



Welcome to the 2023 ADSE Exercise Guide!

This session is intended to help you be familiar with the different features and functions of SAP Analytics Cloud (SAC for short). After this exercise, you should be well equipped to conduct your own analyses using reliable datasets that you will find on your own.

LOG-IN INSTRUCTIONS, will be provided to you by your facilitators.

HOW TO ACCESS THE ADSE 2023 EXERCISE STORY

1. From your home page, click on **Catalog**
2. Select the tile **!ADSE 2023 Training Exercise [START HERE]**
3. Alternatively, you can use the Search Icon on the top right to search for **!ADSE 2023 Training Exercise [START HERE]**

BEFORE ANYTHING ELSE, it is important that that you save a version of this file as your own. Saving a file will allow you to track your individual progress during this exercise. If you do not save your work now, you will be unable to proceed.

HOW TO SAVE:

1. Make sure you are on **Edit mode** on the top right part of the screen (not View mode)
2. Locate the toolbar at the top of your screen and look for the File option.
3. Click File > Save > Save as...
4. Save your file under the following folder directory: **Public > !SAVE YOUR FILE HERE > Your Country > Your University**
5. Name your file as **MMDDYYYY_YOUR FIRST NAME**
6. Example: **01312023_JOHN**

Reminder: any files not saved under the university folder will be deleted periodically.

As a reminder, SAC is meant to be accessed via Google Chrome or Microsoft Edge. Other internet browsers may cause you to not follow along some of the exercises.

Note: as this is a training tenant, files may be deleted without warning periodically. Please do not upload any important information using the shared access accounts.

Name	Description	Type
Brunei		Folder
Cambodia		Folder
Indonesia		Folder
Lao PDR		Folder
Malaysia		Folder
Myanmar		Folder
Philippines		Folder
Singapore		Folder
Thailand		Folder
Viet Nam		Folder

What to expect from today's training session?

In order to give you a better understanding of how SAC can be used for the ADSE competition, we will be working on a real life Covid 19 dataset made publicly available via Kaggle.

In our analysis, we will be answering questions such as:

1. Which countries performed best/worst in the pandemic?
2. When were the peak periods?
3. What affects positivity rate of a country?
4. How did SEA countries manage Covid 19?

While this training has a structure we want you to follow, we also give you freedom to go beyond the exercise and conduct additional data exploration by yourself.

Dataset source: <https://www.kaggle.com/datasets/imdevskp/coronavirus-report>

Training exercise agenda

1. Data Modeling - Learn the basics of uploading datasets and data transformation.
2. Smart Discovery - Automate insight generation with the help of machine learning.
3. Designer Mode - Grasp the core functionalities of creating charts and graphs on SAC.
4. Data Explorer Mode - Discover insights to your dataset
5. Styling - Be able to create more aesthetically pleasing dashboards.

~ around 20-30 minutes per exercise.

What is this gray box?

Over the course of this training, guided instructions will be detailed out in this gray colored box to help you follow along.

Make sure to follow along all the details to ensure everything goes smoothly. A few discrepancies may occur when it comes to rounding numbers or decimal placements, based on how well you follow along.

15:06

February 10, 2023

In this exercise, you will learn how to:

- Upload dataset
- Check for data issues
- Correct data via transformation
- Saving your model
- Applying calculations

!! Use the guided instructions to the left to follow the exercises**!!** Please remember to save often**How to save your work:**

1. Click File > Save > Save as...
2. Name your file as MMDDYY_YOUR INITIALS
3. Example: 01312023_AAA

This section will familiarize you with **Data Modeling** to upload data.**What is Data Modeling?**

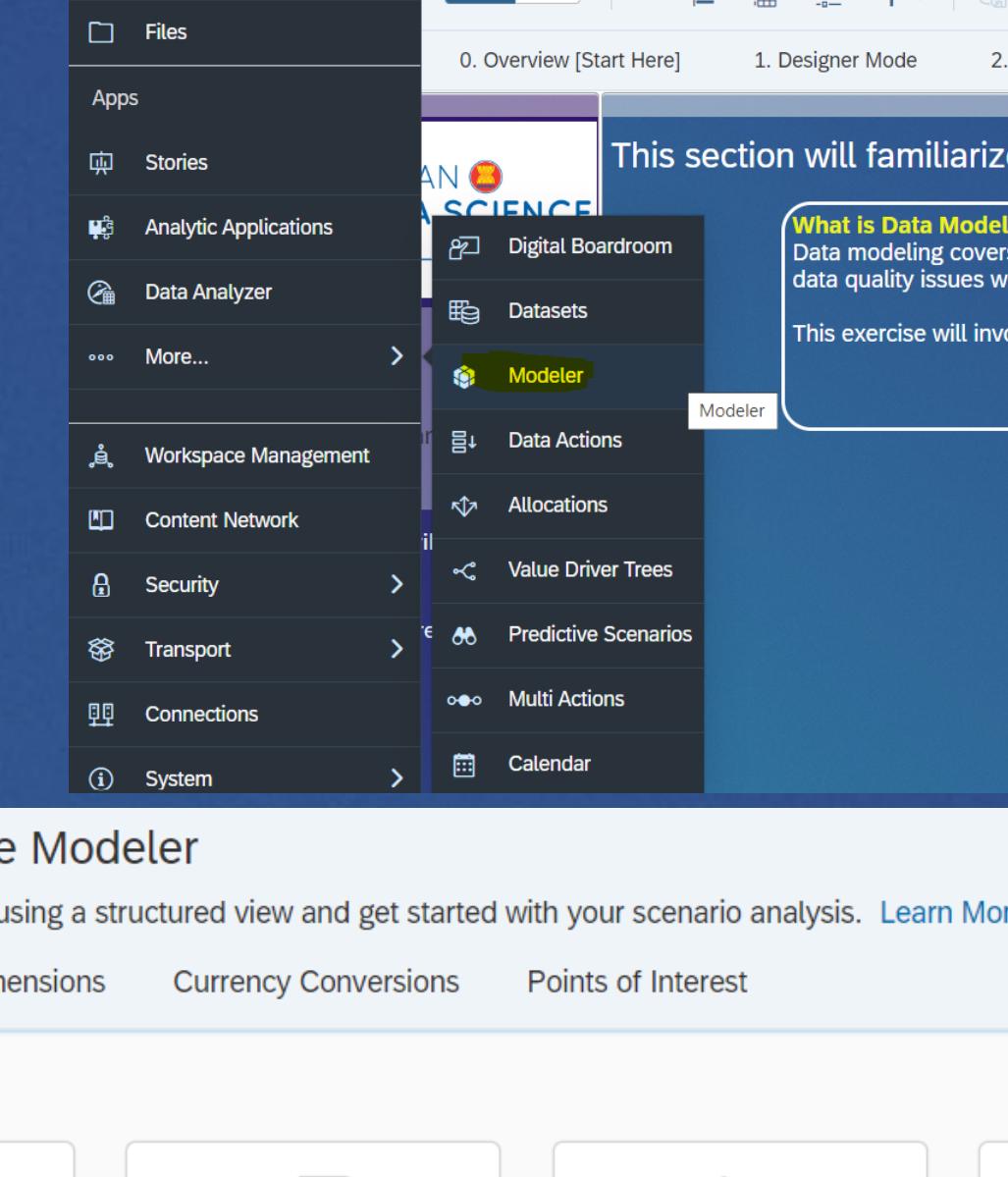
Data modeling covers the entire activity of uploading datasets into SAC, assigning measures and dimension categories, and fixing and data quality issues with the dataset.

This exercise will involve uploading a 13MB file into SAC. Please make sure you have downloaded this prior to the enablement session.

Data Modeling will take place outside this story.Therefore, make sure to **SAVE** your work at this point.You may use this link to access Data Modeler:
<https://aseandse.ap11.hcs.cloud.sap/sap/fpa/ui/app.html#/modeler>

Or follow the steps below

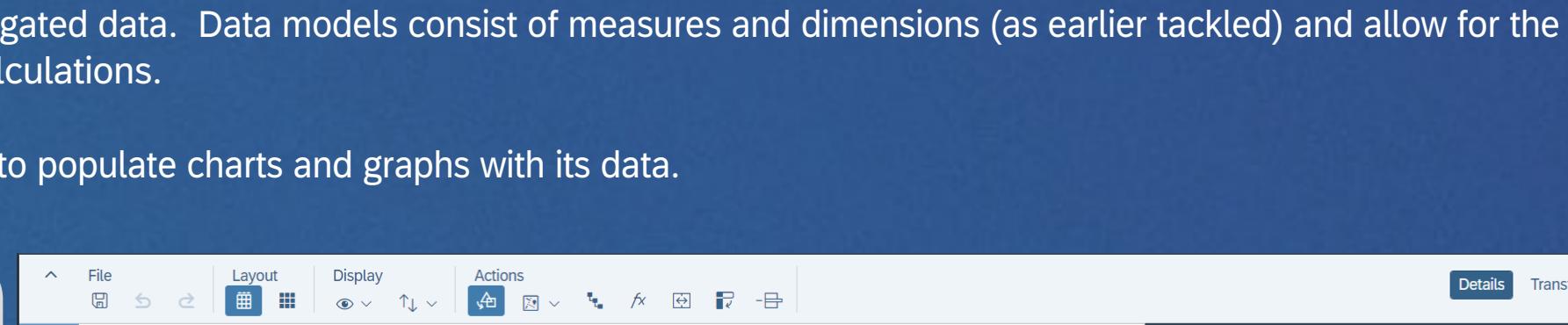
1. Expand the horizontal toolbar to the upper left
2. Under **Apps**, Click **Modeler** (you may find it under More...)

**There are many options for us to create models on SAC.**For this exercise, click From a CSV or Excel File and upload the dataset **ADSE Dataset for Data Modeling (Exercise 1)** which should be saved on your local device.

Give the system a few minutes to upload depending on your internet speed.

Welcome to the ModelerPrepare your raw data using a structured view and get started with your scenario analysis. [Learn More...](#)

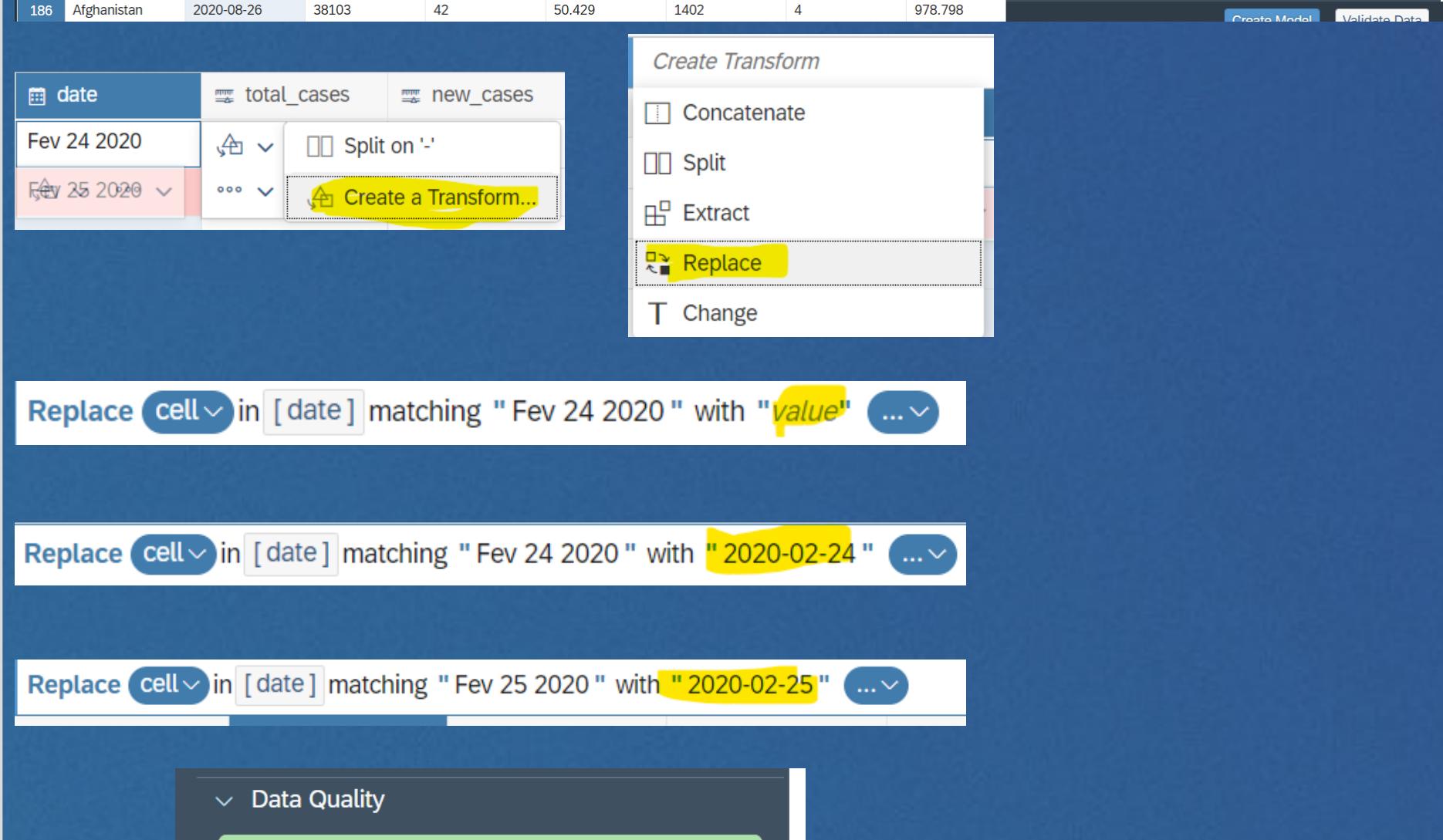
Models Public Dimensions Currency Conversions Points of Interest

**LEARNING:** What is a model?

In SAC, a model is a representation of large amounts of aggregated data. Data models consist of measures and dimensions (as earlier tackled) and allow for the construction of hierarchies, metadata, geolocations, and calculations.

Once a model is created and maintained, it can then be used to populate charts and graphs with its data.

Upon loading, you will now see the dataset displayed similar to Excel format. In this view, we will be able to observe if there are any data quality issues detected in the dataset and correct them before moving on.

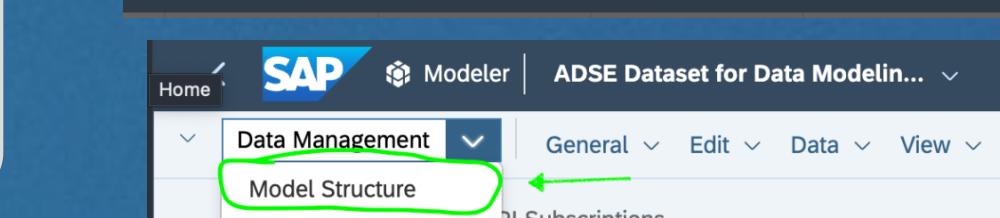
(If it does not automatically load, you may find the uploaded file under **Draft Data**)Note that we have encountered data quality issues in the **date** column due to a typo/misformat.In order to correct this, we need to replace **Fev 24, 2020** to the format of **2020-02-24**, similar to the other date entries.

1. Click on **Fev 24, 2020**
2. Open the transform icon
3. Click **Create a Transform...**
4. Click **Replace**

Replace cell in [date] matching "Fev 24 2020" with "value"

Replace cell in [date] matching "Fev 24 2020" with "2020-02-24"

Replace cell in [date] matching "Fev 25 2020" with "2020-02-25"



You will now observe all data quality issues have been fixed. We can now proceed to saving.

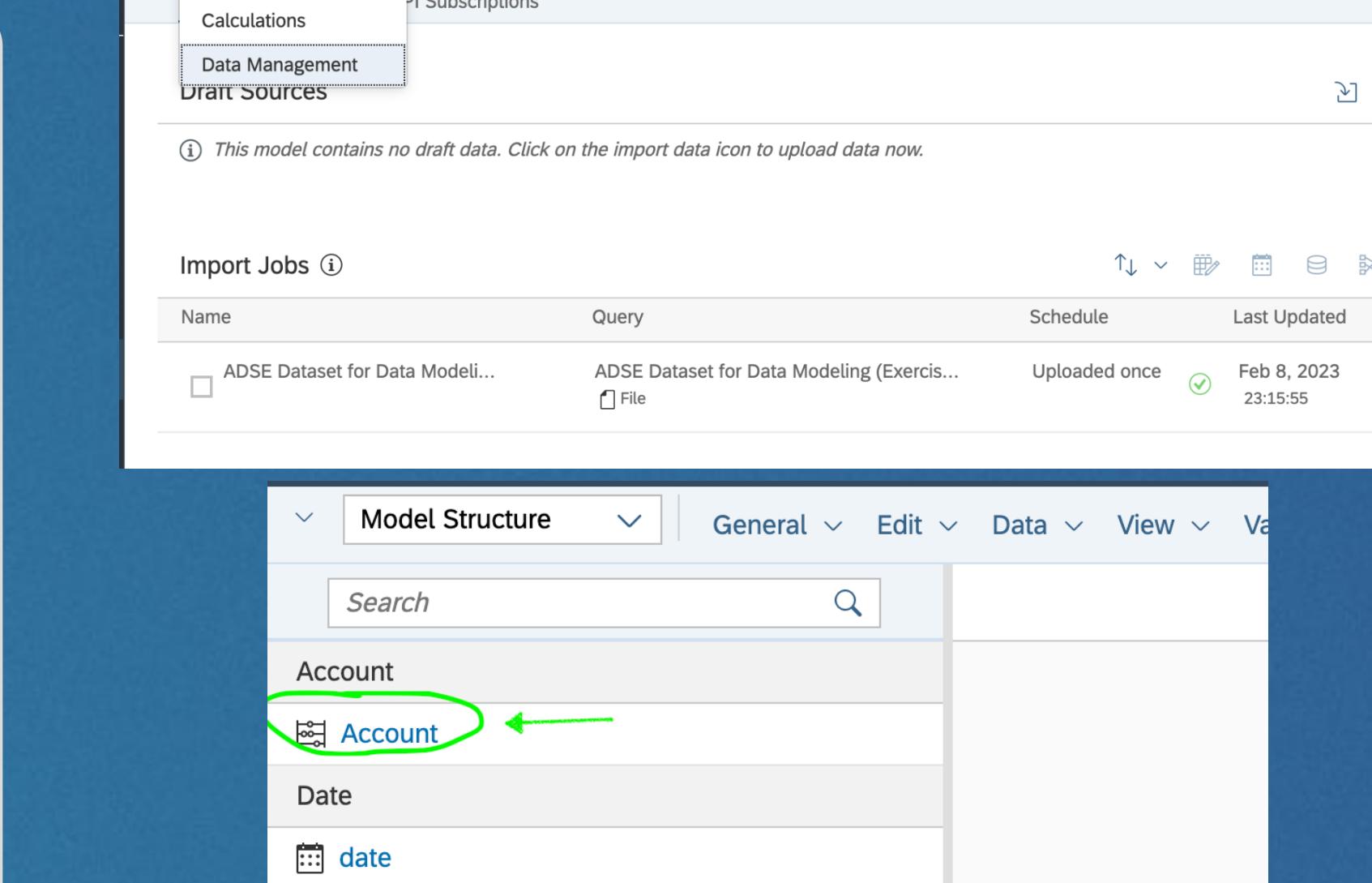
10. Click **Create Model**.

11. Save using the same format as our story earlier, using the file path
Public > ISAVE YOUR FILE HERE > Your country > Your university folder
(MMDDYYYY_your first name)
Example: 01312023_John

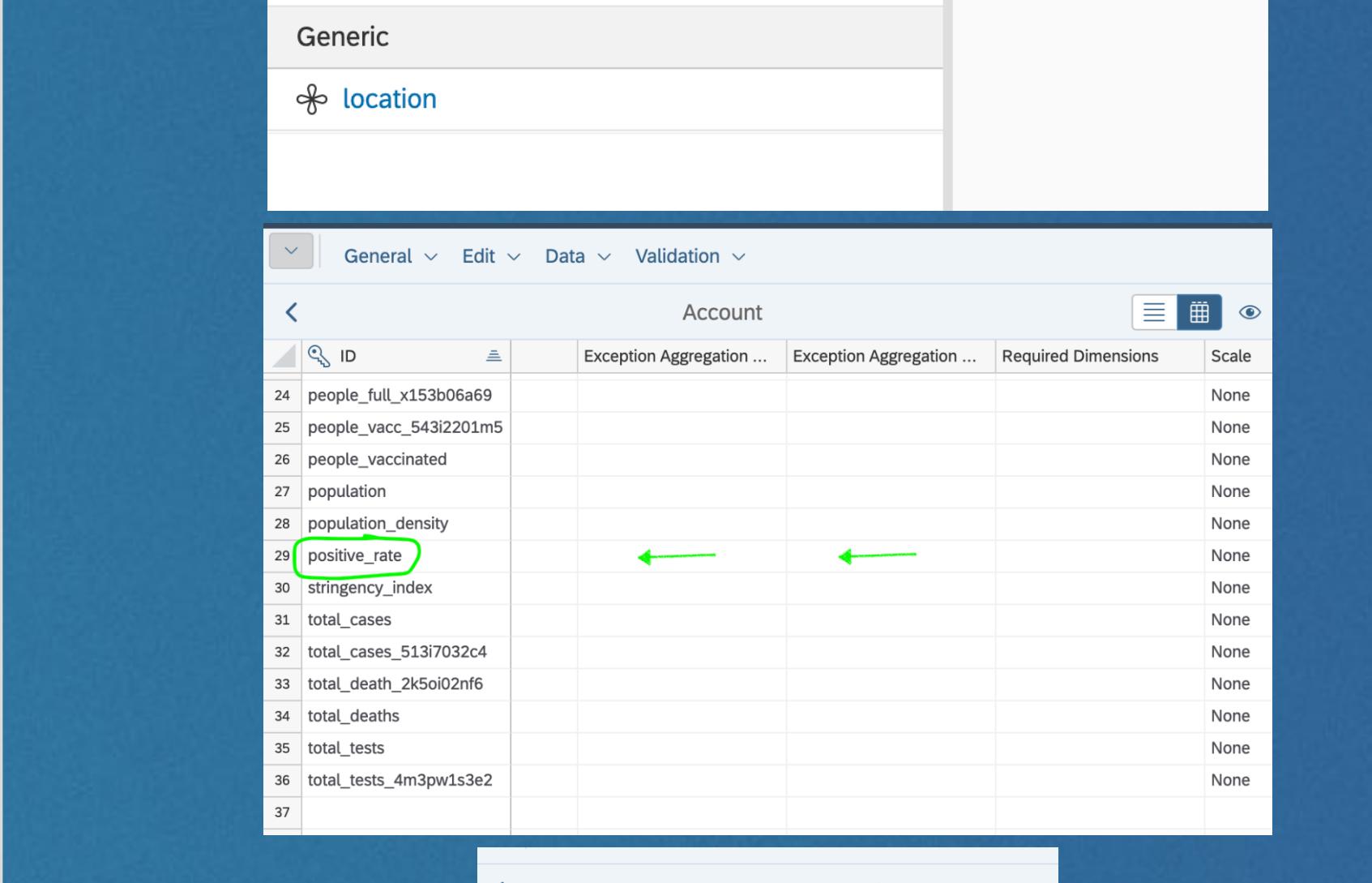
Remember your model name.

Upon saving, we automatically enter a new window Here, we can do further adjustments to the model structure.

1. Click on the dropdown button on Data Management, and switch to **Model Structure**.



2. In the new screen, click **Account**



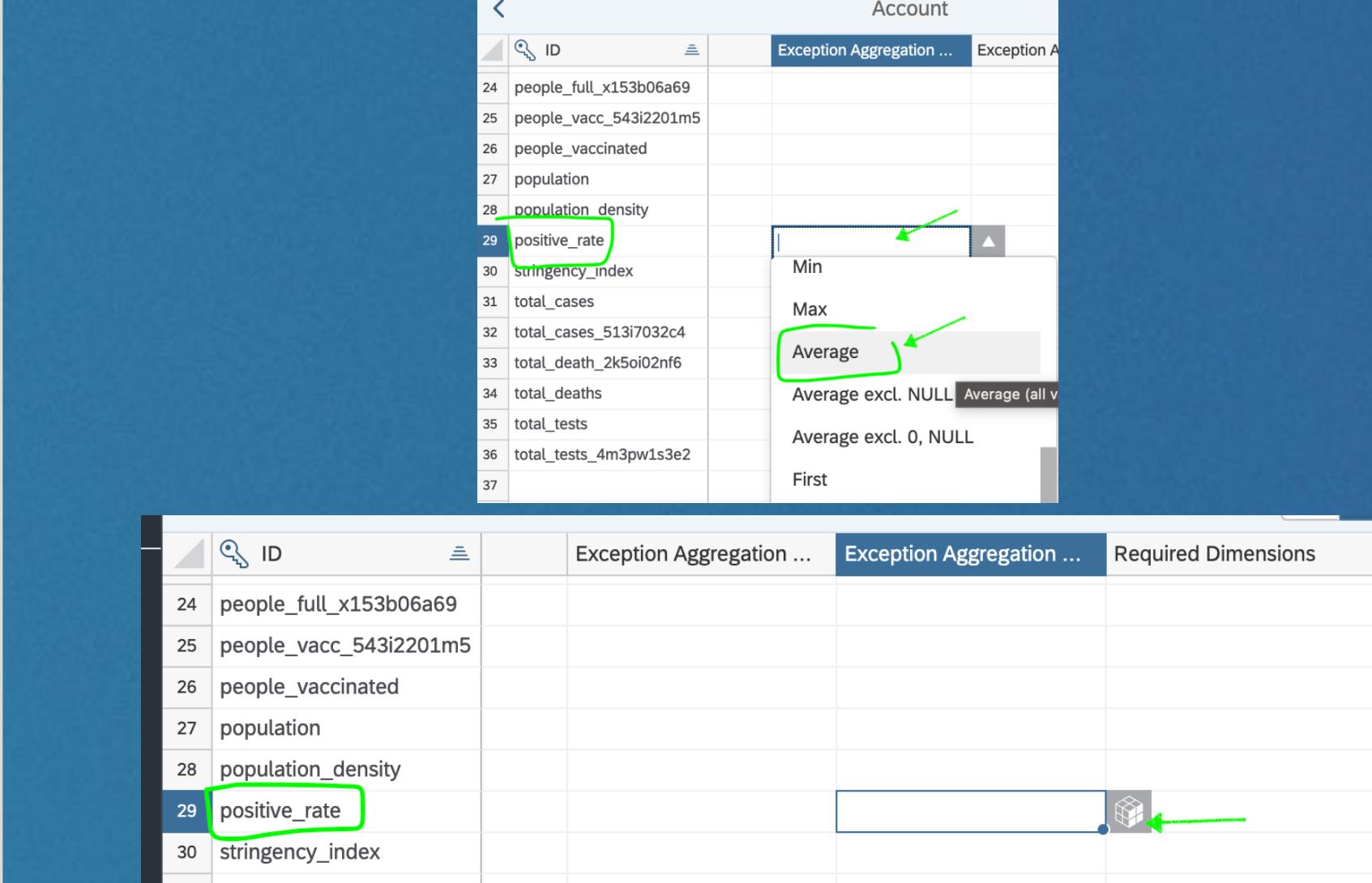
3. The account screen is a summary of all measures in our model. Here we can adjust the settings behind the measures, to make sure they are appropriate.

Scroll and find "positive_rate" which represents the percentage of all COVID-19 diagnostic tests that are positive on a given day. This metric indicates whether enough tests are being done to detect most new cases.

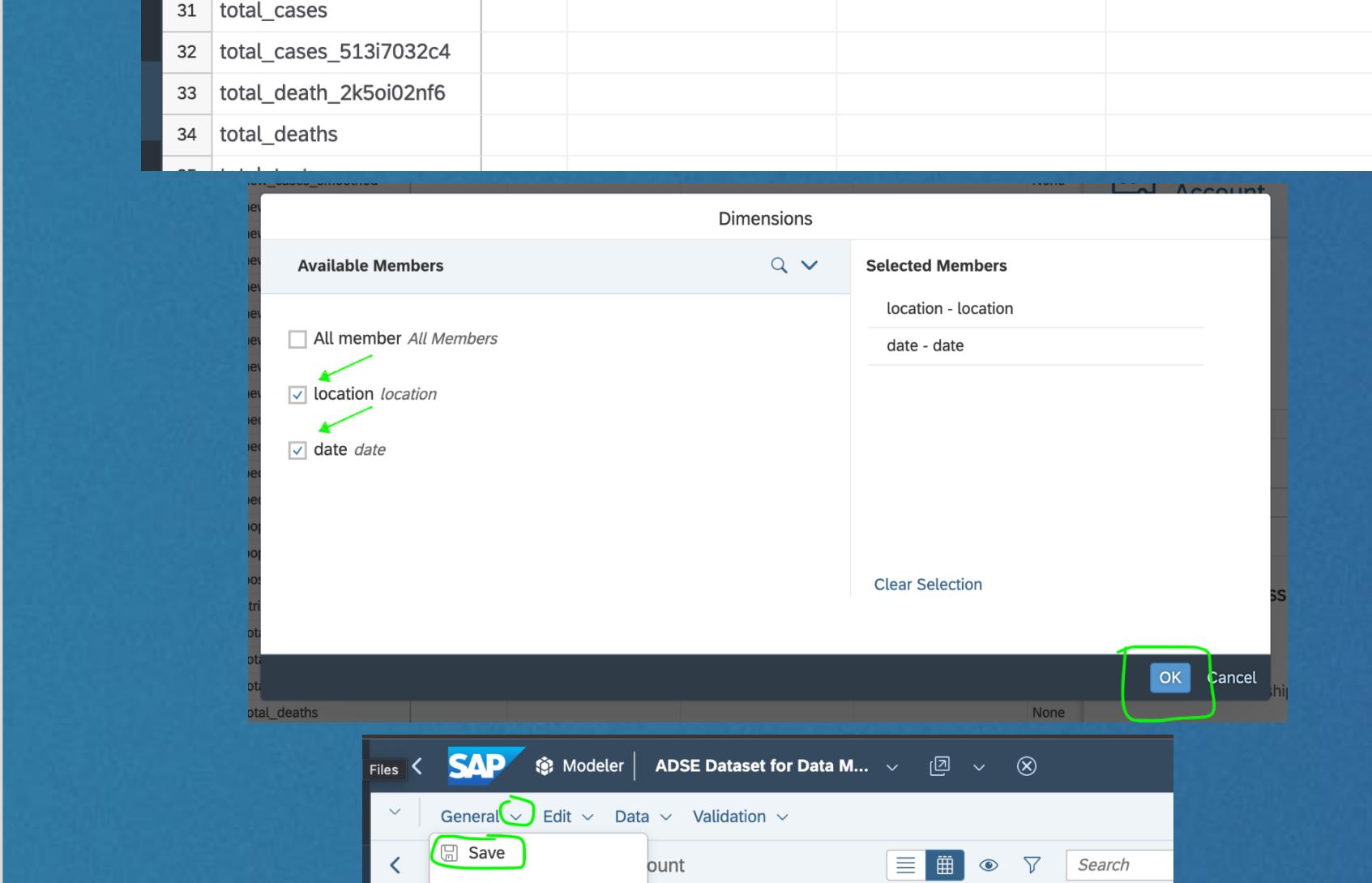
By default, SAC categorizes as a sum of all values. Since positivity rate is a recorded as a percentage in decimal form, we want to make sure SAC recognizes it as an **average** instead.In order to do this, we will be assigning **Exception Aggregation Formula Types** and **Exception Aggregation dimensions**.

4. On the same row as positive_rate, click on the cell under **Exception Aggregation Type** column (you may need to scroll to the right).
5. Among the different formula calculations, select **Average**

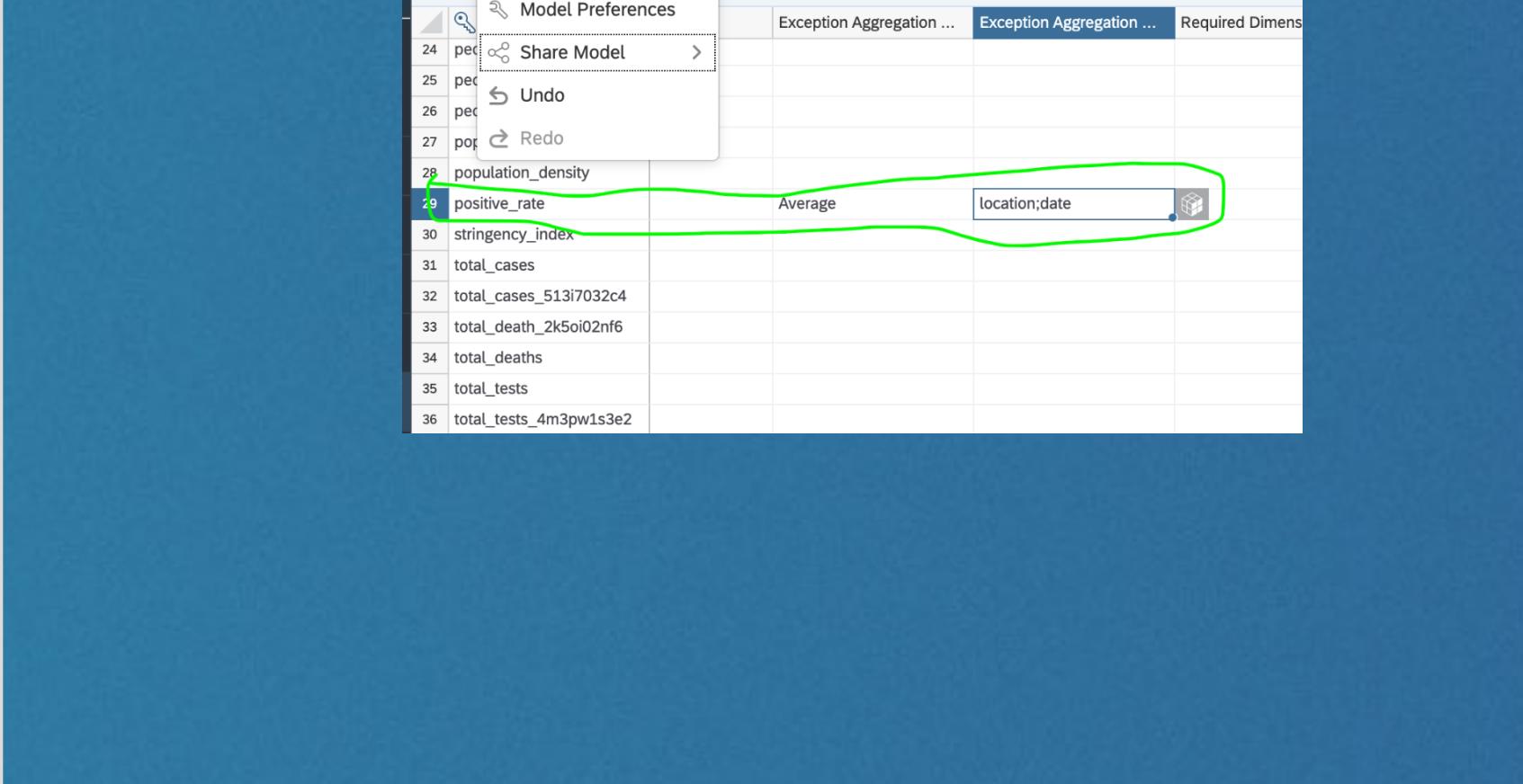
(We are asking SAC to calculate average positivity rate instead of the sum of all rates.)



6. On the same row as positive_rate, click on the cell under **Exception Aggregation Dimensions**
7. Click on the model icon shaped like a cube that shows up to the right of the cell.

8. In the window that appears, tick the boxes for **Location** and **Date**

(We are asking SAC to compute the average across all country locations, and all dates).



9. Your model should now look as follows:

10. Save your work by selecting the dropdown under **General**, and click **Save**.

You are now done with this section.

15:06

February 10, 2023

In this exercise, you will learn how to:
 - Create charts
 - Populate charts
 - Select fields (measures & dimensions)
 - Rank data
 - Duplicate charts

!! Use the guided instructions to the left to follow the exercises

!! Please remember to save often

How to save your work:
 1. Click File > Save > Save as...
 2. Name your file as
 MMDDYYYY_YOUR INITIALS
 3. Example: 01312023_AAA

This section will familiarize you with **Designer mode** to build your graphs.

What is Smart Discovery ?

Smart Discovery analyzes your data by running a machine learning algorithm to help you explore your data in a specific context, to uncover new, or unknown relationships between columns within a dataset. Smart Discovery always gives an overview of your data by automatically building charts to begin discovering more about your data.

Go back to the Training Exercise story by searching for your saved story from **Exercise 0**.

- Find **Smart Discovery** in the upper toolbar (it could be under ...)
- Ensure the selected data set is **Covid19 Test Data for Smart Discovery**
- If not, you may click the pencil icon to select it from the drop down box.
- Alternatively, you can locate your Data Model from exercise 1.

Now it's time to restrict our analysis.

For your first Smart Discovery, let's analyze positivity rate.

- Under target, select **positive_rate**.
- Under entity, select **location**.
- Click **Run**

This process may take a few minutes.

LEARNING: What is a target?

In Smart Discovery, the target variable is what we are finding correlations to between our other available data. Using machine learning, SAC find the top influencers towards this variable.

Analyzing the Smart Discovery Results

You should be presented with 3 new tabs full of charts and details on the Covid 19 positivity rate. Let's break down some of these areas.

Predictive Forecast

SAC comes with predictive capabilities that allow us to forecast outcomes based on historical patterns

Look at the predictive forecast at the bottom left area. We can observe the trend of positivity rate over the last 2 years. We can even observe a forecast of how positivity rate is expected to increase in the next few months, based on trend. Note the upper and lower confidence intervals for this trend.

1. Click on Forecast

SAC gives an automated predictive score to all forecasts. A score of 5 is the highest possible one, which makes the forecast more acceptable and reliable.

How can we improve forecast quality?

- > Additional rows of data helps improve data accuracy
- > Additional date points allow for further forecasting (general rule: 5 years worth of data to predict 1 year ahead).

We can also experiment with other predictive forecasting algorithms to see if can get better results.

- Click on ...
- Click **Add > Forecast > Advanced Options > Linear Regression**.

Is the forecast better or worse?

Smart Insights

SAC comes with predictive capabilities that highlight key contributors behind a graph or data point.

In the Positivity rate by location graph (bottom left), you can observe a system-generated comment mentioning Oman has the highest positivity rate, and on which date.

1. Click View More to see additional insights.

The newly opened tab gives us additional context as to the changes in positivity rate in Oman.

2. Right click on other countries on the chart, click Smart Discovery, and observe their changes as well

Key Influencers

Under Key Influencers, we can observe which columns in our corresponding dataset have the highest affect on positivity rate. These may be signaled as a weak or strong influence.

What are the top influencers for positivity rate?

Type on Zoom chat.

When finished, feel free to run additional Smart Discoveries using other target measures this time.

Simulation

Under Simulation, we can experiment with the outcome on positivity rate by modifying parameters.

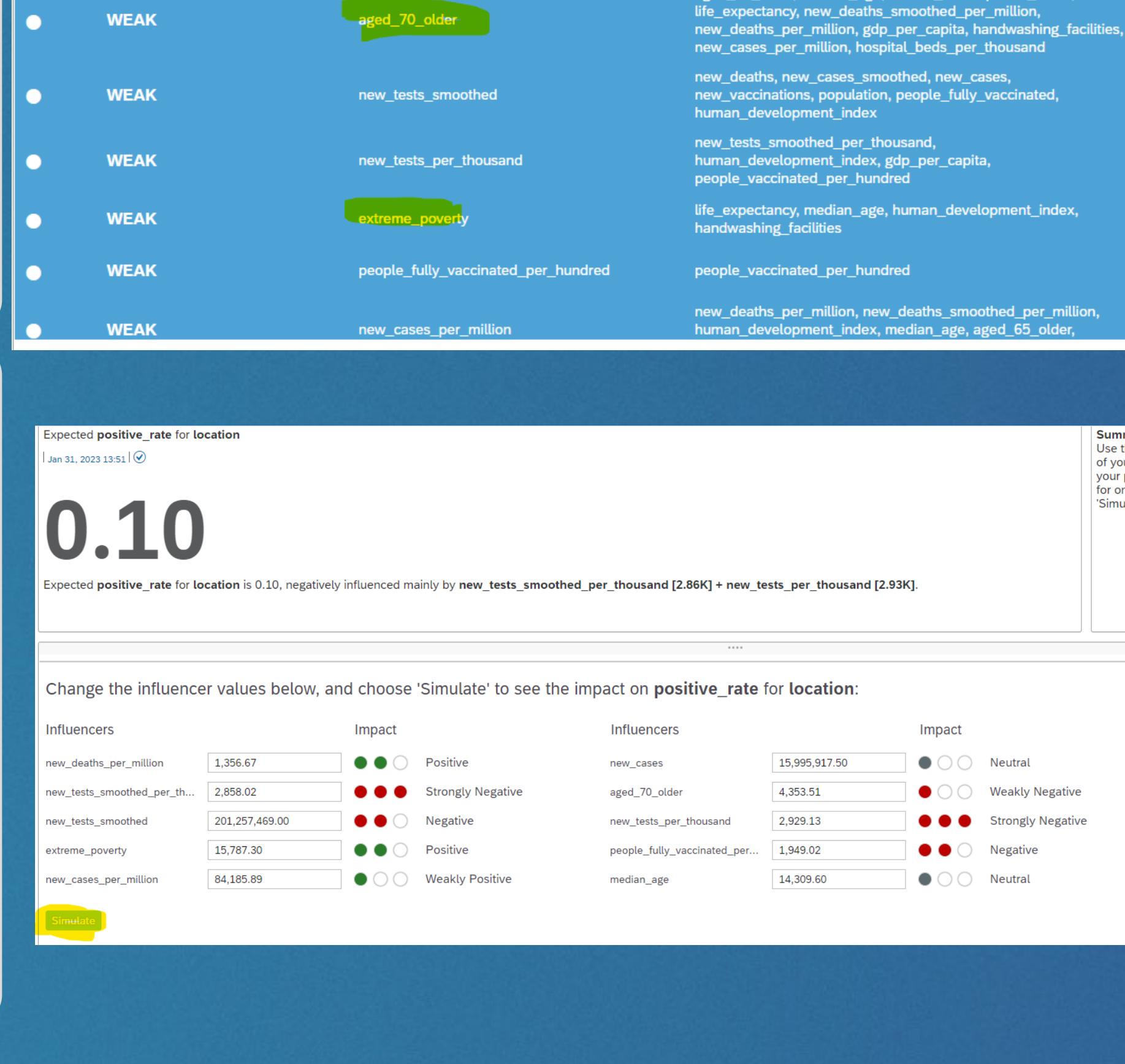
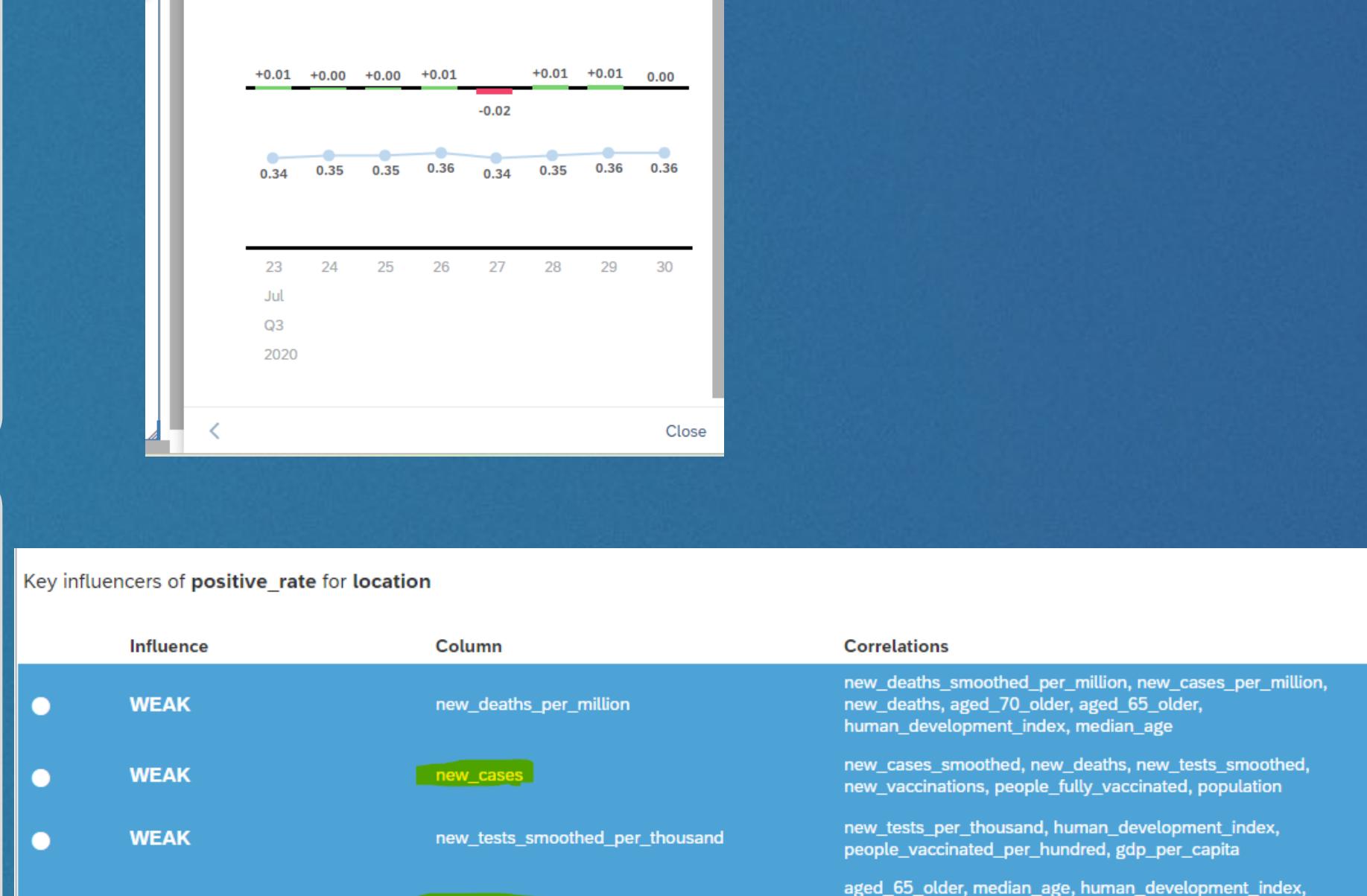
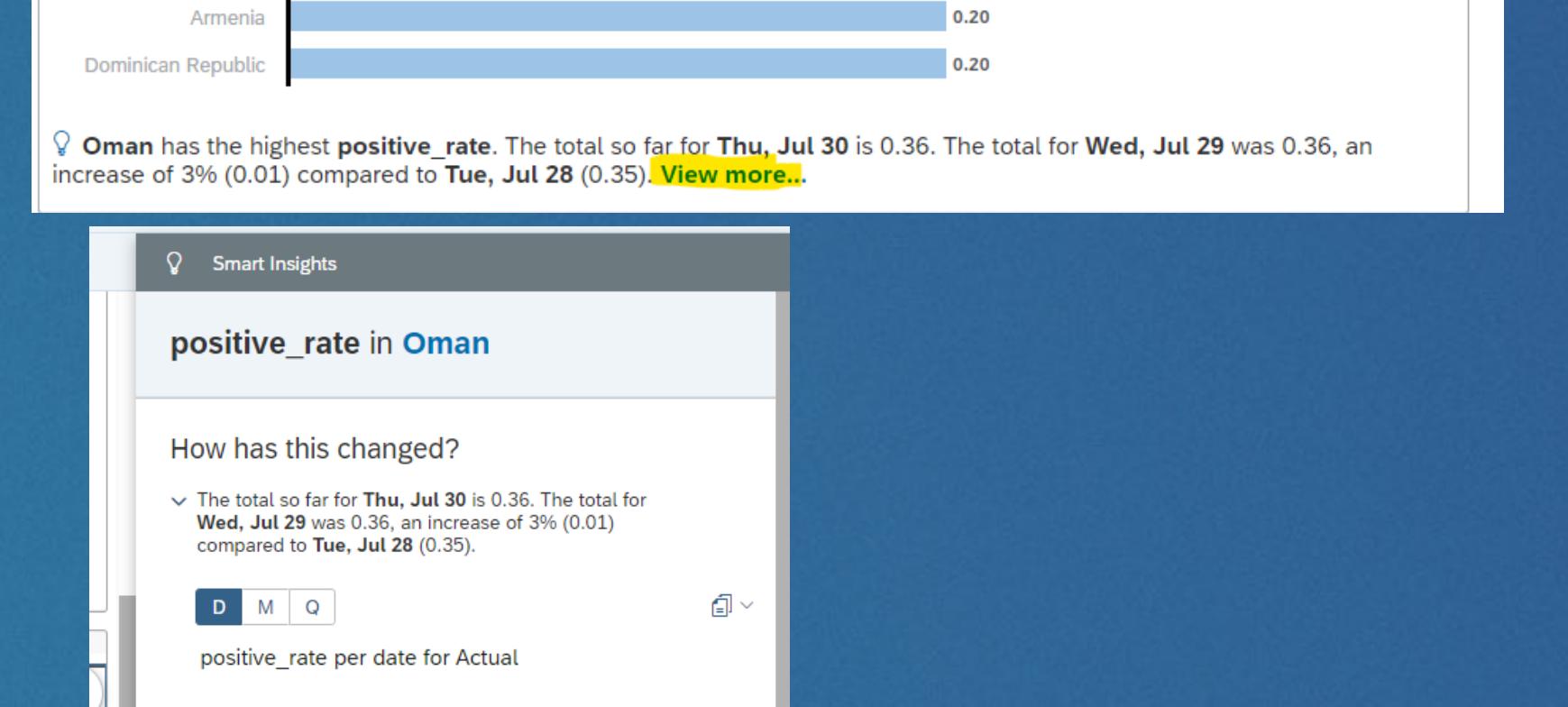
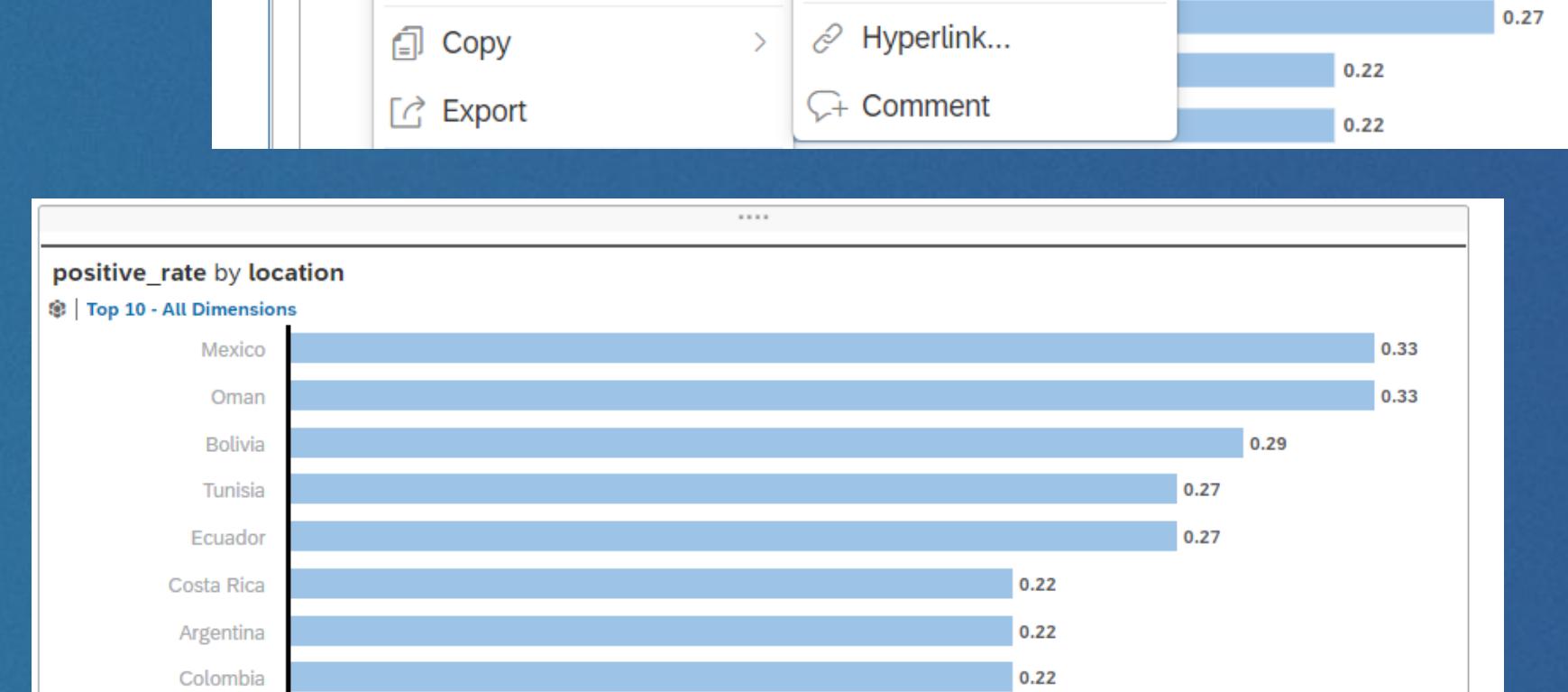
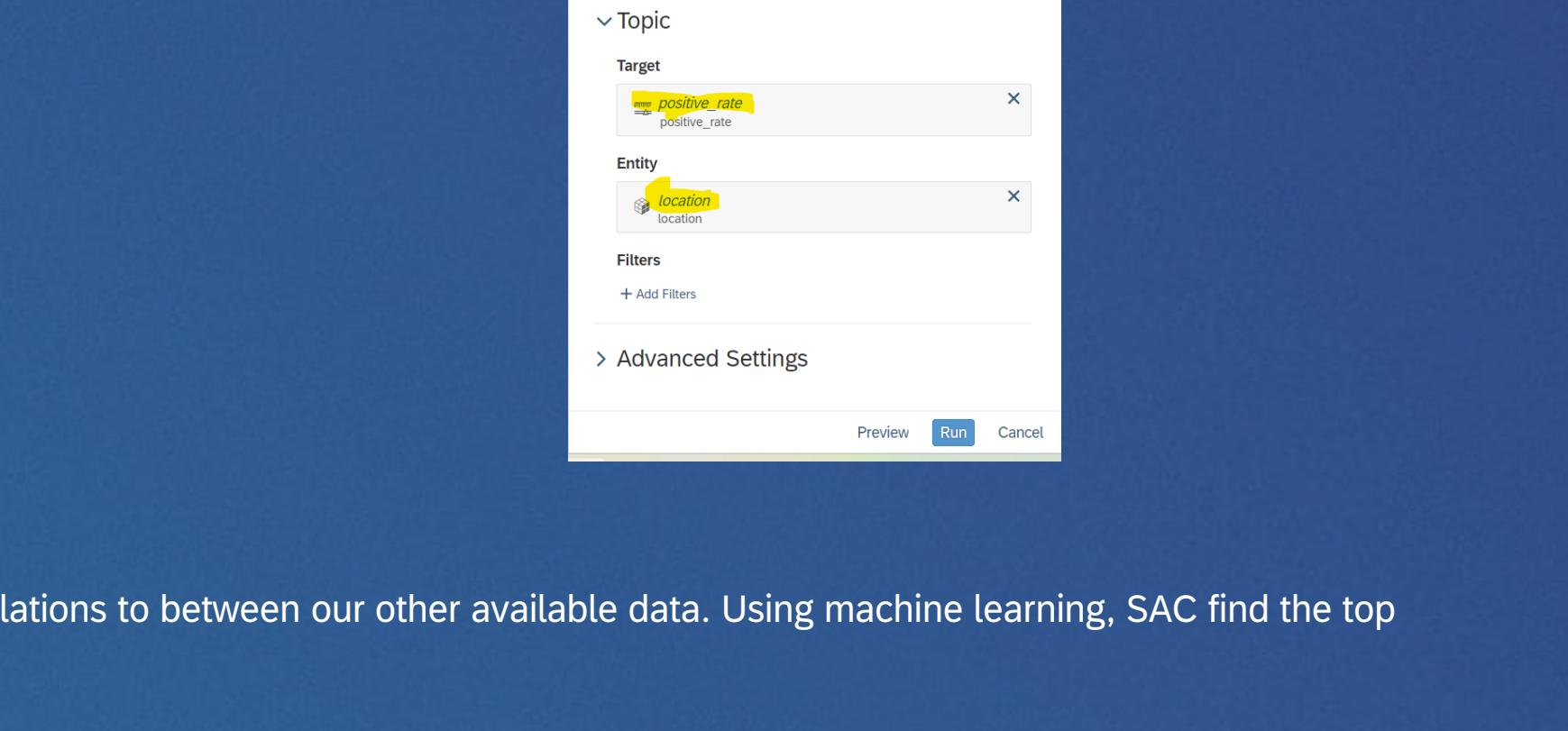
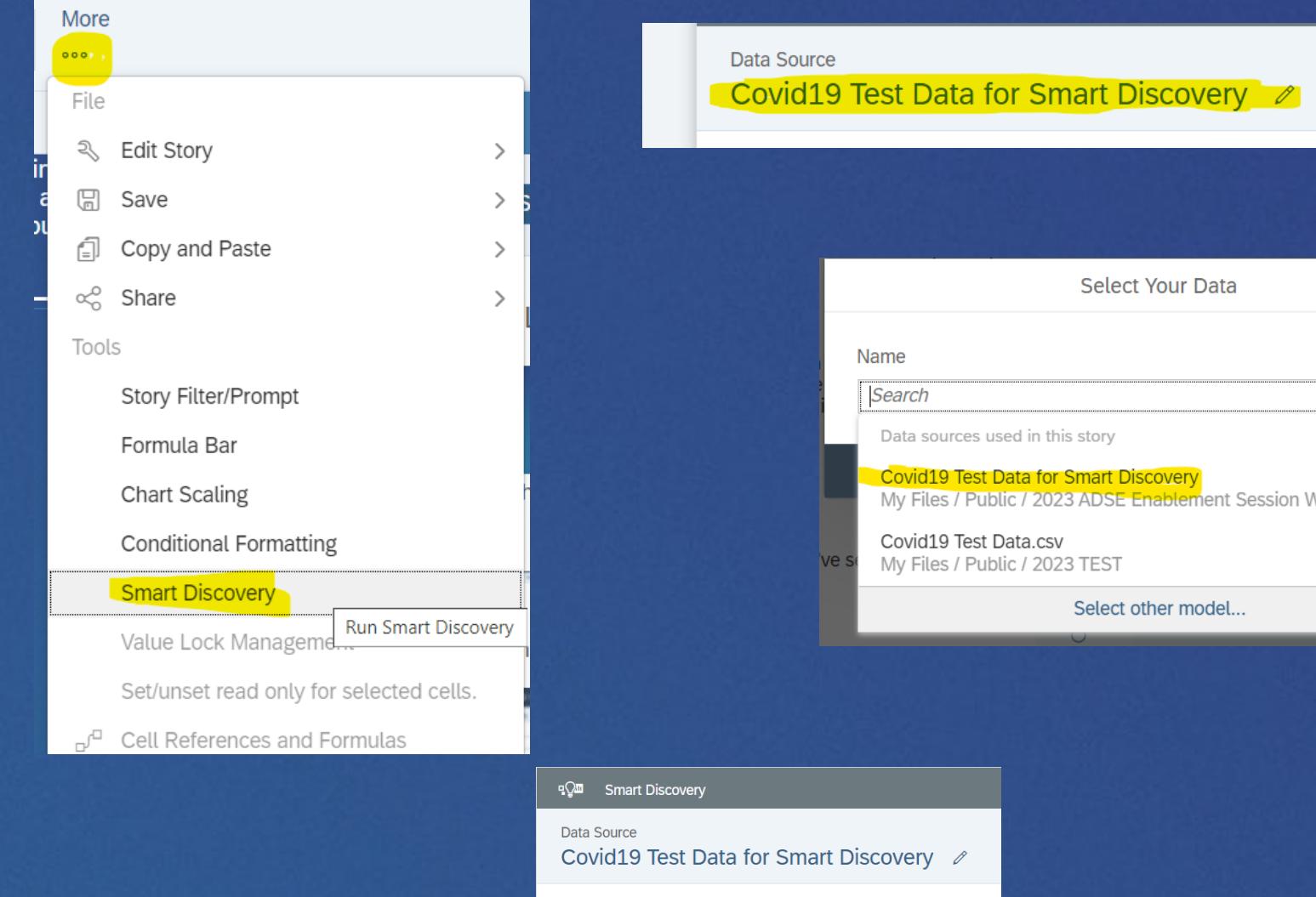
For example, what happens if a country has a higher 70+ years population? What if they have lower poverty line?

Experiment by yourself and find out.

What is the lowest positivity rate you can come up with by changing the different parameters?

Type on Zoom chat.

When finished, feel free to run additional Smart Discoveries using other target measures this time.



You are now done with this section.

In this exercise, you will learn how to:

- Create charts
- Populate charts
- Select fields (measures & dimensions)
- Rank data
- Duplicate charts

!! Use the guided instructions to the left to follow the exercises

!! Please remember to save often

How to save your work:

1. Click File > Save > Save as...
2. Name your file as MMDDYYYY_YOUR INITIALS
3. Example: 01312023_AAA

Let's create our first chart!

1. Click on the yellow chart to the right
2. Make sure we are on **Designer** mode at the upper right tab, and under **Builder**
3. Make sure the selected data model is **Covid 19 Test Data**
4. Select **Line Graph** under the Trend category
5. Select **Total Confirmed Cases** as your measure
6. Select **Date** as your dimension
7. To the right of the **Date dimension**, select Set Hierarchy icon (3rd icon)
8. **Set Hierarchy to level 4**

Which date had the highest cases?

Type your answers in the Zoom chat.

LEARNING: What is a measure and dimension?

Measures and dimensions are core concepts within analytics tools such as SAC. All data within a dataset must be categorized as either a measure or dimension.

Measures contain numeric, quantitative values that you can measure or count. Measures can be aggregated (sum, average, median, etc.). In this dataset, number of Covid cases, deaths and recoveries are measures. Other examples of measures: literacy rate, poverty rate, GDP, salary, costs, population, etc.

Dimensions contain qualitative values (such as names, dates, or geographical data). You can use dimensions to categorize, segment, and reveal the details in your data. In this dataset, country/location and date are considered dimensions. Other examples: gender, educational status, marital status, profession, income bracket.

Note: age can be a measure, but age bracket is a dimension

Let's create another chart.

1. Click on empty chart to the right
2. Make sure we are on **Designer** mode at the upper right tab, and under **Builder**
3. Make sure the selected data model is **Covid 19 Test Data**
4. Select the **Bar/Column** graph under Comparison category
5. Measure: **Total Confirmed Cases**
6. Dimension: **Country/Region**
7. Click on the chart, and click on the "..." icon to its upper right
8. Hover your mouse over **Rank** → **Country/Region** → **Top 5**

Which are the top 5 countries with the highest cases?

Type your answers in the Zoom chat.

Let's repeat the steps for Recovered cases

1. Click on second chart above "Which countries have the highest COVID19 cases?"
2. Click on the "..." icon to its upper right
3. Hover your mouse over **Copy** → **Duplicate**
4. Align the chart to be parallel to the one above (you can drag the chart)
5. Make sure we are on **Designer** mode at the upper right tab, and under **Builder**
6. Under measures, deselect **Total Confirmed Cases**
7. Under measures, select **Deaths**
8. Click on the chart, and click on the "..." icon to its upper right
9. Hover your mouse over **Rank** → **Country/Region** → **Top 5**

Which are the top 5 countries with the highest deaths?

Type your answers in the Zoom chat.

What has been the trend of COVID19 since 2020?

Subtitle

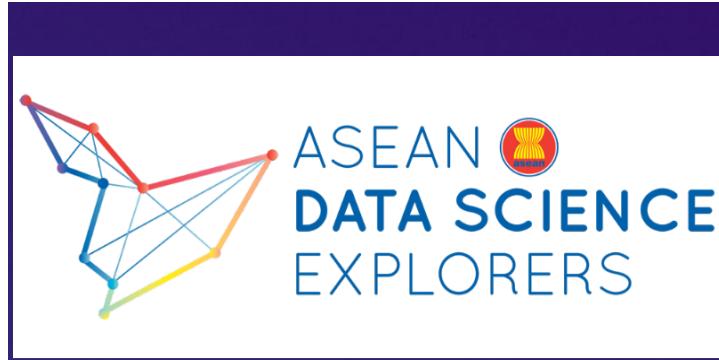
More measures are required to build a Line chart

Which countries have the highest COVID19 cases?

1 Filter

More measures are required to build a Bar/Column chart

You are now done with this section.



15:06

February 10, 2023

In this exercise, you will learn how to:

- Open data explorer
- Perform ad hoc analysis
- Filter and select specific countries
- Sort from highest to lowest
- Copy a chart to your story

!! Use the guided instructions to the left to follow the exercises

!! Please remember to save often

How to save your work:

1. Click File > Save > Save as...
2. Name your file as MMDDYYYY_YOUR INITIALS
3. Example: 01312023_AAA

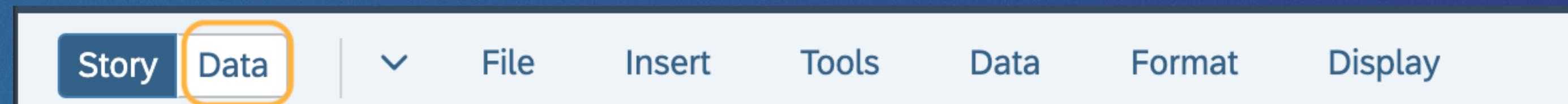
This section will familiarize you with **Data Explorer mode** to answer questions on the fly

What is Data Explorer?

The Explorer helps you perform ad-hoc exploration and gain insights into your data that would normally be time-consuming if using traditional chart-making.

How do I open Data Explorer?

At the upper left of your screen, click on **DATA** to open the Data Explorer interface.



Upon opening Data Explorer, you can now play around by selecting a combination of measures and dimensions and visualize the data.

💡 Try visualizing data for your own country by selecting it under the Country/Region tile

Measure: Total Confirmed Cases

Country/Region: [Your Country]

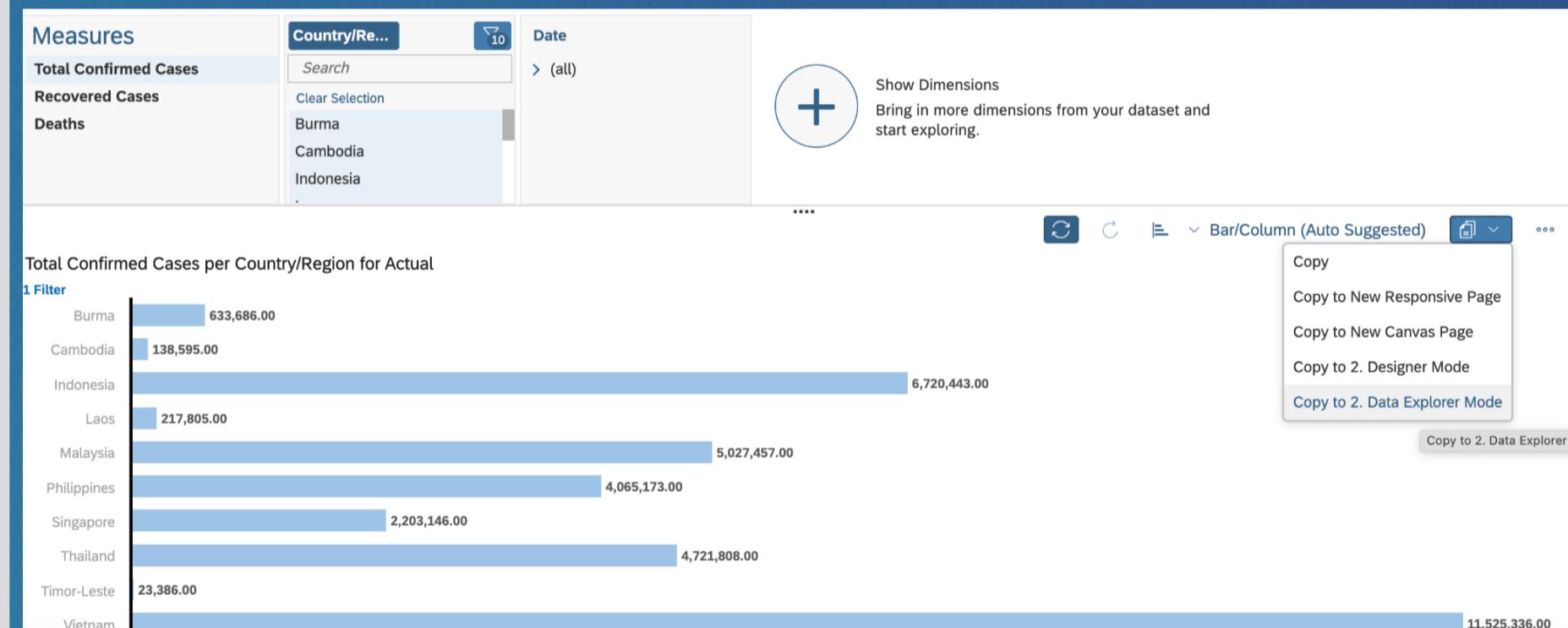
Date: Click on Date



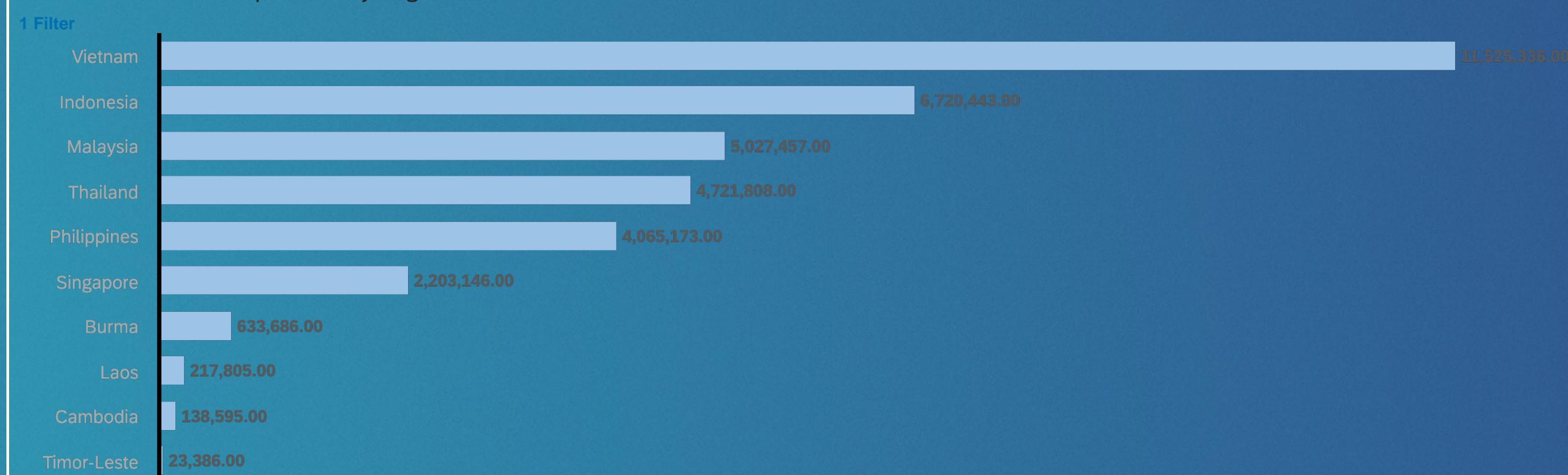
Let's use Data Explorer mode to compare COVID19 cases across SEA countries

1. Measure: Total Confirmed Cases.
2. Click on **Country/Region** to make sure it is highlighted.
3. Search and select the SEA countries below (Brunei not in dataset).
 - [Burma](#)
 - [Cambodia](#)
 - [Timor-Leste](#)
 - [Indonesia](#)
 - [Laos](#)
 - [Malaysia](#)
 - [Philippines](#)
 - [Singapore](#)
 - [Thailand](#)
 - [Vietnam](#)
4. Copy the generated chart to the **Data Explorer Mode** page.
5. Find the chart back in **Story** mode. It may be on the left side of the screen.
6. You may drag the border of the chart box to move it. You may resize it by dragging the bottom right corner.
7. Click on the "..." icon to its upper right.
8. Hover your mouse over **Sort ➔ Total Confirmed Cases ➔ Highest to Lowest**.

Your graph should look like the one below.
Which country has the highest confirmed cases? Type on Zoom chat.



Total Confirmed Cases per Country/Region for Actual



You are now done with this section.

In this exercise, you will learn how to:

- Open styling panel
- Change background Color
- Change fonts
- Change number format and scale

!! Use the guided instructions to the left to follow the exercises

!! Please remember to save often

How to save your work:

1. Click File > Save > Save as...
2. Name your file as
MMDDYYYY_YOUR INITIALS
3. Example: 01312023_AAA

This section will familiarize you with **Styling mode** to create visually pleasing charts.

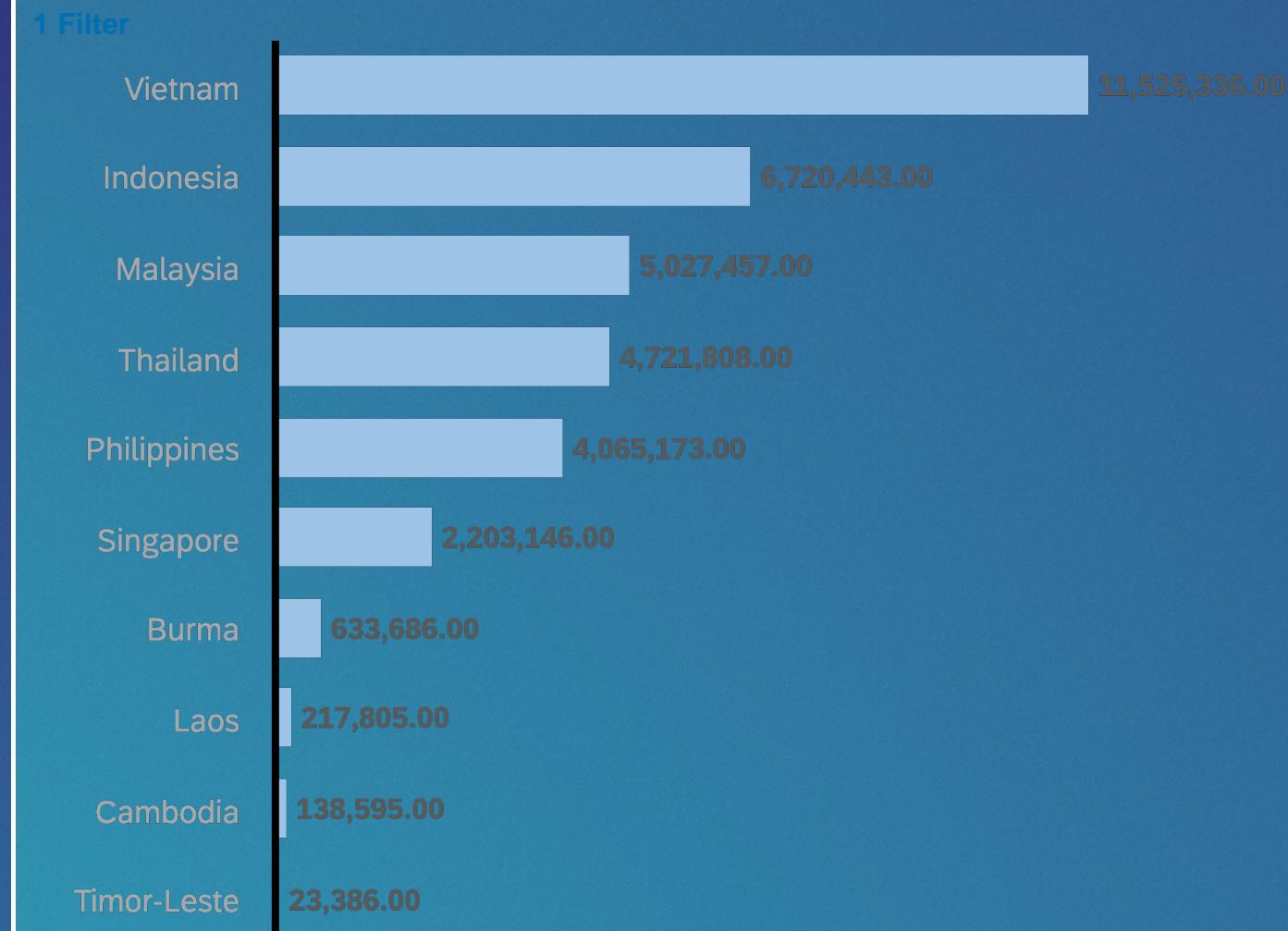
How is Styling mode different from Builder mode?

Builder mode primarily affects the content and data populating a given chart. Styling affects the look and feel, formatting, colors, fonts & sizes that you can adjust to aesthetically enhance your storyboards.

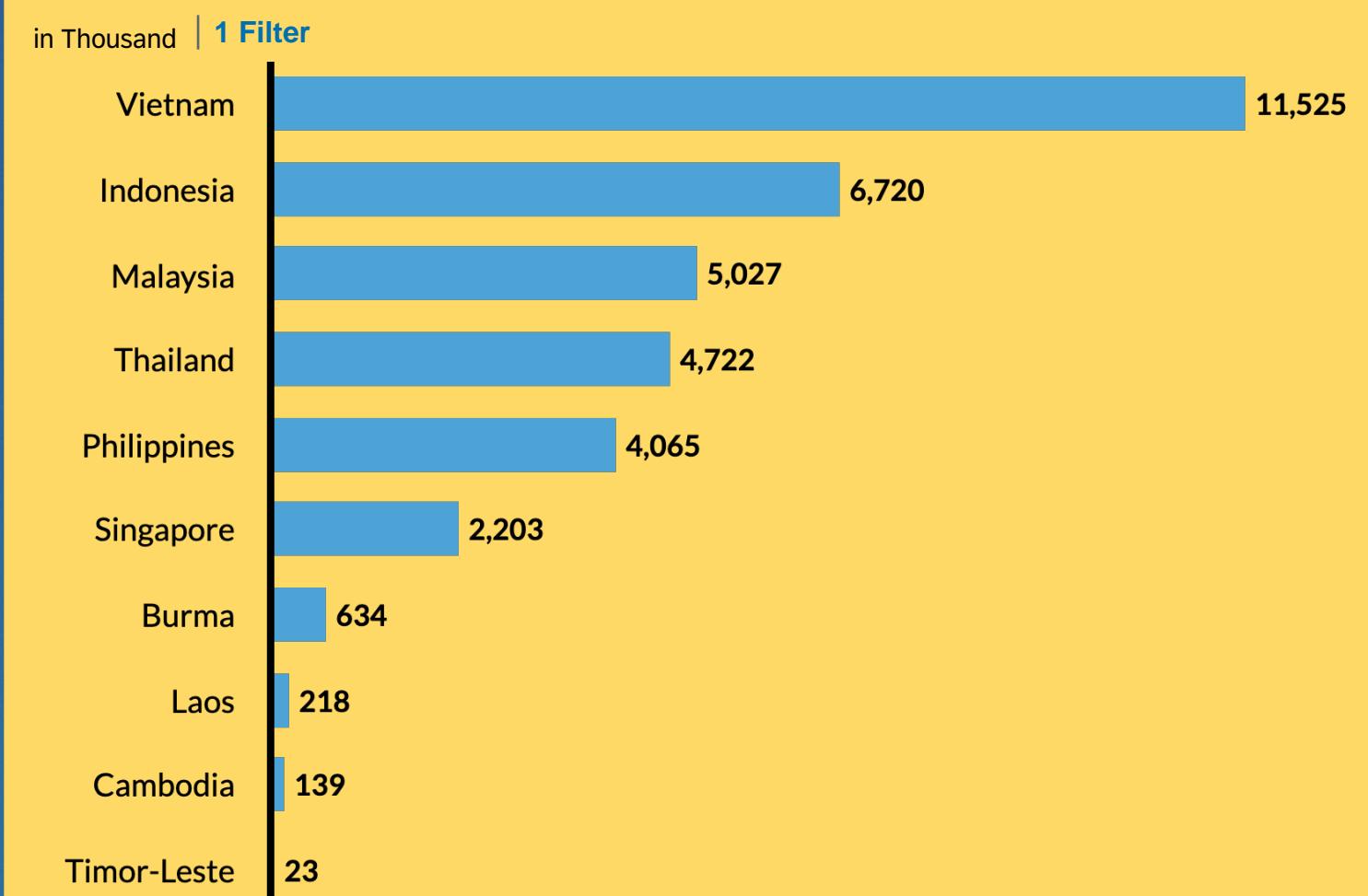
Note: Bar graph colors are changed in Builder mode, not Designer mode.

In this exercise, we will transform the graph from the previous exercise, to the one on the right.

Total Confirmed Cases per Country/Region for Actual



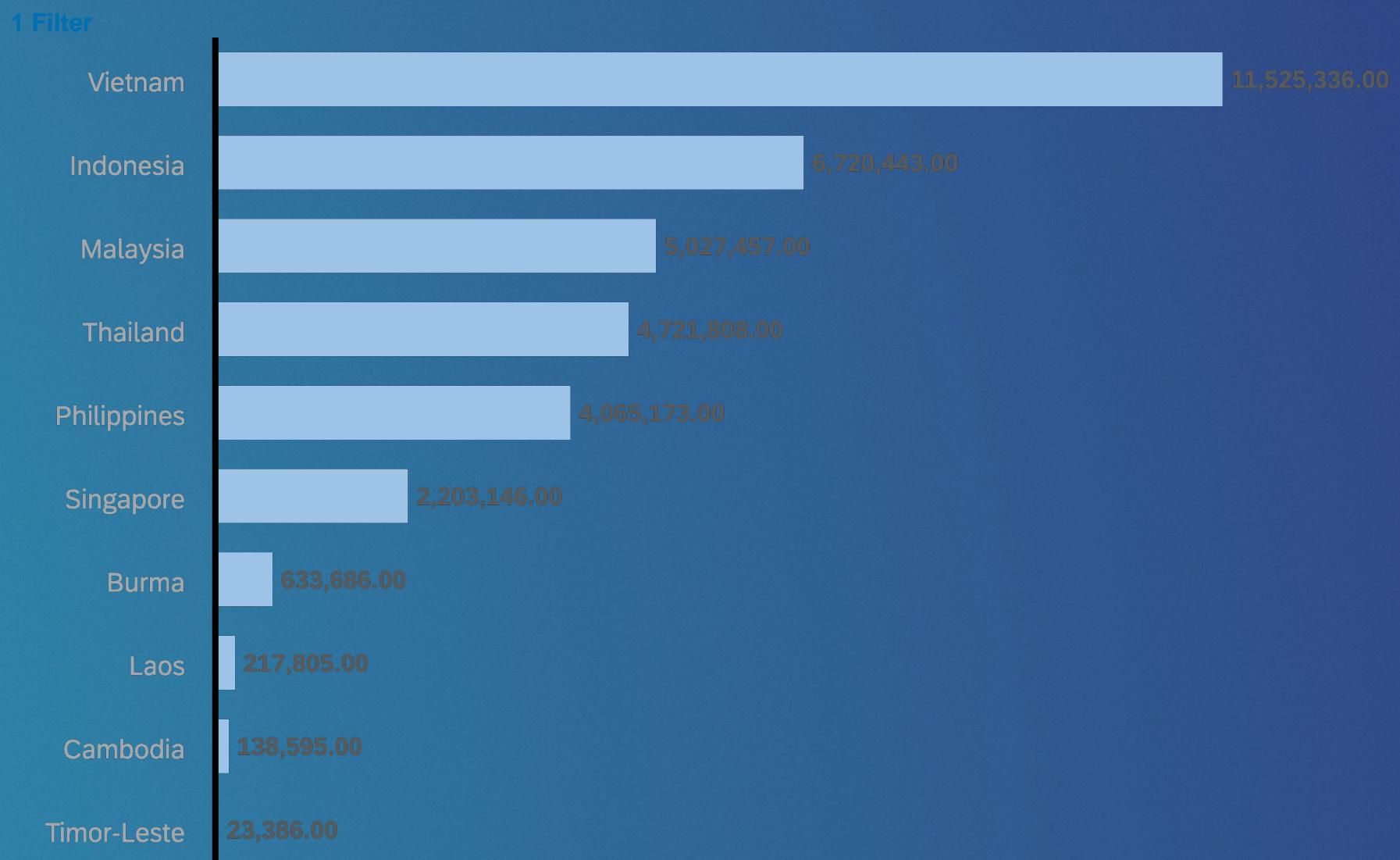
Breakdown of COVID19 cases across Southeast Asian countries



Let's first give the graph a background color to make it more visible.

1. Click on the chart on the left ←.
2. Click on **Designer** on the upper right of the screen ↗.
3. Instead of Builder mode (screw and wrench icon blue header) click on **Styling** (paintbrush icon with orange header)
4. Choose a good contrast **background color** (such as yellow)
5. You may choose to add **borders** to make your chart look cleaner. Play around with the colors, corner radius, and line width.
6. Font: Pick a good clean font, such as Lato, to make the text more legible. You may want to increase the font size of the Chart Title to 18+.
7. Number Format: Change the Scale to **thousand**, to decrease the amount of digits displayed on screen. You may also choose to reduce decimal places to **zero**.

Total Confirmed Cases per Country/Region for Actual



You are now done with this section.

Overview of positive_rate for location

Total positive_rate for location

0.09

Total number of location

117

Minimum value

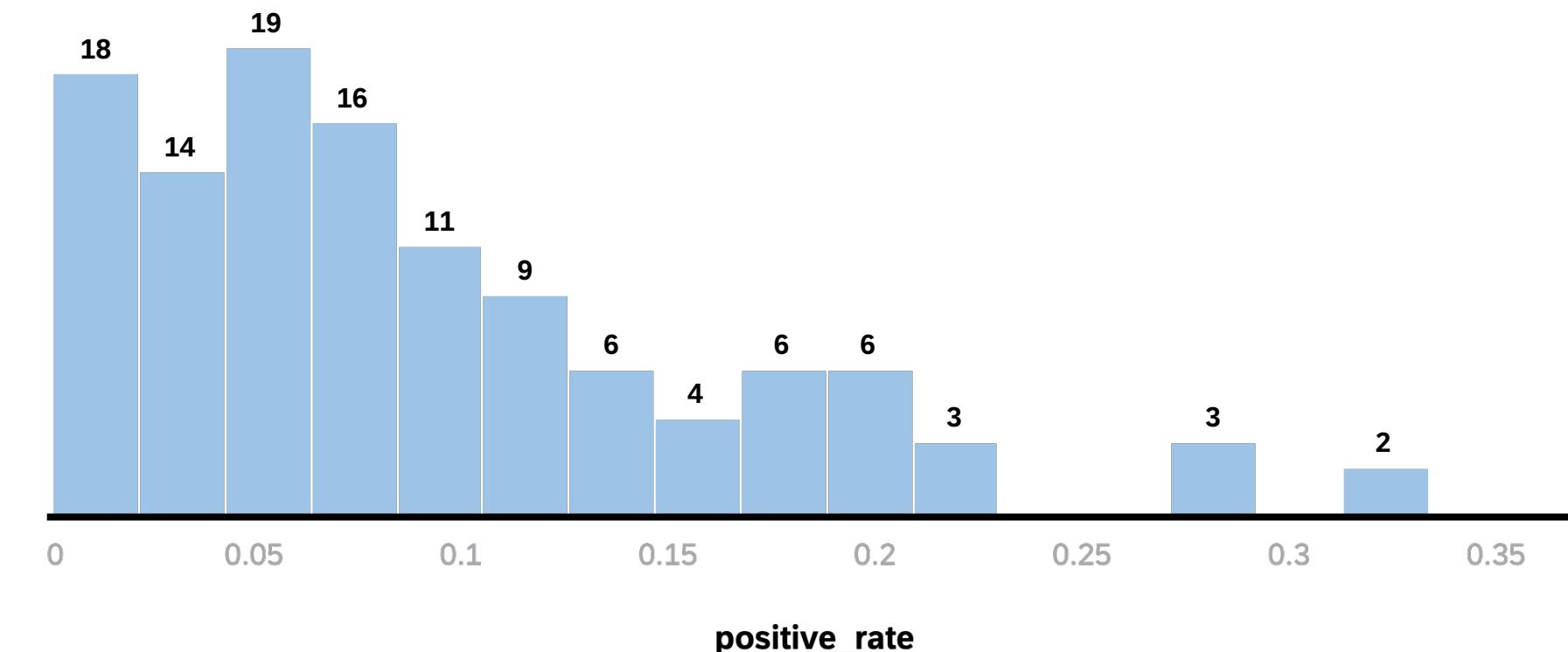
+0.00

Maximum value

0.33

Distribution of location by positive_rate

16 Bins



positive_rate over time

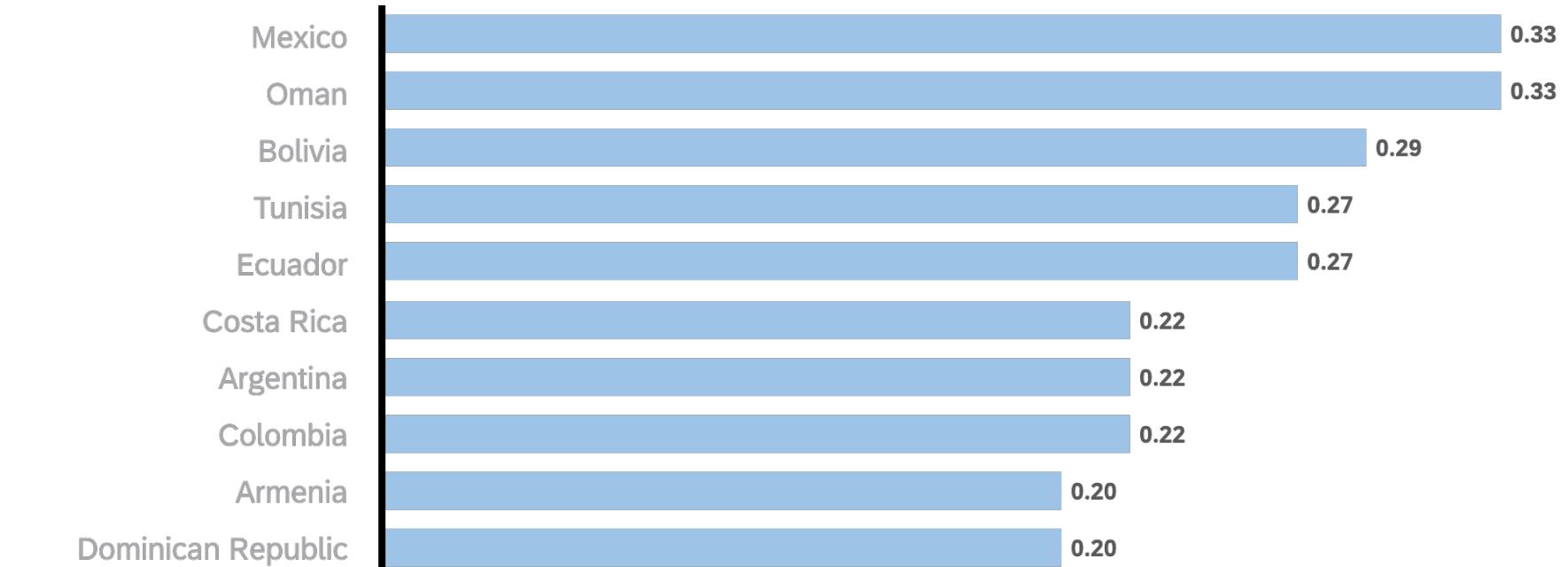
Forecast Loading...

1M 3M 6M YTD 1Y All



positive_rate by location

Top 10 - All Dimensions



💡 Oman has the highest positive_rate. The total so far for Thu, Jul 30 is 0.36. The total for Wed, Jul 29 was 0.36, an increase of 3% (0.01) compared to Tue, Jul 28 (0.35). [View more...](#)

Key influencers of positive_rate for location

About this Smart Discovery

This Smart Discovery analyzed **positive_rate** for **location** from **Covid19 Test Data for Smart Discovery**. It identified **10** columns as key influencers. These key influencers are based on a snapshot of your data from **Jan 31, 2023**.

We aggregate all the relevant measures and dimensions to the level of the entity.

The underlying machine learning model indicates that the analysis quality is good.

How to interpret the charts

The charts below are based on your live data. In the first chart, select a key influencer to analyze the impact it has on **positive_rate** for **location**. In the second chart, select an additional key influencer to understand the relationship between both key influencers, and their impact on **positive_rate** for **location**.

Key influencers of positive_rate for location

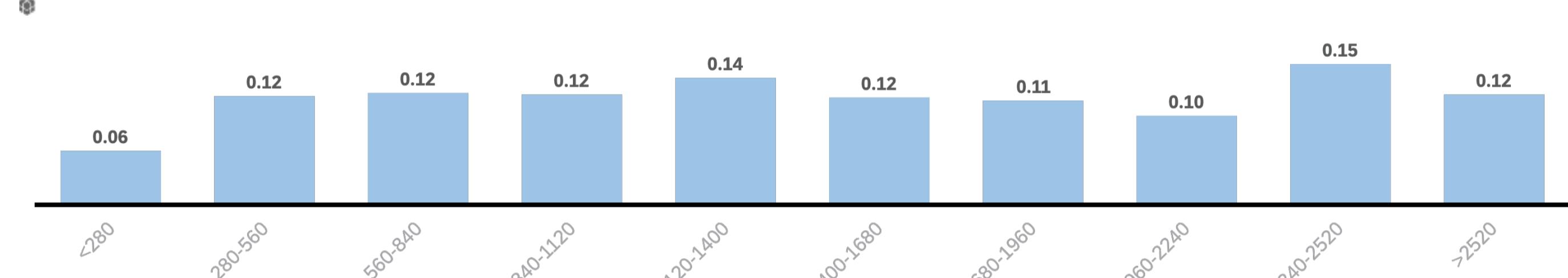
Influence	Column	Correlations
● WEAK	new_deaths_per_million	new_deaths_smoothed_per_million, new_cases_per_million, new_deaths, aged_70_older, aged_65_older, human_development_index, median_age
● WEAK	new_cases	new_cases_smoothed, new_deaths, new_tests_smoothed, new_vaccinations, people_fully_vaccinated, population
● WEAK	new_tests_smoothed_per_thousand	new_tests_per_thousand, human_development_index, people_vaccinated_per_hundred, gdp_per_capita
● WEAK	aged_70_older	aged_65_older, median_age, human_development_index, life_expectancy, new_deaths_smoothed_per_million, new_deaths_per_million, gdp_per_capita, handwashing_facilities, new_cases_per_million, hospital_beds_per_thousand
● WEAK	new_tests_smoothed	new_deaths, new_cases_smoothed, new_cases, new_vaccinations, population, people_fully_vaccinated, human_development_index
● WEAK	new_tests_per_thousand	new_tests_smoothed_per_thousand, human_development_index, gdp_per_capita, people_vaccinated_per_hundred
● WEAK	extreme_poverty	life_expectancy, median_age, human_development_index, handwashing_facilities

Select a key influencer from **List A** to see how it has an impact on **positive_rate** for **location**

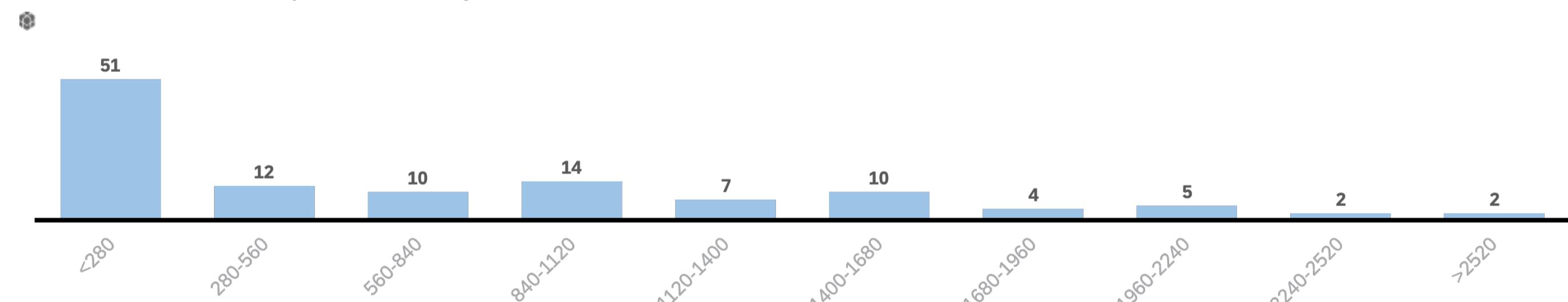
List A

- new_deaths_per_million bin
- new_cases bin
- new_tests_smoothed_per_thousand bin
- aged_70_older bin
- new_tests_smoothed bin
- new_tests_per_thousand bin
- extreme_poverty bin
- people_fully_vaccinated_per_hundred bin
- new_cases_per_million bin
- median_age bin

Average positive_rate for location by new_deaths_per_million bin



Distribution of location by new_deaths_per_million bin

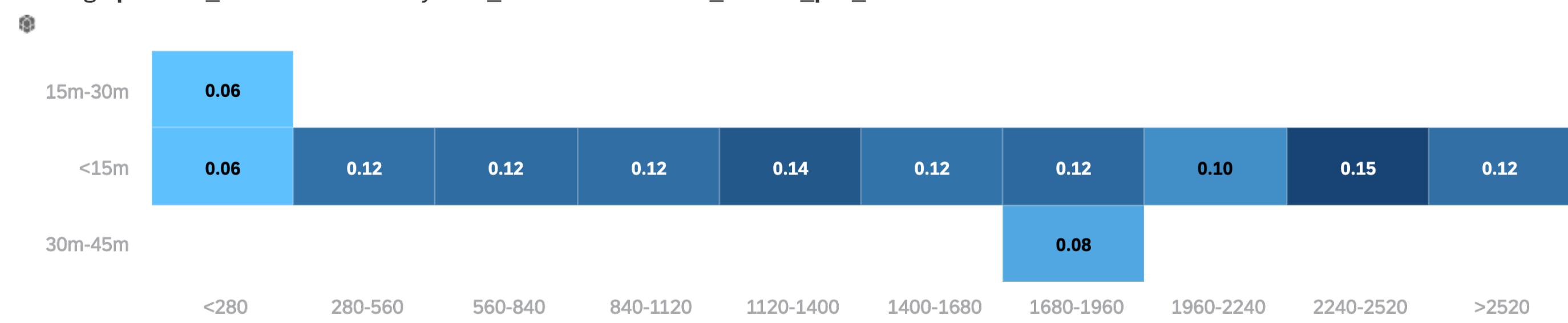


Select a different key influencer from **List B** to see how the relationship between it and **new_deaths_per_million bin** has an impact on **positive_rate** for **location**

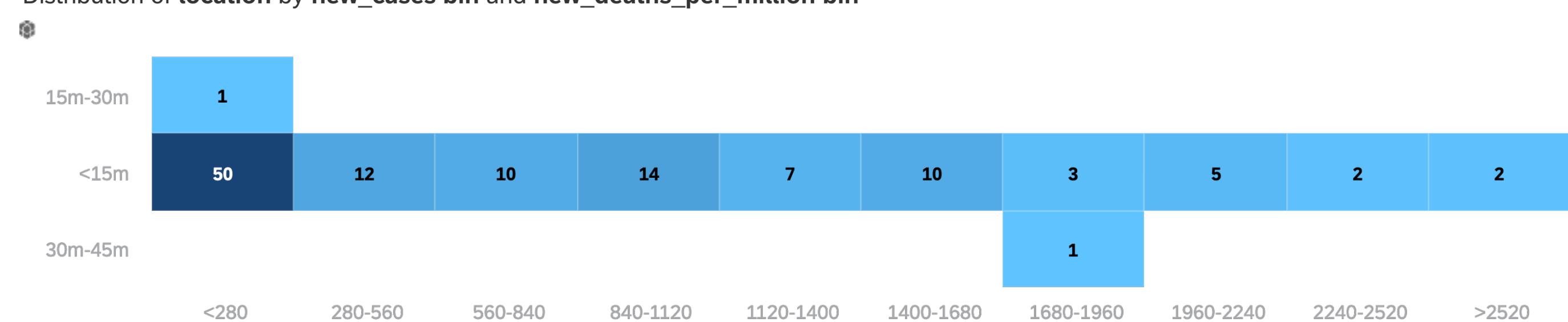
List B

- new_deaths_per_million bin
- new_cases bin
- new_tests_smoothed_per_thousand bin
- aged_70_older bin
- new_tests_smoothed bin
- new_tests_per_thousand bin
- extreme_poverty bin
- people_fully_vaccinated_per_hundred bin
- new_cases_per_million bin
- median_age bin

Average positive_rate for location by new_cases bin and new_deaths_per_million bin



Distribution of location by new_cases bin and new_deaths_per_million bin



How my influencers have an impact on **positive_rate** for location

Expected **positive_rate** for location

| Jan 31, 2023 13:51 | 

0.08

-20%

Expected **positive_rate** for location is 0.08, negatively influenced mainly by **new_tests_smoothed_per_thousand [5.35K]** +

Summary

Use the Simulation view to discover how changing the values of your key influencers could have an impact on the value of your **positive_rate** for location. Simply specify a new value for one or more of your key influencers, and choose the 'Simulate' button.

Change the influencer values below, and choose 'Simulate' to see the impact on **positive_rate** for location:

Influencers

new_deaths_per_million

1,356.67



Positive

new_cases

8,135,896.85



Neutral

new_tests_smoothed_per_thou:

5,345.45



Strongly Negative

aged_70_older

4,353.51



Weakly Negative

new_tests_smoothed

201,257,469.00



Negative

new tests per thousand

2.929.13



Strongly Negative

Simulate