Tableau Tutorial

In this tutorial, we will see how we can use the tool Tableau to create visualisations going from simple line charts to complex dashboards.

Download & Install
Follow the instructions at the following address to create an account, then download and install Tableau. Note that with your KU Leuven address, you can get a student license which allows you to use Tableau for free.
🔗: https://www.tableau.com/university-students

Dataset
We will use a dataset called 'Bechdel Test'. This dataset published by FiveThirtyEight contains a list of 1,615 films from 1990 to 2013 measured against the Bechdel Test. The test, based on Alison Bechdel's comic book 'Dykes to Watch out For' (1985) is a simple indicator of representation of women in media. It asks whether a work of fiction features (1) at least two named female characters (2) who talk to each other (3) about something other than a man.

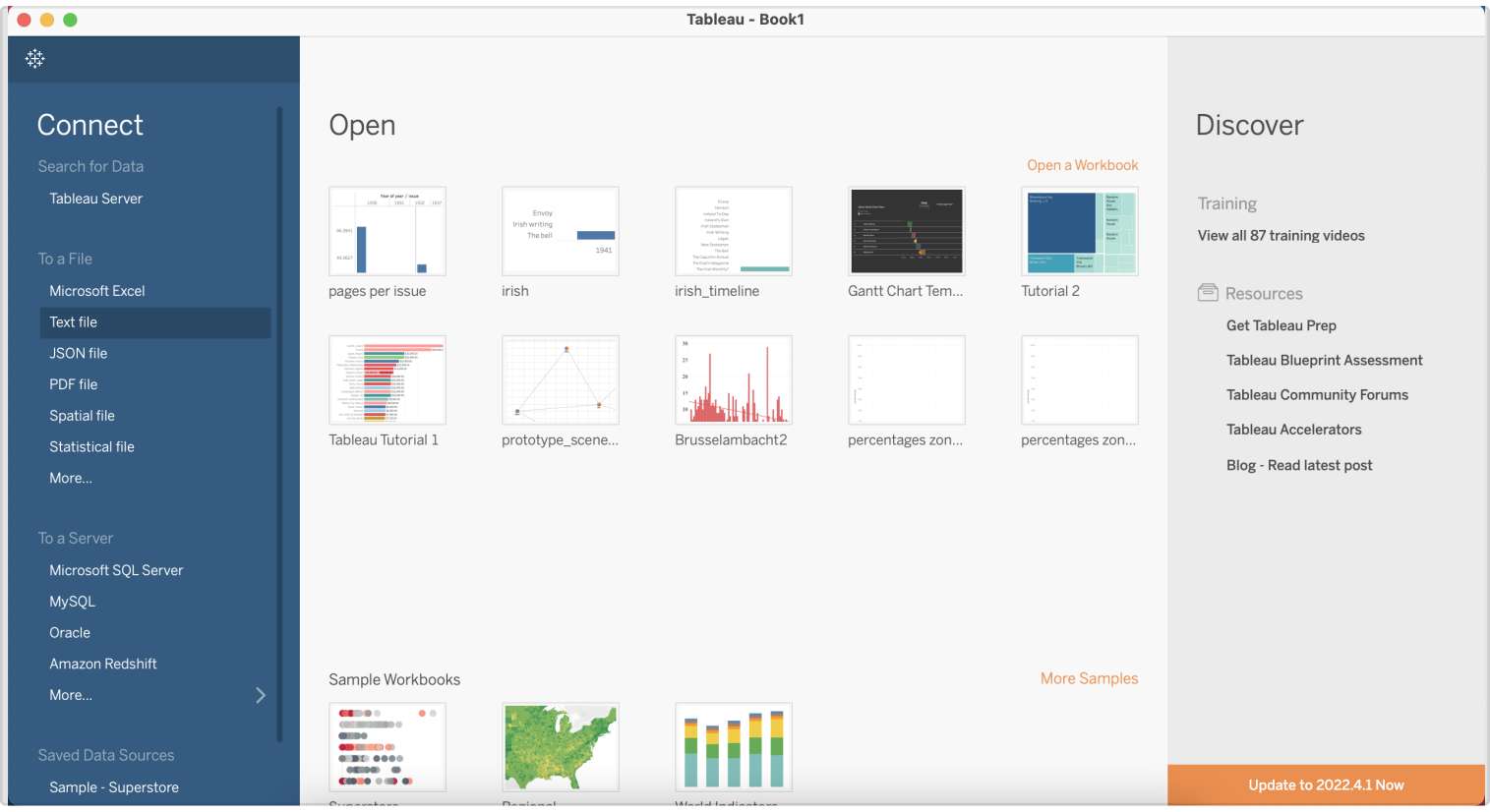
🔗: https://docs.google.com/spreadsheets/d/1_Tg-xQZeb4wgG6WxaDfcQH3G1-rMw7gAu8XJpJm0Cww/edit#gid=947183866

📍 This tutorial was made as part of a series on information visualisation for the humanities developed for the KU Leuven Faculty of Arts.

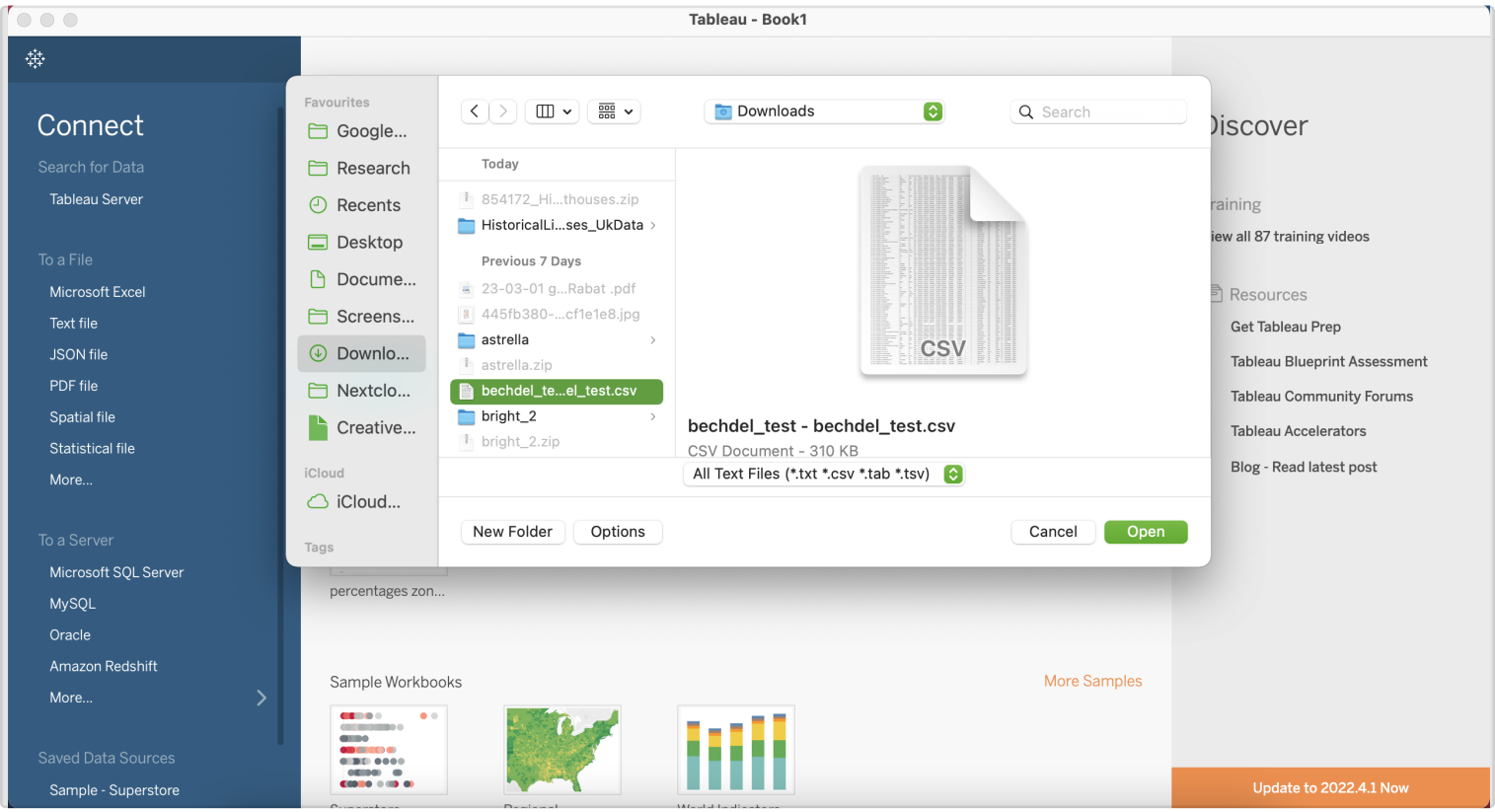
🔗 https://www.arts.kuleuven.be/digitalhumanities/english

Loading & Formatting Data

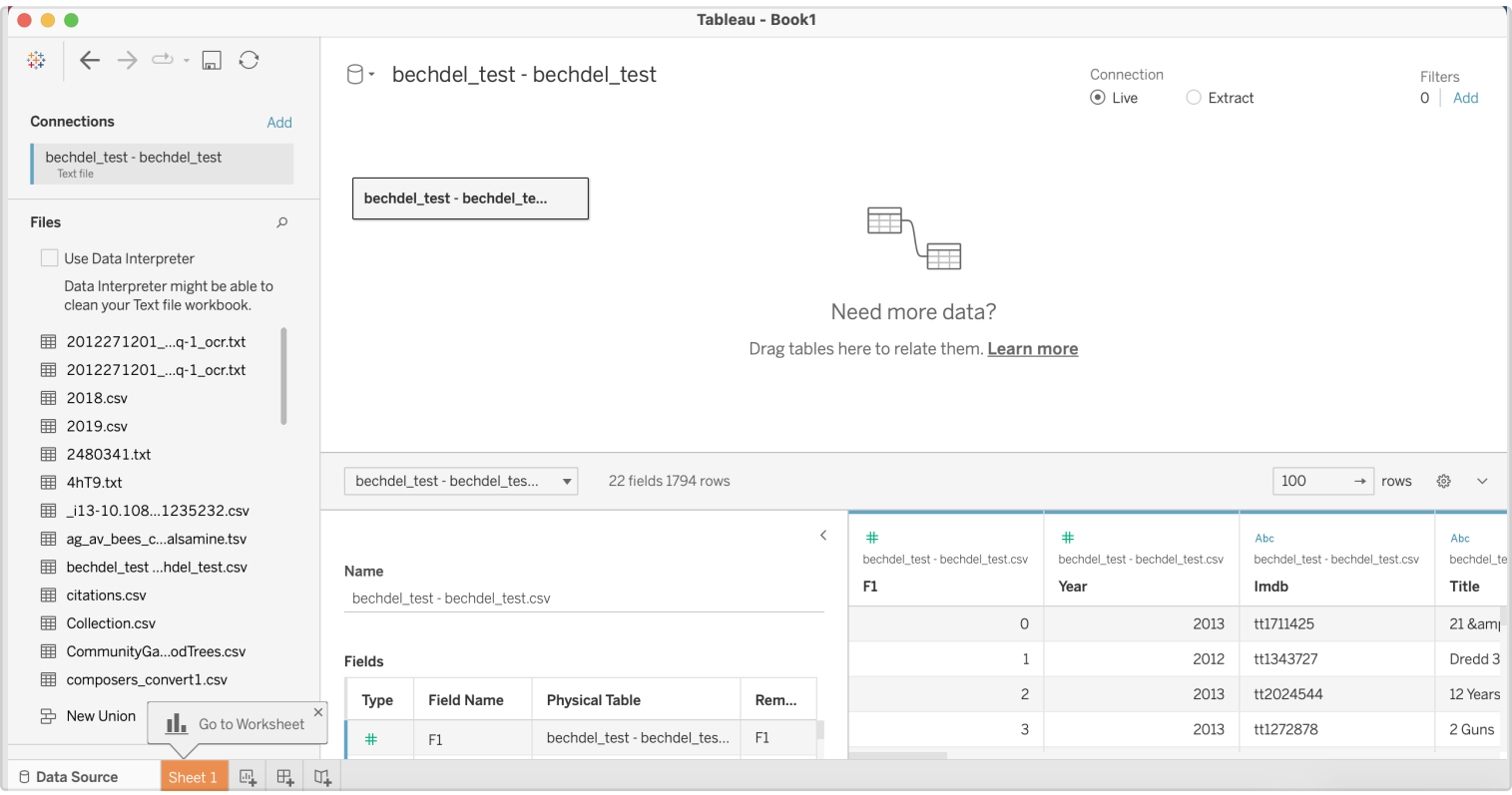
When you first open Tableau, you will see a list of recent projects in the center pane. On the left, the 'Connect' pane is what will allow you to link to your data files. You can upload an Excel file, connect to a database and so on. In this tutorial, we will use a CSV file, which we can upload using the 'Text file' button.

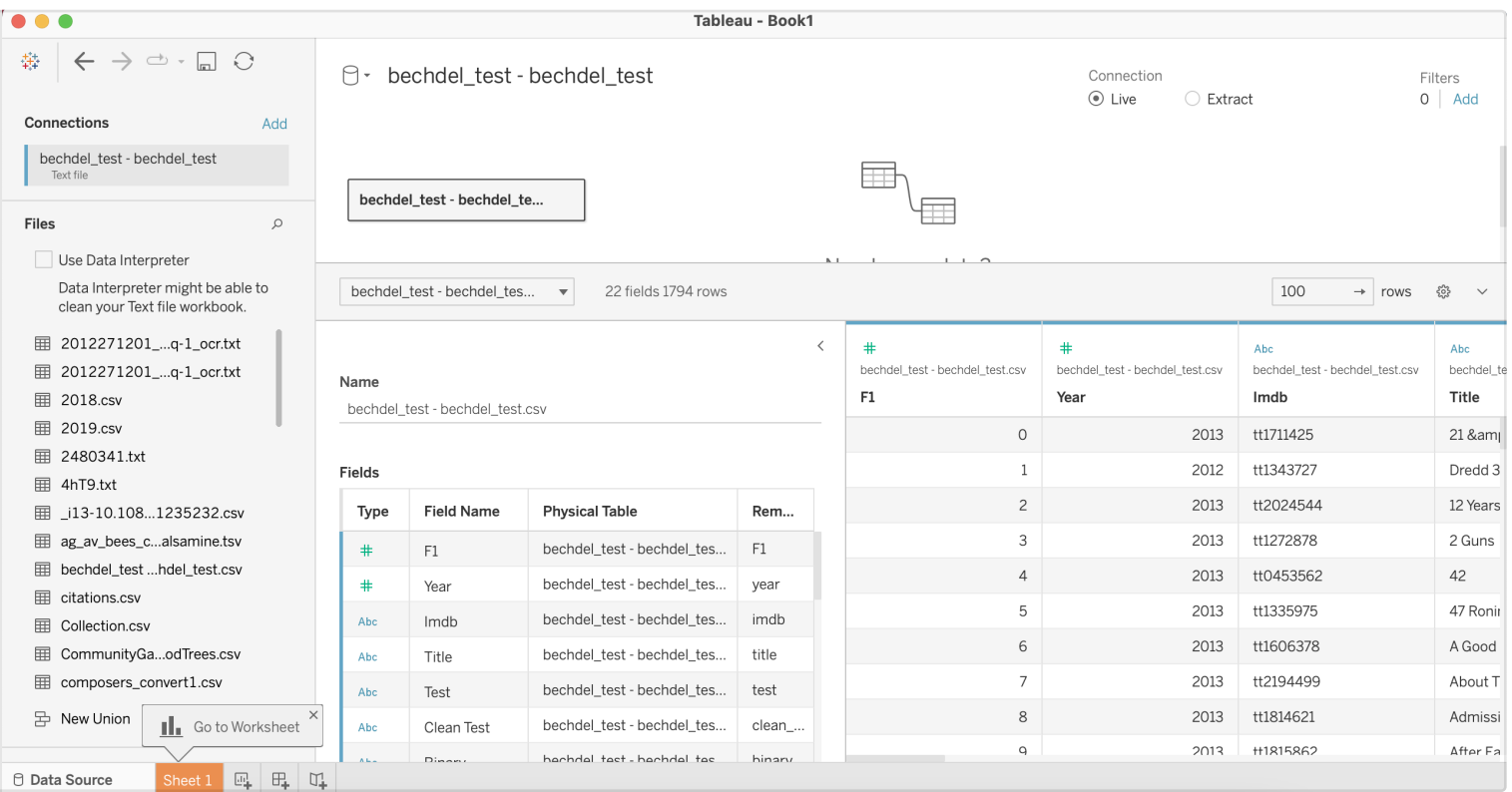


After clicking ‘Text file’, you will be prompted to pick a file from your computer. Select the file ‘bechdel-test_bechdel_test.csv’ that you downloaded from this link. Click ‘Open’ to open it in Tableau.

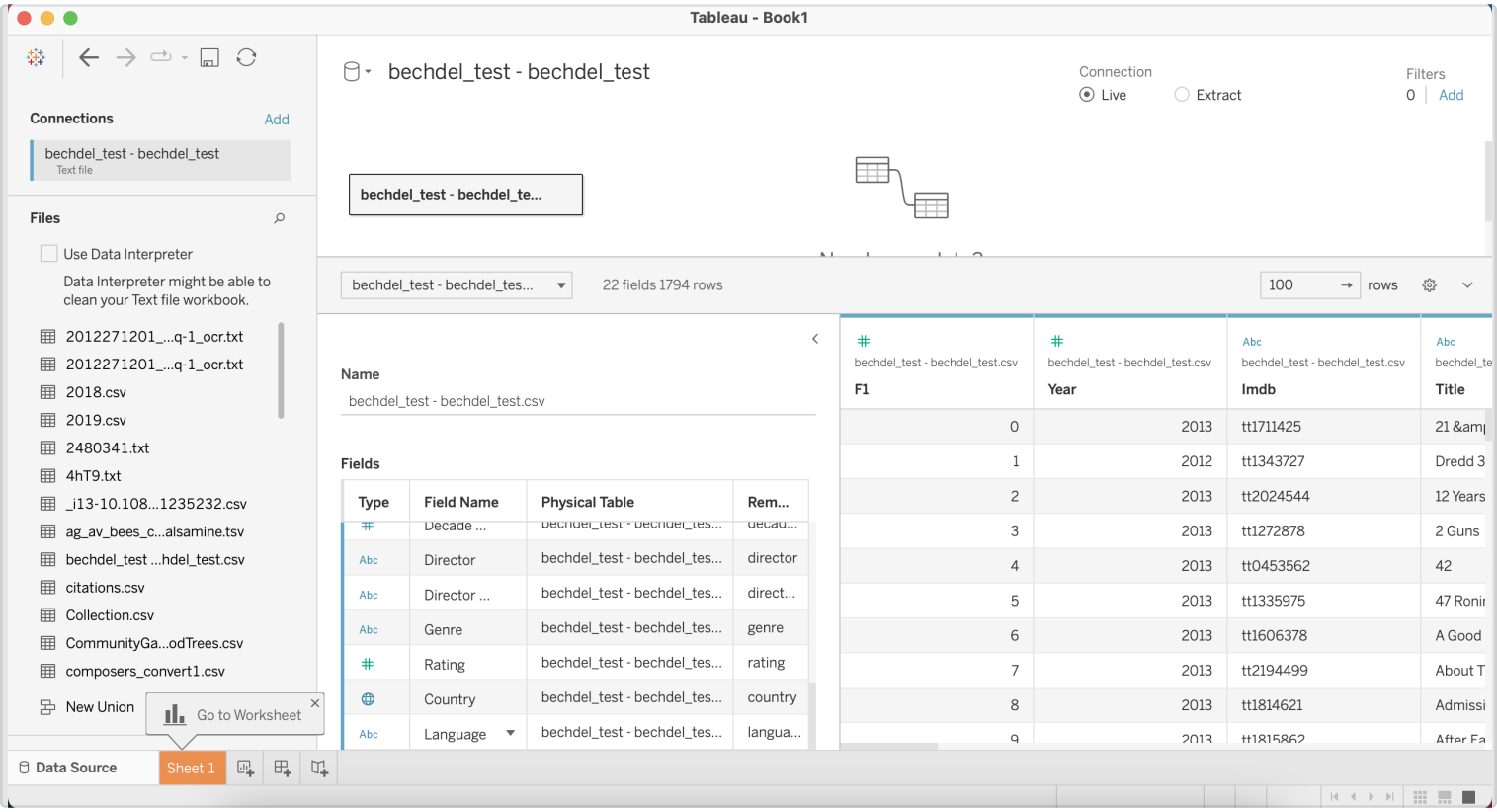


Now the fun begins! After you have opened the file, Tableau will take you to the 'Data Source' view where you will have a first summary of the contents of your data file. We will focus on the lower center pane where we can see the columns of the input file.

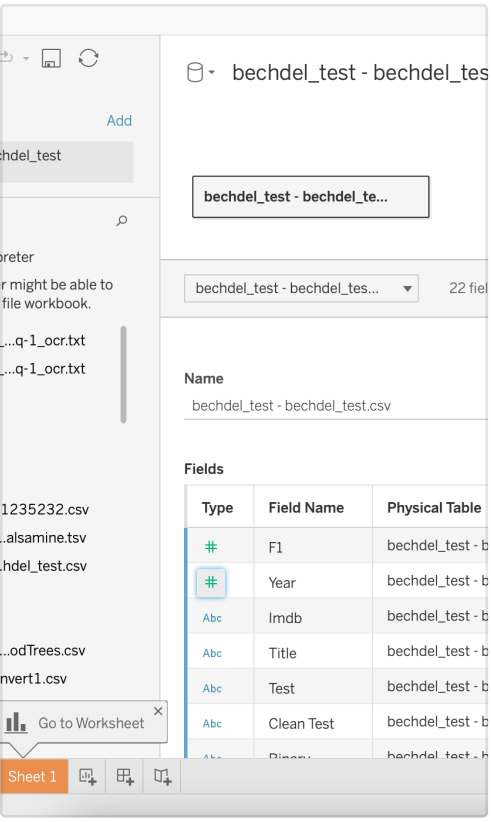


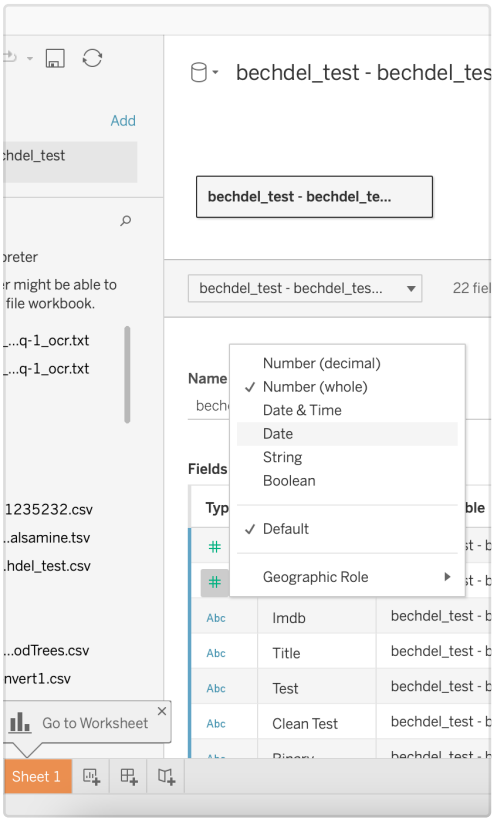


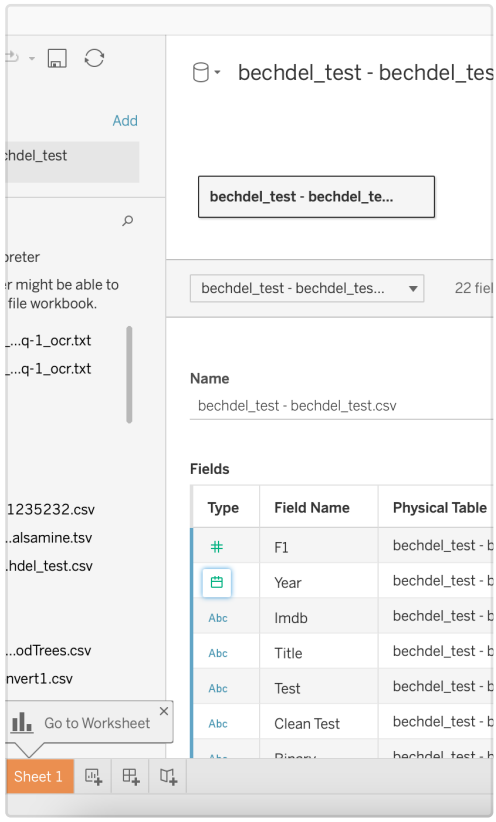
As you can see, Tableau will automatically try to infer the type of each column. For example, the ‘director’ name is a text field (or ‘string’), where as the rating is a numerical value (a rating out of 10). We can see that the country field was also correctly categorised as location data (see the globe icon).



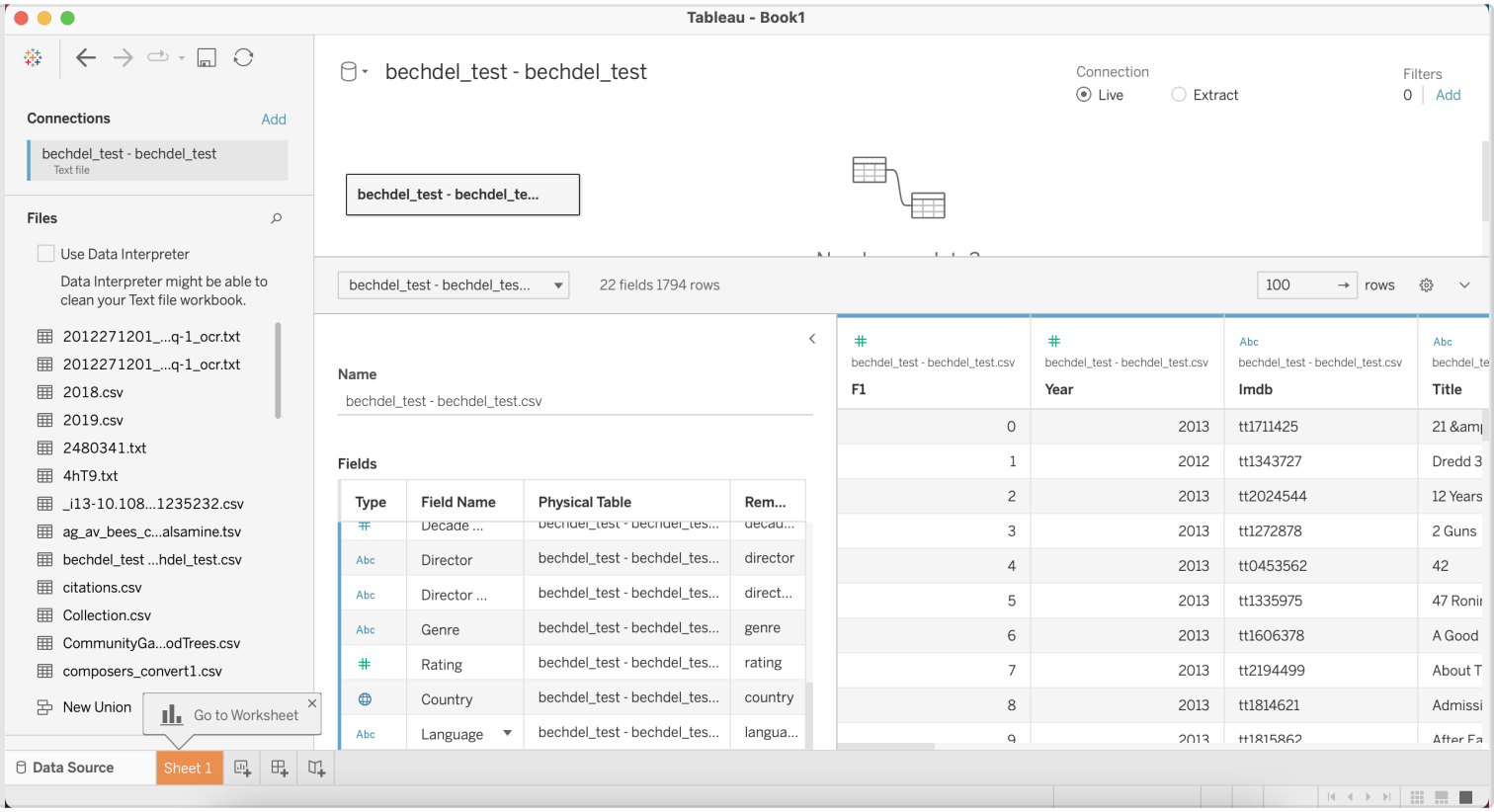
Sometimes, Tableau can make a mistake and assign an incorrect data type to a field. Let's see how we can correct that. Below, you can see that the field 'Year' was categorised as a numerical value. This is technically true, but we want the year to be viewed as a temporal attribute, which will allow us to do different operations with it.
We can change it by clicking on the 'dash' icon next to Year and selecting 'Date'. Tableau will therefore change the type to a Date and show a calendar icon instead of the dash.







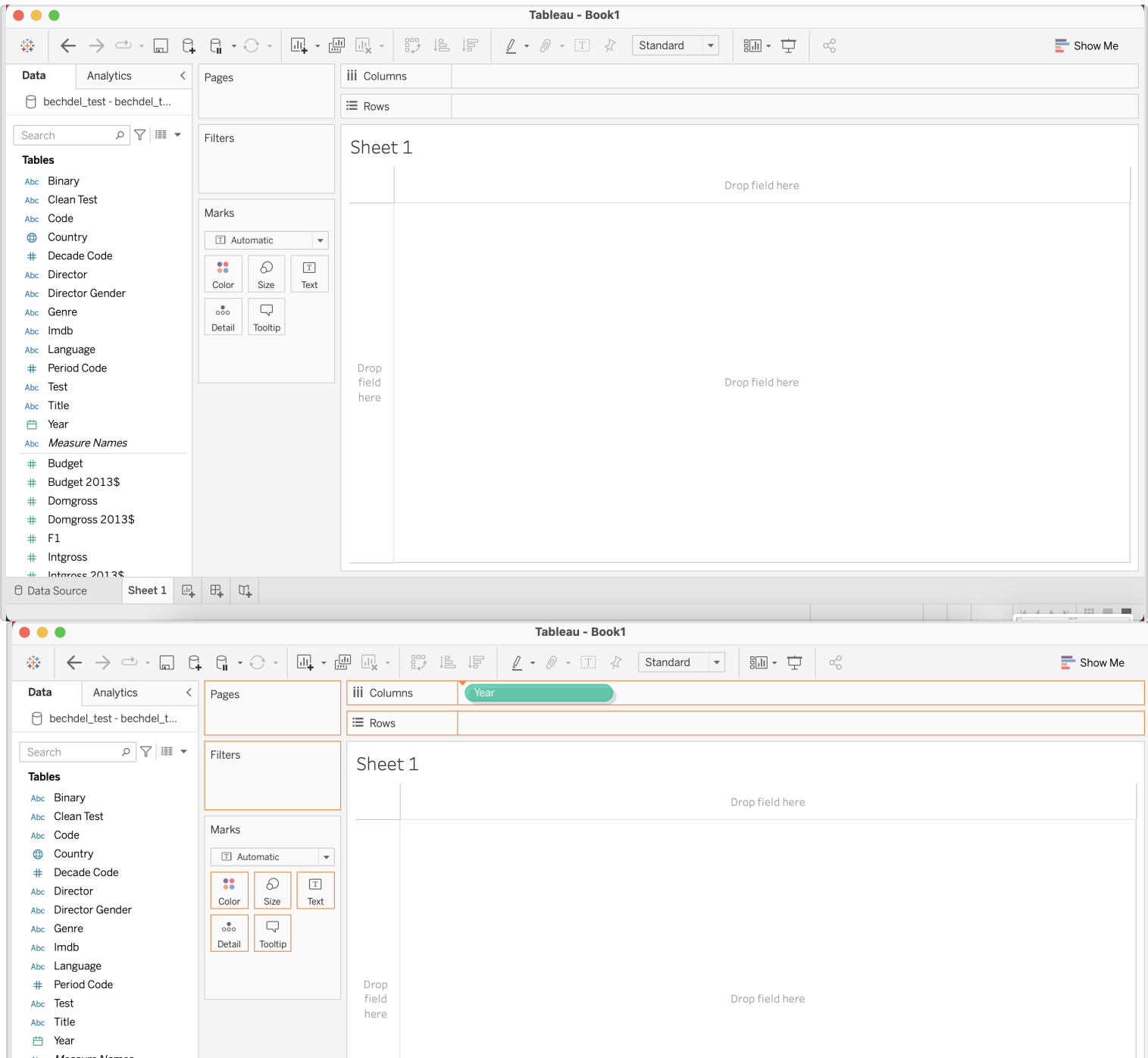
Now that our data is formatted properly, we can get to the next step and start exploring it in a worksheet. We can do that by clicking the orange button ‘Sheet 1’



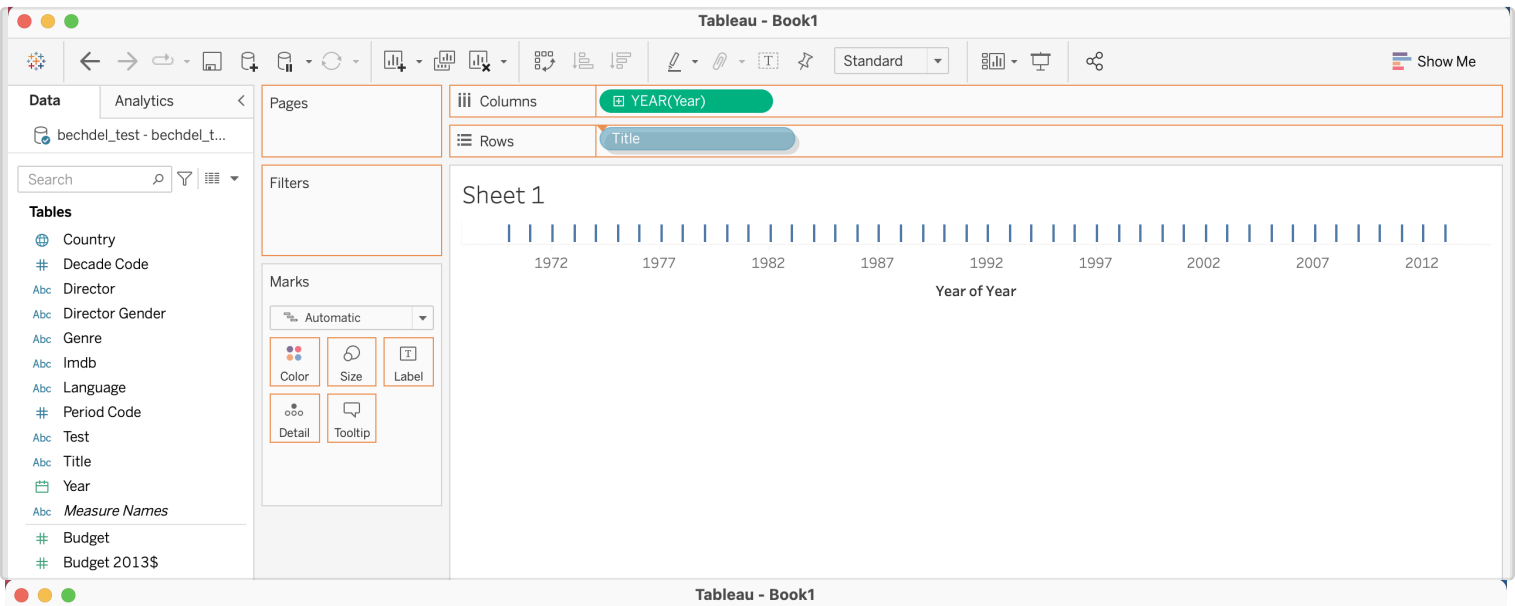
Line & Bar Charts

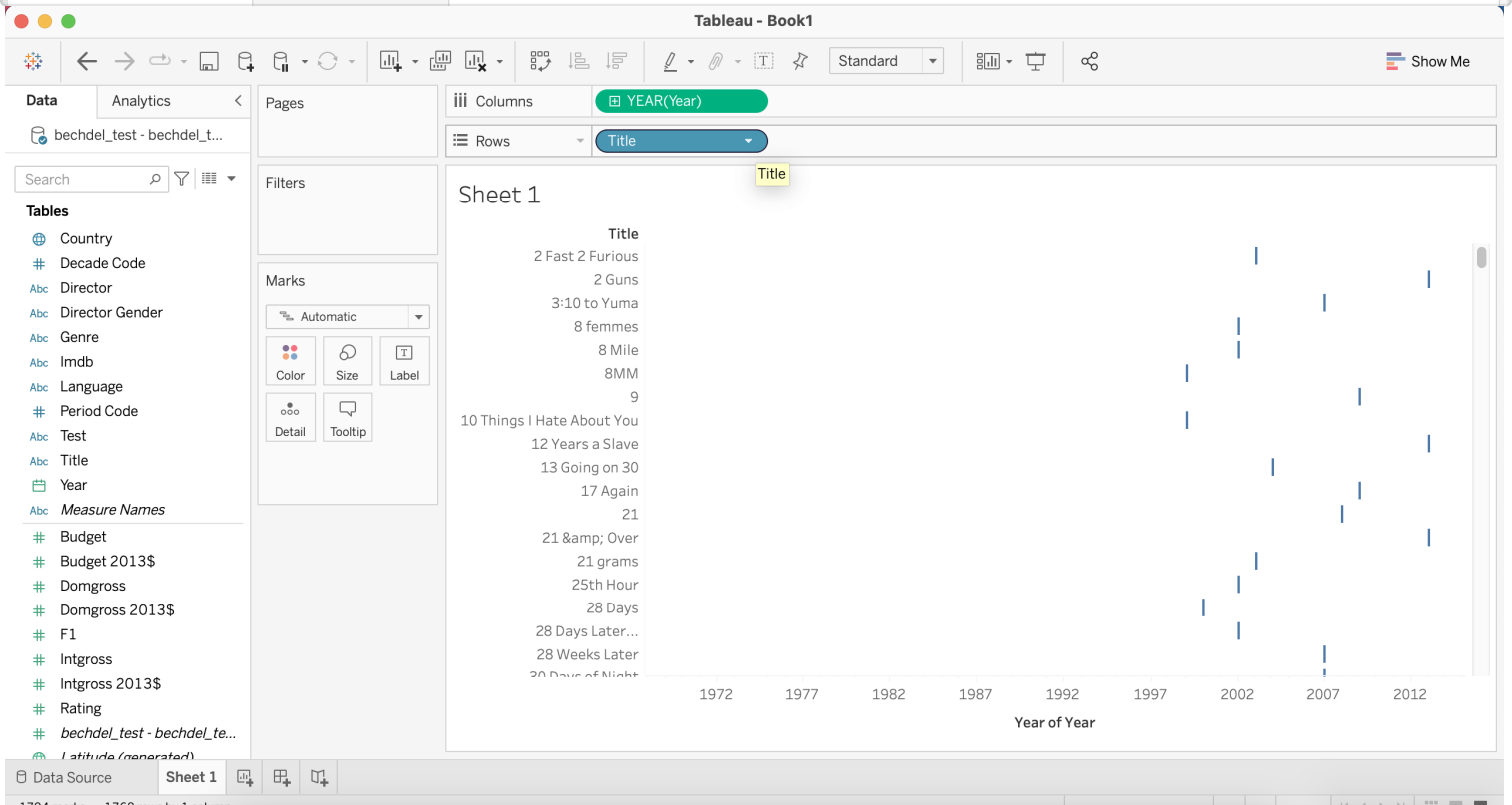
When you create a new worksheet, you will see a blank canvas with control panels on the left and top areas. The main area of the canvas ('Sheet 1') is where you can drag and drop fields to create your visualization.

On the left-hand side, in the "Data" panel, you can see all your fields listed. We want to create a distribution chart that will show us how what happens throughout the years. To do that, we will click on the 'Year' button and drag it to the 'Columns' area.

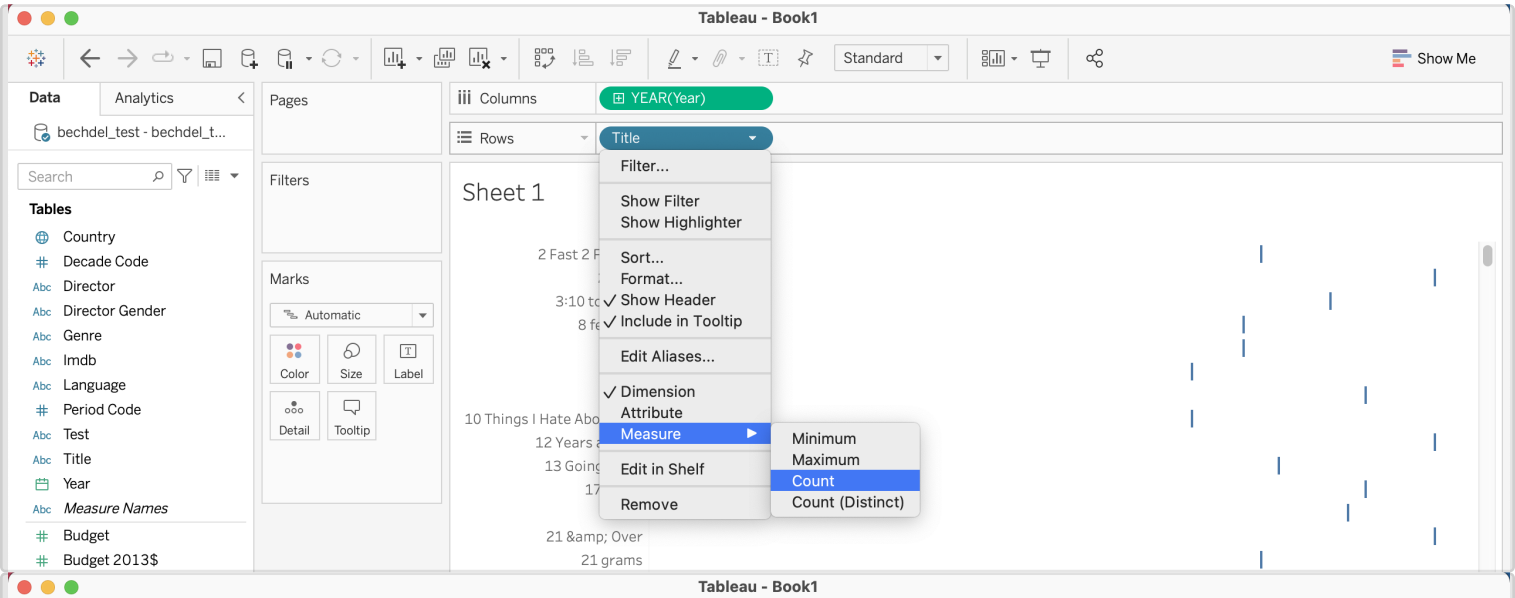


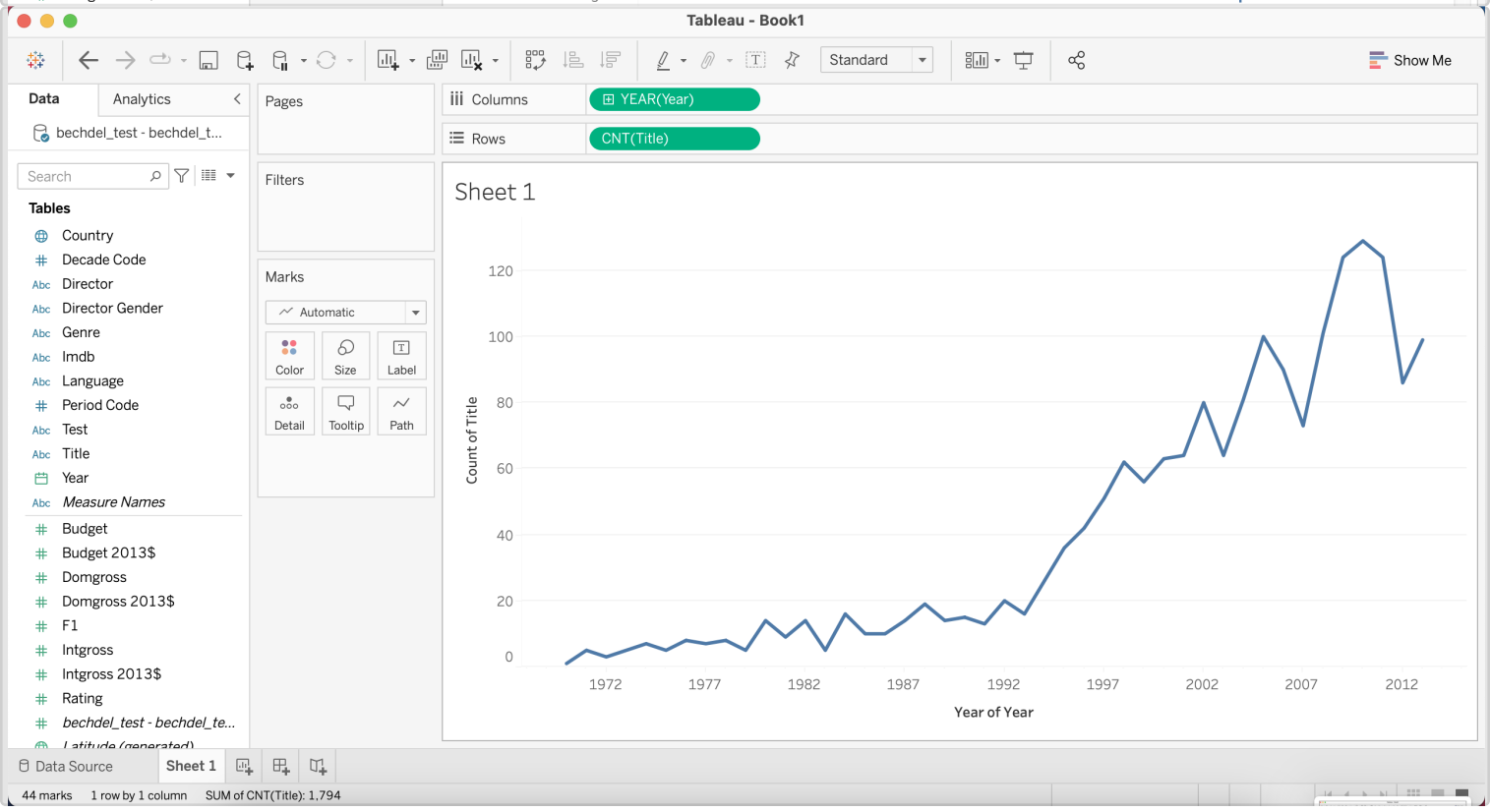
A chart appears that shows a number of short lines distributed through the years. This is not very helpful, we need to select a value to display against the years in order to get any meaningful representations. We will select the 'Title' button and drop it on the row area. This time, we see that each row shows a different movie and each column shows a year. We see a small mark for the year for which each movie was produced.



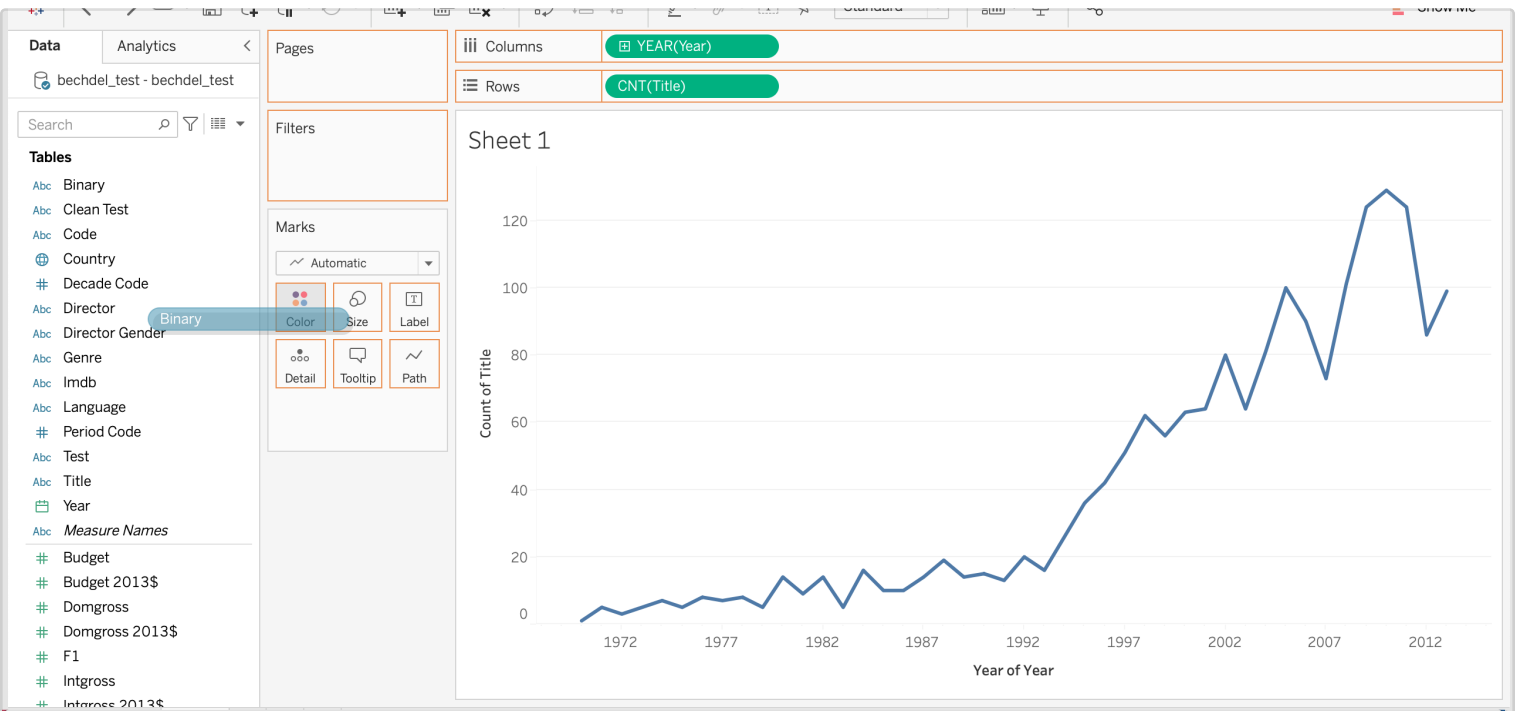


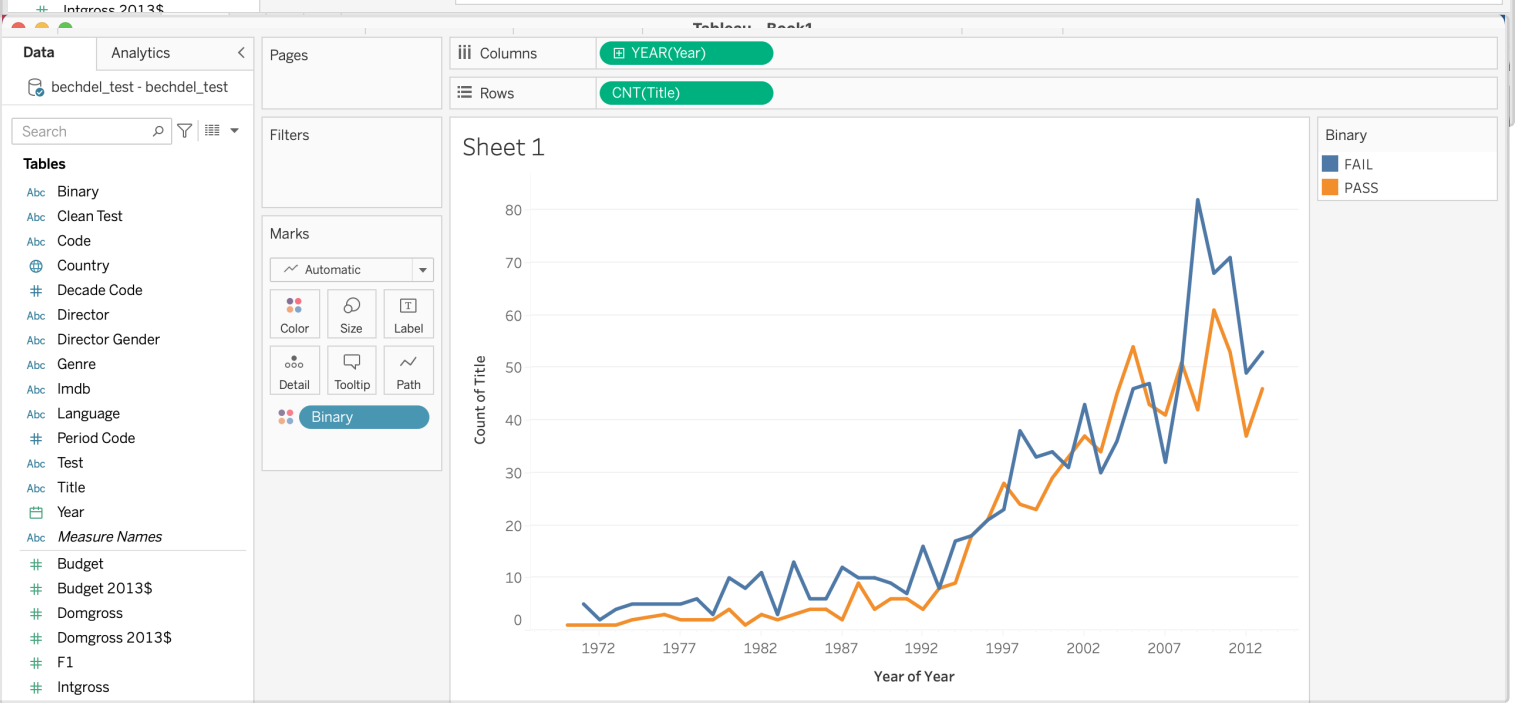
In order to get to a distribution, we need to tell Tableau to look beyond the title of each movie, and to instead count the number of titles for each year. If we click on the blue 'Title' button (called a blue pill in Tableau lingo), we can click on 'Measure' and then 'Count'. This is automatically change the visualisation to show a line chart displaying the count of different titles per year.



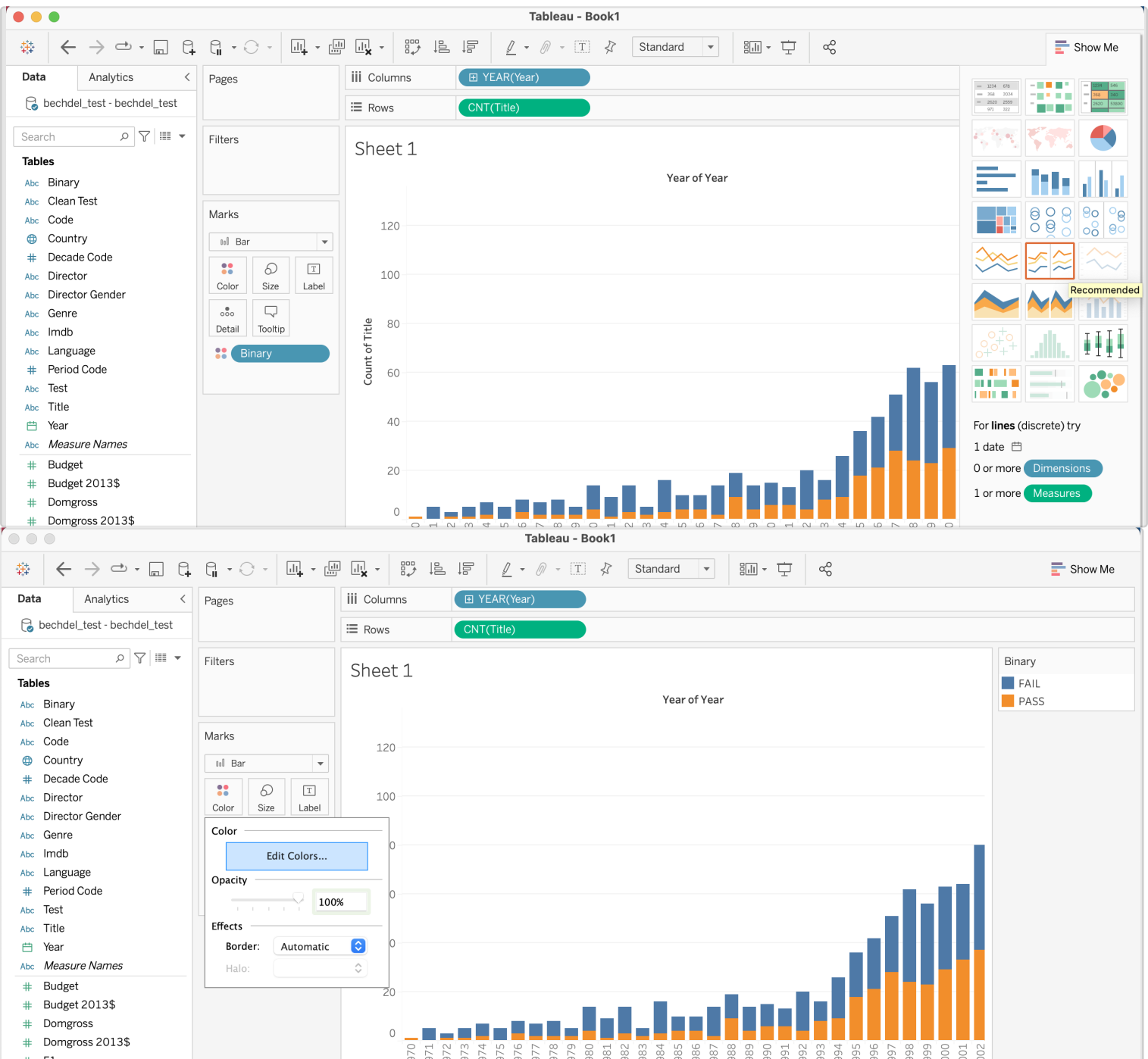


As you can see, Tableau will automatically try to infer the type of each column. For example, the ‘director’ name is a text field (or ‘string’), where as the rating is a numerical value (a rating out of 10). We can see that the country field was also correctly categorised as location data (see the globe icon).

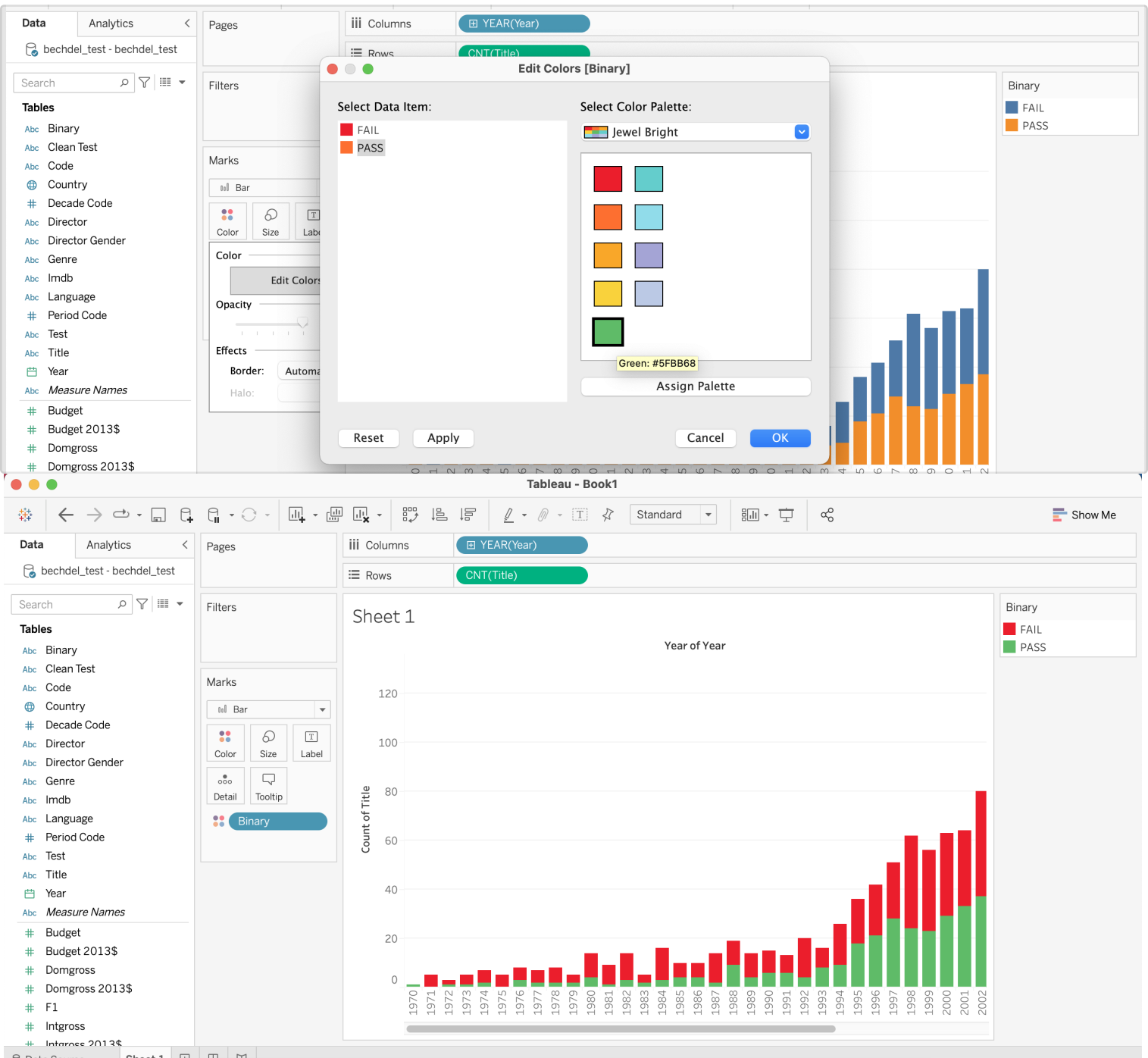




As you can see, Tableau will automatically try to infer the type of each column. For example, the ‘director’ name is a text field (or ‘string’), where as the rating is a numerical value (a rating out of 10). We can see that the country field was also correctly categorised as location data (see the globe icon).



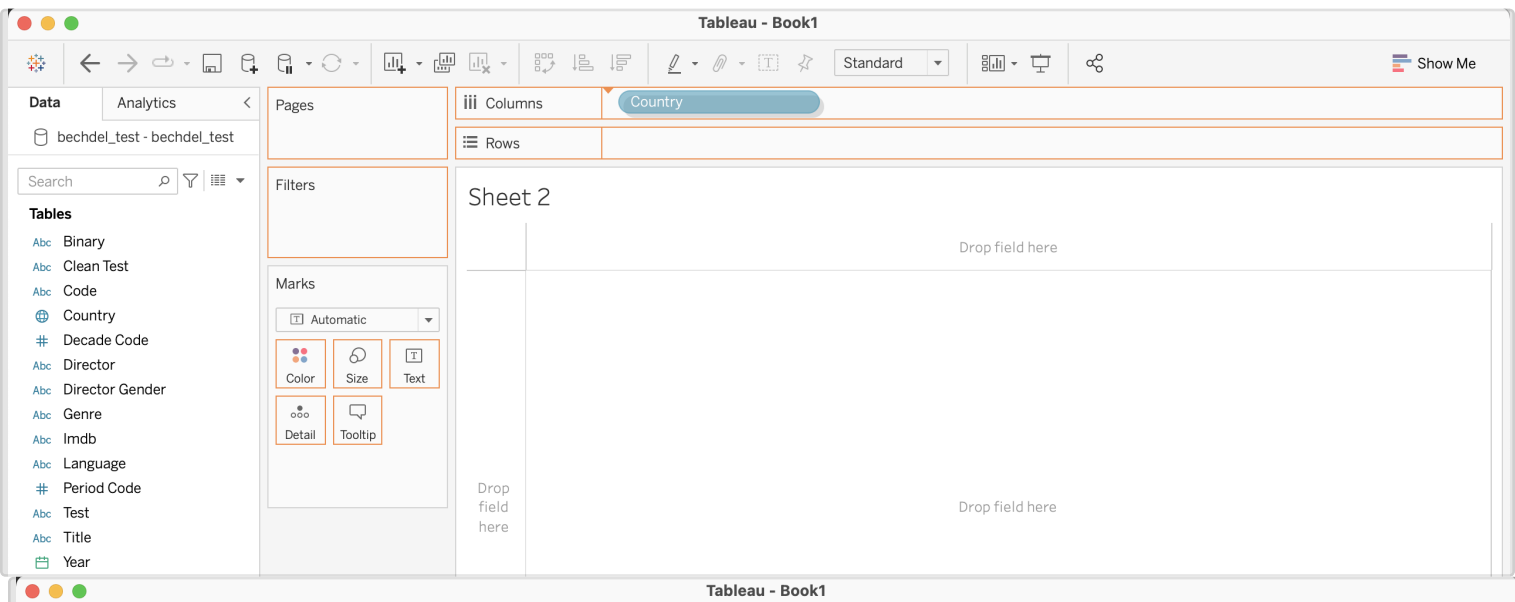
As you can see, Tableau will automatically try to infer the type of each column. For example, the ‘director’ name is a text field (or ‘string’), where as the rating is a numerical value (a rating out of 10). We can see that the country field was also correctly categorised as location data (see the globe icon).

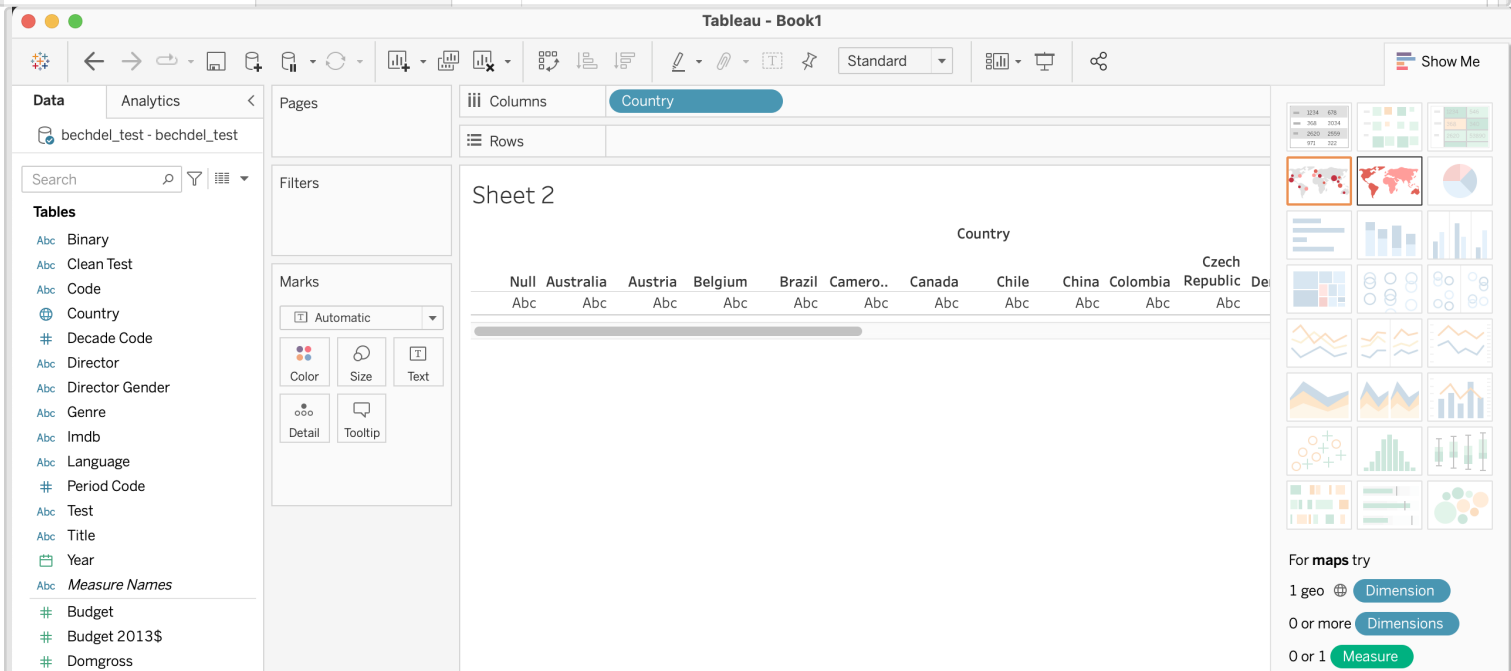


Map Visualisation

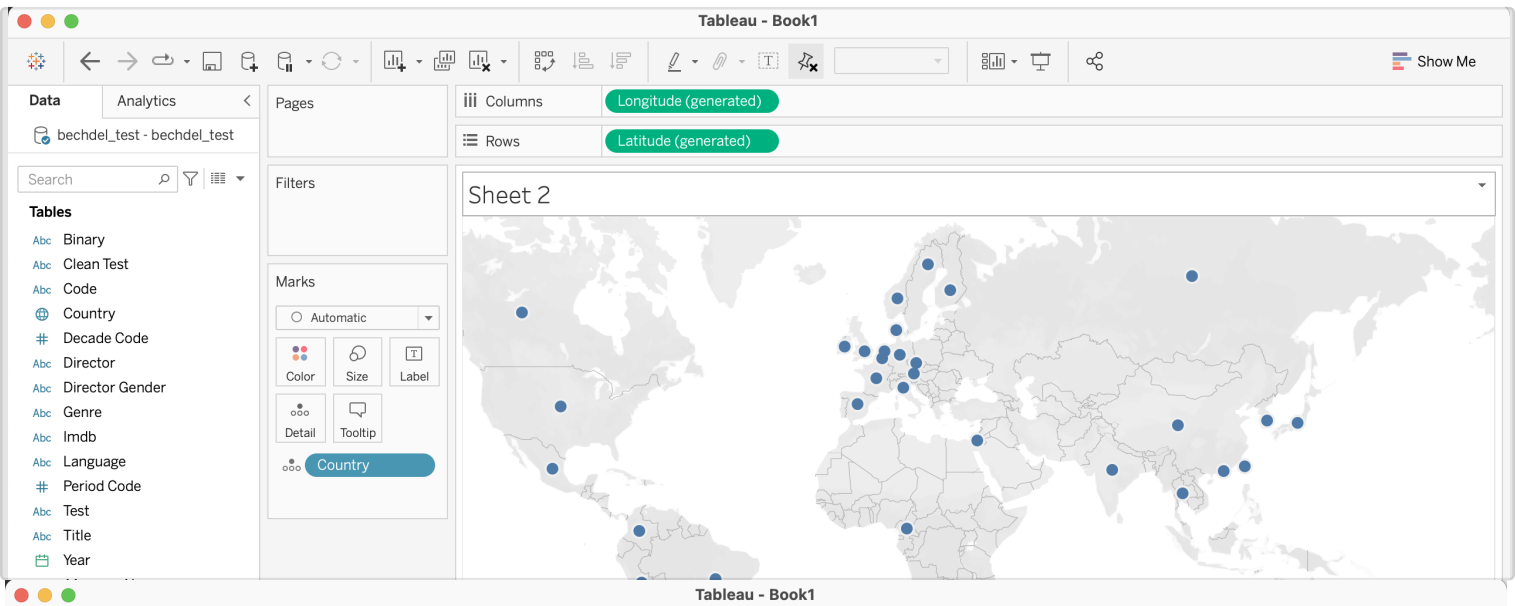
You can easily create a map on Tableau if you have a field that is categorized as location data. In this dataset, we saw that we have a 'Country' field. We will use it to create a map visualisation and explore the spatial attributes in our dataset.

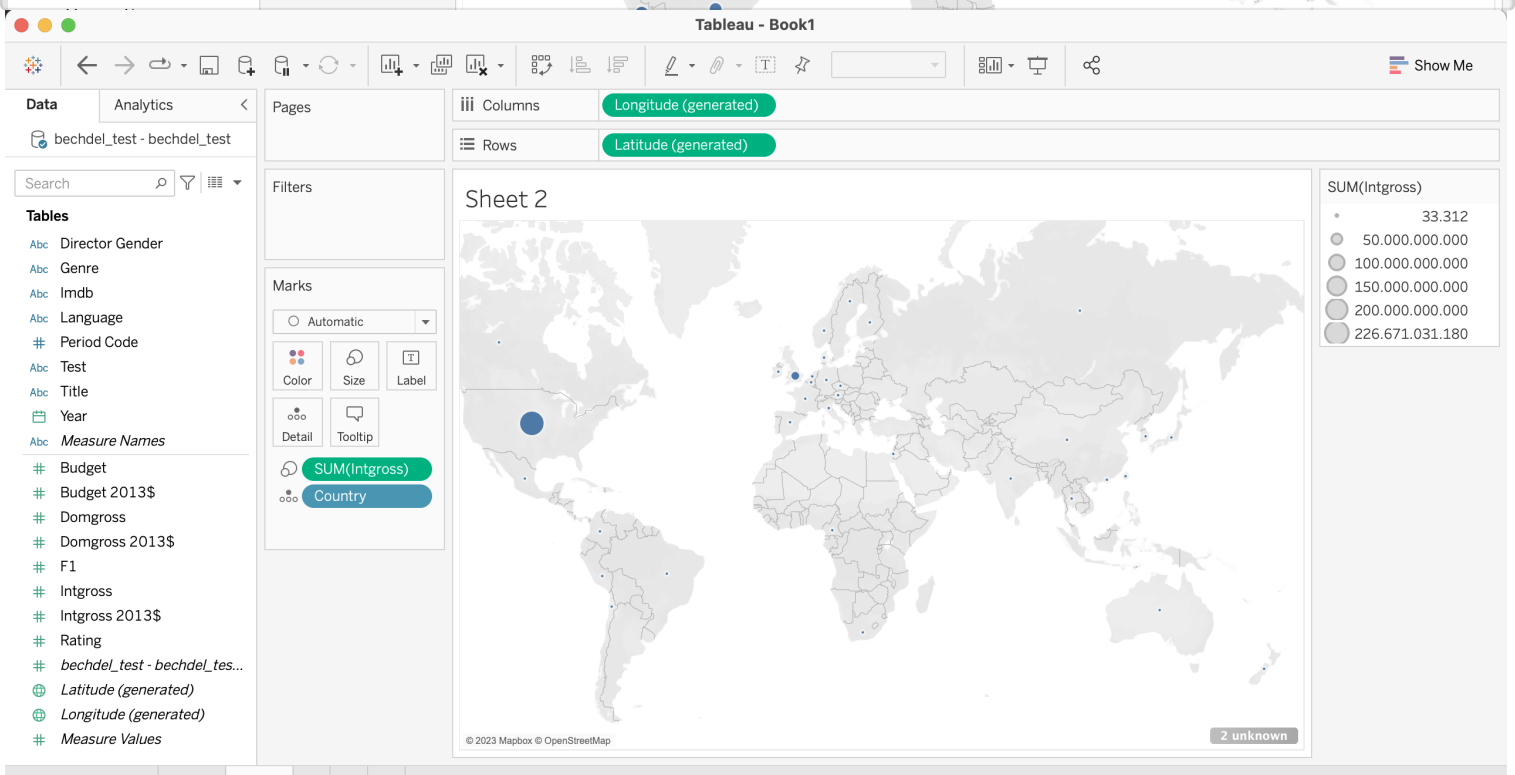
We will click on 'Country' and drag it to the columns area. Tableau will show us a table with all the countries in our dataset. This is not what we want, so we can use the 'Show me' button that has preset visualisation types. We see that we can select from one of two map options. The one framed in orange is the one that Tableau recommends for our data. Let's click it.



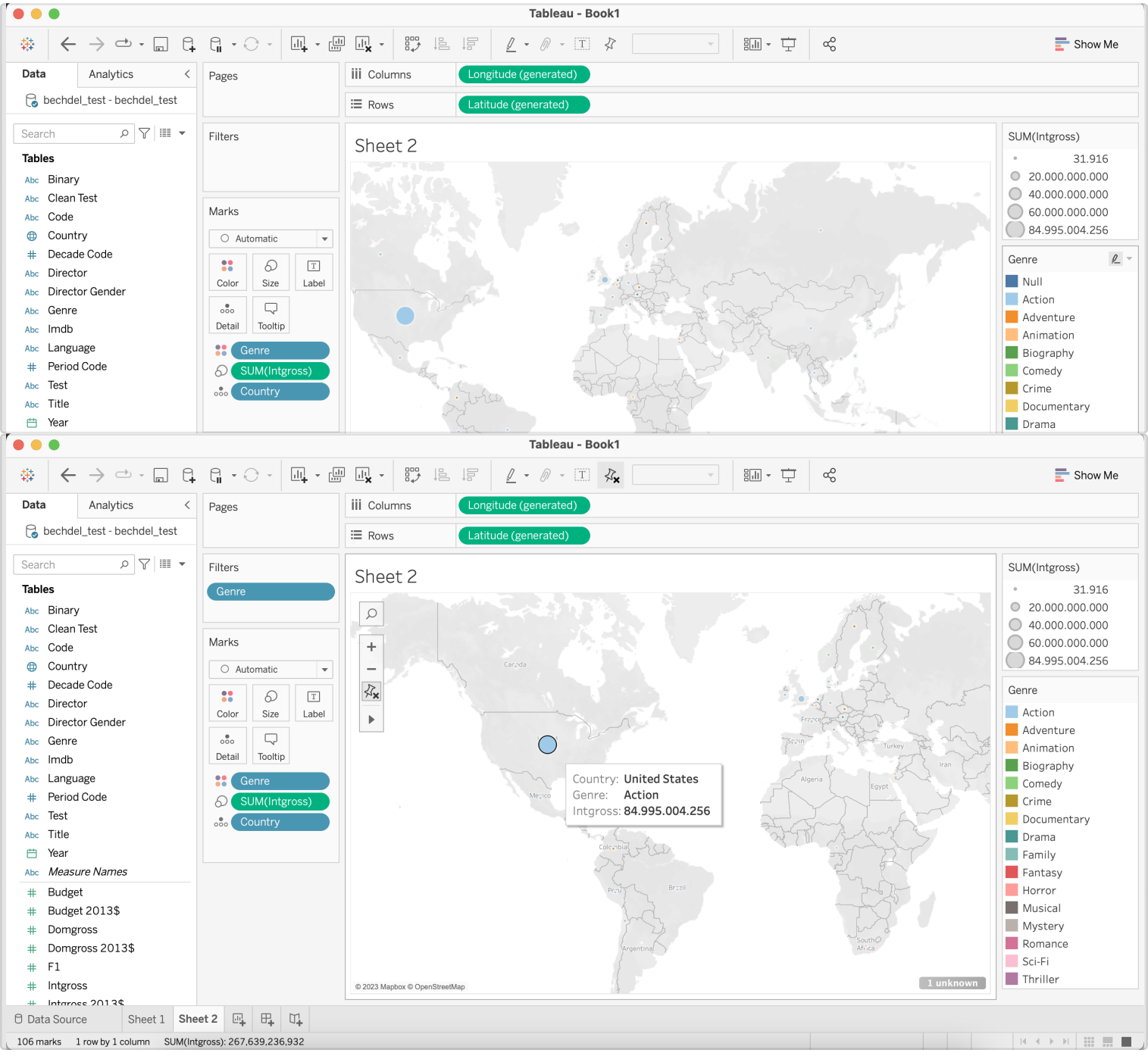


We immediately see a map in the main pane. We will explore the distribution of international gross per country and genre. Let's then pick the 'Intgross' field and drag it to the 'Size' mark as below. When we do that, we ask Tableau to make each dot's size proportional to the value of its international gross. We can see that the largest dot is in the USA, which is not surprising considering the importance of the USA film industry.





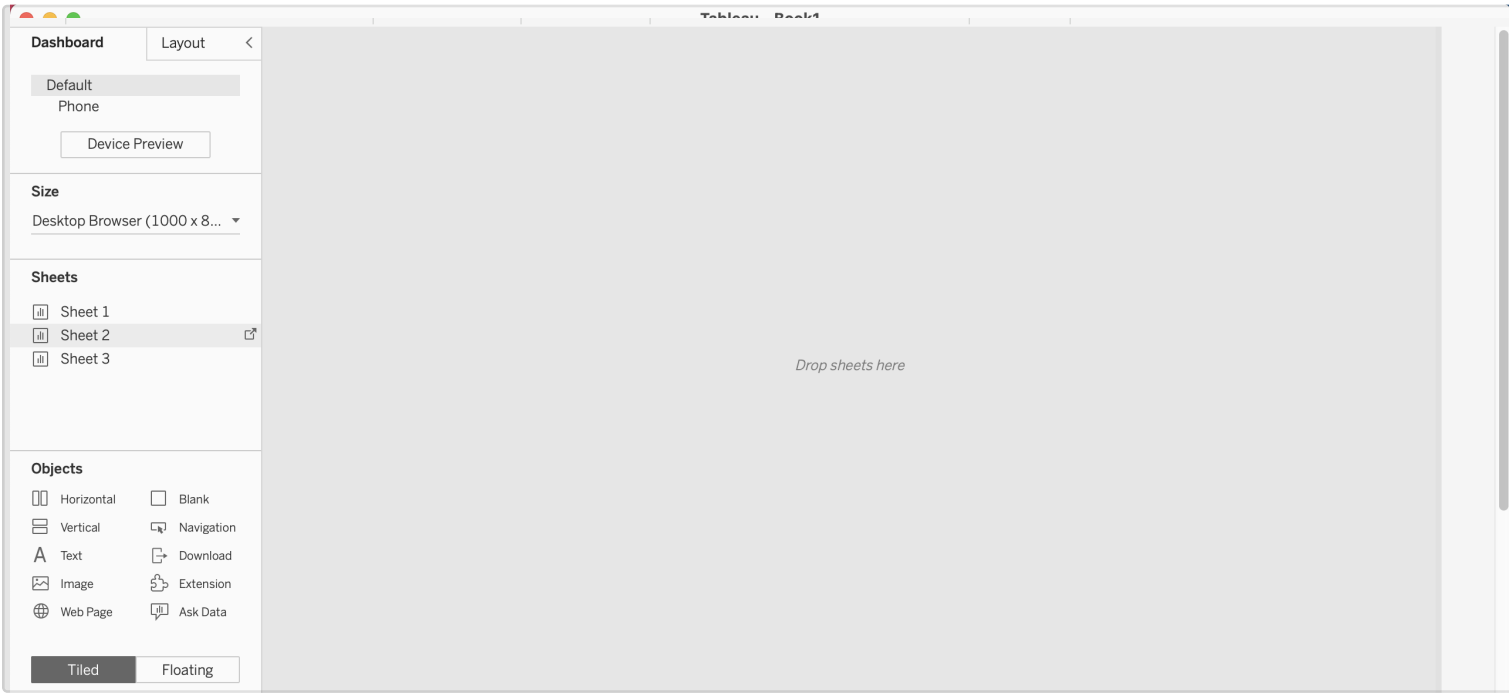
We can add another field to see how these parameters coincide. For instance, if we drag the ‘genre’ field to the ‘Color’ mark, each of the dots will keep their size proportions, and now also be colored depending on the movie genres.



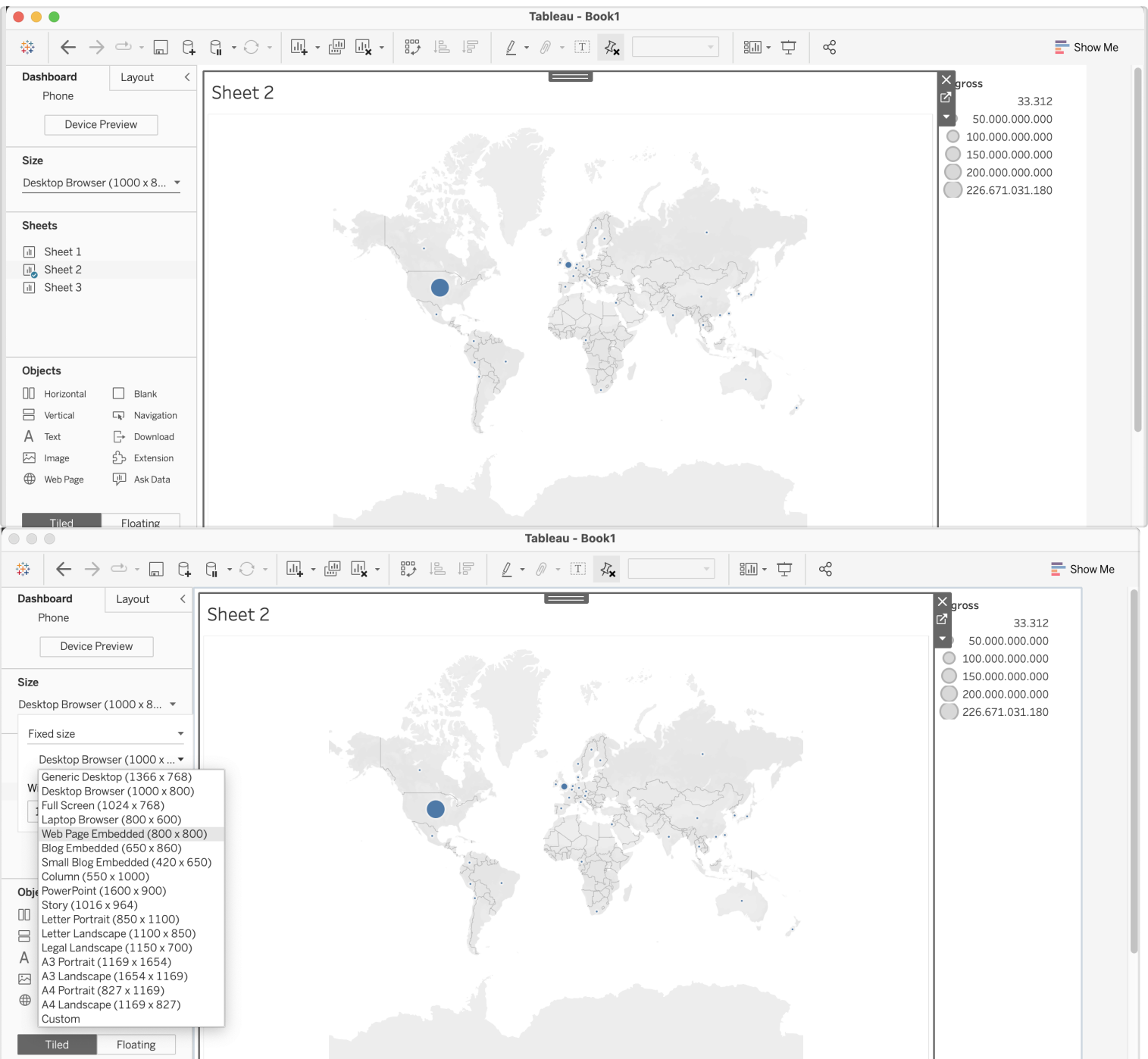
Dashboards & Linked Views

Using this map and the previous bar chart, we can create our custom dashboard that allows us to cross check this data. We do that by first creating a new Dashboard using the button below.

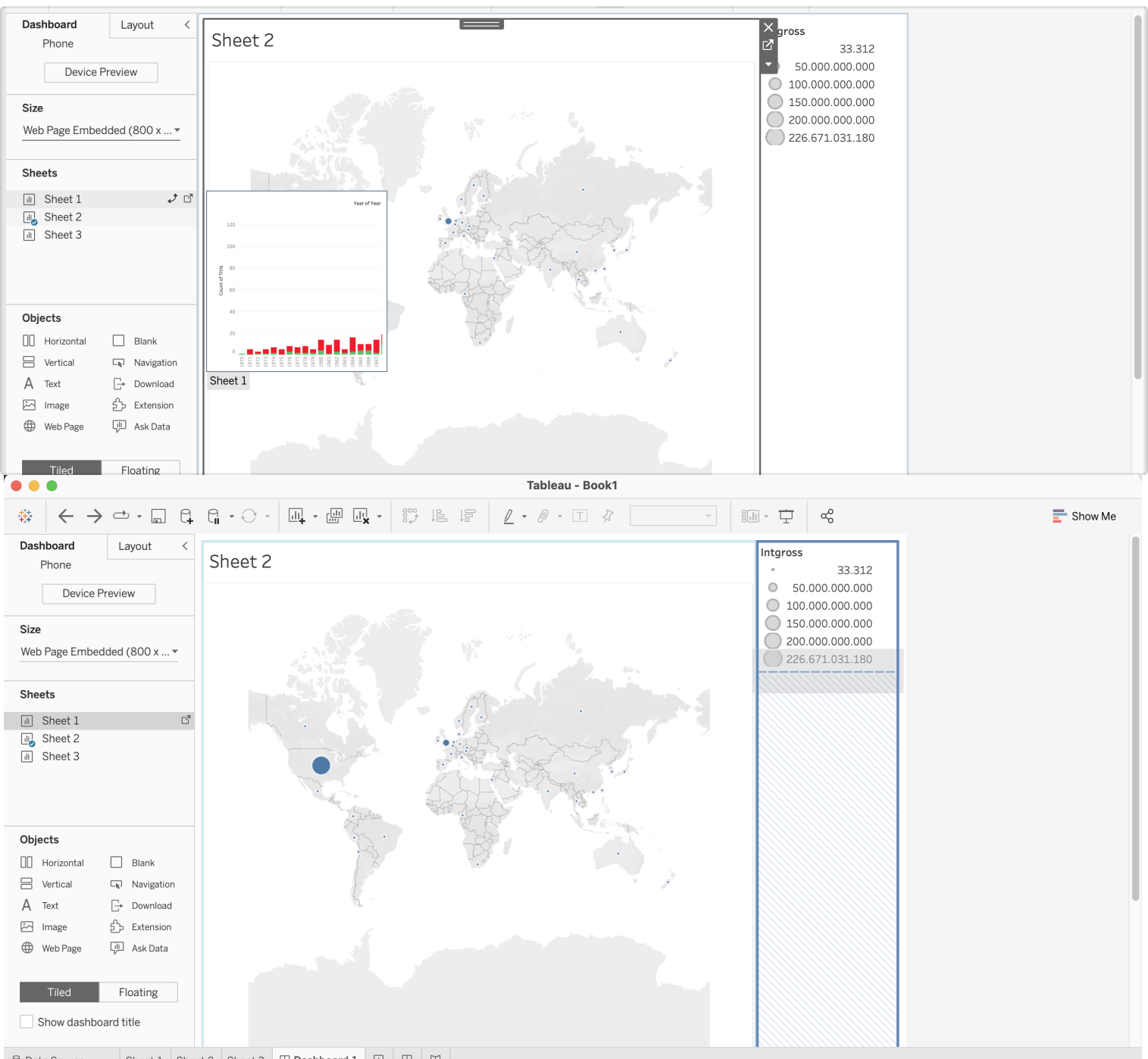
We then can start filling our dashboard. We will do that by dragging ‘Sheet 2’ into the main area.



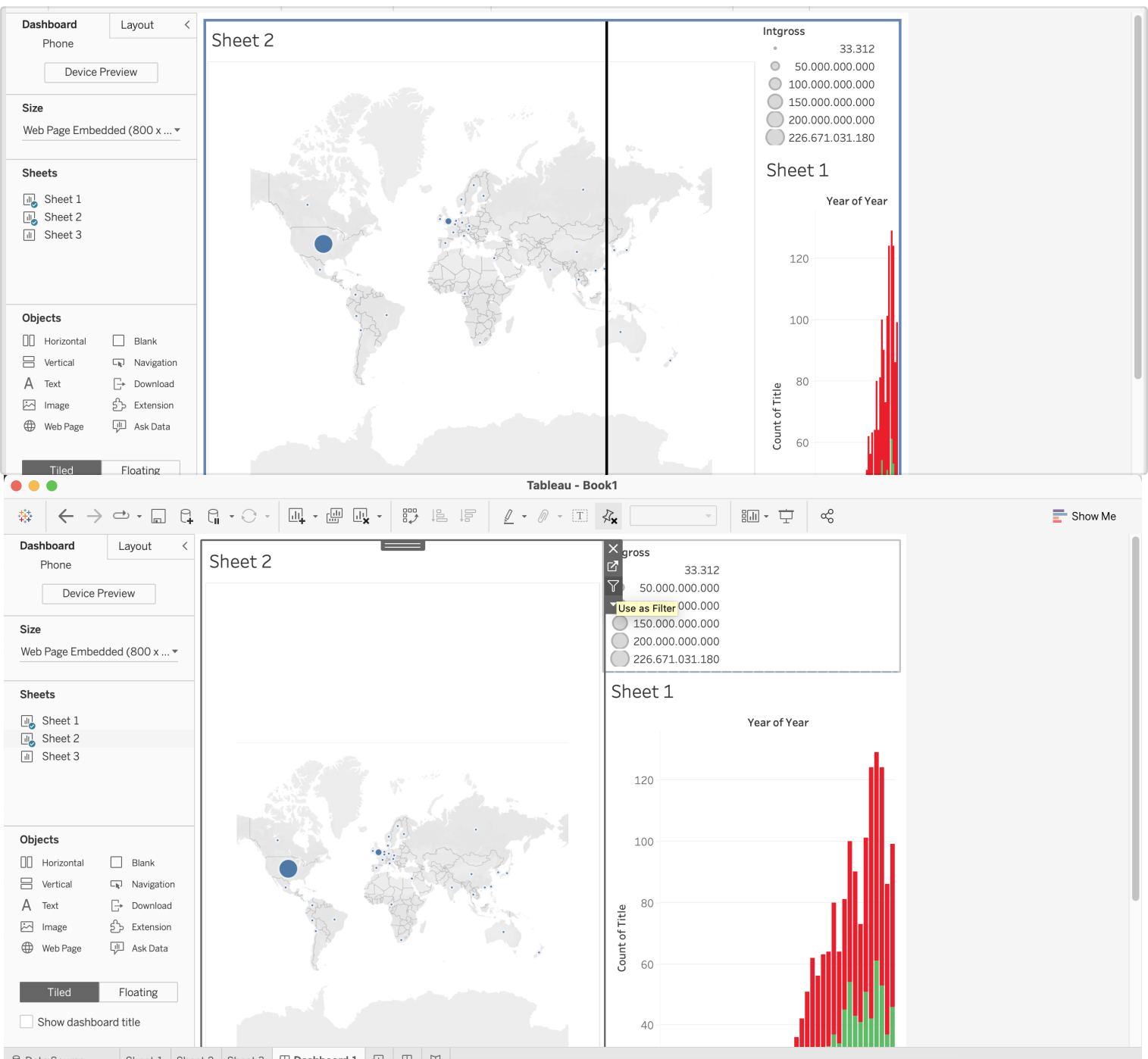
This will bring in our map view and its legend into the center. If the size is too big or too small for your screen, you can fix that by clicking the ‘Size’ button and choosing your preferred option. We can use the option ‘Web Page Embedded’ to make a dashboard that will fit in a standard web page.



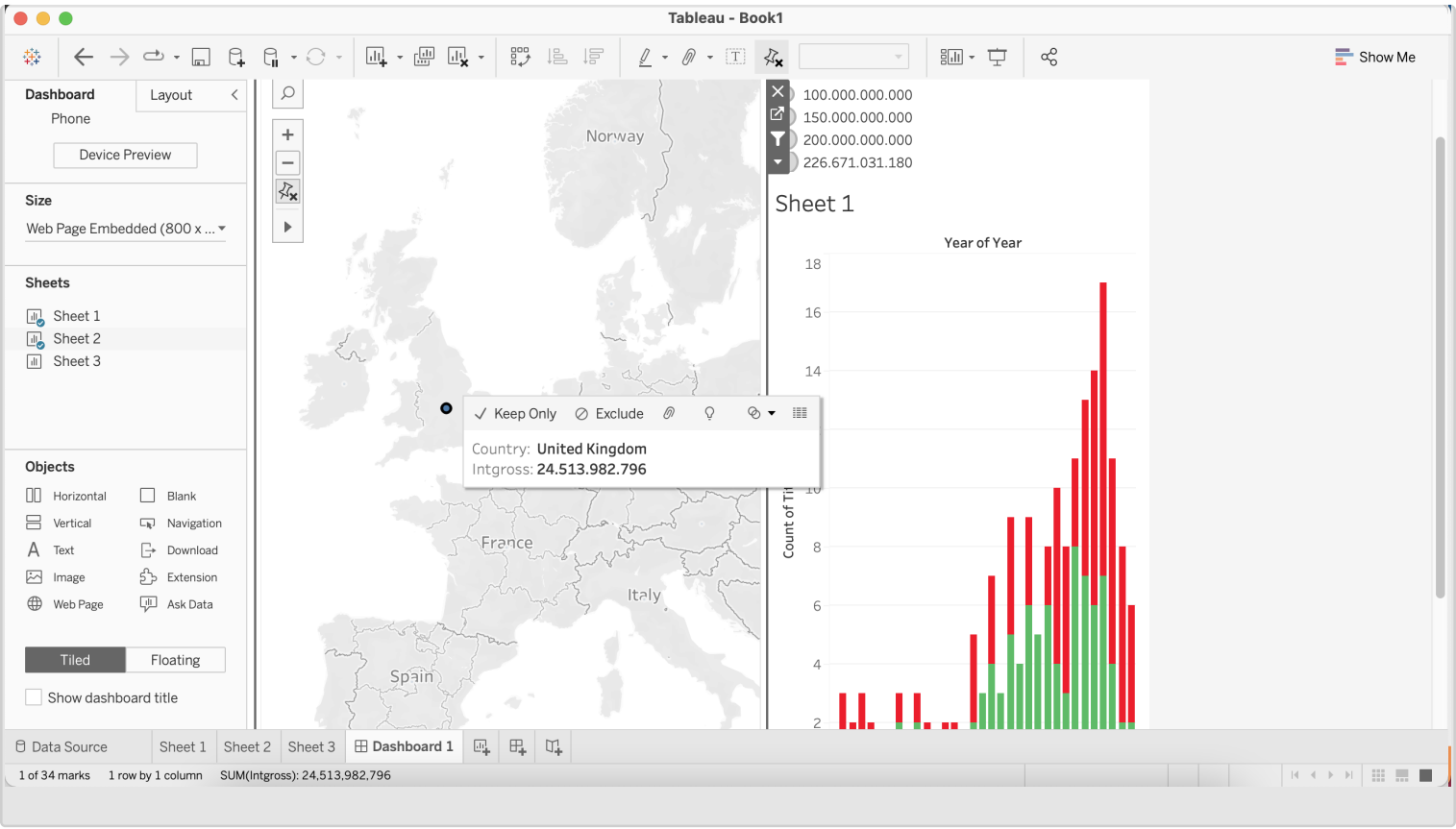
Next, we add in our bar chart by clicking ‘Sheet 1’ and dragging it to the main view. We can see by moving the mouse that Tableau will suggest possible areas to drop it. We can put it in the bottom right corner under the legend.



Again, we can fix the size by resizing some of the components. To give the bar chart more space, we resize the area of the map to be less wide.
Finally, we will connect these two charts by going to the top right of the map component and clicking the 'filter' option that is displayed as a small sieve icon. This will tell Tableau that the map should act as a filter for the rest of the dashboard.

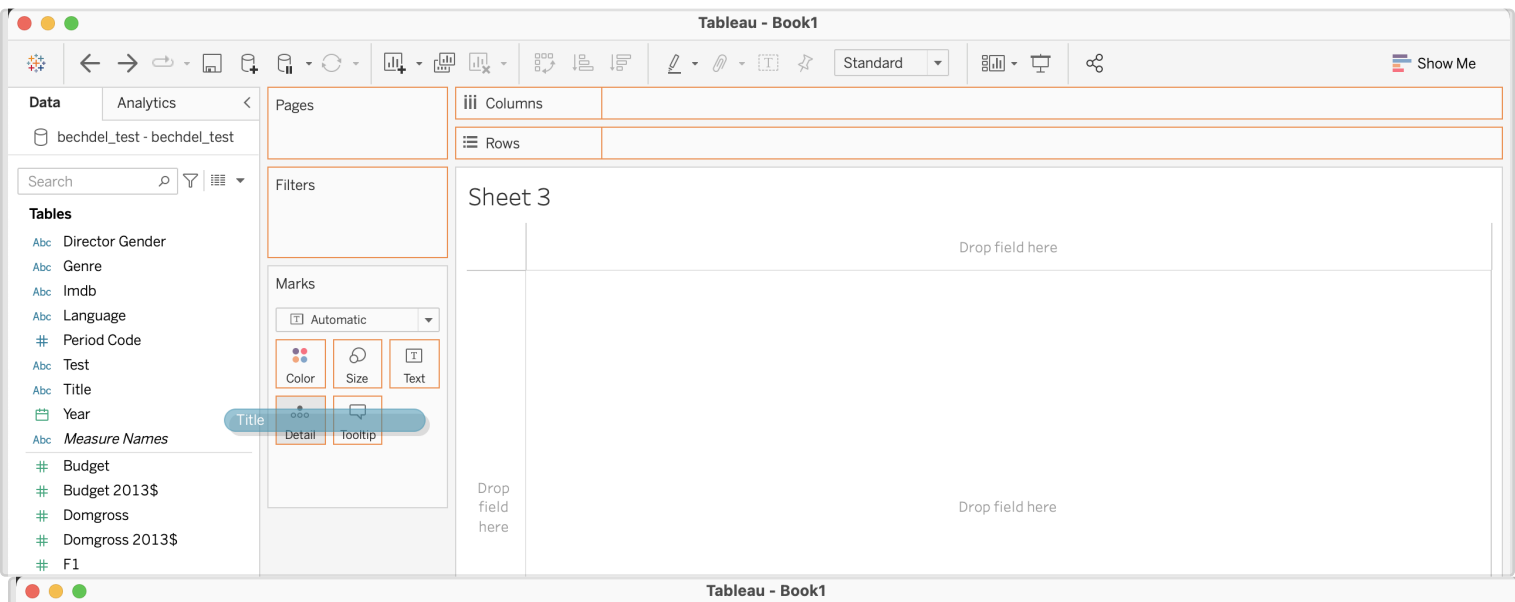


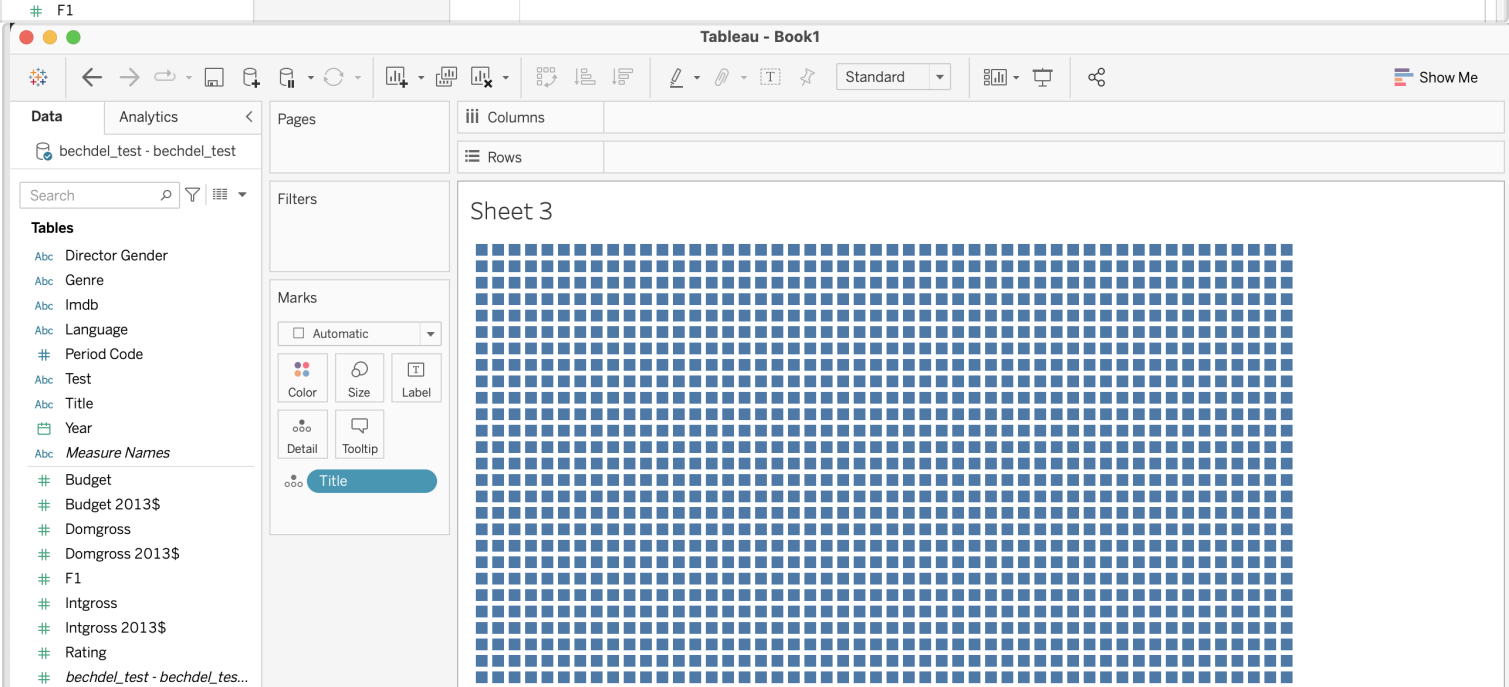
As a result, any interaction with the map will propagate across the dashboard. If we click the dot on the United Kingdom for instance, the bar chart will update to only show the distribution of movies from the UK.



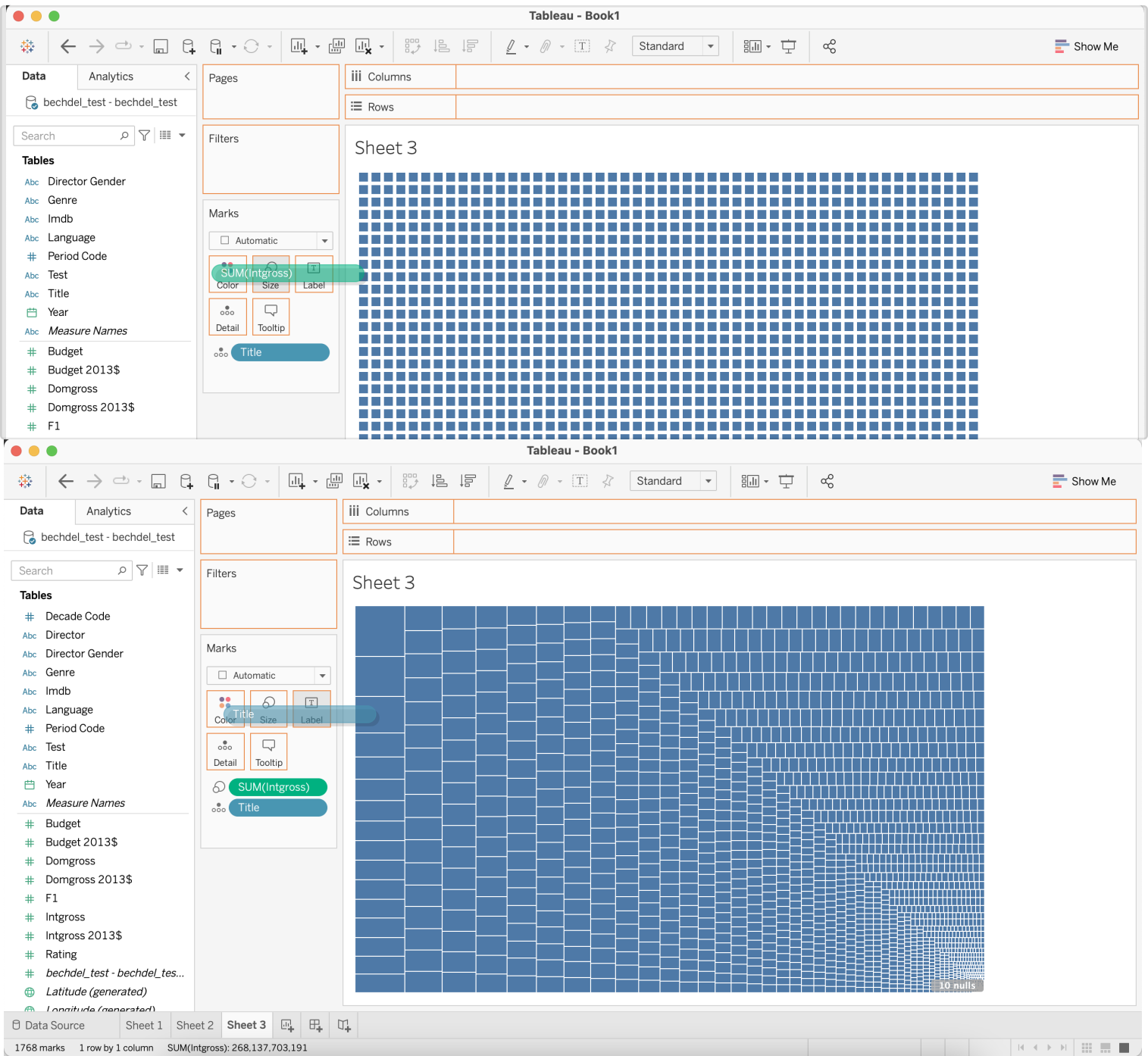
Tree Maps & Word Cloud

Tree maps are visualisations that use size and color to show hierarchical data. We can create one in Tableau to show which movies grossed the most, and which of those are marked as passing the Bechdel Test.
First, let us drag the title in the detail mark.

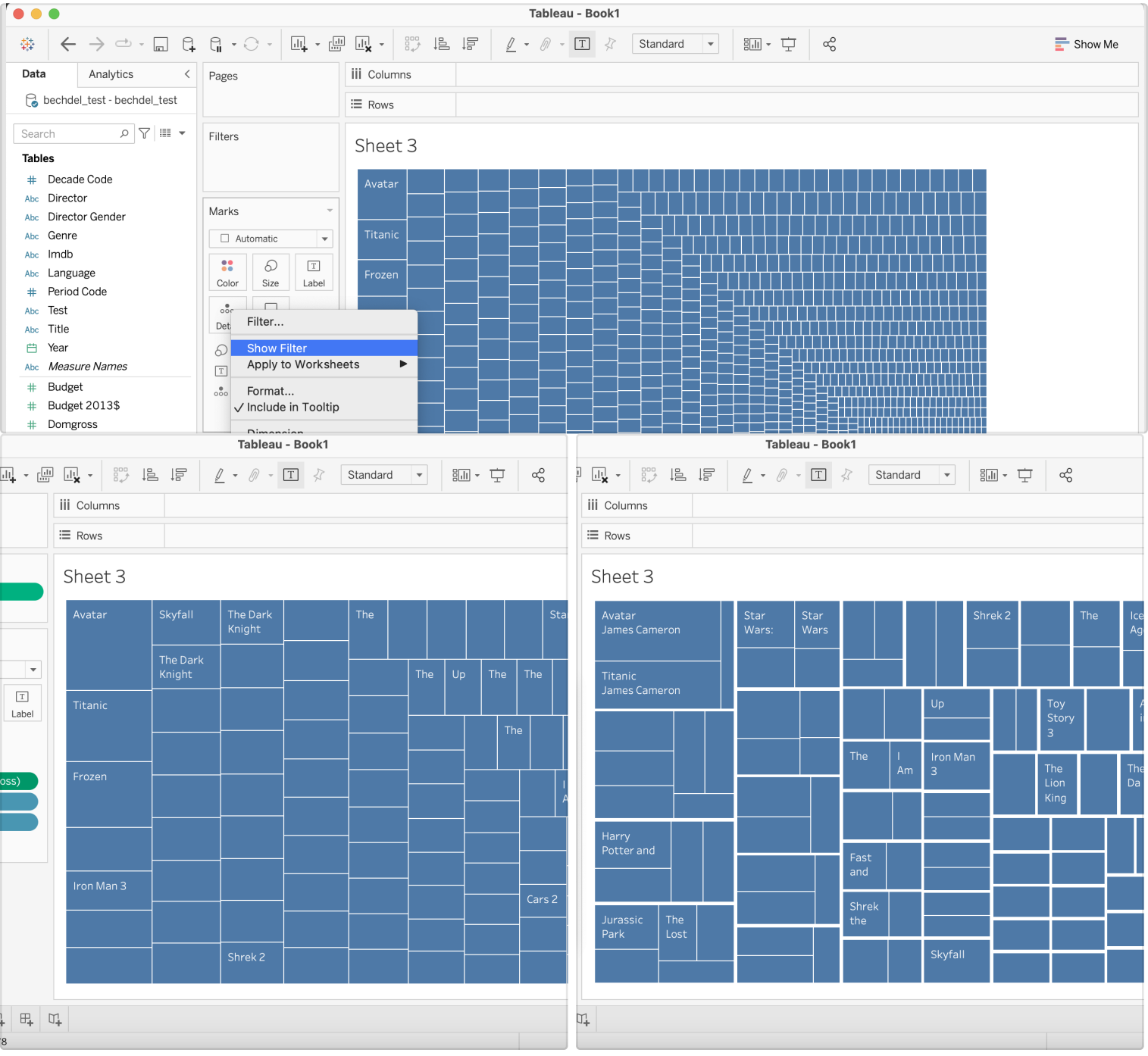




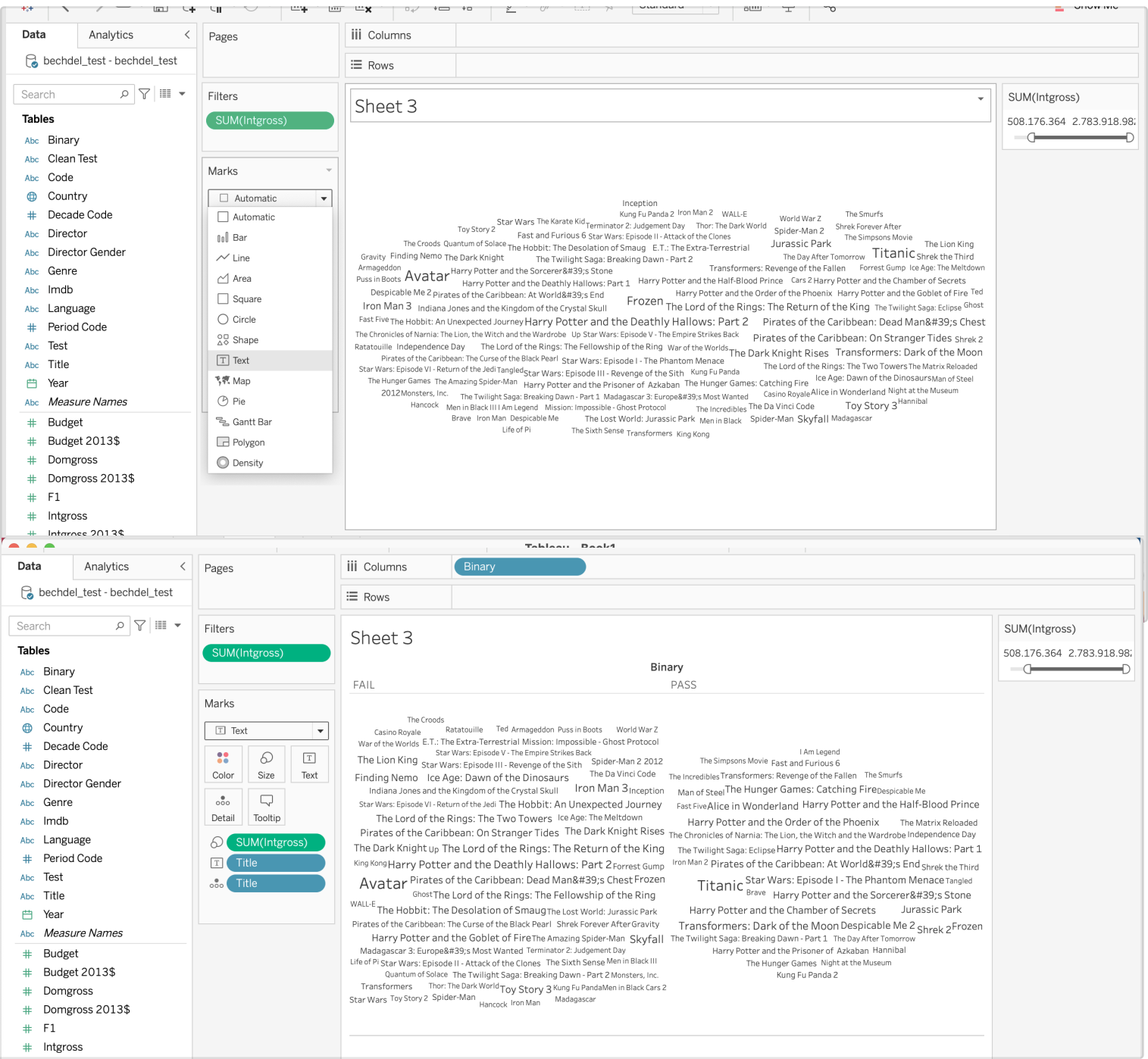
For now, all squares are the same size. We can change that by dragging the ‘Intgross’ field into the ‘size’ mark. We can also drag the ‘title’ field into the ‘label’ mark in order to see which of these squares stands in for which movie.



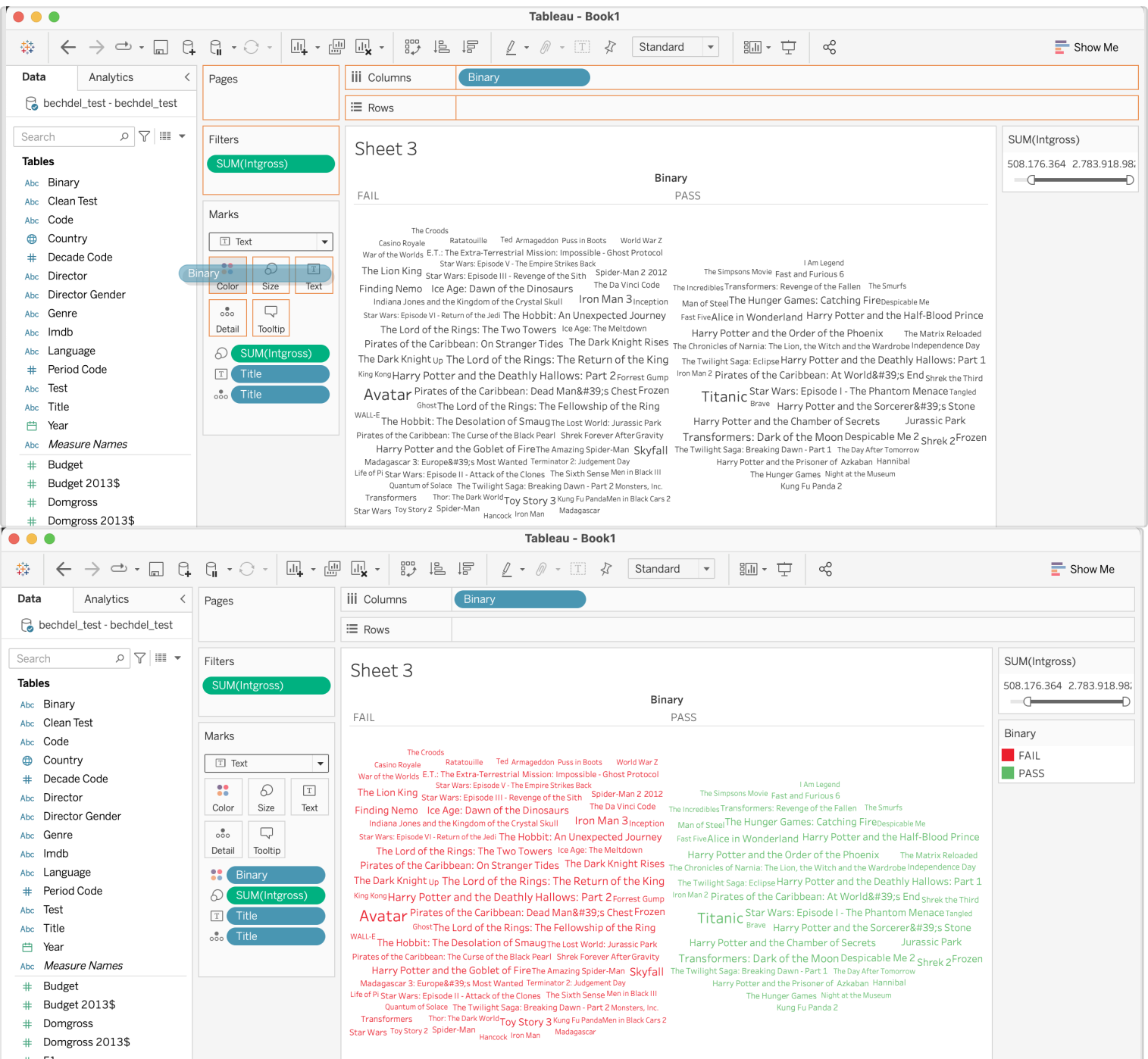
We can filter some of the movies because we want to focus on the highest grossing ones. We click on the inverted triangle on the right of the 'Intgross' pill and pick 'Show Filter' from the menu. This displays a filter on the right of the screen that shows the range of values in that field. We move the left handle to about 500Million. This updates the visualisation to show the subset of movies that grossed a value higher than that. To see more information on the label, we can also add additional fields to the 'label' mark, such as director for instance.



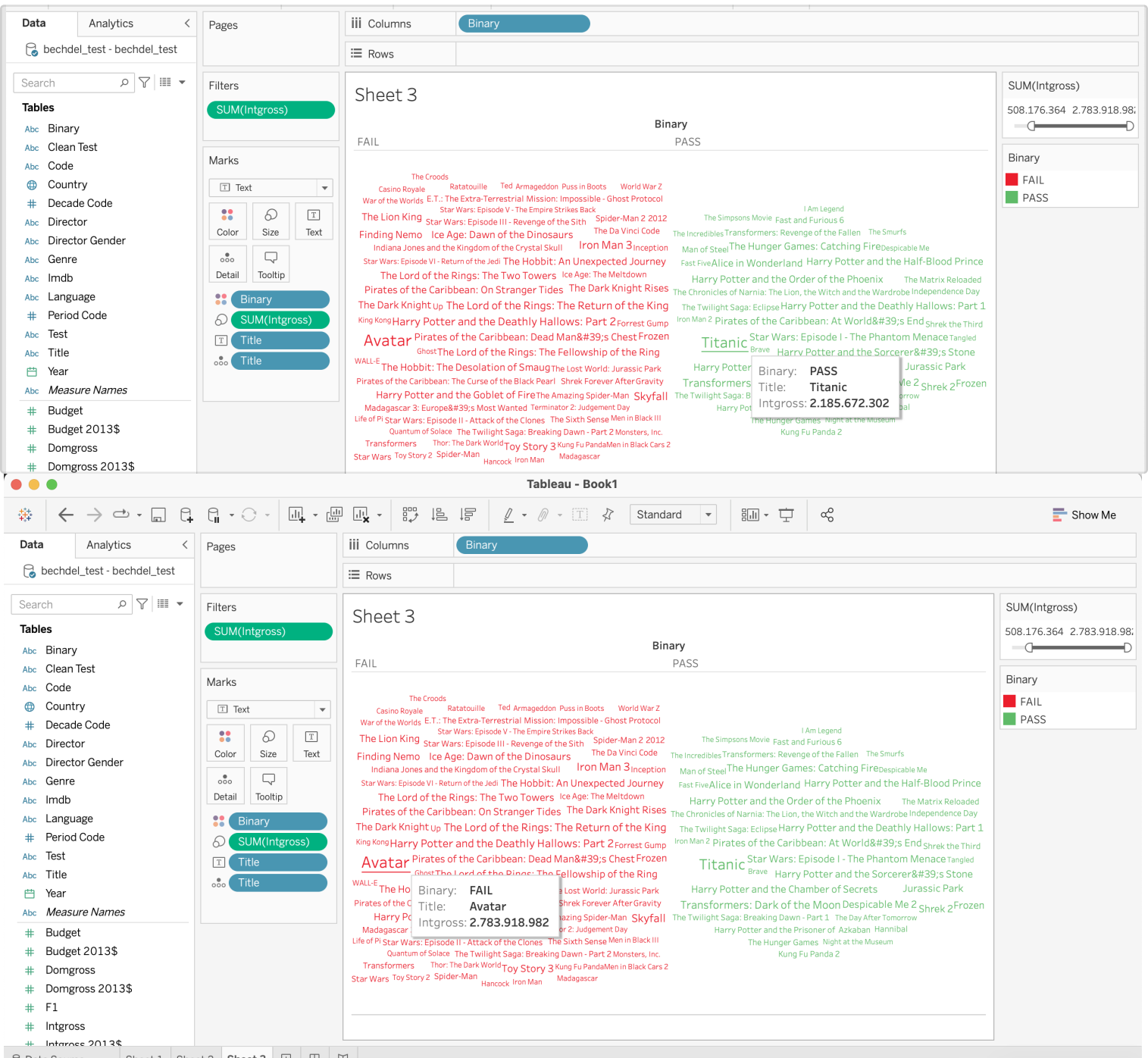
If we want, we can change this treemap into another similar type of visualisation, such as a word cloud. We do so by clicking on the menu in the ‘Marks’ panel, and selecting Text instead of the automatic value. We immediately see a large word cloud that shows the size of the title proportionately to the grossing value.
To explore how these highest grossing movies fare with regards to the Bechdel test, we drag the ‘binary’ field to the columns area.



We can use multiple channels to show the same variable in order to make it clearer. In this case, we drag the binary field to the color mark where we use the same green/red color scheme as before.
This will give us our final word clouds which are categorised by fail/ pass status and colored accordingly.



We can see based on this representation that the highest grossing movie that passes the Bechdel test is the Titanic. If we click on it, we can read more information about it, such as its grossing value. Similarly, the highest grossing movie that fails the Bechdel test is Avatar.



Learn More

In this tutorial, we have seen how to use Tableau to explore a dataset, starting from uploading and formatting, to experimenting with different types of representations and interactions.
If you want to learn more, there are fantastic additional resources on the web you can use. Here are a few favorites:

Miriam Posner's Getting Started with Tableau Public
http://miriamposner.com/classes/dh201w21/tutorials-guides/data-visualization/getting-started-with-tableau-public/

Jennifer Murray's Visualizations for the Digital Humanities Using Tableau
https://digitalcommons.unf.edu/cgi/viewcontent.cgi?article=1078&context=library_facpub

Digital Humanities at Berkeley: Content analysis with Tableau
https://digitalhumanities.berkeley.edu/content-analysis-tableau

In need more technical support? You can use these Tableau resources:

Tableau Community Forum
https://community.tableau.com/s/explore-forums

Official Tableau Tutorials
https://help.tableau.com/current/guides/get-started-tutorial/en-us/get-started-tutorial-home.htm