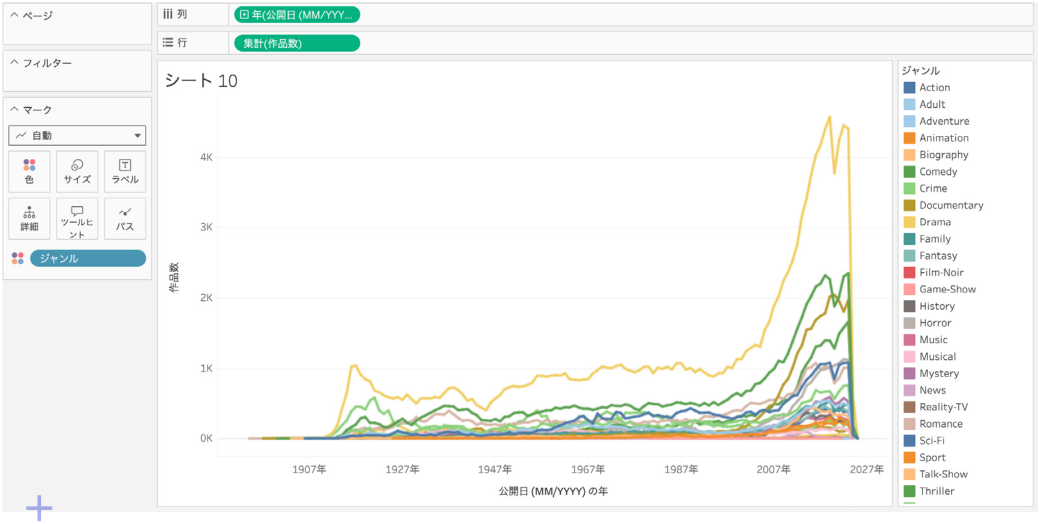
このビューから、さまざまなことがわかります。評価のばらつきが最も小さいのは、メリル・ストリープの出演作品です。モーガン・フリーマンとニコラス・ケイジは、評価の高い作品にも、いまひとつな作品にも出演しています。3人の中で、最も評価の高い作品に出ているのは、モーガン・フリーマンです。最低評価の映画に出演しているのは、ニコラス・ケイジです（2020年の作品『アース・フォール JIU JITSU』で、スコアは10点満点で2.9点）。

このビューに平均値のリファレンスラインを追加することもできます。出演作品の評価の平均が最も高い俳優は誰でしょうか。また、[投票数] などの測定値を [サイズ] シェルフに追加してみてもいいでしょう。

グラフを作成する – ジャンルの変遷を探ってみる

このセクションでは、映画のジャンルに注目して、簡単なダッシュボードを作成します。

ジャンルのトレンドがどう変化しているかを探ってみましょう。新しいシートを作り、[公開日 (MM/YYYY)] を [列] シェルフにドラッグします。次に [作品数] を [行] シェルフにドラッグします。これで、映画の作品数を時系列で示すグラフが作成されます。ここで、[ジャンル] を [色] シェルフにドラッグすると、下図のようなグラフができます。



2000年以降、ドラマが急増していますね。IMDbでは現在、ドラマの作品数が最も多く、次にドキュメンタリーとコメディが続きます。凡例でジャンルの色をクリックすると、該当するジャンルがハイライトされます。初期の年代では [Unknown] というジャンルが多くなっていますが、これは多くの作品がどのジャンルにも分類されないことを示しています。凡例で [Unknown] を右クリックすると、グラフから除外できます。

[ジャンル] を [色] フィールドに追加すると、たくさんの色で表示されます。なかには重複している色もあります。個人的なデータ探索であればそのままでもいいのですが、ビジュアライゼーションを共有するなら、色数を減らすことをおすすめします。Tableauのホワイトペーパー『優れたビジュアライゼーションの条件』では、ビジュアライゼーションのベストプラクティスを紹介しています。

画面の下にあるシートタブをダブルクリックして、このシートに「ジャンルの変遷」という名前を付けます。

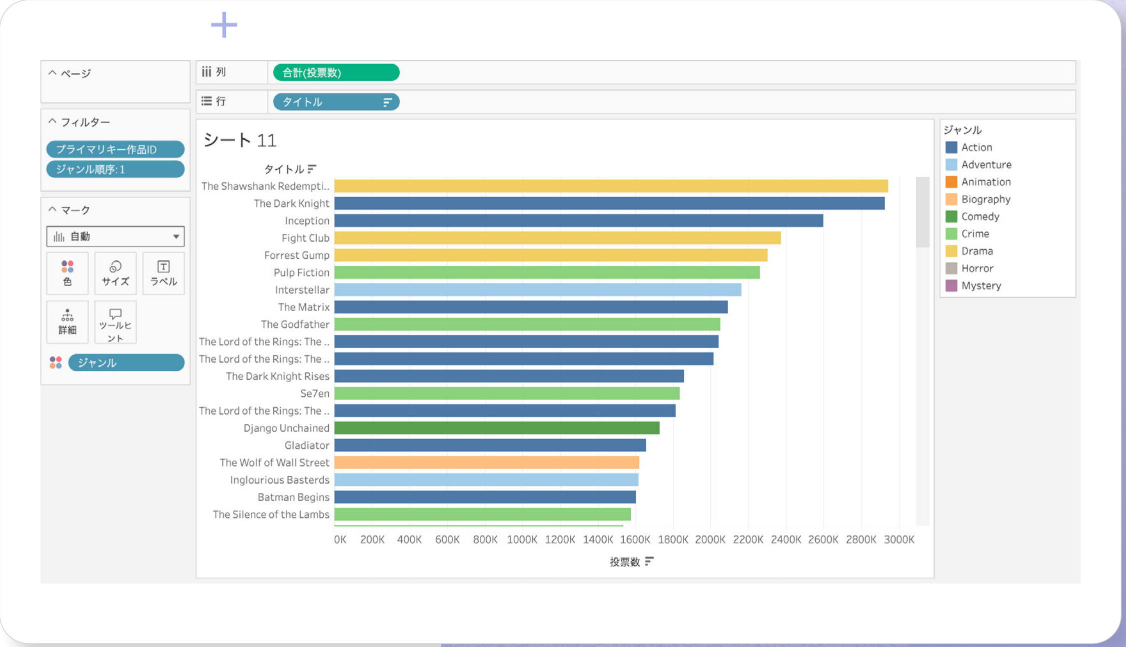
グラフを作成する – 投票数の多い映画を確認する

今度は棒グラフを作成して、最も投票数の多い映画を調べてみましょう。

まず、上位100本の映画に絞り込みます。新しいシートを作り、［データ］ウィンドウから［プライマリキー作品ID］を［フィルター］シェルフにドラッグします。［フィルター］ダイアログが開いたら［上位/下位］をクリックします。右の図のように詳細を入力して、上位100作品のみが表示されるようにします。

次に［投票数］を［列］シェルフに、［タイトル］を［行］シェルフにドラッグします。ツールバーのソートボタンをクリックすると、投票数の多い映画を確認できます。上位に入っているのはどの作品でしょうか。

［ジャンル順序］を［フィルター］シェルフにドラッグしてフィルターを「1」に設定し、［ジャンル］を［色］シェルフにドラッグすると、ジャンルの情報を表示できます。右の図のようなグラフになります。

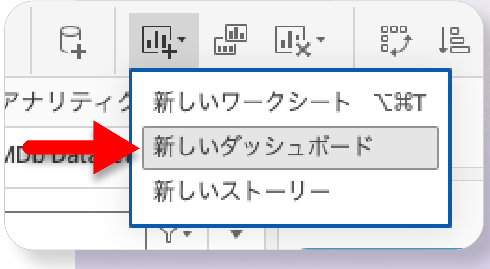


画面の下にあるシートタブをダブルクリックして、このシートに「トップ投票数映画」という名前を付けます。

ダッシュボードを作成する – 2つのグラフをダッシュボードにまとめる

最後に、2つの「ジャンル」グラフを1つのダッシュボードにまとめて、年別、ジャンル別に、最も投票数の多い映画を探索できるようにしましょう。

ツールバーで［新しいダッシュボード］ボタンをクリックして、ダッシュボードを作成します。

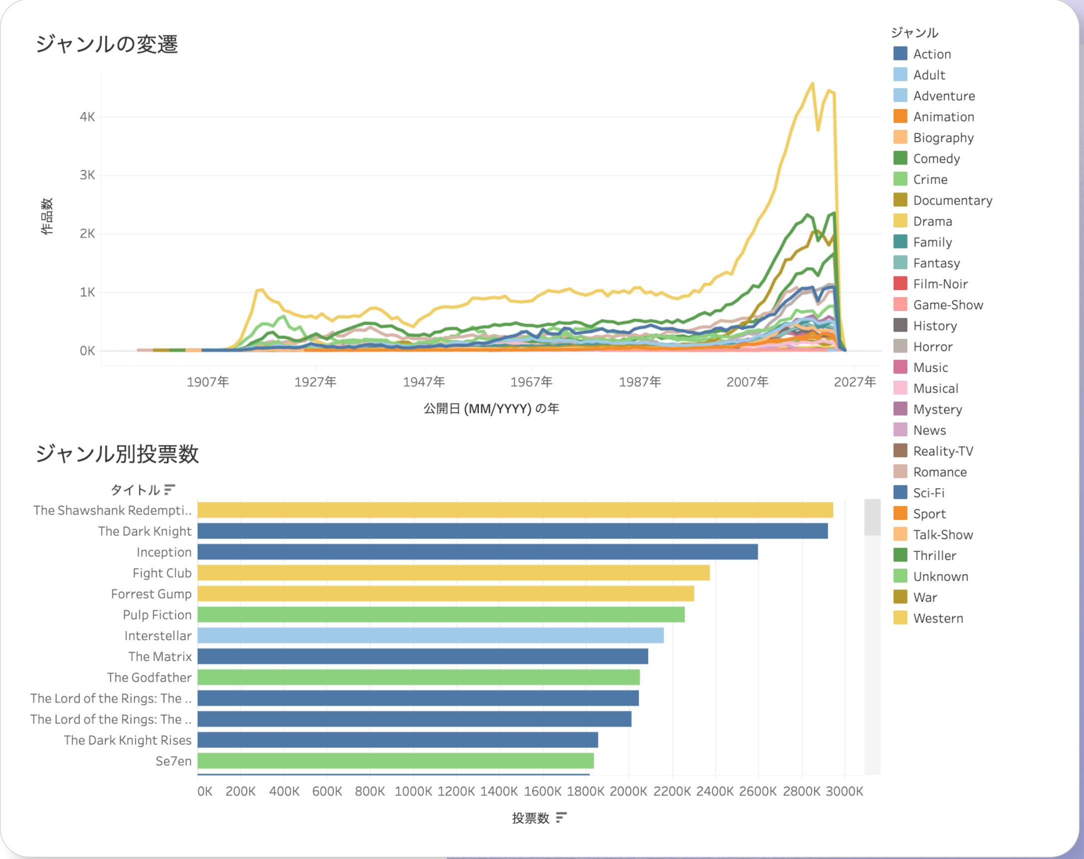


画面の左に、ワークブック内にあるすべてのシートのリストが表示されます。

［ジャンルの変遷］シートをダッシュボードのキャンバスにドラッグします。グラフと凡例が表示されます。

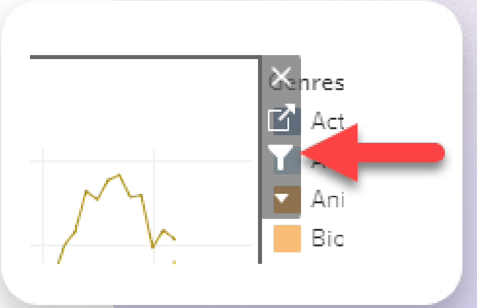
次に［トップ投票数映画］シートをダッシュボードにドラッグします。マウスボタンを押している間、グラフを配置する場所がプレビューで表示されます。2つ目のグラフは、折れ線グラフの下にドロップしてください。すると、右の図のように表示されます。

これでダッシュボードが作成されました。2つのシートを連動させてみましょう。



折れ線グラフを使ってフィルタリングできるようにします。折れ線グラフをクリックすると、棒グラフにその年の上位作品が表示されるようにしてみましょう。

ダッシュボードには、簡単にフィルターを追加できます。折れ線グラフのシートをクリックして、右上の［フィルターとして使用］をクリックします。



折れ線グラフでいずれかの線をクリックして、棒グラフをフィルタリングしてみましょう。……あれ？棒グラフが消えてしまいました。

ここで、もう1つやらなければならないことがあります。［トップ投票数映画］シートに戻ってください。［フィルター］シェルフに［アクション］という新しいフィルターが追加されています。これは、先ほどダッシュボードで作成したフィルターアクションです。

［アクション］フィルターは、上位100作品のフィルターの前に適用する必要があります。そこで、コンテキストフィルターを作成します。［アクション］フィルターを右クリックして、［コンテキストに追加］を選択します。フィルターの色が青からグレーに変わり、またグラフが表示されるようになります。コンテキストフィルターの詳細については、こちらを参照してください。

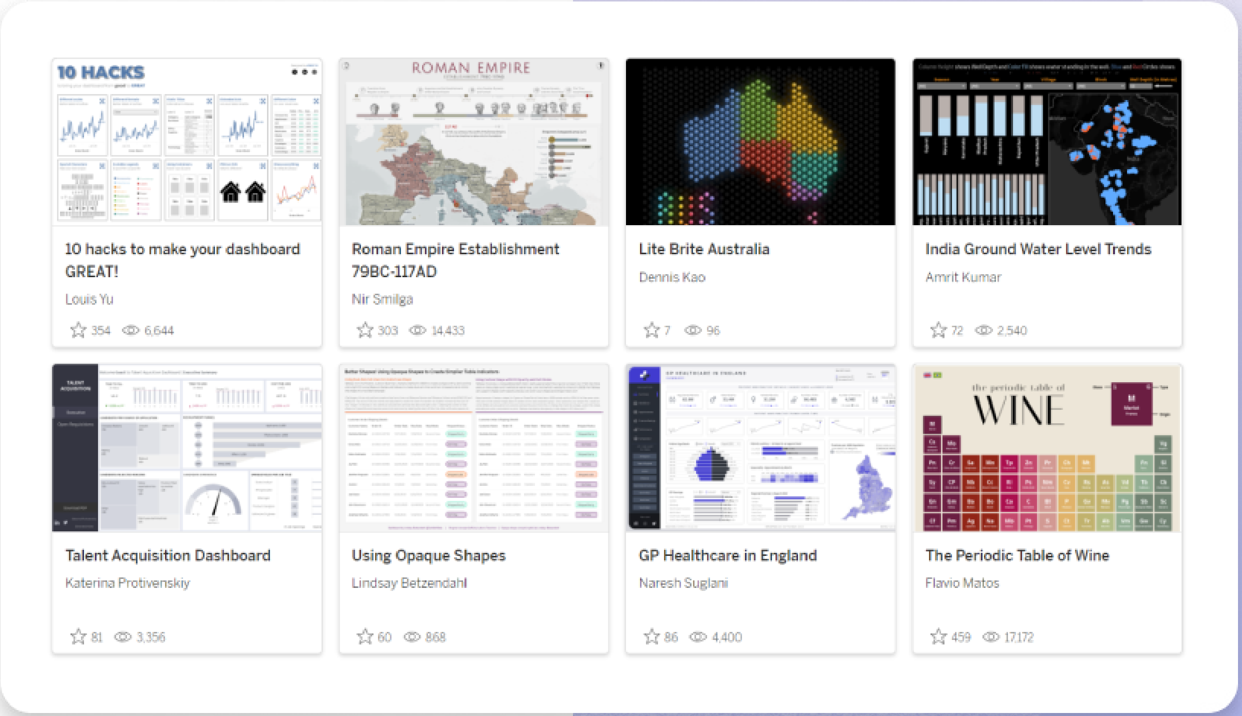
ダッシュボードに戻り、フィルターを試してみましょう。ダッシュボードがインタラクティブになり、ジャンルの折れ線グラフ上で公開年に当たる部分をクリックすると、その年の投票数の多い映画が表示されるようになりました。

書式を設定する

これでグラフとダッシュボードを作れるようになりました。ドラッグ&ドロップ操作で、さらにいろいろなストーリーを発見できます。

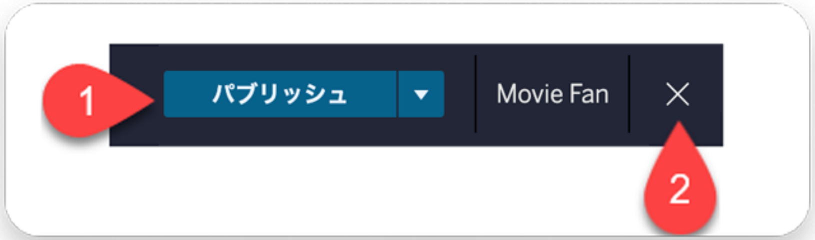
共有したいストーリーを見つけたら、書式を設定して、見せ方を工夫しましょう。このeBookでは、書式設定の詳細は割愛しますが、色、注釈、レイアウト、背景など、さまざまなカスタマイズが可能です。詳しくは、Tableauのヘルプページを参照してください。

Tableauの「今日のViz」では、Tableau Publicに最近投稿されたVizを紹介しています。ぜひ、書式設定のアイデア探しに活用してみてください。また、映画をテーマにしたVizを検索することもできます。



ワークブックを保存し、表示する

作業内容は、画面の右上にある［パブリッシュ］ボタン（1）でいつでも保存できます。その後［X］ボタン（2）をクリックしてエディターを終了し、作成したVizを確認できます。



ローカルで作業している場合は、以下のメニューオプションからワークブックをパブリッシュしてください。

- Tableau Desktop Public Editionで、［ファイル］＞［Tableau Publicに保存］を選択します。
- Tableau Desktopで、[サーバー] > [Tableau Public] > [Tableau Publicに保存]を選択します。

これでワークブックがプロファイルに表示されます。

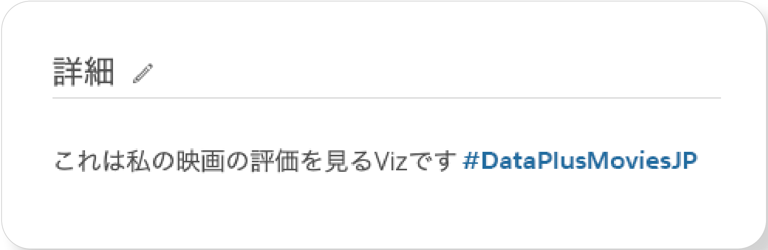
IMDbのデータセットを楽しんでいただけましたか？ IMDbの商用ライセンスについては、imdb-licensing-support@imdb.comにお問い合わせください。

ワークブックを共有する

Vizにハッシュタグ「#DataPlusMoviesJP」を追加する

1. パブリッシュしたビューを下までスクロールし、[詳細] の隣にある編集ボタンをクリックします。

2. 作成したVizにタイトルを付け、説明を追加します。**当社があなたのワークブックを見つけられるように、説明にハッシュタグ#DataPlusMoviesJPを追加してください。**下図のようになります。



Data+Moviesチャレンジにご参加いただき、ありがとうございました。楽しんでいただけましたか？

次のステップ

Tableauについてもっと詳しく学べる、さまざまなリソースを用意しています。

1. Tableauの無料トライアルを始める
2. Tableau Publicで映画をテーマにしたダッシュボードをチェック
3. Tableau Communityに参加。どなたでも参加できます
4. 無料のトレーニングビデオを視聴
5. DataFamニュースレターに登録
6. Tableau Publicについてもっと詳しく知りたい方はこちら

付録 – データフィールドリスト

データセットの各フィールドの意味を説明します。[データ］ウィンドウ内のフィールド名にカーソルを合わせると、説明が表示されます。

1. Movie

映画1作品ごとに1件のレコードがあります。

フィールド名	説明と補足
色	モノクロ、カラー、またはパートカラー
国	映画の主たる製作国
配給会社	映画を配給した会社
ジャンル	作品のジャンル ・1つの作品に複数のジャンルが適用されることがあります。 ・［ジャンル順序］フィールドと組み合わせて、ジャンルの順序を 1～3まで決定できます。
ジャンル順序	1つの作品に複数のジャンルが適用されることがあります。 このフィールドで、適用されるジャンルの並び順を決定します。
IMDbの作品ID	この作品に対するIMDbでの一意のID

IMDbの作品画像URL	作品のサムネイル画像のURL
IMDbの作品URL	IMDb.comの作品ページのURL
言語	作品で主に使われている言語
プライマリキー作品ID	この作品に対する一意のID
製作会社	映画の製作会社
公開年	映画が公開された年
公開日 (MM/YYYY)	映画が初めて公開された日 (わかる場合)
キャッチコピー	映画の宣伝に使われたキャッチコピー
タイトル	映画のタイトル
あらすじ	映画のあらすじ
IMDbでの評価	投票にもとづく映画の評価
作品数	ビューで作品の総数をカウントするために使用
投票数	作品に対する投票数
上映時間 (分)	映画の上映時間 (分)

2. Cast and Crew

映画における配役や役割ではなく、人物そのものに関する情報を示します。

フィールド名	説明と補足
逝去日	この人物が亡くなった日（わかる場合）
IMDbの人物ID	この人物に対するIMDbでの一意のID
IMDbの人物URL	IMDb.comの人物ページのURL
故人	故人であるかどうか
名前	一般的に知られている人物名。通常、クレジットで最も多く使われている名前です。IMDbで使用する人物名の定義については、IMDbのヘルプサイトを参照してください。
プライマリキー人物ID	この人物に対する一意のID
人数	データに含まれる厳密な人数。各人物のIDを1回だけカウントします。人物を重複してカウントしたくない場合に使用します。

3. Person’s Role in the Movie

1作品で1人の人物が演じた役と職種ごとに1件のレコードがあります。たとえばトム・ハンクスは、出演作の役ごとと職種ごとに、複数のレコードがデータベースに登録されています。

フィールド名	説明と補足
クレジット表示順序	出演者リストに表示される人物または役名の順番
職種	作品で担った役割（俳優、監督、編集など）
役名	俳優が演じた役の名前
登場人物の数	ビューで人物の総数をカウントするために使用

職種

職種には、俳優や監督に加え、美術監督、助監督、作曲家、編集、プロデューサー、脚本家、本人として出演する個人も含まれます。

4. Award Nominations

以下のフィールドで、作品がどの映画賞にノミネートされたか、
またはどの映画賞を受賞したかを調べることができます。

フィールド名	説明と補足
受賞カテゴリ	作品が受賞したカテゴリ
授賞イベント	授賞式などのイベント
受賞名	賞の名前
受賞年	作品がノミネートされた年
IMDbのノミネートID	映画賞のノミネートに対する一意のID。このIDで、その賞のノミネート作品を特定できます。
受賞歴あり	賞を受賞したかどうか
ノミネート数	ビューで賞にノミネートされた数をカウントするために使用

5. Age Rating

以下のフィールドで、データセット内の映画に対する各国の鑑賞年齢制限を調べることができます。

フィールド名	説明と補足
オーストラリア年齢制限	オーストラリアにおける鑑賞年齢制限
ブラジル年齢制限	ブラジルにおける鑑賞年齢制限
ドイツ年齢制限	ドイツにおける鑑賞年齢制限
イギリス年齢制限	イギリスにおける鑑賞年齢制限
インド年齢制限	インドにおける鑑賞年齢制限
スウェーデン年齢制限	スウェーデンにおける鑑賞年齢制限
アメリカ年齢制限	アメリカにおける鑑賞年齢制限