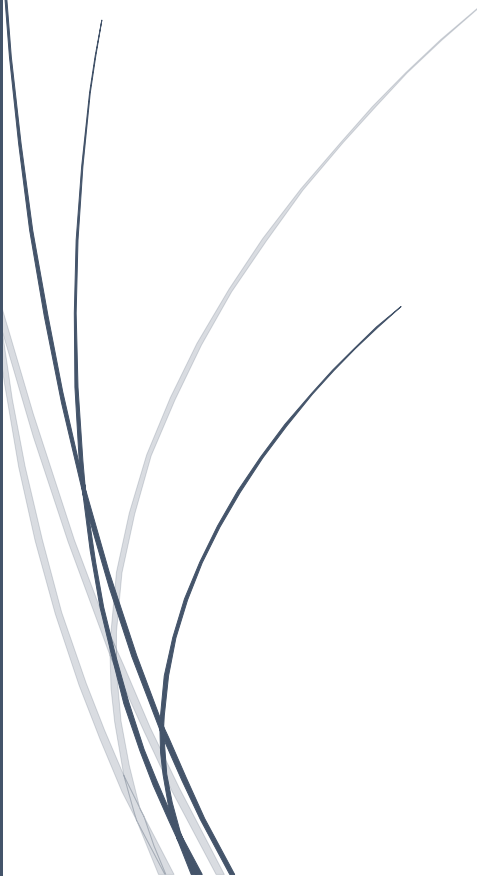


A dark blue vertical bar on the left side of the page. A blue arrow points to the right from the bar, containing the date.

6/11/2023

# BÁO CÁO LAB 03

Trực quan hóa dữ liệu

Several thin, curved lines in dark blue and light grey originate from the bottom left corner and curve upwards and to the right.

[20120268@student.hcmus.edu.vn](mailto:20120268@student.hcmus.edu.vn)  
ĐẠI HỌC KHOA HỌC TỰ NHIÊN

## Mục lục

I. Thành viên nhóm .....	2
II. Thông tin về bài làm.....	2
1. Công cụ.....	2
2. Dataset .....	2
III. Đánh giá mức độ hoàn thành.....	2
IV. Giới thiệu về Tableau .....	3
1. Tableau là gì?.....	3
2. Các sản phẩm Tableau .....	3
2.1. Tableau Desktop .....	3
2.2. Tableau Server .....	3
2.3. Tableau Online.....	4
2.4. Tableau Public .....	4
2.5. Tableau Reader .....	4
2.6. Tableau Mobile .....	5
3. Ưu điểm và nhược điểm của Tableau.....	5
3.1. Ưu điểm .....	5
3.2. Nhược điểm .....	6
V. Các tính năng mà Tableau hỗ trợ .....	7
1. Kết nối đa nguồn dữ liệu .....	7
1.1. Kết nối với các cơ sở dữ liệu SQL phổ biến (MySQL, SQL Server, Oracle hay Postgresql) .....	7
1.2. Kết nối với Google Drive .....	10
1.3. Kết nối với các tập tin Excel, CSV, JSON, TEXT và nhiều định dạng tập tin khác .....	10
1.4. Dịch vụ web.....	10
2. Trực quan dữ liệu.....	10
2.1. Biểu đồ.....	10
2.2. Calculated Field (trường tính toán) .....	16
2.3. Parameter .....	19

## I. Thành viên nhóm

Họ tên	MSSV	Mức độ hoàn thành
Hoàng Mạnh Cường	20120261	100%
Phan Dương Quốc Đạt	20120268	100%
Lâm Quốc Chung	20120258	100%
Phùng Anh Khoa	19120548	100%

## II. Thông tin về bài làm

### 1. Công cụ

Tableau Desktop Professional 2023

### 2. Dataset

Nguồn: <https://www.kaggle.com/datasets/jahaidulislam/world-happiness-report-2005-2021>

Dataset mà tụi em sử dụng để trực quan đã được lọc ra, chỉ lấy số liệu từ năm 2015-2021. Thầy có thể xem rõ hơn ở file Dataset.xlsx được đính kèm trong thư mục **docs**

## III. Đánh giá mức độ hoàn thành

- Giới thiệu về Tableau: 100%
- Giới thiệu về các tính năng mà Tableau hỗ trợ: 100%
- Áp dụng Tableau để trực quan bộ dữ liệu đã chọn: 100%

## IV. Giới thiệu về Tableau

### 1. Tableau là gì?



Tableau là một phần mềm phân tích, trực quan hóa dữ liệu mạnh mẽ.

Tableau cho phép người dùng kết nối, tiền xử lý, phân tích và biểu diễn dữ liệu từ nhiều nguồn khác nhau. Ngoài ra, người dùng có thể dùng Tableau để tạo ra các biểu đồ, đồ thị, báo cáo có thể tương tác được, bản đồ địa lý, dashboard,... để khám phá và chia sẻ thông tin.

Tableau được sử dụng rộng rãi trong các lĩnh vực như kinh doanh, tài chính, tiếp thị, y tế, giáo dục và nhiều ngành công nghiệp khác.

Ví dụ: nếu bạn có bất kỳ dữ liệu nào như file excel, Big Data, SQL hoặc dữ liệu đám mây, và bạn muốn phân tích dữ liệu đó dưới dạng biểu đồ, bạn có thể sử dụng Tableau.

### 2. Các sản phẩm Tableau

#### 2.1. Tableau Desktop

Tableau Desktop là một ứng dụng giúp bạn khám phá gần như bất kỳ loại dữ liệu có cấu trúc nào và tạo ra các biểu đồ, dashboard, báo cáo có thể tương tác được. Bạn có thể kết nối với gần như bất kỳ nguồn dữ liệu nào từ bảng tính đến data warehouse và hiển thị thông tin dưới nhiều góc nhìn đồ họa khác nhau. Tableau được thiết kế để sử dụng một cách dễ dàng nên bạn sẽ làm việc nhanh hơn bao giờ hết.

#### 2.2. Tableau Server

Tableau Server là một ứng dụng BI (Business Intelligence) hoạt động thông qua trình duyệt mà bất kỳ ai cũng có thể sử dụng. Đây là một giải pháp thay thế những phần mềm BI truyền thống, vốn chạy rất chậm chạp.

Mọi người chia sẻ, quản lý dữ liệu dựa trên điện toán đám mây và cần trả phí cho không gian lưu trữ trên cloud.

Tableau Server thường là nơi chia sẻ các báo cáo phân tích của doanh nghiệp, tuy nhiên những dữ liệu trên cloud được bảo mật cẩn thận và cấp quyền nếu muốn truy cập. Độ an toàn cao trong Tableau Server khiến nó phù hợp cho việc chia sẻ dữ liệu trong một tổ chức hay doanh nghiệp

### 2.3. Tableau Online

Đây là một công cụ chia sẻ trực tuyến của Tableau. Các chức năng của nó tương tự như Tableau Server, nhưng dữ liệu được lưu trữ trên cơ sở dữ liệu đám mây, như Amazon Redshift và Google BigQuery. Tableau Online có thể cho phép liên kết đến hơn 40 nguồn dữ liệu khác nhau.

### 2.4. Tableau Public

Khi sử dụng Tableau Public, các dữ liệu phân tích sẽ không được lưu trữ trên máy tính cá nhân mà được đăng trên cloud công cộng của Tableau, mọi người dùng đều có quyền truy cập. Chính vì vậy khả năng bảo mật của Tableau Public không cao, chỉ phù hợp với sinh viên trong quá trình học tập, trao đổi.

### 2.5. Tableau Reader

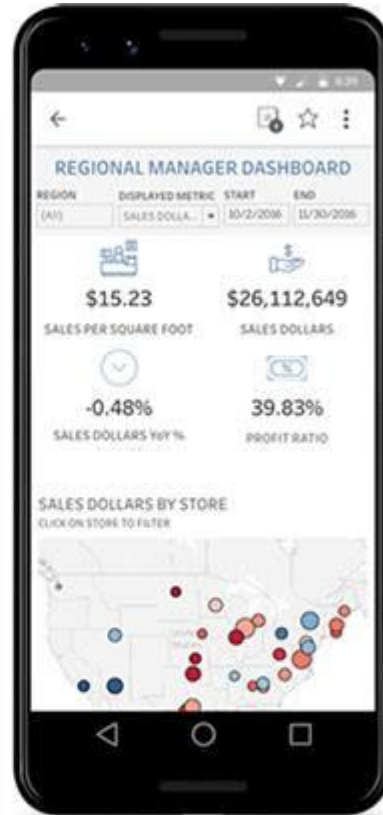
Phiên bản này dành riêng cho người có nhu cầu xem báo cáo mà không can thiệp trực tiếp vào quá trình phân tích, xử lý dữ liệu. Tableau Reader chỉ bao gồm các tính năng đơn giản. Tuy nhiên, chúng vẫn hiển thị đầy đủ tất cả định dạng báo cáo.

## 2.6. Tableau Mobile



**Desktop Layout**

Desktop layouts can be small and hard to read on a mobile device



**Phone Layout (same dashboard)**

Phone layouts use the full width of the screen and are easier to read

Tableau Mobile cung cấp các bản preview có thể tương tác được. Người dùng có thể kiểm tra dữ liệu, xem các dashboard bất cứ lúc nào từ bất kỳ nơi đâu. Ngoài ra, người dùng có thể lựa chọn, lọc và điều chỉnh dữ liệu thông qua ứng dụng di động.

## 3. Ưu điểm và nhược điểm của Tableau

### 3.1. Ưu điểm

- Điểm mạnh nhất của Tableau là tốc độ của nó. Tableau có thể duyệt qua hàng triệu hàng và phản hồi gần như ngay lập tức, qua đó góp phần cải thiện năng suất làm việc của người dùng.
- Tableau rất dễ sử dụng. Người dùng có thể bắt đầu sử dụng Tableau mà không cần có kinh nghiệm lập trình. Chỉ với kiến thức cơ bản về Microsoft Excel, mọi người có thể dễ dàng học Tableau.
- Dashboard của Tableau có thể tương tác được, do Tableau cung cấp kết quả động (dynamic outcome). Có thể tạo ra các hiển thị hình ảnh phong phú một cách dễ dàng.

dàng. Đồ họa và biểu đồ thông minh và đẹp mắt. Hình ảnh, trang web và tài liệu có thể được thêm vào bảng điều khiển tương tác để kể chuyện dễ dàng. Tất cả những điều này giúp tăng tính tương tác của dữ liệu.

- Tableau có sẵn nhiều dạng thông tin như tên các địa danh, mã bưu chính,... Những thứ này hỗ trợ rất tốt cho việc thể hiện thông tin chi tiết và chính xác trên bản đồ.
- Tableau cho phép người dùng kết nối trực tiếp với cơ sở dữ liệu, data warehouse,... Việc truy cập dữ liệu rất đơn giản, và dữ liệu được cập nhật trực tiếp mà không cần phải thực hiện bất kỳ thao tác nào.
- Hệ thống phân quyền và xác thực có sẵn giúp Tableau giảm thiểu được nguy cơ mất mát dữ liệu.
- Sau khi làm việc với dữ liệu, các kết quả có thể được chia sẻ trực tiếp chỉ với một vài thao tác đơn giản. Dashboard có thể được publish hoặc chia sẻ trực tiếp thông qua web và các thiết bị di động.
- Tableau còn khá mới mẻ trong thị trường BI nhưng thị phần của nó đang tăng lên từng ngày. Nó được sử dụng rộng rãi trong hầu hết các ngành công nghiệp, từ giao thông vận tải đến chăm sóc sức khỏe. Tableau cũng có một cộng đồng người dùng rất lớn, bao gồm cả các công ty lớn. Họ sử dụng Tableau hàng ngày để thu thập thông tin kinh doanh của mình.

### 3.2. Nhược điểm

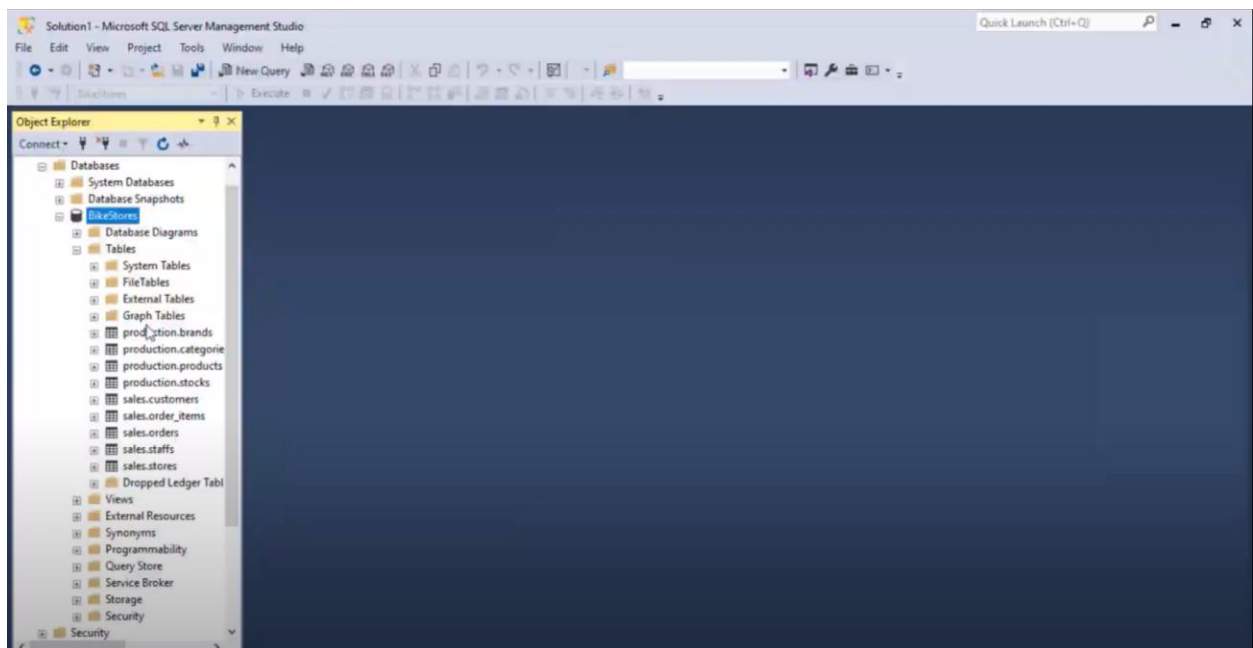
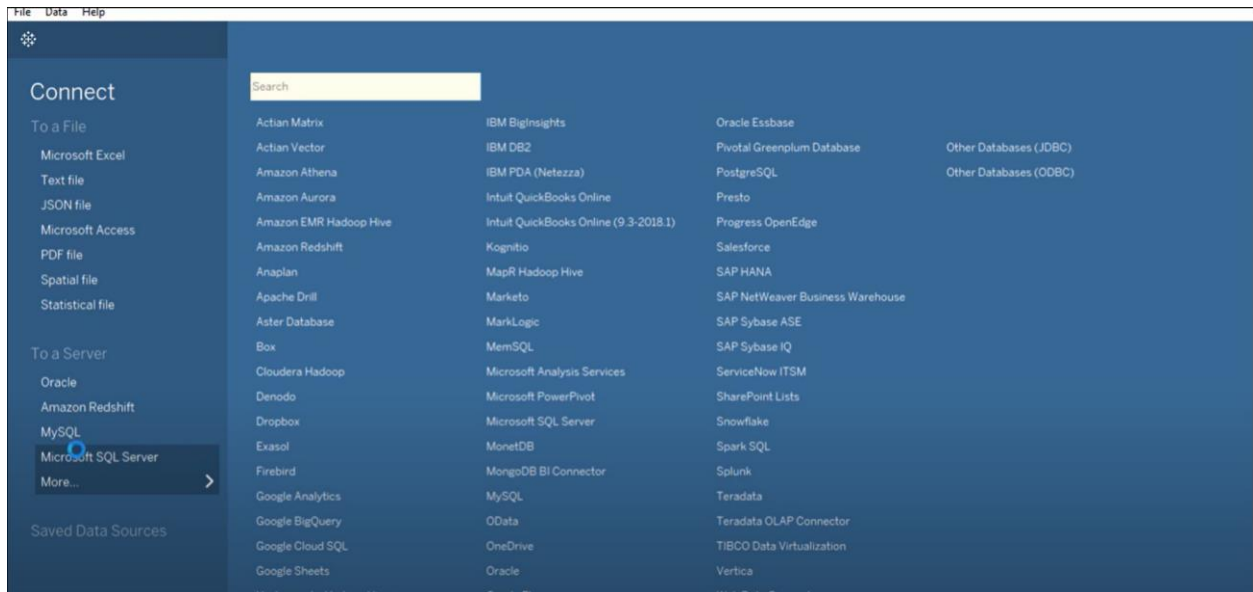
- Tableau có mức giá khá cao, đặc biệt là các phiên bản cao cấp như Tableau Desktop Professional và Tableau Server. Điều này có thể là rào cản đối với các cá nhân hoặc doanh nghiệp có nguồn lực hạn chế.
- Tableau không cung cấp tính năng tự động tạo schedule reports, thứ rất quan trọng đối với các doanh nghiệp.
- Mặc dù Tableau có giao diện người dùng thân thiện, nhưng để tận dụng được tất cả các tính năng và khả năng của nó, người dùng cần có kiến thức và kỹ năng phân tích dữ liệu.
- Tableau có thể gặp khó khăn khi kết nối với các nguồn dữ liệu không cấu trúc hoặc có cấu trúc không chuẩn.
- Một số người dùng cho rằng Tableau có hạn chế về tùy chỉnh so với một số công cụ BI khác. Việc tạo các tùy chỉnh phức tạp và đáp ứng các yêu cầu đặc thù có thể gặp khó khăn.

## V. Các tính năng mà Tableau hỗ trợ

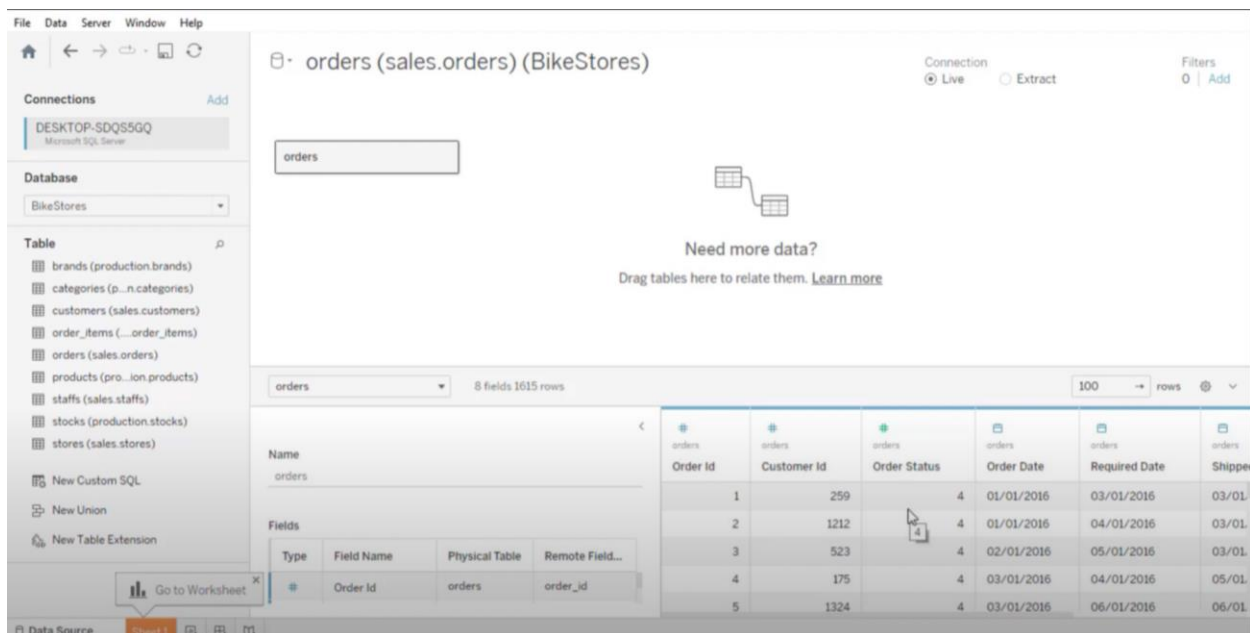
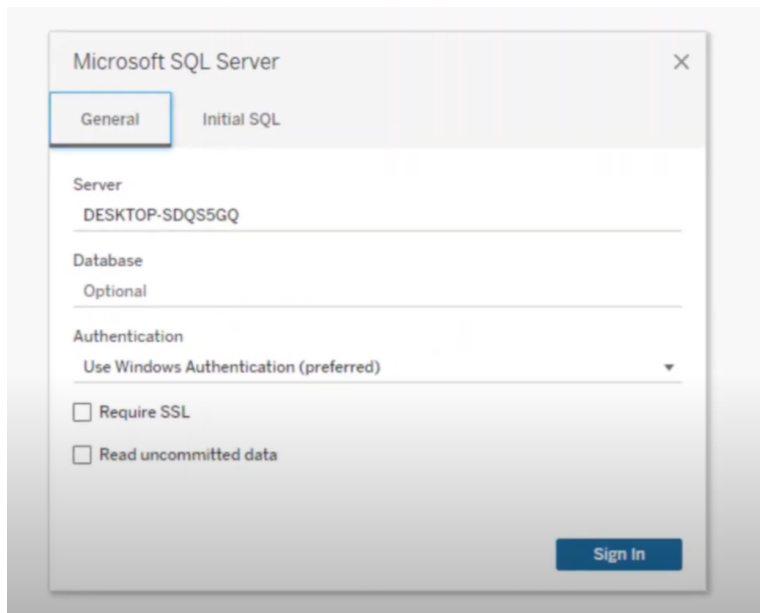
### 1. Kết nối đa nguồn dữ liệu

1.1. Kết nối với các cơ sở dữ liệu SQL phổ biến (MySQL, SQL Server, Oracle hay PostgreSQL)

- Có thể truy vấn và lấy dữ liệu ra từ table trong Database



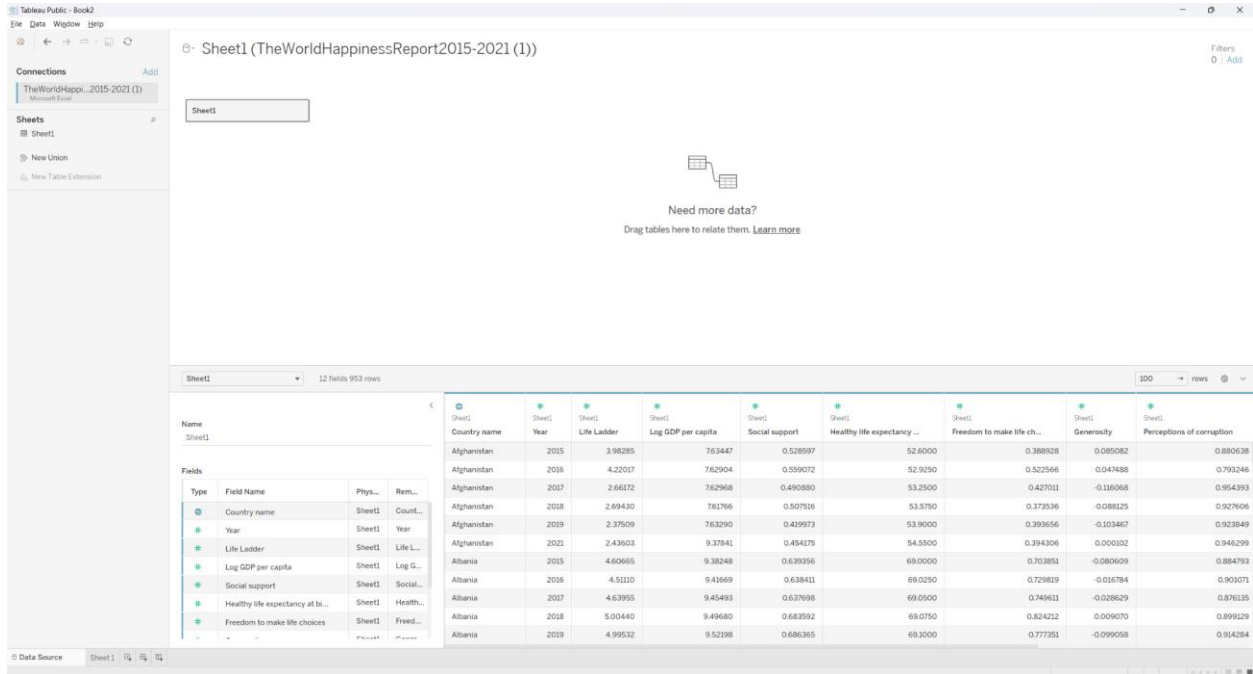




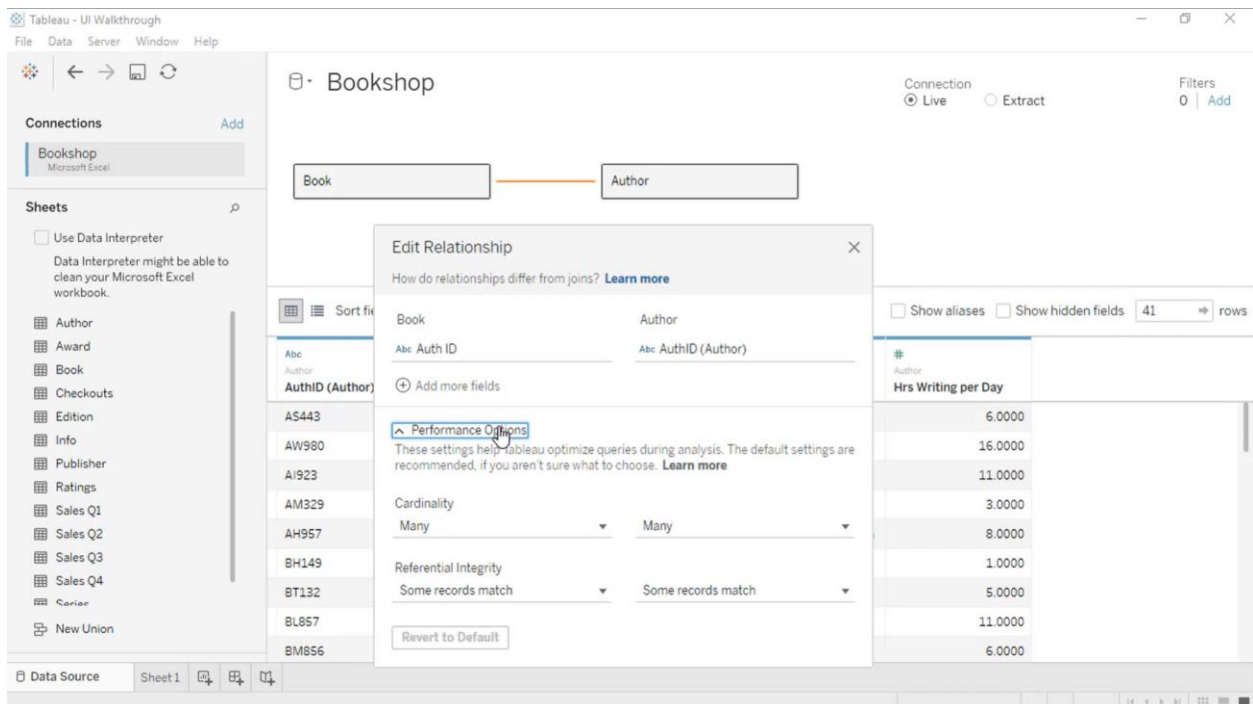
Sau khi connect thành công với database chúng ta có thể sử dụng được tất cả các table của database đó, và có thể thực hiện các thao tác kéo thả để trực quan dữ liệu.

- Xem trước dữ liệu
  - Trong Tableau, data preview cho phép người dùng kiểm tra cấu trúc của dữ liệu đó, chúng ta có thể nhìn lướt qua các cột để hiểu hơn về dữ liệu trước khi đi vào trực quan hóa và phân tích
  - Đầu tiên chúng ta kết nối tới nguồn dữ liệu mà chúng ta đã chọn (SQL Server, Excel ,CSV,...)

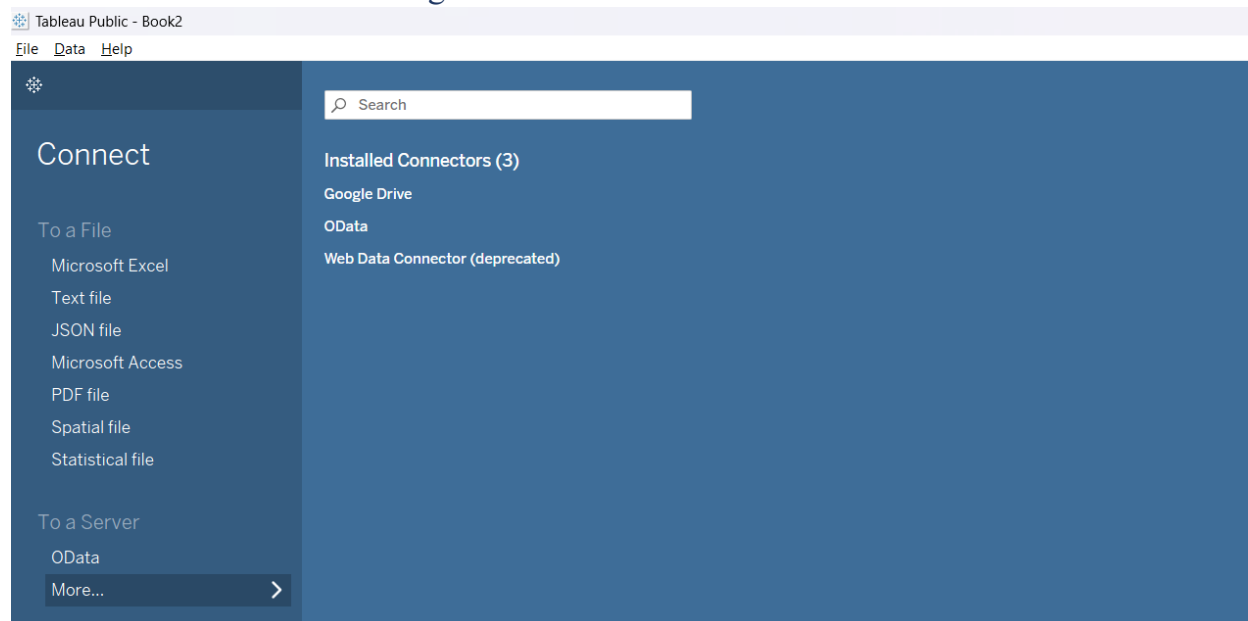
- Sau khi kết nối thành công, trên giao diện của Tableau sẽ hiện ra một bảng dữ liệu nhỏ
- Trong đây sẽ có các cột và các dòng của tập dữ liệu đã được chọn
- Chúng ta cũng có thể thực hiện thao tác trên bảng dữ liệu đây:



- Tạo các mối quan hệ giữa các table



## 1.2. Kết nối với Google Drive



## 1.3. Kết nối với các tập tin Excel, CSV, JSON, TEXT và nhiều định dạng tập tin khác

Tableau sẽ tự động phân tích và hiển thị định dạng của tập tin cấu trúc này, cho phép trực quan hóa dữ liệu một cách dễ dàng.

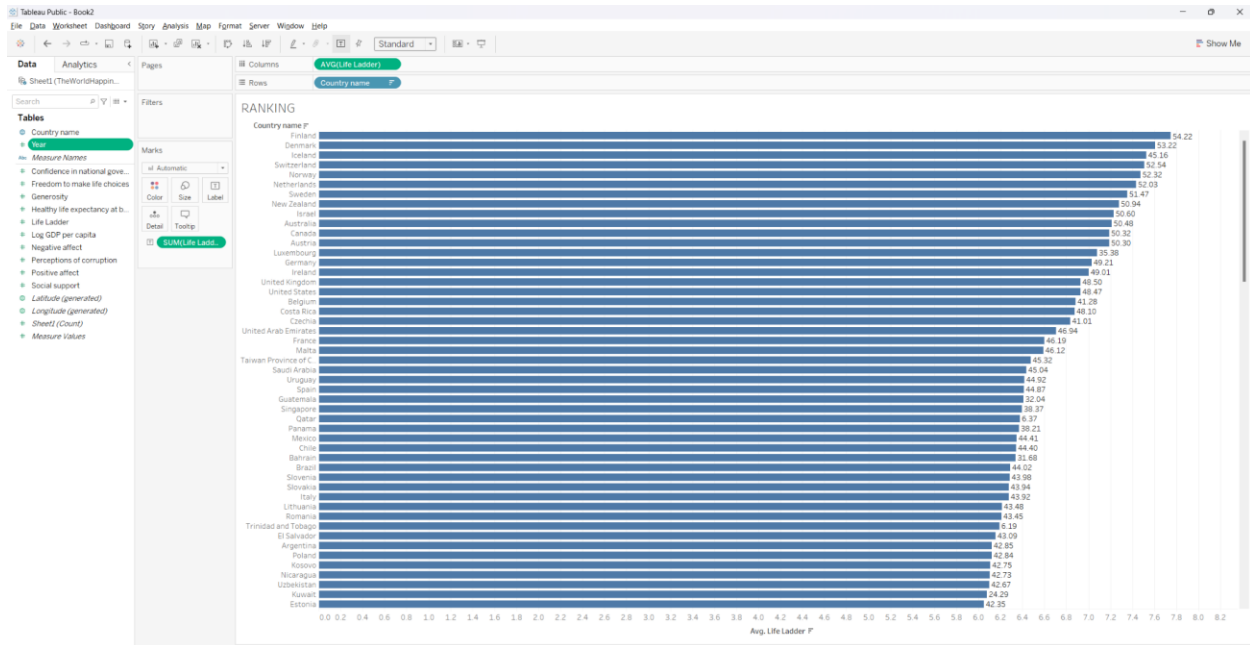
## 1.4. Dịch vụ web

Tableau hỗ trợ kết nối tới các web và API, cho phép lấy dữ liệu trực tiếp từ các nguồn như Google Analytics, Salesforce, Facebook, Twitter và nhiều dịch vụ khác.

# 2. Trực quan dữ liệu

## 2.1. Biểu đồ

Tableau cung cấp nhiều loại biểu đồ như: Bar chart, Scatter plot, Heatmap, Pie Chart, Bubble chart, Line chart, Candlesticks, Biểu đồ bản đồ,...

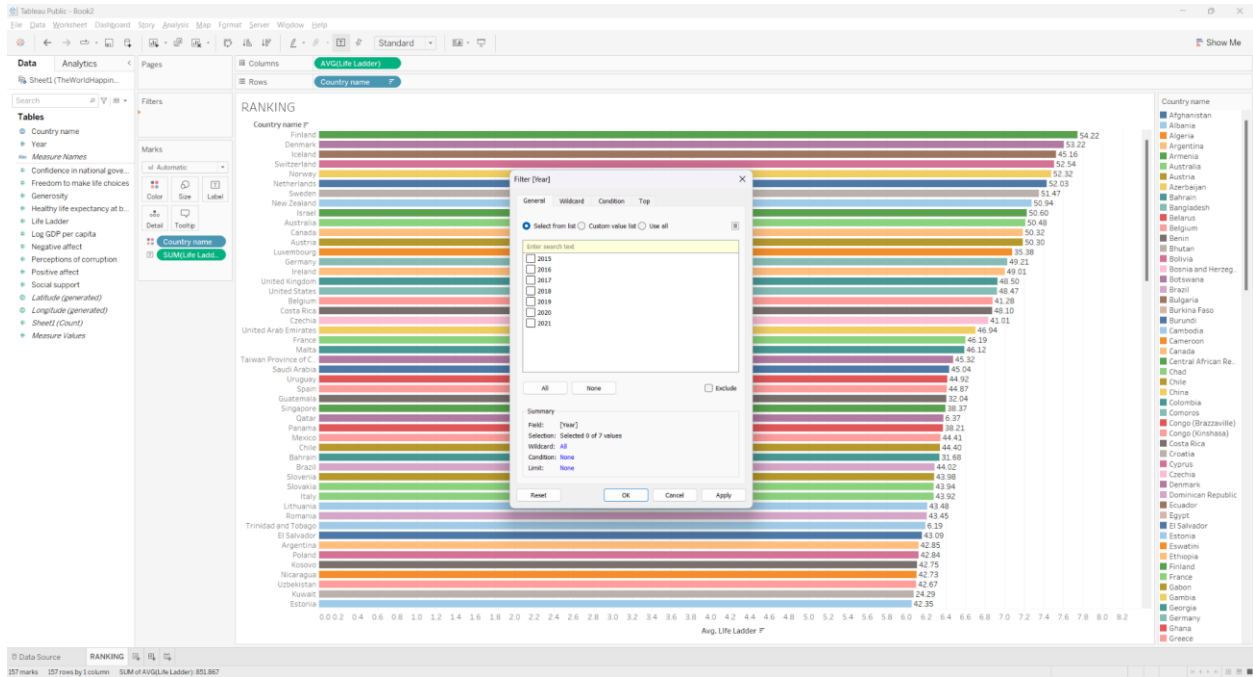


Chúng ta cũng có thể gán màu cho các dòng, ví dụ bây giờ chúng ta sẽ gán màu cho từng quốc gia đi, thì đơn giản chỉ cần giữ ctrl rồi click cột Country sau đó thả vào Color trong bảng Marks, sẽ thu được kết quả như sau:



Tuy nhiên vẫn còn rất khó nhìn phải không? Việc hiển thị tất cả các quốc gia như thế này cũng không phải là một ý trực quan hay, vậy giờ tôi sẽ lọc nó theo từng năm, để

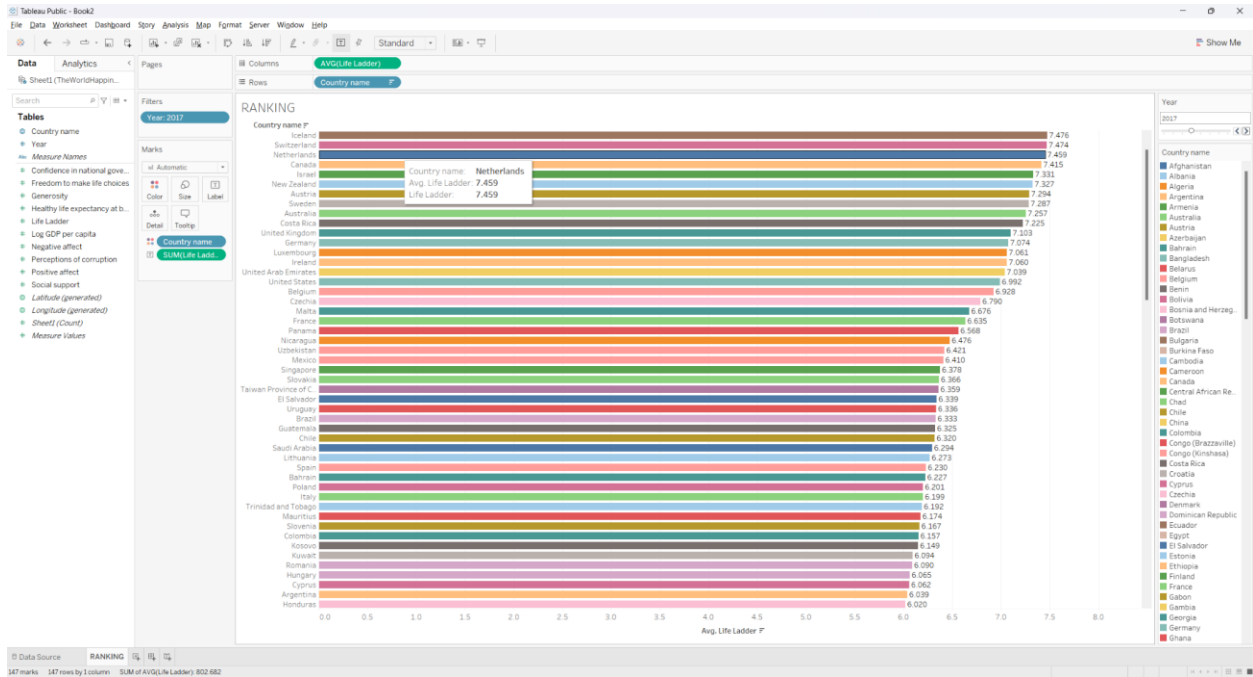
làm việc này ta cần một filter “Year”, bây giờ ta kéo cột Year vào bảng Filter ở trên Bảng Mark



Sau đó tick vào tất cả các năm ta muốn sort và bấm OK

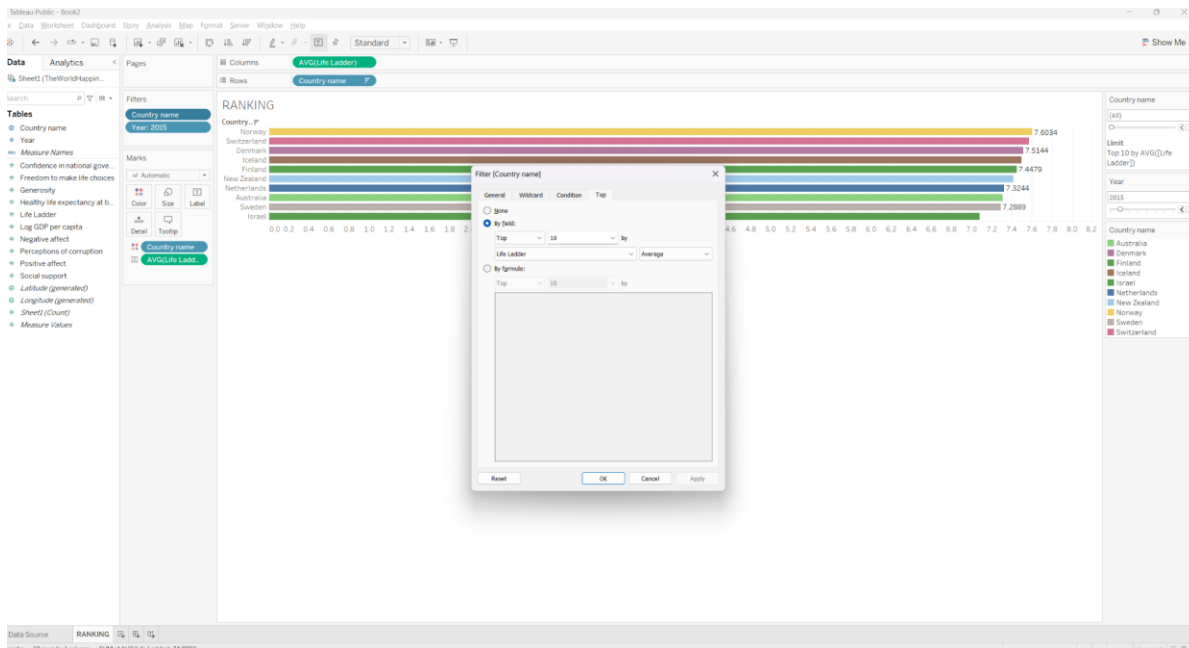
The image shows a close-up of the "Filter [Year]" dialog box. The "General" tab is selected. Under "Select from list", all years from 2015 to 2021 are checked. The "Summary" section shows "Field: [Year]", "Selection: Selected 7 of 7 values", "Wildcard: All", "Condition: None", and "Limit: None". The "OK" button is highlighted.

Click chuột phải vào Filter Year sau đó chọn Show Filter, sau đó Filter sẽ được hiển thị trên thanh bên phải của màn hình:

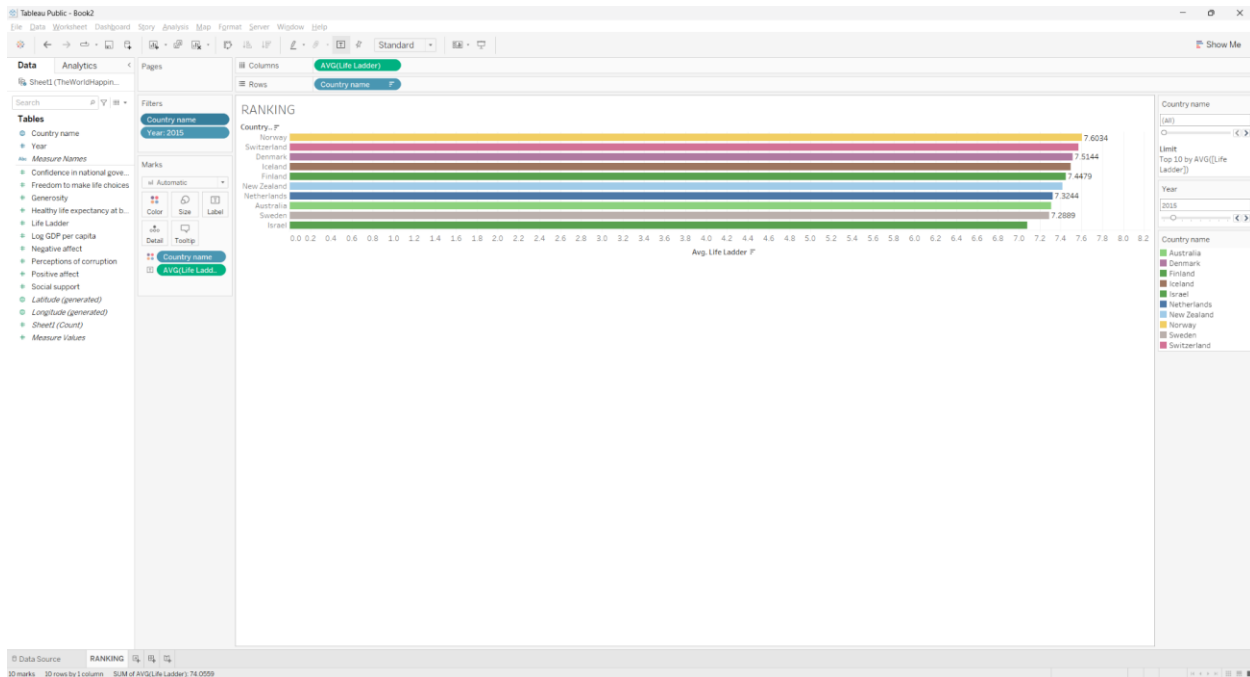


Giờ chúng ta có thể Sort theo từng năm của cái Filter

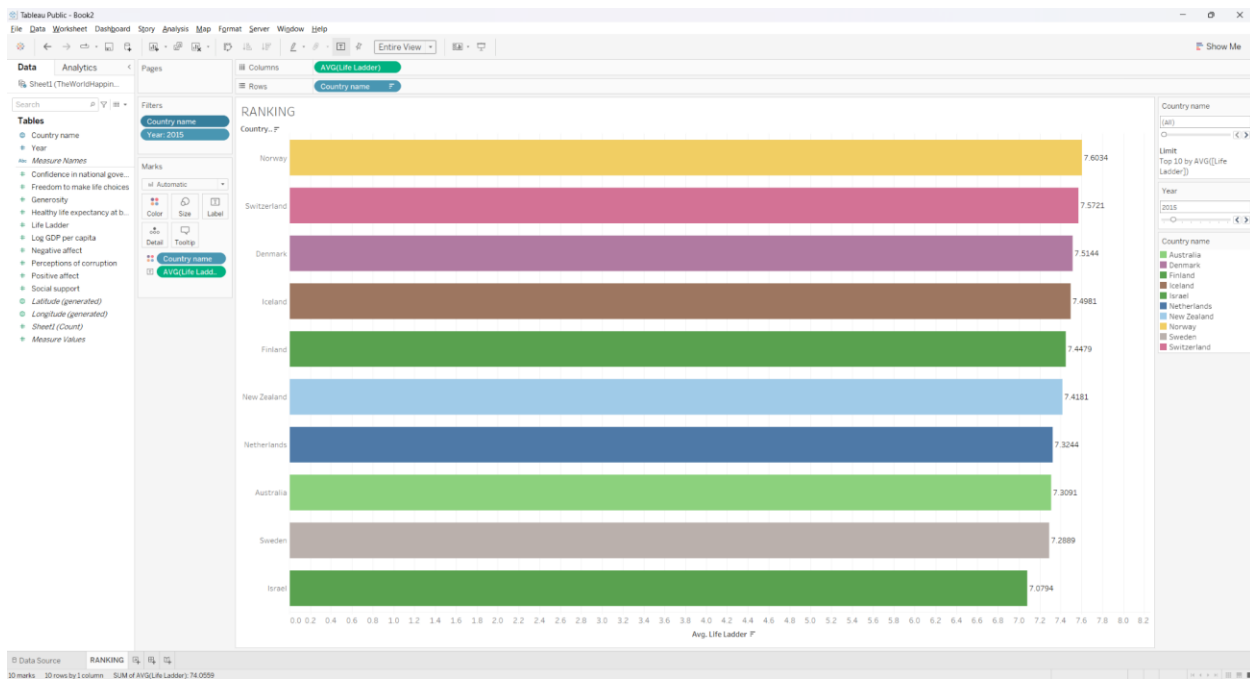
Nhưng vẫn hiển thị qua nhiều quốc gia bây giờ giả sử ta chỉ muốn hiển thị 10 quốc gia top đầu thôi thì làm như nào ? Tiếp tục tạo một cái Filter “Country Name” bằng cách kéo cột Country Name vào trường Filter



Click Ok và kết quả sẽ là như thế này:

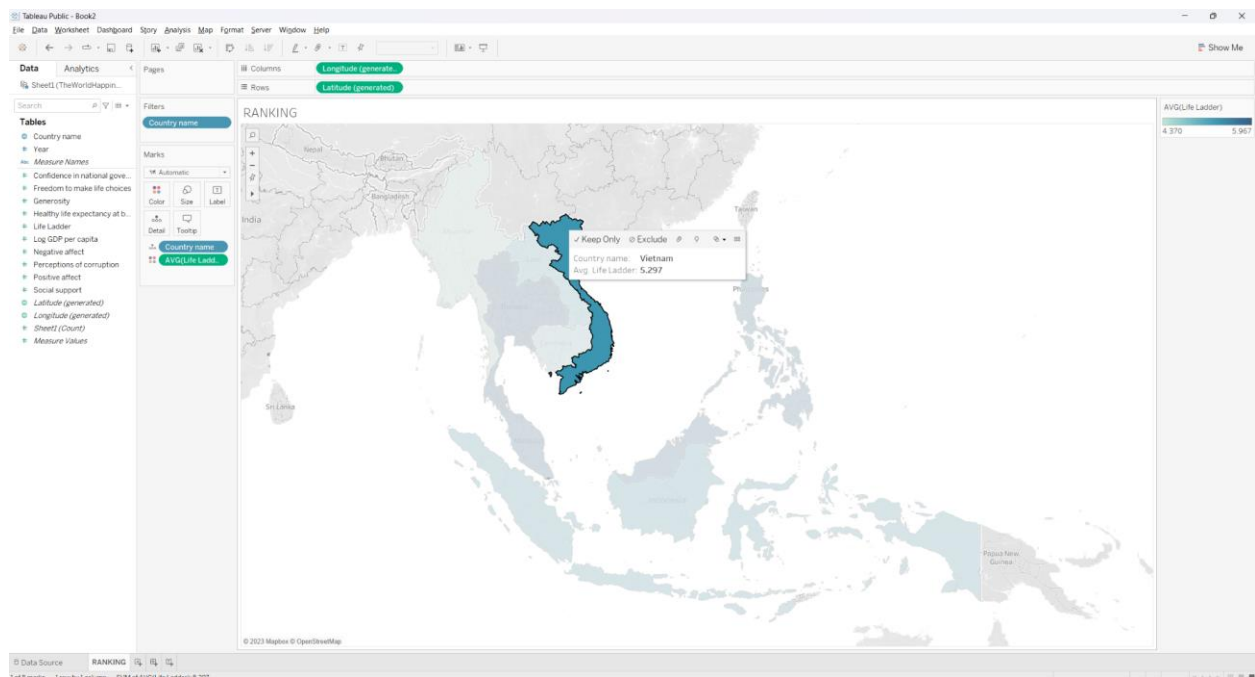
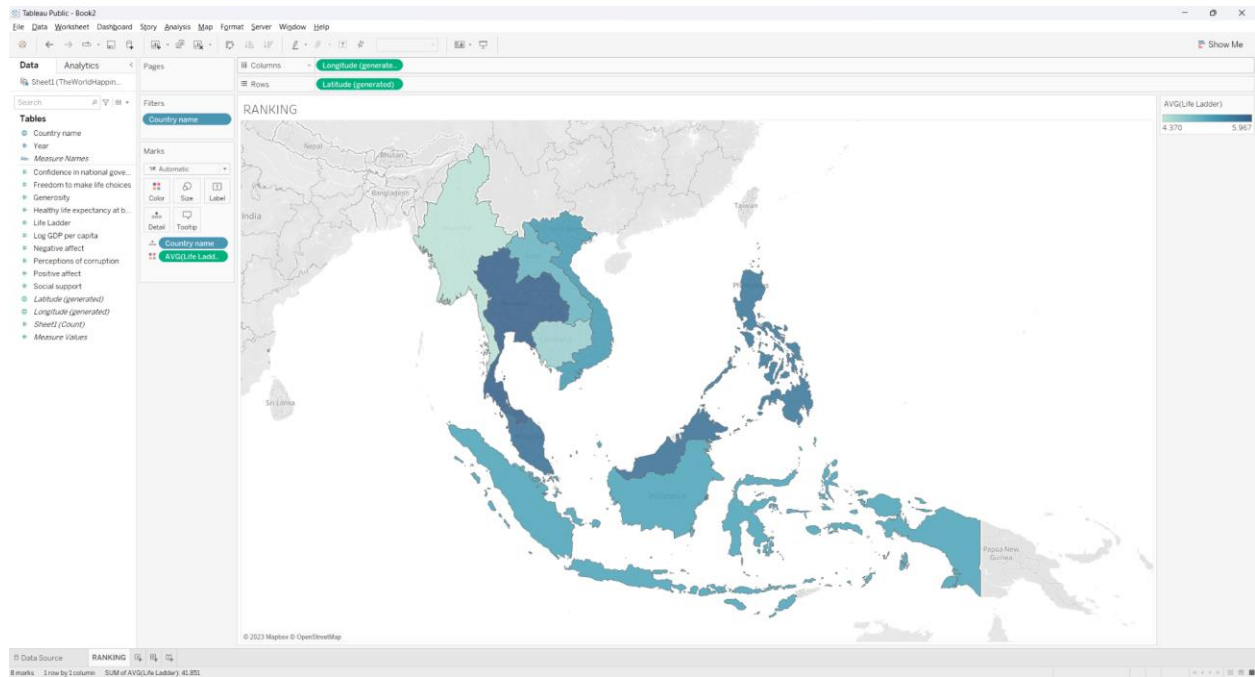


Đổi view nhìn qua Entire View:



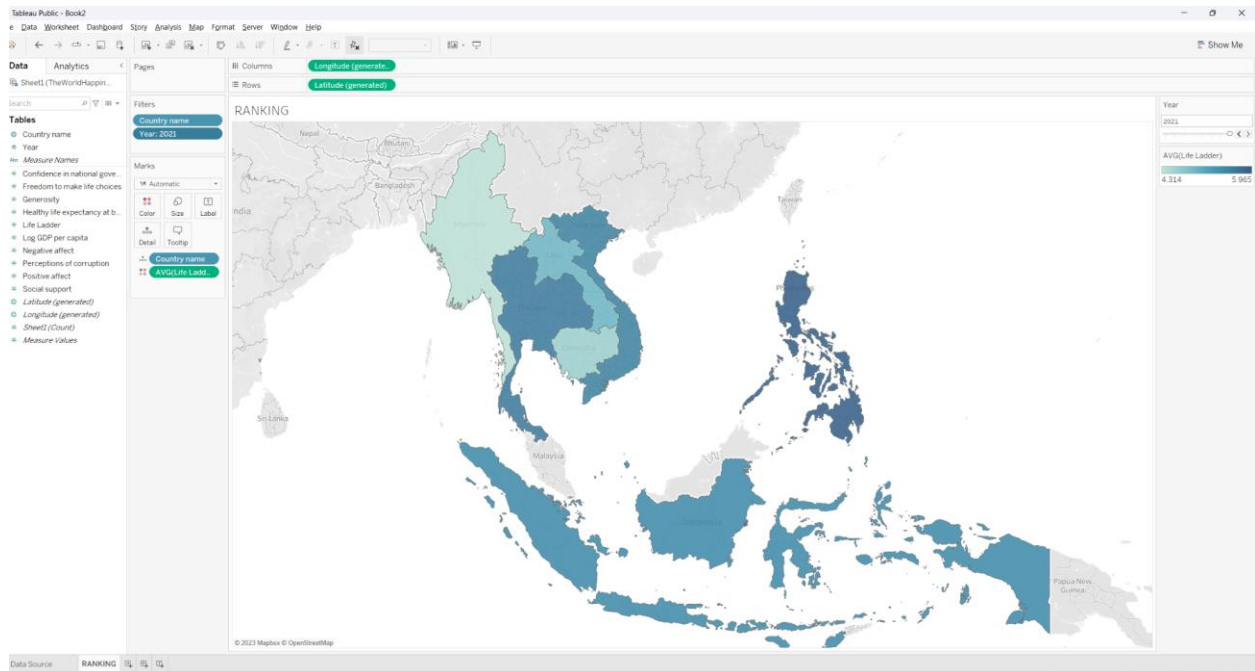
Ok dễ nhìn hơn nhiều rồi, chúng ta cũng có thể làm tương tự với những nước xếp cuối.

Tương tự đây là biểu đồ thế giới giữa Việt Nam và các người anh em hàng xóm:



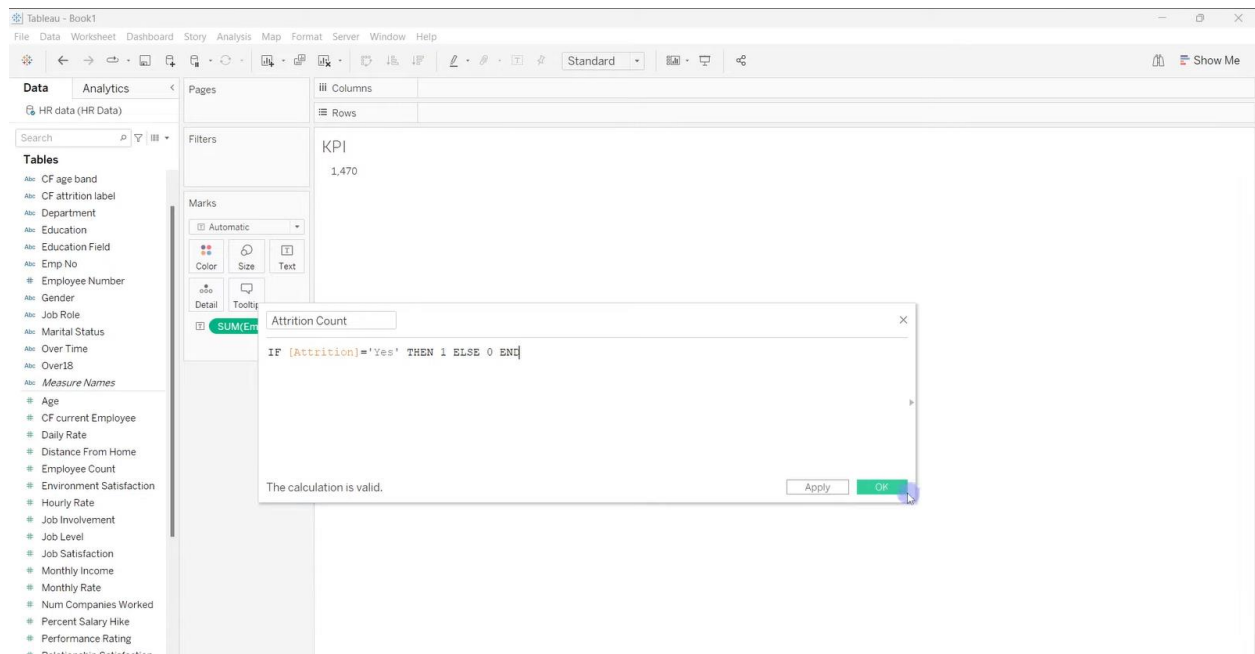


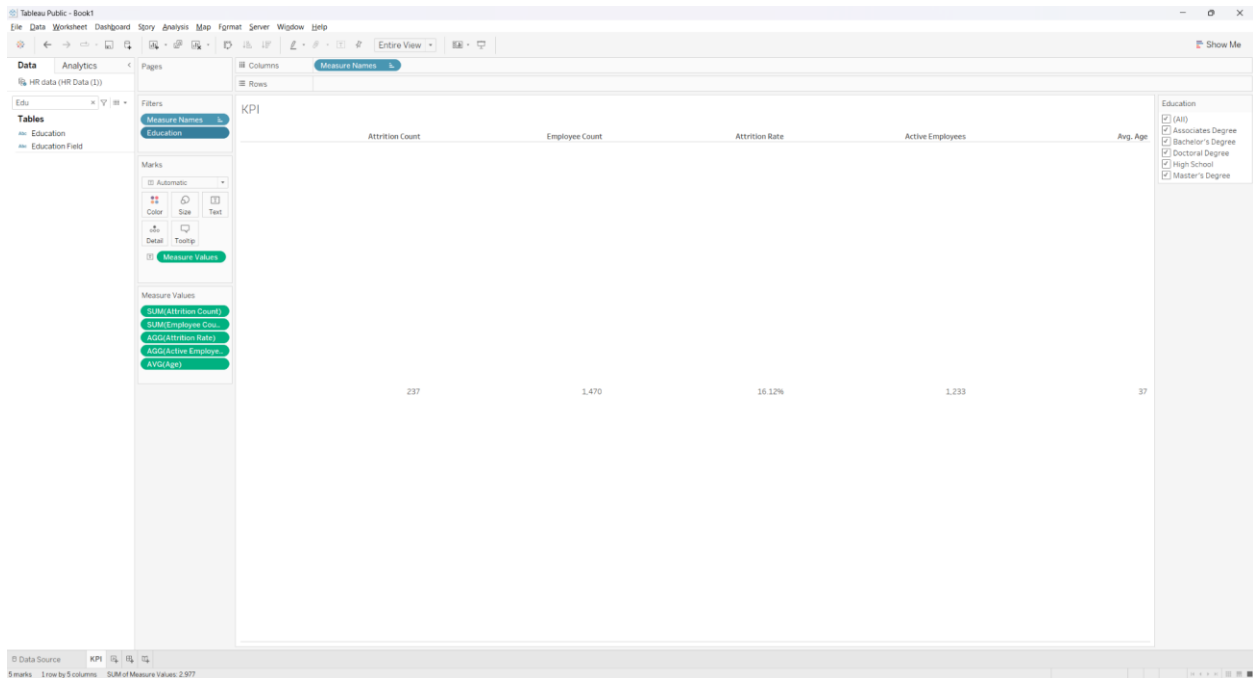
Biểu đồ này cũng có thể làm theo các năm luôn, làm theo giống cách ở trên



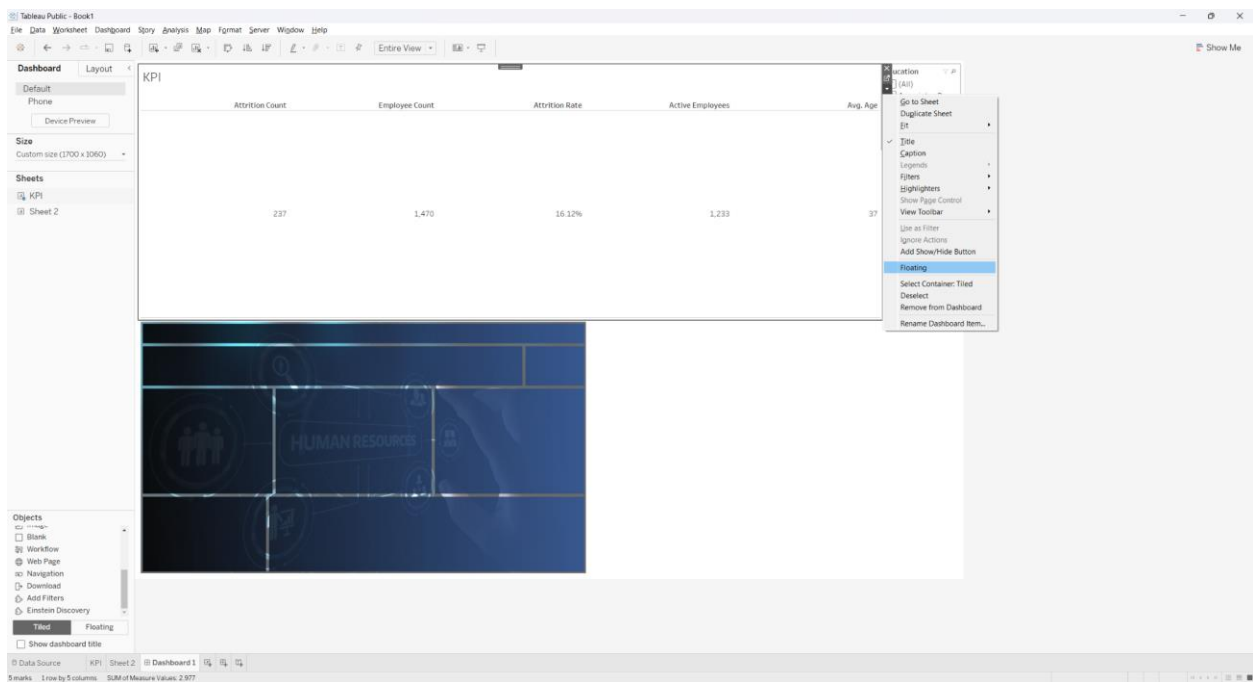
## 2.2. Calculated Field (trường tính toán)

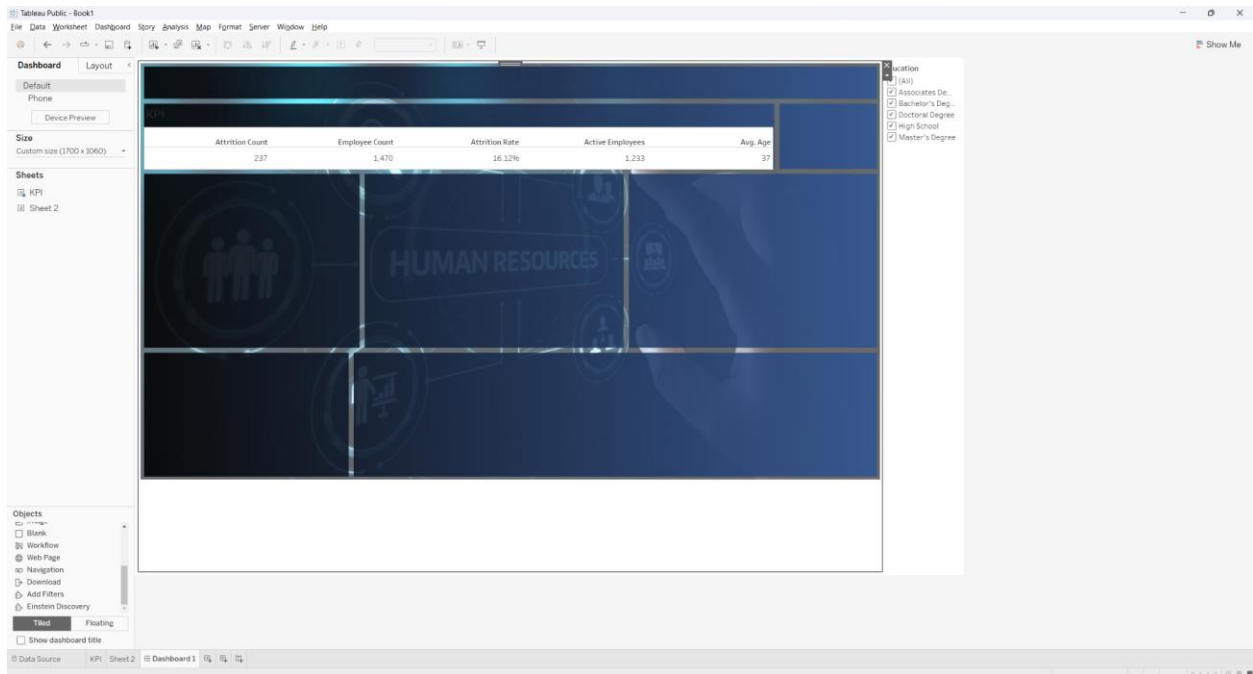
Tableau còn có chức năng tạo trường tính toán





Giờ ta tạo thử một cái dashboard rồi đưa vào: thì mặc định nó sẽ xếp cạnh nhau nhưng nếu chúng ta muốn xếp chồng lên cái background thì có thể chọn floating của cái item đó:





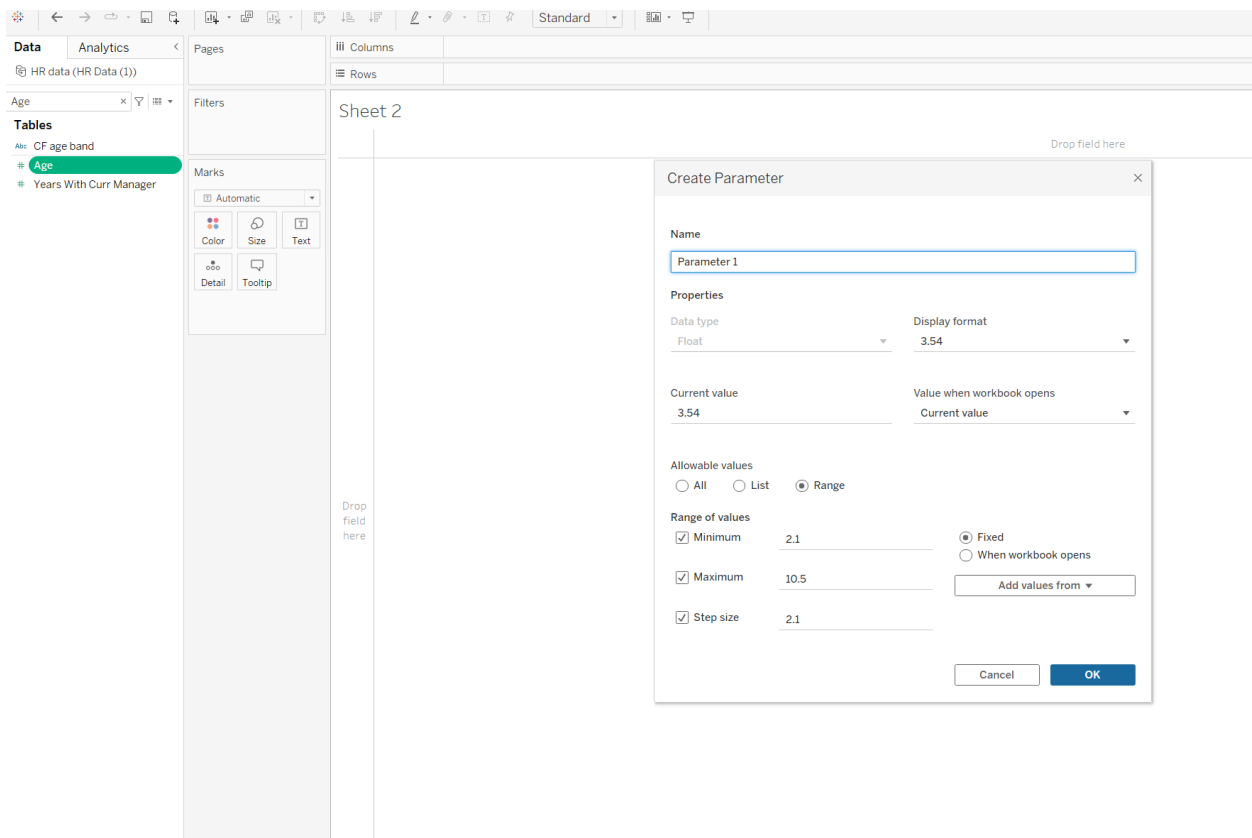
Chúng ta có thể Click chuột phải vào cái sheet đây để customize nó:



### 2.3. Parameter

Tableau còn giúp chúng ta tạo các parameter, các parameter sẽ giúp chúng ta tạo ra các yếu tố tùy chỉnh trong table của chúng ta, giúp cải thiện tính linh hoạt và tương tác của biểu đồ trong việc trực quan và phân tích dữ liệu. Parameter còn giúp chúng ta thực hiện các phân tích động, tính toán linh hoạt dựa trên các thông số thay đổi theo thời gian.

Ví dụ:



Drop field here

### Create Parameter

Name

Size

---

Properties

Data type: Float

Display format: 2

---

Current value: 2

Value when workbook opens: Current value

---

Allowable values

☐ All ☐ List ☒ Range

Range of values

☒ Minimum: 2

☒ Maximum: 10

☒ Step size: 1

☒ Fixed

☐ When workbook opens

Add values from

Cancel OK

### Edit Bins [Age]

New field name: Age (bin)

Size of bins: Size

Suggest Bin Size

Range of Values:

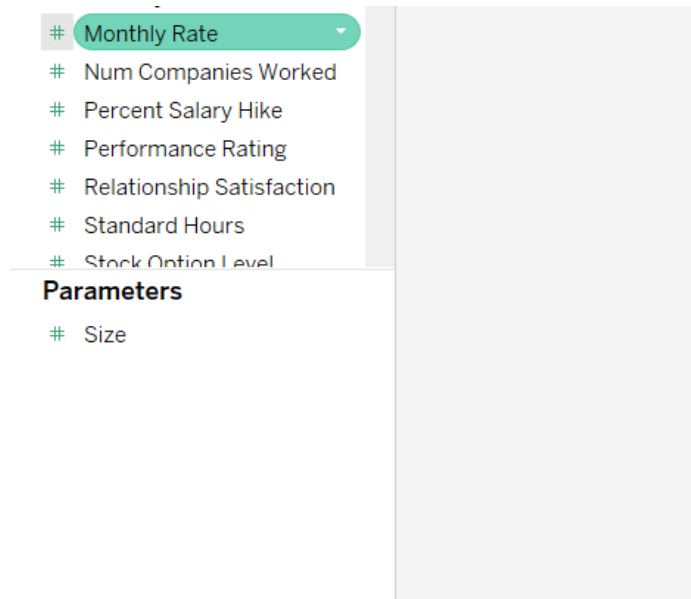
Min: 18 Diff: 42

Max: 60 CntD: 43

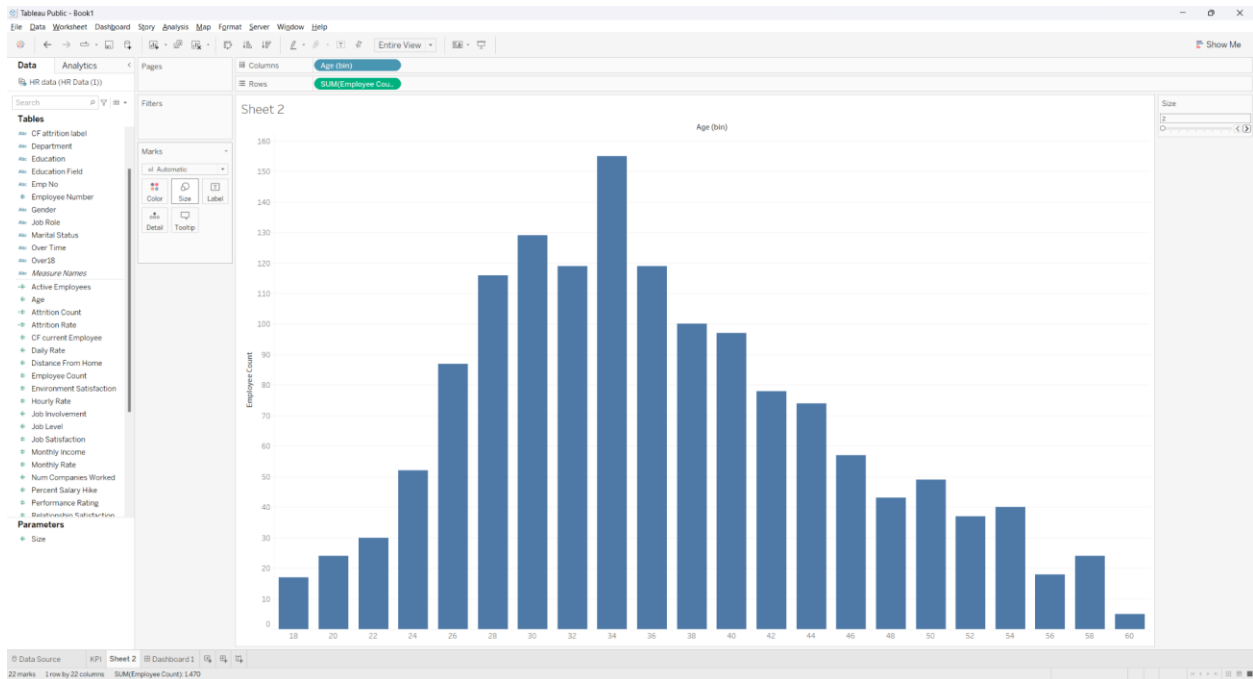
OK Cancel

Drop field here

Đây là Parameter mà chúng ta vừa mới tạo ra, bây giờ thử áp dụng nó:

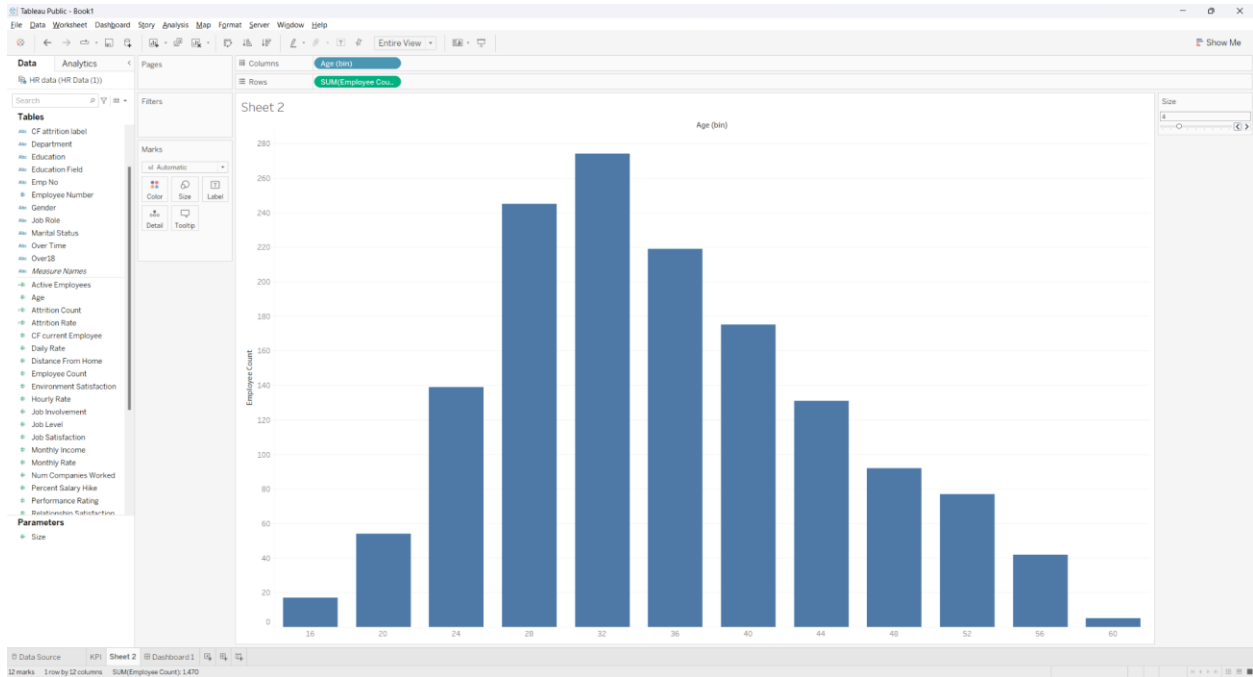


Đầu tiên Click chuột phải để add Parameter đó vào sheet trước thì mới có thể sử dụng, ở đây Size là khoảng cách độ tuổi, và khi chúng ta thay đổi Size thì khoảng cách tuổi tăng lên



Đây là khi tăng Size lên 4 có thể thấy các cột đã bị ít đi vì khoảng cách giữa các độ tuổi tăng lên, thì điều này giúp cho người phân tích dữ liệu biết được là những độ tuổi nào

sẽ có số lượng nhân viên nghỉ việc là nhiều nhất, ở đây thêm một Label cho từng cột nữa để nhìn rõ hơn số người:



Thì chúng ta chỉ cần thêm kéo table số lượng Employee và dùng hàm SUM lại với nhau, kết quả sẽ được và đây là một trong những tính chất của việc dùng Parameter

