



## Hospital Survey Report:

Basically, This is a Kaggle Dataset on Healthcare and Provider system survey around the world where each survey describe about the Hospital rating, patient experience score and demographic details where it consist of mainly 7 .csv files and 1 data\_dict file ( in both csv and .xlsx format).

Mainly File name used in this dataset to create visualisation are: ↗

1. measures
2. national
3. questions
4. reports
5. state
6. responses
7. states\_cont

💡 I change the names on my own requirement as I already mentioned the changes named to which below

```
data_dict = data_names
measures = Measure ID
national = Measure ID
questions = Measure ID
reports = Release Period
responses = Release Period, state
state = Release Period, state, Measure ID
states_cont = state

there are the unique_id which connect each dataset to one another now i can easily visualize them
sorry for inconvenience because i am using tableau as i don't have power bi access
and i am now on mac but ik power bi well
```

Python

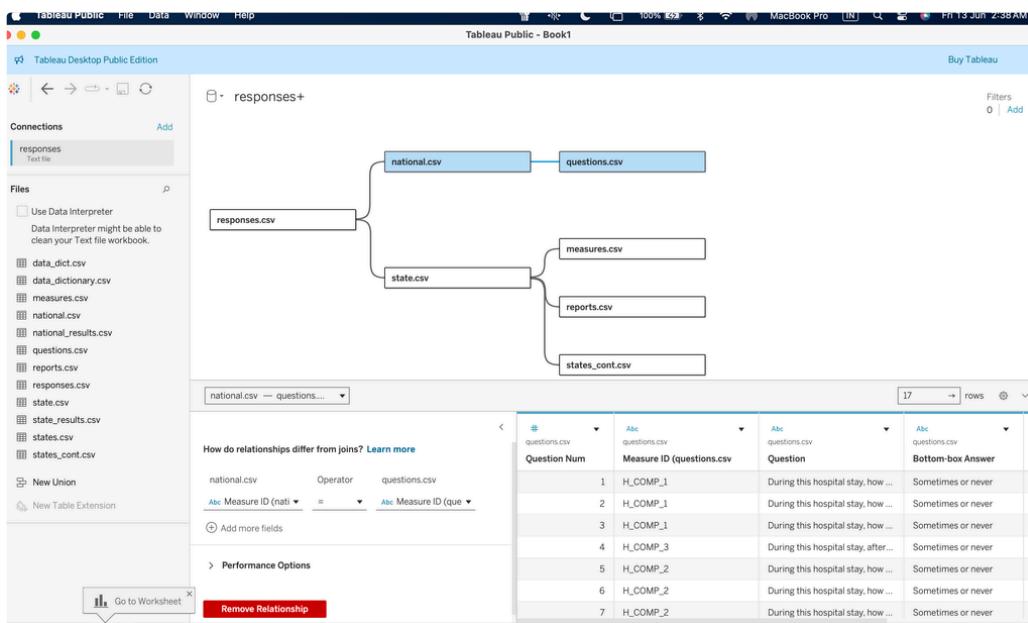
I already mentions above also the unique id and primary id which belong to which table as a relation between them --- below I am going to mention a table for clear explanation ( few file have one unique\_id or other have much more than 2

.csv file	unique_id 1	unique_id 2
measures	Measure ID	
national	Measure ID	
reports	Release period	
questions	Measure ID	

responses	Release period	state
state	Release Period	state
states_cont	state	

state.csv file also contains Measure ID which I am unable to mentioned above that why I am writing here

- 1. Data Cleaning:** I did data cleaning through pandas as I have already command on it and use it for another reason too, so that I can visualize the csv file just running the commands and check how many columns are present and what are their names etc as I am attaching the link so that you can access the .ipynb file to see the data cleaning process there [Hospital Survey.ipynb](#)
- 2. Star Schema (Data Modeling):** which we can say that applying flowchart in simple term or in technical way we can say that Data modelling or Data linking with each other based on their unique id as I already mention about the unique id so it become much more easy for you to understand when you see the below image



- 3. DAX Formulas:** In tableau, we also implement sql formula by creating [creating calculated field](#) which I also implemented so that we can implement statistical analysis through these formulas and then implement it into sheet to create visualisation by using metrics as per data requirement

I am mentioning all the formulas which I used in this report :

- **kpi\_1 = avg(top box percent)** but before creating it, I implement [Top-box Percentage] / 100 formulas because it showed me the data in 7xxx.xx format which is not accurate percent 2 digit format
- **kpi\_2 = avg(response rate)** for this too I also implement [Response Rate (%) ] / 100 because it also give me the same error like above
- **viz\_1:** for this I use Dax formulas because here in viz as per assessment I have to create heatmaps or etc others maps rather than normal maps – for this as a **Treemap** the formulas are used as **sum(top box % correct)** in size and **sum(top box % correct)** in color and **state(state.csv)** in text so that it can create the treemap easily for better visualisation for Top Box % based on state
- **viz\_2:** response rate based on date for which I implement **sum(response rate % )** in size and **state** in columns and **released period** in rows to implement heatmaps

- **viz\_3:** First I filter sum response rate here because it collide while preparing the completed survery chart because there is use of sum response rate in columns too , Hope so now it will be cleared then I used **completed survery** in details and **max(survery status)** in color
- **viz\_4:** Ahh! this one is critical as it took my lot of time to implement a graph for this analysis but I will explain you in easy way basically here we have to count the top box % by taking a assumption of target 75 for which I have to create calculated field with target name and put value 75 there but for visualising it. we have to understand that count top box % , I implement **max(target)** in detail, **top box % correct (bin)** in columns and **CNT(top box % correct)** in rows

that's it where viz\_1, viz\_2, viz\_3 and viz\_4 are DAX formulas method

4. **KPI:** well above I already explain it in DAX but KPI is a tile table in the tableau which we have to implement using formulas and implement as a tile

Now I am giving the screenshot of my report and also the link of my [Tableau](#) so that It will be easy to see the working status of the report apart from that I am also going to attach  Hospital\_survery\_v2024.3.twbx here so that it become much more easy to see the full report analysis

Thank-you for reading this

– Adrika Panwar