### **BIA WORKSHOP – POWER BI USING DAX LAB 3 - Data analytics and DAX**

### Prepared exclusively for 2023 BIA Graduating Seniors (with an INFO 3300 prereq)

## May 2023

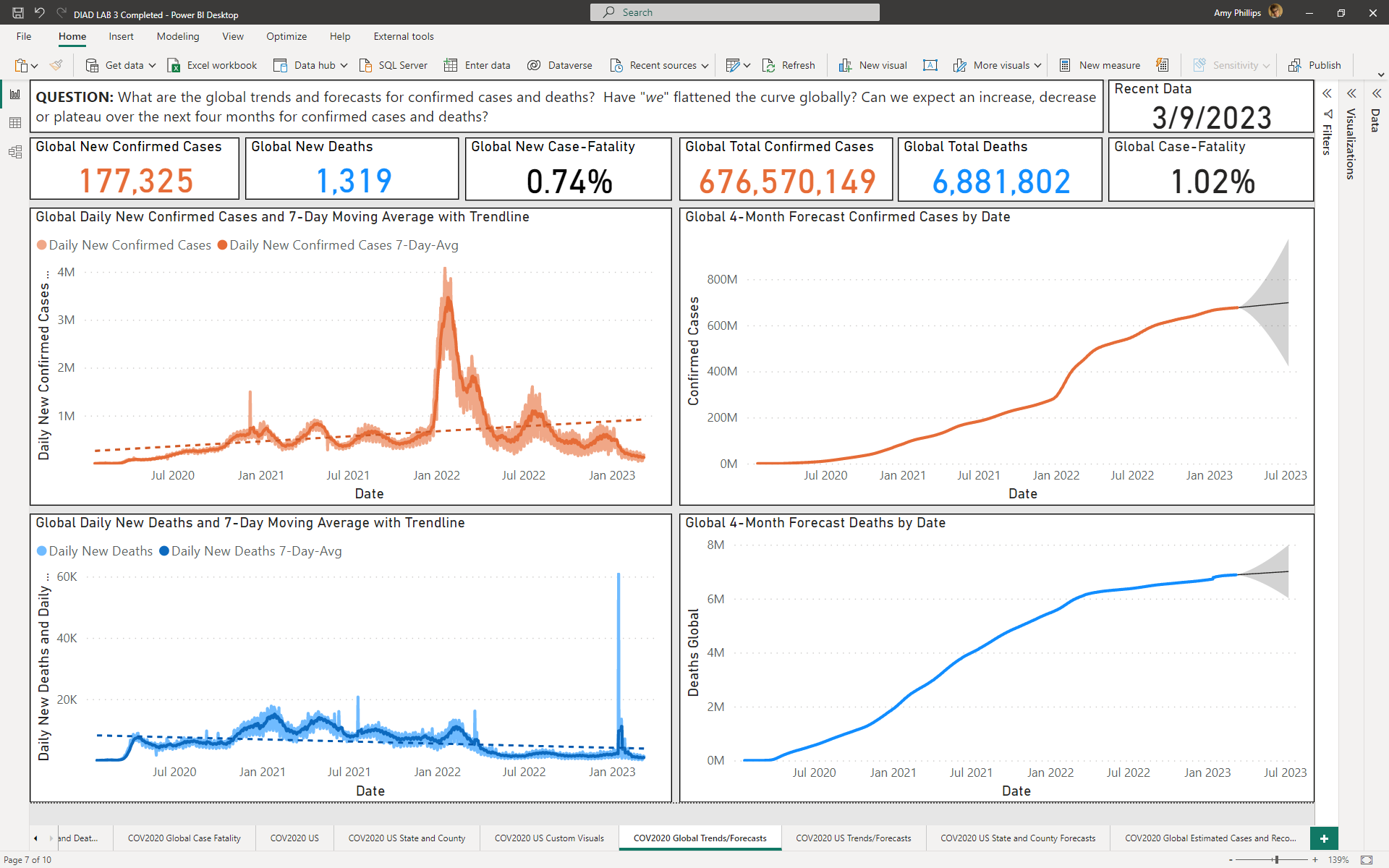
## **LAB RESOURCES**

|  |  |
| --- | --- |
|  | DIAD LAB 2.pbix |

**LAB PURPOSE AND LEARNING OBJECTIVES:** For this lab, we focus on **using the analytics and (what-if) parameters option**.

## **LESSON 1: THE ANALYTICS OPTIONS USING THE GLOBAL DATASET**

1. Create a **new report/page** (e.g., dashboard), we will **try to replicate** (mainly using the COV2020\_CSSE dataset):



1. **Start with a question** (include as a textbox at the top of the dashboard)
   1. **QUESTION:** What are the global trends and forecasts for confirmed cases and deaths? Have “*we*” flattened the curve globally? Can we expect an increase, decrease or plateau over the next four months for confirmed cases and deaths?

**NOTE 1:** To calculate the **Daily New Confirmed Cases 7-Day Avg** and **Daily New Deaths 7-Day Avg**, use the following two DAX measures (apply to the **COV2020\_CSSE** dataset)[[1]](#footnote-1):

Daily New Confirmed Cases 7-Day-Avg Global = AVERAGEX(DATESBETWEEN(COV2020\_CSSE[Date], MAX('COV2020\_CSSE'[Date])-7, MAX('COV2020\_CSSE'[Date])), [Daily New Confirmed Cases Global])

Daily New Deaths 7-Day-Avg Global = AVERAGEX(DATESBETWEEN(COV2020\_CSSE[Date], MAX('COV2020\_CSSE'[Date])-7, MAX('COV2020\_CSSE'[Date])), [Daily New Deaths Global])

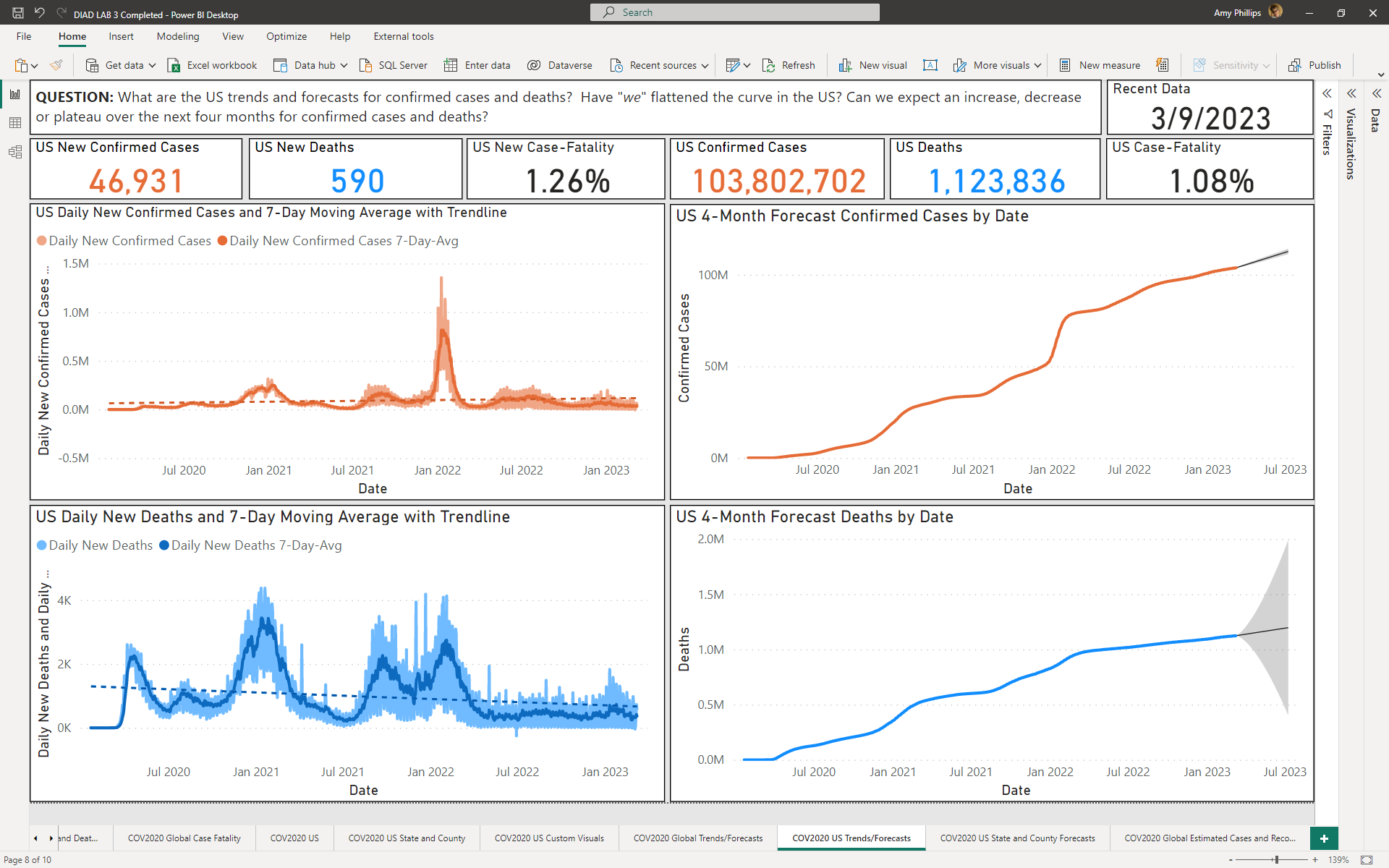
**NOTE 2:** To calculate the **New** **Case-Fatality**, use the following DAX measure (apply to the **COV2020\_CSSE** dataset):

Daily New Case-Fatality Global = DIVIDE([Daily New Deaths Global],[Daily New Confirmed Cases Global])

1. **Rename** the report/page **COV2020 Global Trends/Forecasts**
2. **Save** your Power BI file as **DIAD LAB 3**

## **LESSON 2: THE ANALYTICS OPTIONS USING THE US DATASET**

1. Create a **new report/page** (e.g., dashboard), we will **try to replicate:**



1. Start with a question (include as a textbox at the top of the dashboard)
   1. **QUESTION:** What are the US trends and forecasts for confirmed cases and deaths? Have "*we*" flattened the curve in the US? Can we expect an increase, decrease or plateau over the next four months for confirmed cases and deaths?

**NOTE 1: Duplicate** the **COV2020 Global Trends/Forecasts** dashboard to apply to the **COV2020\_CSSE\_US** data

**NOTE 2:** To calculate the **Daily New Confirmed Cases 7-Day Avg**, use the following DAX measure (apply to the **COV2020\_CSSE\_US** dataset):

Daily New Confirmed Cases 7-Day-Avg US = AVERAGEX(DATESBETWEEN(COV2020\_CSSE\_US[Date], MAX('COV2020\_CSSE\_US'[Date])-7, MAX('COV2020\_CSSE\_US'[Date])), [Daily New Confirmed Cases US])

Daily New Deaths 7-Day-Avg US = AVERAGEX(DATESBETWEEN(COV2020\_CSSE\_US[Date], MAX('COV2020\_CSSE\_US'[Date])-7, MAX('COV2020\_CSSE\_US'[Date])), [Daily New Deaths US])

**NOTE 3:** To calculate the **Case-Fatality US**, use the following DAX measure (apply to the **COV2020\_CSSE\_US** dataset):

Case-Fatality US = DIVIDE([Deaths Recent Value US],[Confirmed Recent Value US])

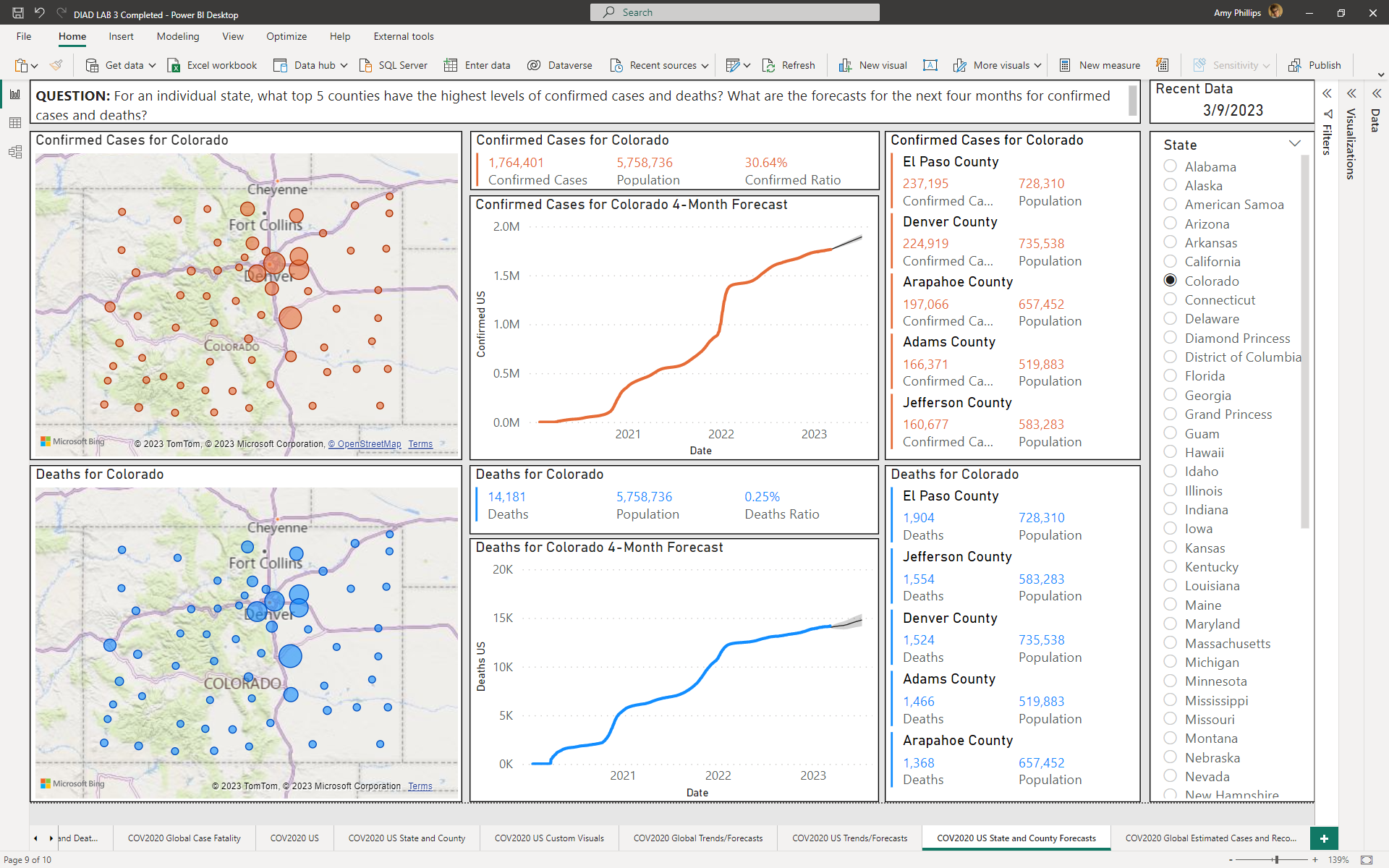
**NOTE 4:** To calculate the **New** **Case-Fatality US**, use the following DAX measure (apply to the **COV2020\_CSSE\_US** dataset):

Daily New Case-Fatality US = DIVIDE([Daily New Deaths US],[Daily New Confirmed Cases US])

1. Rename the report/page **COV2020 US Trends/Forecasts**
2. **Save** your Power BI file

## **LESSON 3: MORE ANALYTICS OPTIONS USING THE US DATASET**

1. Create a **new report/page** (e.g., dashboard), we will **try to replicate:**



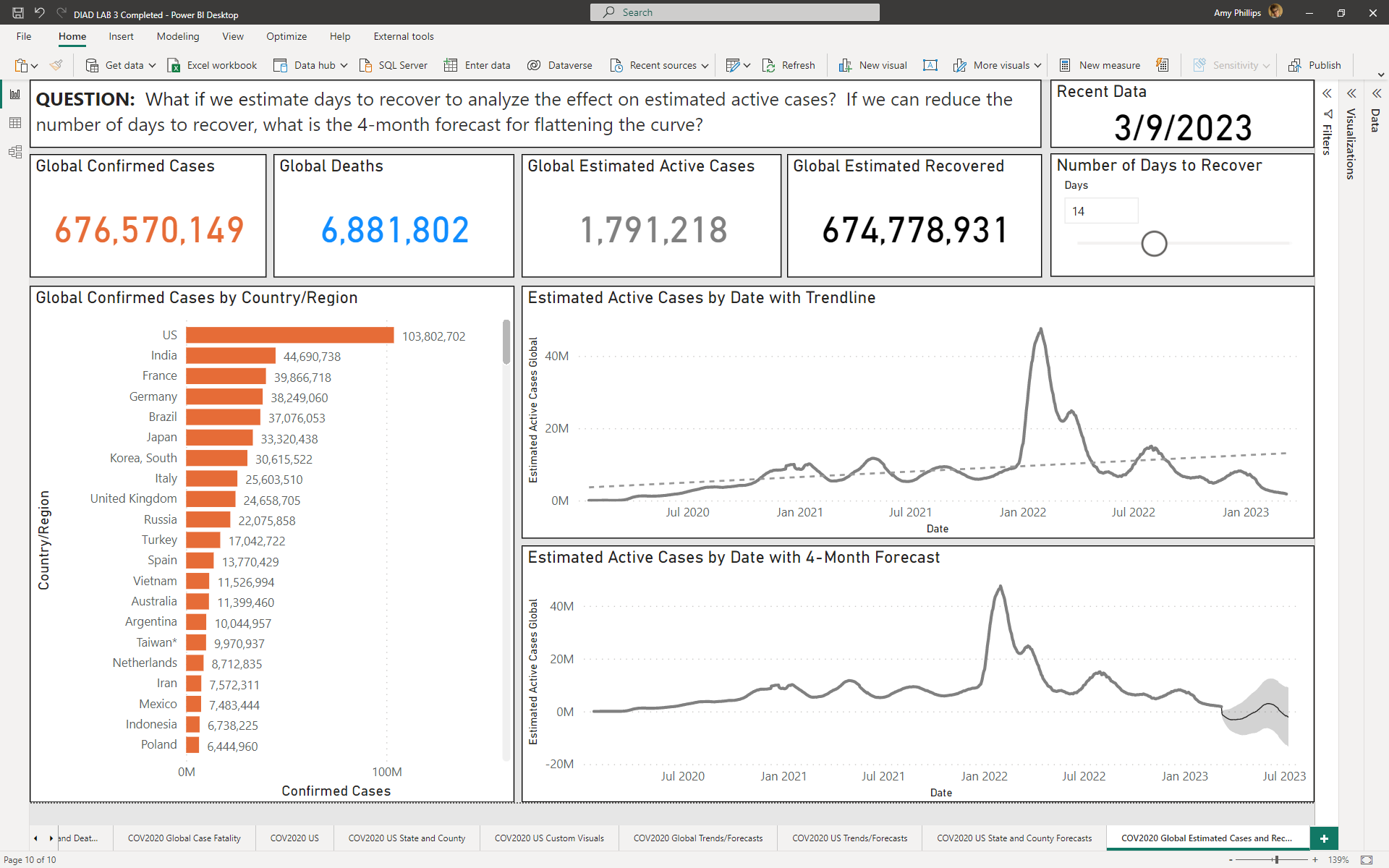
**NOTE: Duplicate the COV2020 US State and County report from Week 7 apply forecasts to that report**

1. Start with a question (include as a textbox at the top of the dashboard)
   1. **QUESTION:** For an individual state, what top 5 counties have the highest levels of confirmed cases and deaths? What are the forecasts for the next four months for confirmed cases and deaths?
2. **Add a 4-month forecast to the Confirmed Cases and Deaths line charts.**
3. Rename the report/page **COV2020 US State and County Trends/Forecasts**
4. **Save** your Power BI file

## **LESSON 4: USING A FORECAST AND WHAT-IF PARAMETER**

**CONTEXT:** In Part IV, we will use a dynamic measure, Recovery in Days, selection in a slicer. This will allow us to change the estimated days to recover and see how different numbers affect the estimated active cases… a classic **what-if scenario** (e.g., prescriptive). To apply this example, we might consider evaluating different treatments and how they might affect hospital capacity (data that we do not have).

1. Create a **new report/page** (e.g., dashboard), we will **try to replicate:**

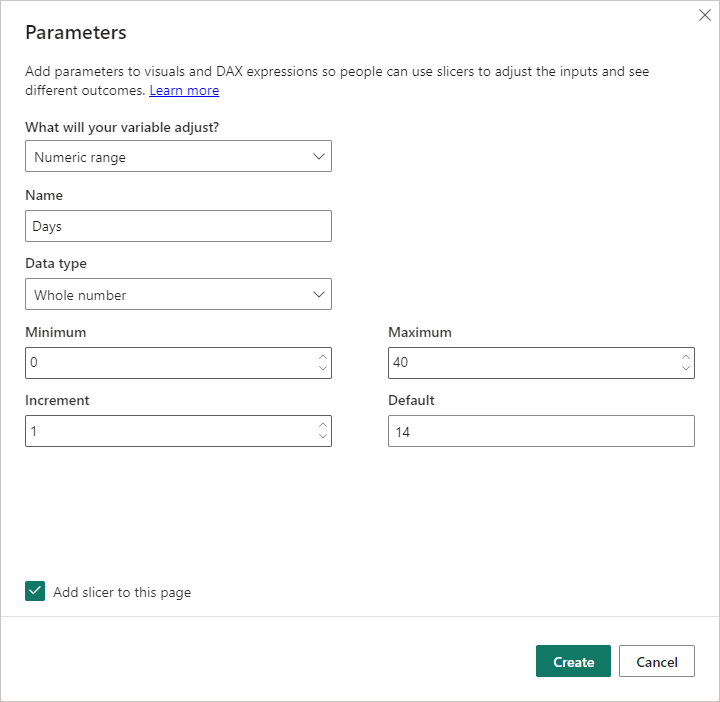


1. Start with a question (include as a textbox at the top of the dashboard)
   1. **QUESTION: What if we estimate days to recover to analyze the effect on estimated active cases? If we can reduce the number of days to recover, what is the 4-month forecast for flattening the curve?**

**NOTE 1**: The Stacked Bar chart provides the number of confirmed cases (not using a Filter this time), since the bar chart acts as a “slicer” **and** a quantitative visual.

**NOTE 2:** Great resource to understand the (obvious) waves. ***COVID waves: What causes a spike in coronavirus cases?*** <https://www.hopkinsmedicine.org/health/conditions-and-diseases/coronavirus/first-and-second-waves-of-coronavirus>

1. In **Modeling** tab, select **New parameter, Numeric range**, use the following[[2]](#footnote-2):



1. Add a **DAX measure** to the **COV2020\_CSSE** dataset to calculate the estimated recovered cases by calculating the previous period using DATEADD that offsets the time period by subtracting the current date with Days (e.g., Days to Recover).

Estimated Recovered Global = CALCULATE ([Confirmed Recent Value Global] + [Deaths Recent Value Global], DATEADD ('COV2020\_CSSE'[Date], - [Days Value], DAY)) - [Deaths Recent Value Global]

**NOTE:** Tracking recovery cases globally may be logistically challenging. At one point, due to data reliability issues, Johns Hopkins removed the recovered cases from their dataset (specifically for the US). Instead of tracking the actual recovered cases, we can calculate an estimation for the recovered cases.

1. Add a **DAX measure** to the **COV2020\_CSSE** dataset to estimate active cases by subtracting Total Confirmed with Estimated Recovered

Estimated Active Cases Global = [Confirmed Recent Value Global] - [Estimated Recovered Global]

**NOTE**: To illustrate how this technique is relevant to our current analysis let’s say that in Italy we have only 10,000 available hospital beds and that 20% of all active cases need to be hospitalized (these numbers are fictitious and serve for illustration purposes only). Now, let’s say we need to evaluate different treatments and how they will affect our hospital capacity. Each treatment reduces or increases the recovery time according to the line chart (5, 10, 15, 20, 25, 30 days). We can then take advantage of our What-If technique to see that we can select the treatments that will have 10 or fewer days to recovery (data we do not have). To apply this example, we might consider evaluating different treatments and how they might affect hospital capacity (again, data that we do not have).

1. **Rename** the report/page **COV2020 Global Estimated Cases and Recovery**
2. **Save** your Power BI file

1. Rolling averages are commonly used when you have a measure that has a strong variation over time. When you plot the measures, you see highs and lows, it is tough to understand the overall trend of the measure. Creating a rolling average gives a smoother line that makes it easier to understand the overall trend. [↑](#footnote-ref-1)
2. Why 40 days? On average, COVID-19 patients with mild symptoms should begin to see improvement after two weeks. Those with more severe symptoms may feel unwell for up to six weeks. Of course, the recovery period can differ from person to person. COVID-19 is a new virus and scientists are still studying its effects to determine the average recovery time. <https://www.medicalnewstoday.com/articles/long-covid> and <https://www.webmd.com/lung/covid-recovery-overview#1> [↑](#footnote-ref-2)