



Department of Computer Science and Engineering (Data Science)
Experiment 2
(Data Visualization)

NAME: RISHABH PATIL
SAPID: 60009200056
DIV/BATCH: K/K2
DATE: 24/11/2021

Aim: Perform time series aggregation, apply filters on a given dataset, create line and area charts.

Theory:

Time aggregation:

Time aggregation is the aggregation of all data points for a single resource over a specified period (the granularity). Data aggregations in Resource Time Series reports are of the time aggregation type.

The result of the aggregation is one data point that reflects a statistical view of the collected and aggregated data points. For example, average, minimum, maximum, sum, or count. Typically, multiple aggregated data points are presented in a report for a given reporting period.

Data Filtering:

Data filtering is the process of choosing a smaller part of your data set and using that subset for viewing or analysis. Filtering is generally (but not always) temporary – the complete data set is kept, but only part of it is used for the calculation.

Line Charts:

A line chart displays information as a series of data points called 'markers' connected by straight line segments. It is a basic type of chart common in many fields.

Z2D to QQQ Exchange Rates



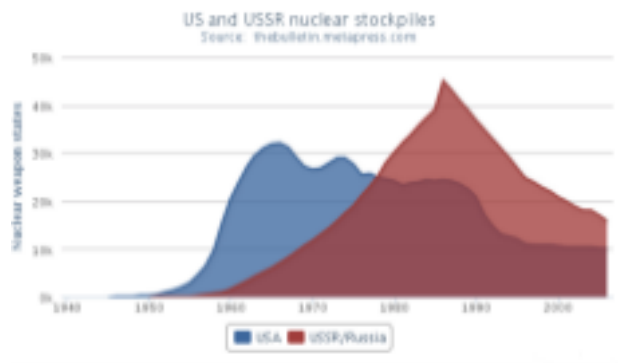
Wildlife Population





Area Charts:

An area chart is like a line chart in terms of how data values are plotted on the chart and connected using line segments. In an area chart, however, the area between the line segments and the x-axis is filled with color. It helps in comparison of multiple series, relationships between two series' and analysis of various trends.



Lab Assignments to complete in this session

Use the given dataset and perform the following tasks:

Dataset: 1

<https://raw.githubusercontent.com/nytimes/covid-19-data/master/us.csv>

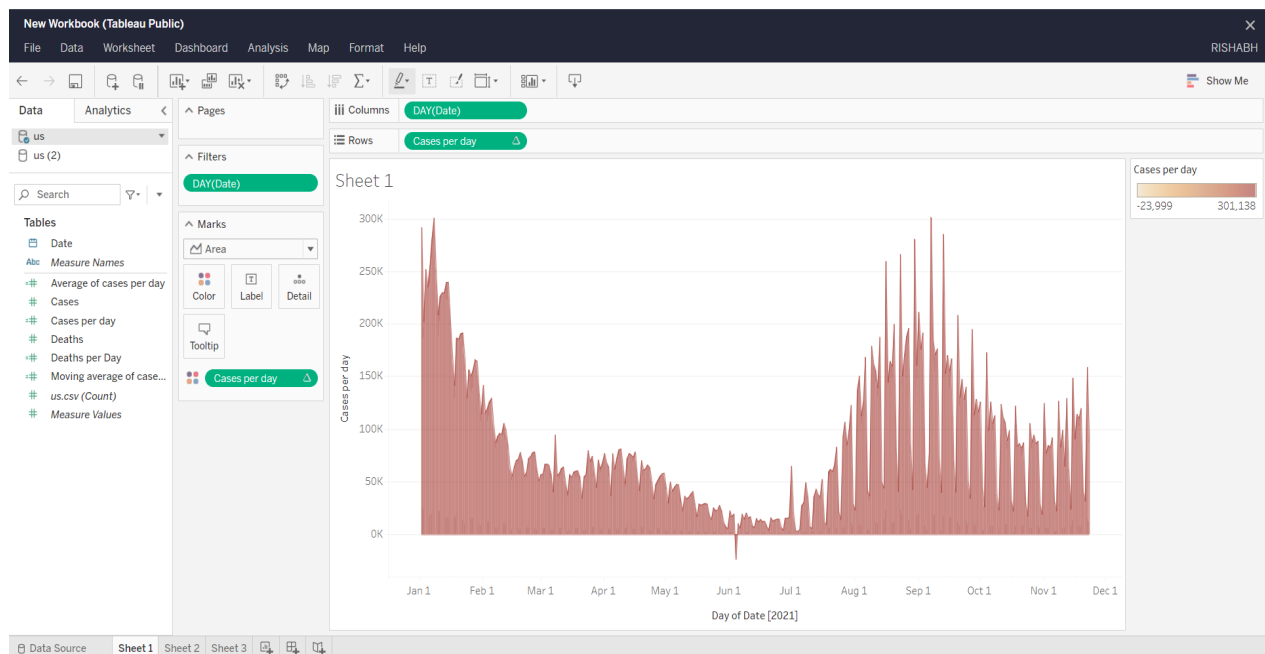
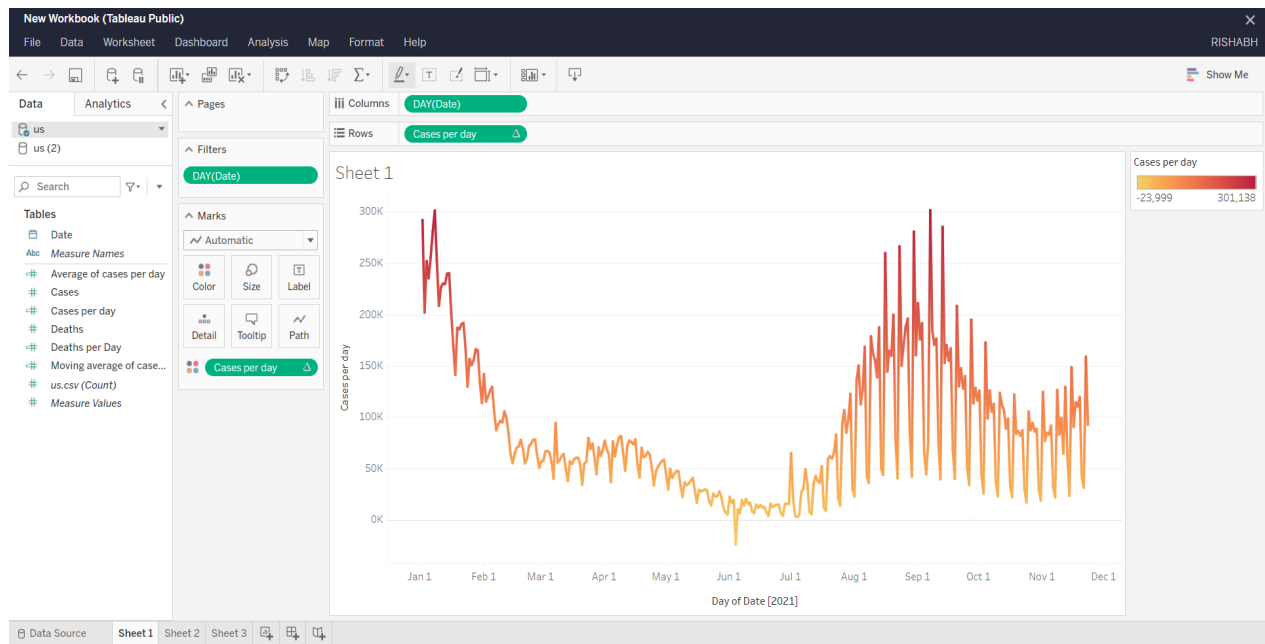
Compatible Web Data Connector for easy importing:

<https://basic-csv-wdc.herokuapp.com/>

- Vaccination in the USA started in full swing in Early Jan 2021. Analyze the trend of Cases per day from Jan 2021 (Vaccines begun) to the current day and report if this policy was useful.
- The USA implemented a policy where Vaccinated Individuals could avoid wearing masks in public from June 2021. The UN wants to view the impact of making such a move in the long term as well as the short term.
- Analysis requires the creation of a calculated field in terms of Cases Per day and a Weekly moving average of cases per day.
- Justify why you have chosen a particular visualization for this dataset. Also discuss the details of your visualization.



a) Visualization



Here, according to the question, we need to analyze the trend of Cases per day from Jan 2021 (Vaccines begun) to the current day and report if the policy was useful. Hence we filtered data accordingly and got the required graph. So, in this graph, we set

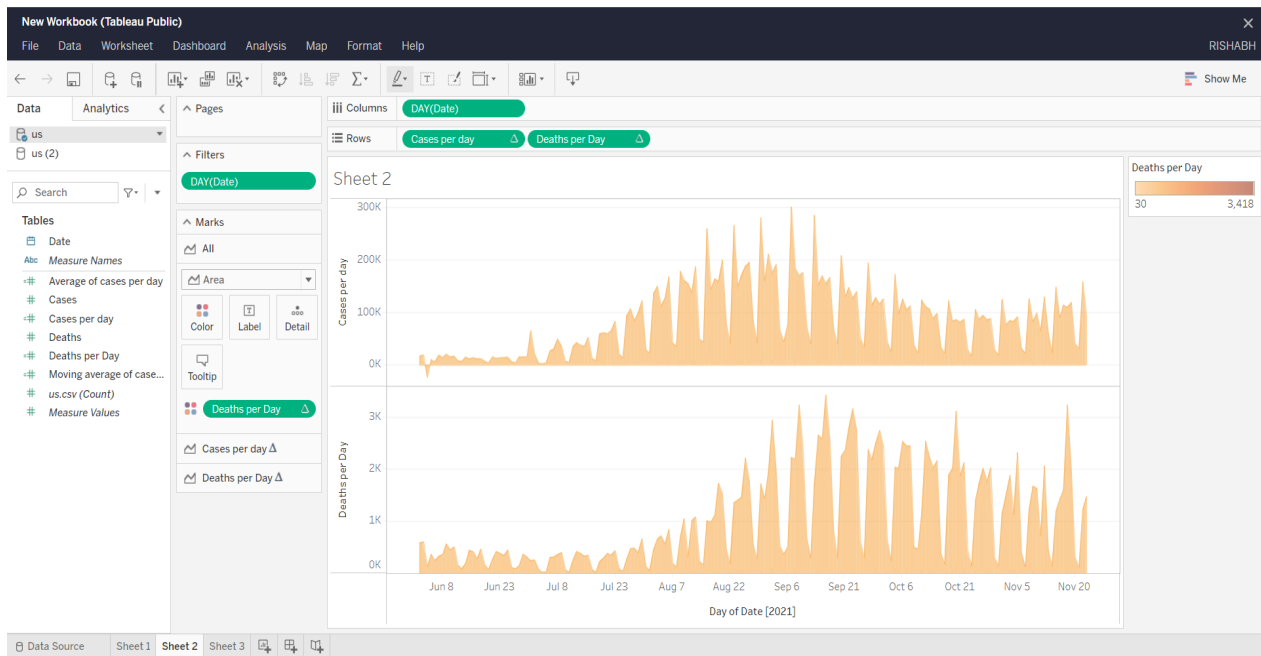
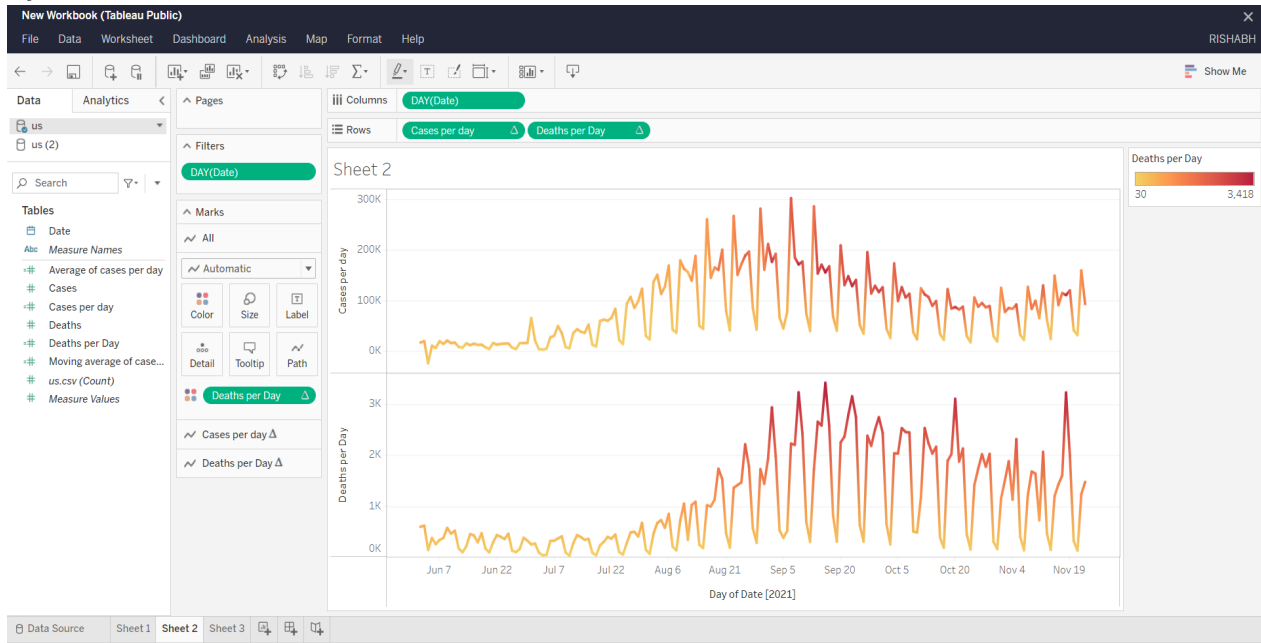
Y-axis: Cases Per Day X-axis: Day of Date

d)i) Justification and Analysis:

Here, using graphs gives clear visualization of the trend of cases per day. We can see that this vaccination policy was useful till May-June but eventually cases per day increased after June.



b) Visualization



Here, according to the question, we need to analyze the impact of implementing a policy where Vaccinated Individuals could avoid wearing masks in public from June 2021, hence we filtered data accordingly and got the required graph So, in this graph, we set

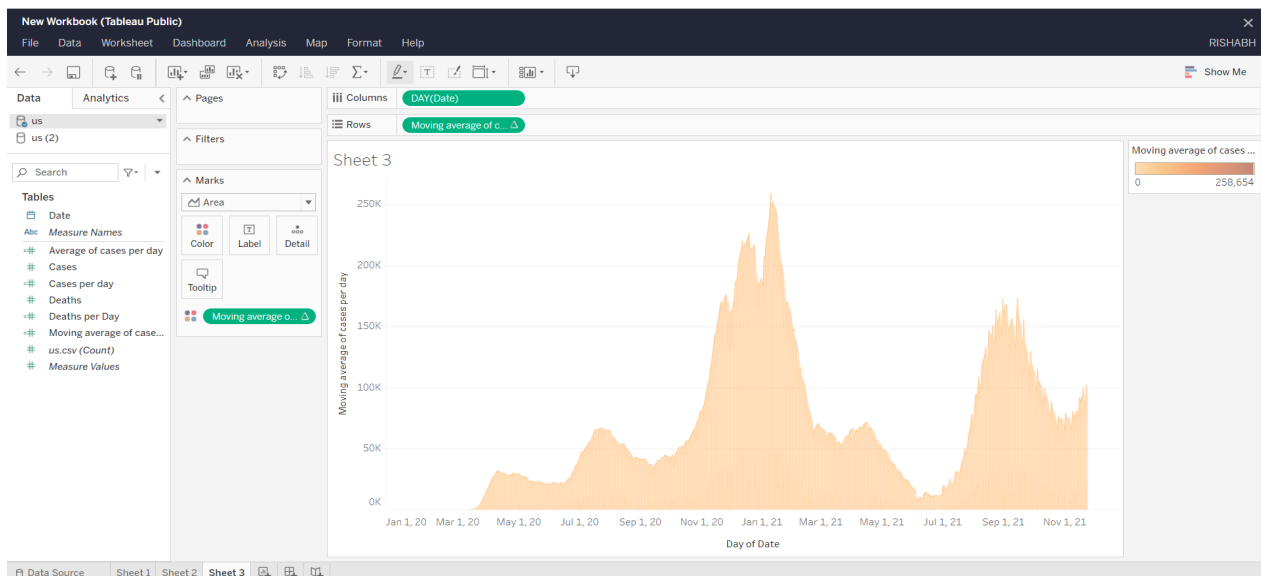
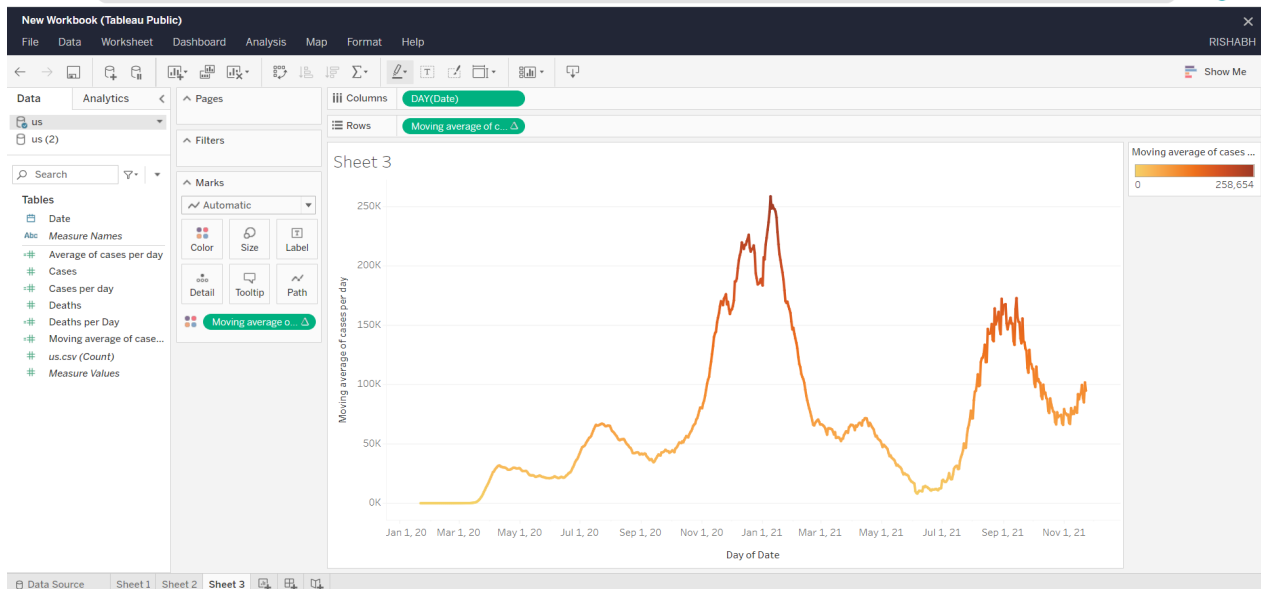
Y-axis: Cases/Deaths Per Day X-axis: Day of Date

d)ii)Justification and Analysis:

Here,using graphs gives clear visualization of the trend of cases and per day from June 2021.We can see that this US implemented policy was useful till Mid-August but eventually cases and per day increased after that,so it was too soon to implement it.



c) Visualization



Here, according to the question, we need to create a calculated field in terms of cases per day and weekly moving average of cases per day and analyze it. We got the required graph in this we set

Y-axis: Avg X-axis: Day of Date

We used the formula: `WINDOW_AVG([cases per day],-7,0)`

d)iii) Justification and Analysis:

Here, using graphs gives clear visualization of the trend of cases per day and weekly moving average of cases per day. We can see that cases were increasing at a high rate till Jan 2021, decreased till June, again increased till September, but eventually decreased after that.