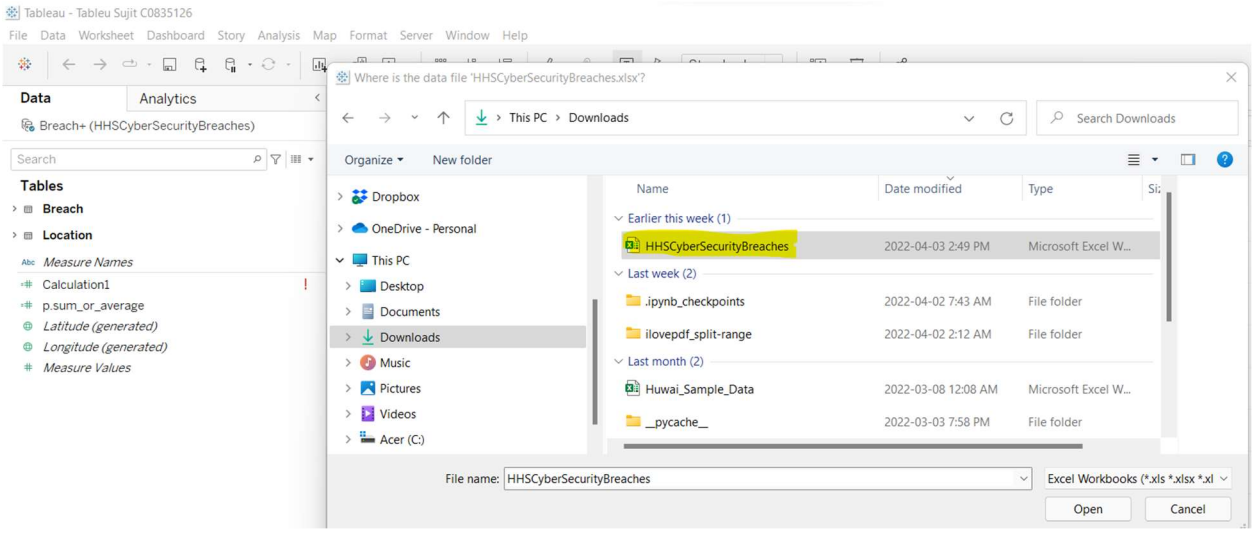
2022W-T2 BDM 2053 - Big Data Algorithms and Statistic 01

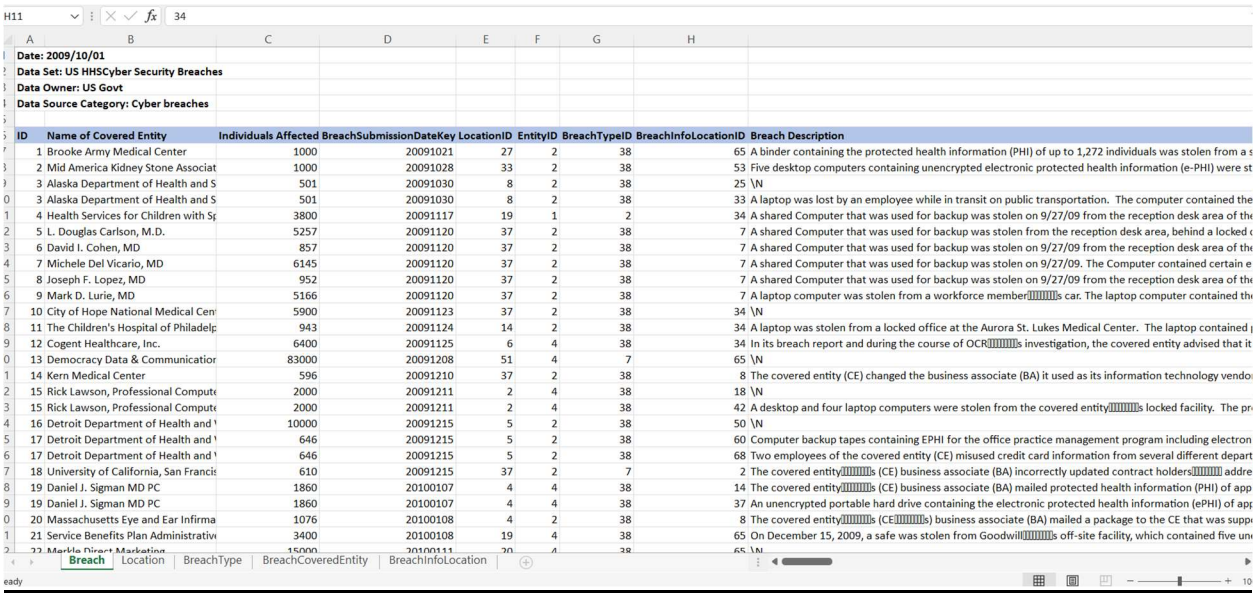
Lab 7

Submitted by: Sujit Khatiwada (C0835126)

Tableau Report Generation

1. Importing the Dataset:

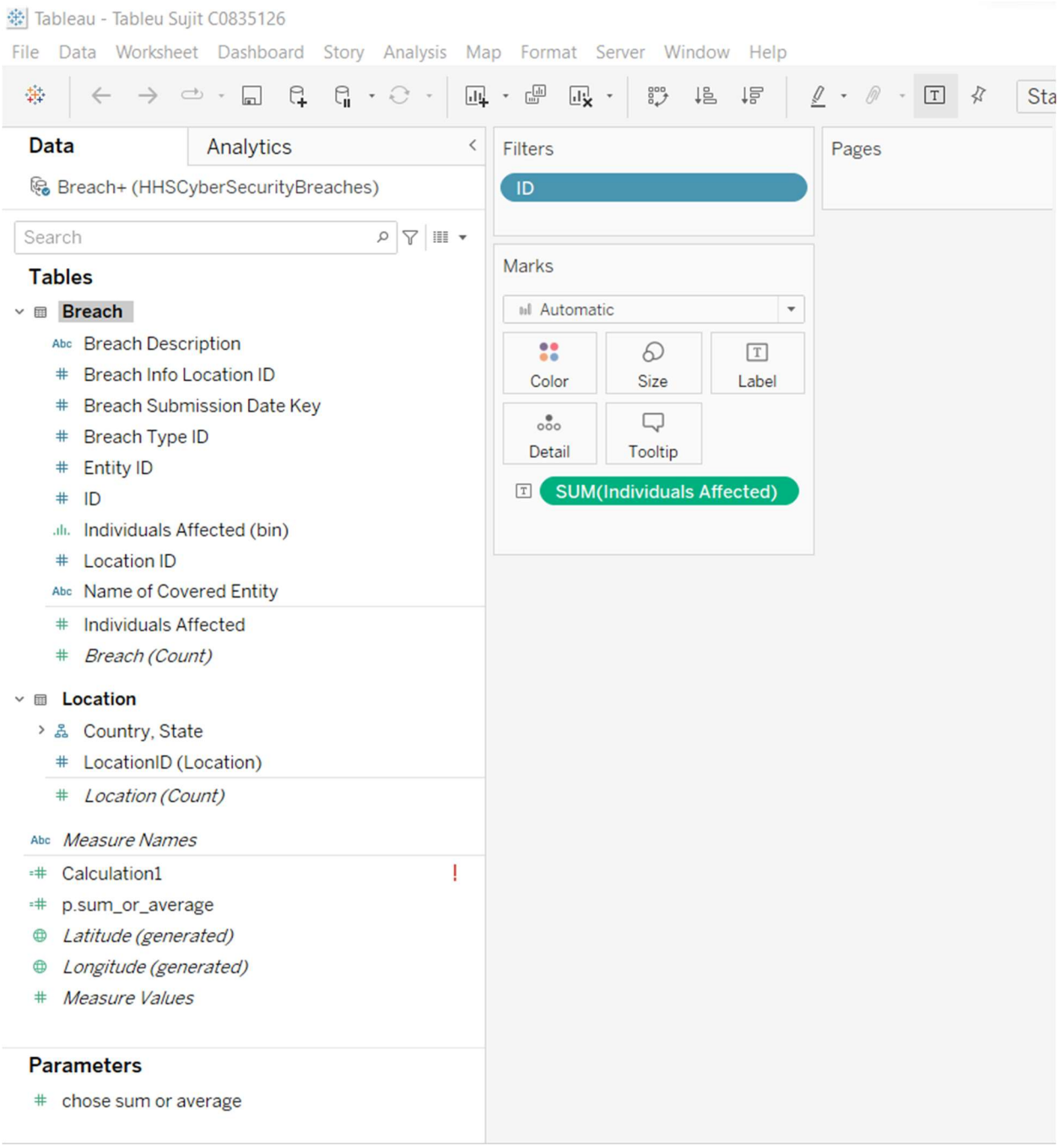




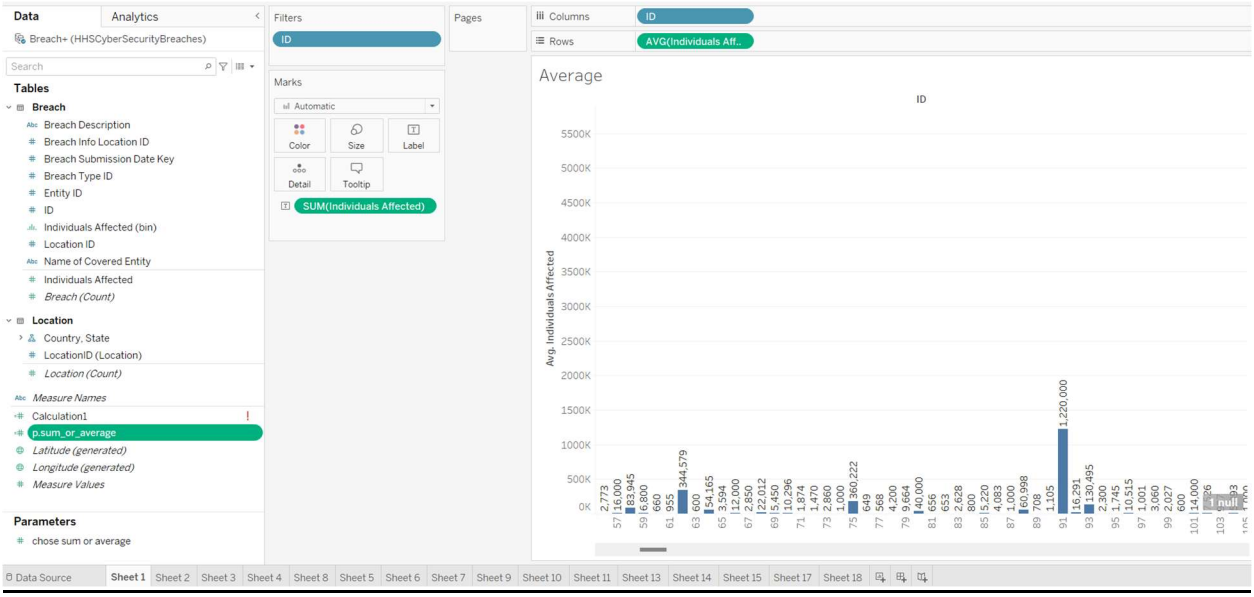
	A	B	C	D	E	F	G	H
1	Date: 2009/10/01							
2	Data Set: US HHSCyber Security Breaches							
3	Data Owner: US Govt							
4	Data Source Category: Cyber breaches							
5								
6	ID	Name of Covered Entity	Individuals Affected	BreachSubmissionDateKey	LocationID	EntityID	BreachTypeID	BreachInfoLocationID	Breach Description
7	1	Brooke Army Medical Center	1000	20091021	27	2	38	65	A binder containing the protected health information (PHI) of up to 1,272 individuals was stolen from a s
8	2	Mid America Kidney Stone Associat	1000	20091028	33	2	38	53	Five desktop computers containing unencrypted electronic protected health information (e-PHI) were st
9	3	Alaska Department of Health and S	501	20091030	8	2	38	25	\N
10	3	Alaska Department of Health and S	501	20091030	8	2	38	33	A laptop was lost by an employee while in transit on public transportation. The computer contained the
11	4	Health Services for Children with Sp	3800	20091117	19	1	2	34	A shared Computer that was used for backup was stolen on 9/27/09 from the reception desk area of the
12	5	L. Douglas Carlson, M.D.	5257	20091120	37	2	38	7	A shared Computer that was used for backup was stolen from the reception desk area, behind a locked
13	6	David I. Cohen, MD	857	20091120	37	2	38	7	A shared Computer that was used for backup was stolen on 9/27/09 from the reception desk area of the
14	7	Michele Del Vicario, MD	6145	20091120	37	2	38	7	A shared Computer that was used for backup was stolen on 9/27/09. The Computer contained certain e
15	8	Joseph F. Lopez, MD	952	20091120	37	2	38	7	A shared Computer that was used for backup was stolen on 9/27/09 from the reception desk area of the
16	9	Mark D. Lurie, MD	5166	20091120	37	2	38	7	A laptop computer was stolen from a workforce member[redacted]s car. The laptop computer contained the
17	10	City of Hope National Medical Cen	5900	20091123	37	2	38	34	\N
18	11	The Children's Hospital of Philadelp	943	20091124	14	2	38	34	A laptop was stolen from a locked office at the Aurora St. Lukes Medical Center. The laptop contained
19	12	Cogent Healthcare, Inc.	6400	20091125	6	4	38	34	In its breach report and during the course of OCR[redacted]s investigation, the covered entity advised that it
20	13	Democracy Data & Communicatior	83000	20091208	51	4	7	65	\N
21	14	Kern Medical Center	596	20091210	37	2	38	8	The covered entity (CE) changed the business associate (BA) it used as its information technology vendo
22	15	Rick Lawson, Professional Compute	2000	20091211	2	4	38	18	\N
23	15	Rick Lawson, Professional Compute	2000	20091211	2	4	38	42	A desktop and four laptop computers were stolen from the covered entity[redacted]s locked facility. The pr
24	16	Detroit Department of Health and \	10000	20091215	5	2	38	50	\N
25	17	Detroit Department of Health and \	646	20091215	5	2	38	60	Computer backup tapes containing EPHI for the office practice management program including electron
26	17	Detroit Department of Health and \	646	20091215	5	2	38	68	Two employees of the covered entity (CE) misused credit card information from several different depart
27	18	University of California, San Francis	610	20091215	37	2	7	2	The covered entity[redacted]s (CE) business associate (BA) incorrectly updated contract holders[redacted] addre
28	19	Daniel J. Sigman MD PC	1860	20100107	4	4	38	14	The covered entity[redacted]s (CE) business associate (BA) mailed protected health information (PHI) of app
29	19	Daniel J. Sigman MD PC	1860	20100107	4	4	38	37	An unencrypted portable hard drive containing the electronic protected health information (ePHI) of app
30	20	Massachusetts Eye and Ear Infirma	1076	20100108	4	2	38	8	The covered entity[redacted]s (CE[redacted]s) business associate (BA) mailed a package to the CE that was suppo
31	21	Service Benefits Plan Administrativ	3400	20100108	19	4	38	65	On December 15, 2009, a safe was stolen from Goodwill[redacted]s off-site facility, which contained five un
32	22	Merkle Direct Marketing	15000	20100111	20	4	38	65	\N

Breach | Location | BreachType | BreachCoveredEntity | BreachInfoLocation

2. Table Structure



3. Calculate avg, medians , minimum, maximum


Data
Analytics
Breach+ (HHSCyberSecurityBreaches)
Search
Tables
Breach
Breach Description
Breach Info Location ID
Breach Submission Date Key
Breach Type ID
Entity ID
ID
Individuals Affected (bin)
Location ID
Name of Covered Entity
Individuals Affected
Breach (Count)
Location
Country, State
LocationID (Location)
Location (Count)
Measure Names
Calculation1
p.sum_or_average
Latitude (generated)
Longitude (generated)
Measure Values
Parameters
chose sum or average
Filters
ID
Marks
Automatic
Color
Size
Label
Detail
Tooltip
SUM(Individuals Affected)
Pages
Columns
ID
Rows
AVG(Individuals Aff..
Average
ID
Avg. Individuals Affected
5500K
5000K
4500K
4000K
3500K
3000K
2500K
2000K
1500K
1000K
500K
0K
2,773
16,000
83,945
6,800
660
955
344,579
600
54,165
3,594
12,000
2,860
22,012
5,450
10,296
1,874
1,470
2,860
1,000
360,222
649
568
4,200
9,664
40,000
656
553
628
800
5,220
4,083
1,000
60,998
708
1,105
1,220,000
16,291
130,495
2,300
1,745
10,515
1,001
969
2,027
600
14,000
1 null
Data Source
Sheet 1
Sheet 2
Sheet 3
Sheet 4
Sheet 8
Sheet 5
Sheet 6
Sheet 7
Sheet 9
Sheet 10
Sheet 11
Sheet 13
Sheet 14
Sheet 15
Sheet 17
Sheet 18

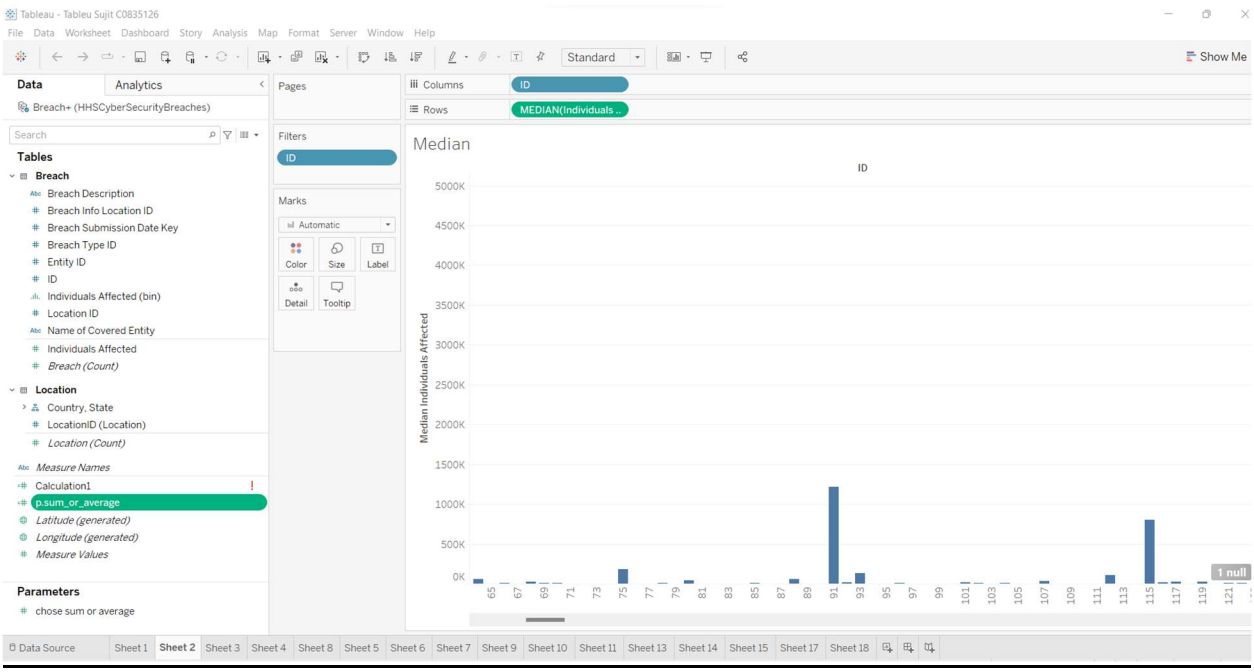
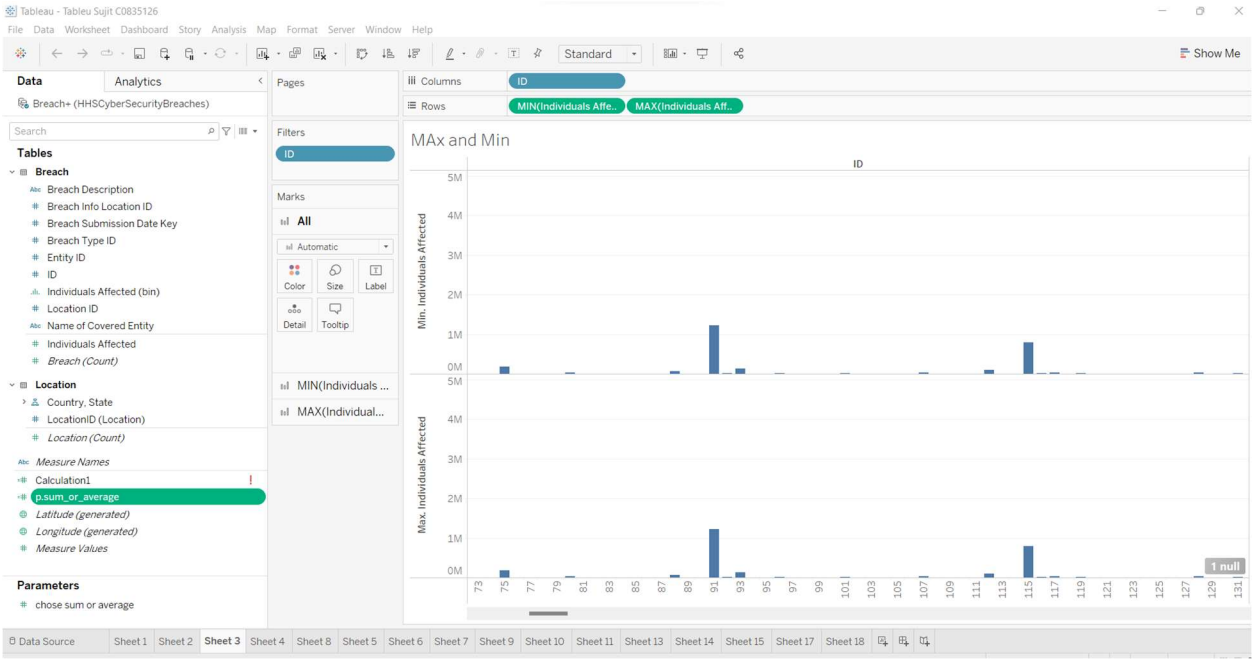
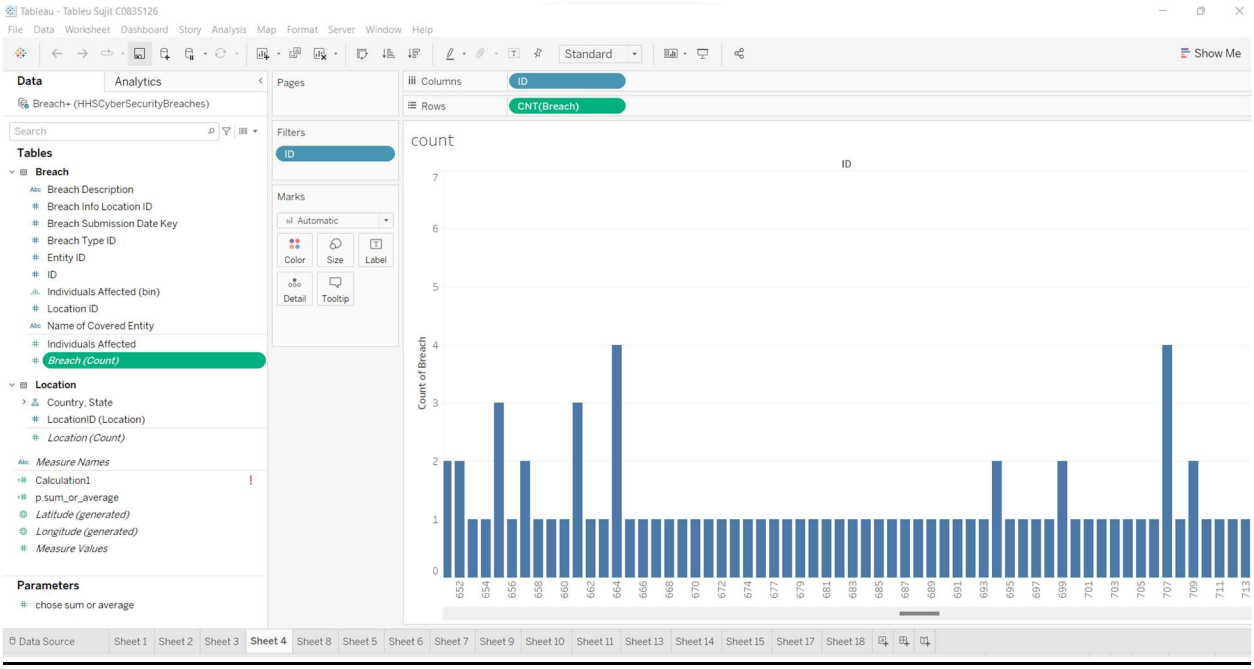


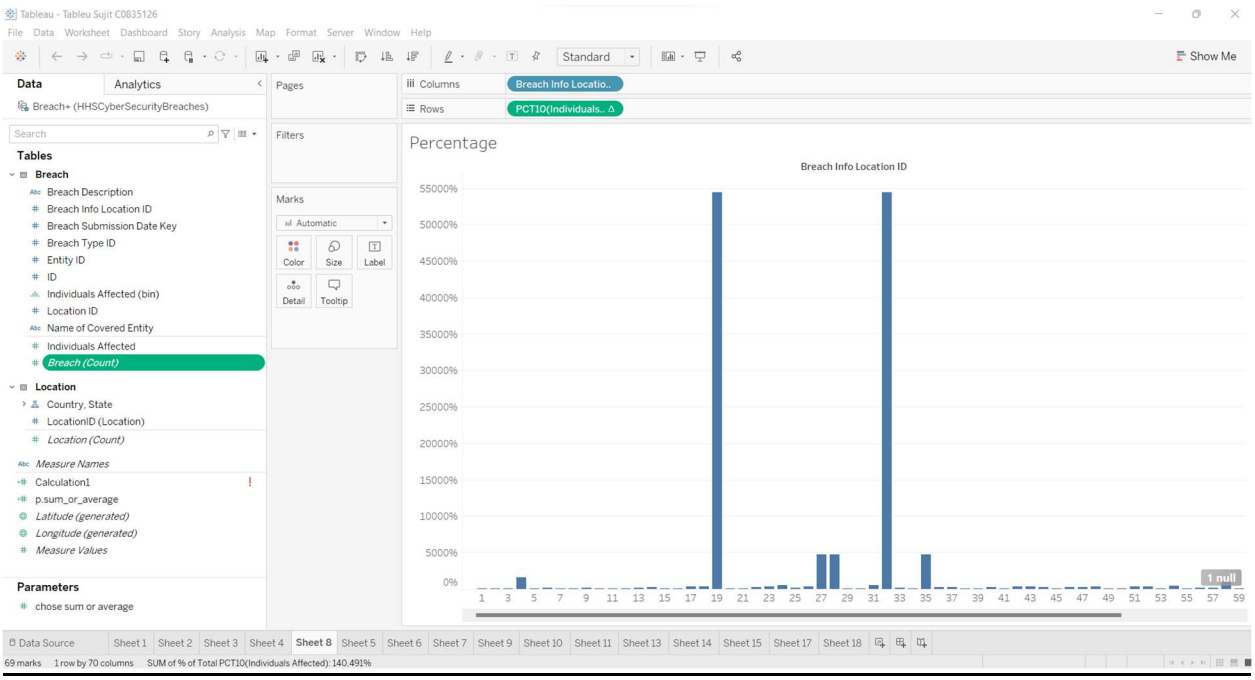
Tableau - Tableu Sujit C0835126
File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Standard
Show Me
Data
Analytics
Breach+ (HHSCyberSecurityBreaches)
Search
Tables
Breach
Breach Description
Breach Info Location ID
Breach Submission Date Key
Breach Type ID
Entity ID
ID
Individuals Affected (bin)
Location ID
Name of Covered Entity
Individuals Affected
Breach (Count)
Location
Country, State
LocationID (Location)
Location (Count)
Measure Names
Calculation1
p.sum_or_average
Latitude (generated)
Longitude (generated)
Measure Values
Parameters
chose sum or average
Pages
Filters
ID
Marks
Automatic
Color
Size
Label
Detail
Tooltip
Columns
ID
Rows
MEDIAN(Individuals ..
Median
ID
Median Individuals Affected
5000K
4500K
4000K
3500K
3000K
2500K
2000K
1500K
1000K
500K
0K
1 null
Data Source
Sheet 1
Sheet 2
Sheet 3
Sheet 4
Sheet 8
Sheet 5
Sheet 6
Sheet 7
Sheet 9
Sheet 10
Sheet 11
Sheet 13
Sheet 14
Sheet 15
Sheet 17
Sheet 18



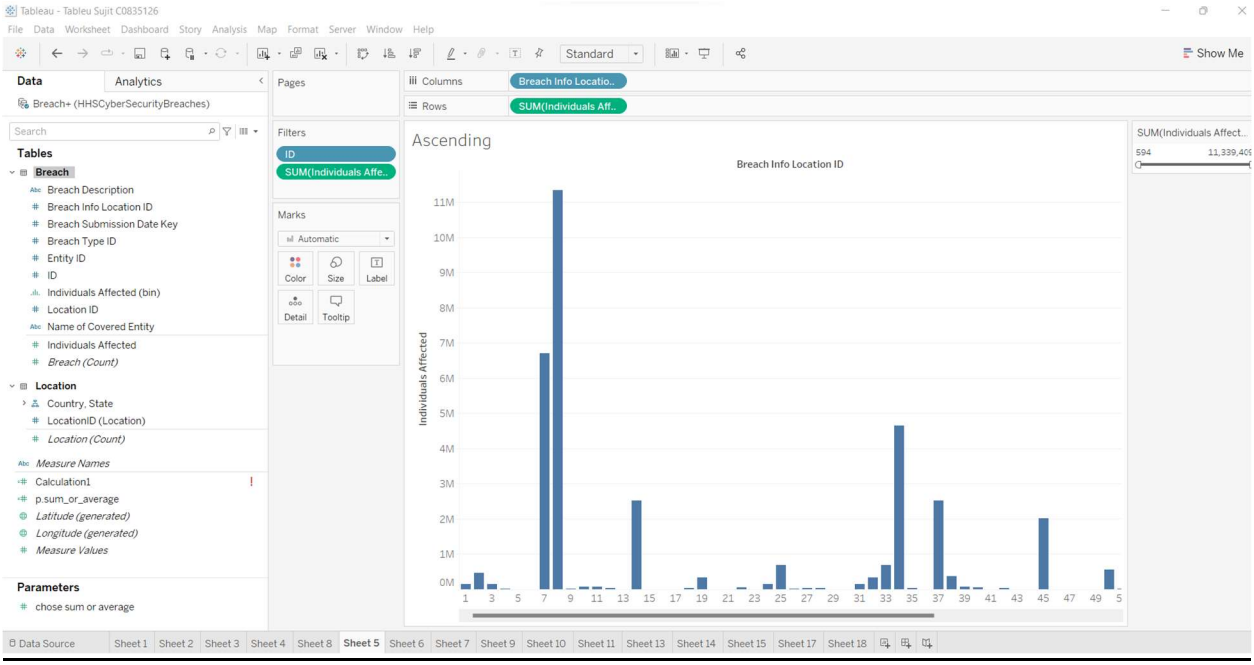
4. Count

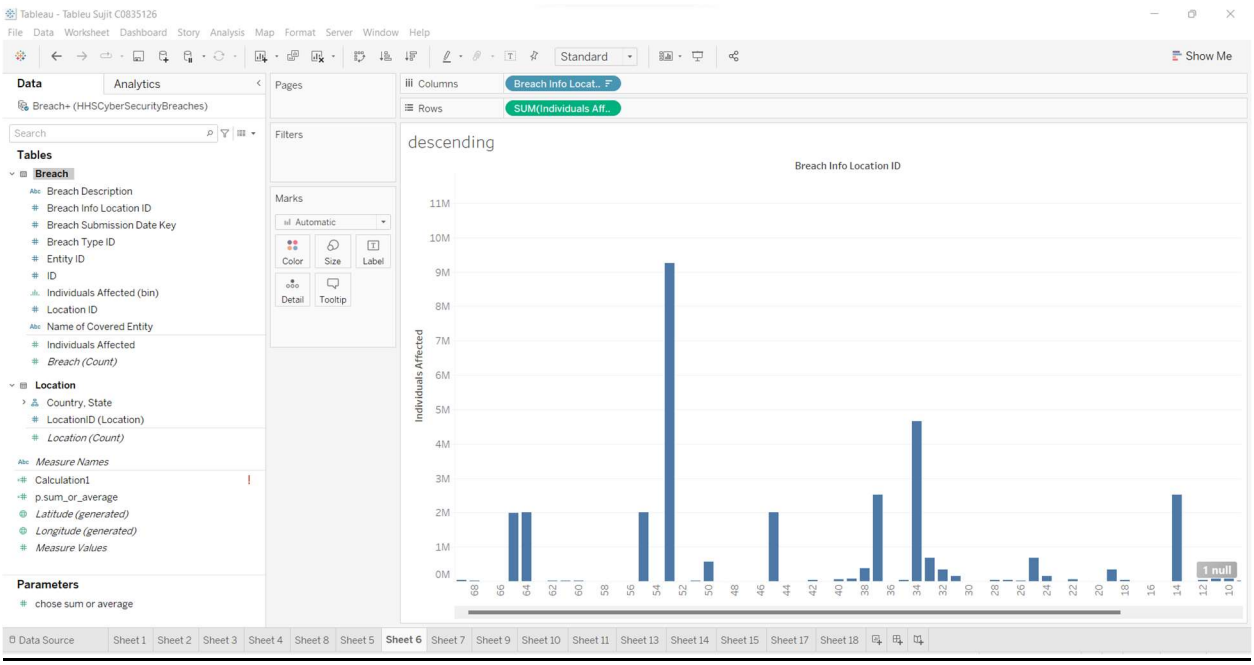


5. Percentages within a table

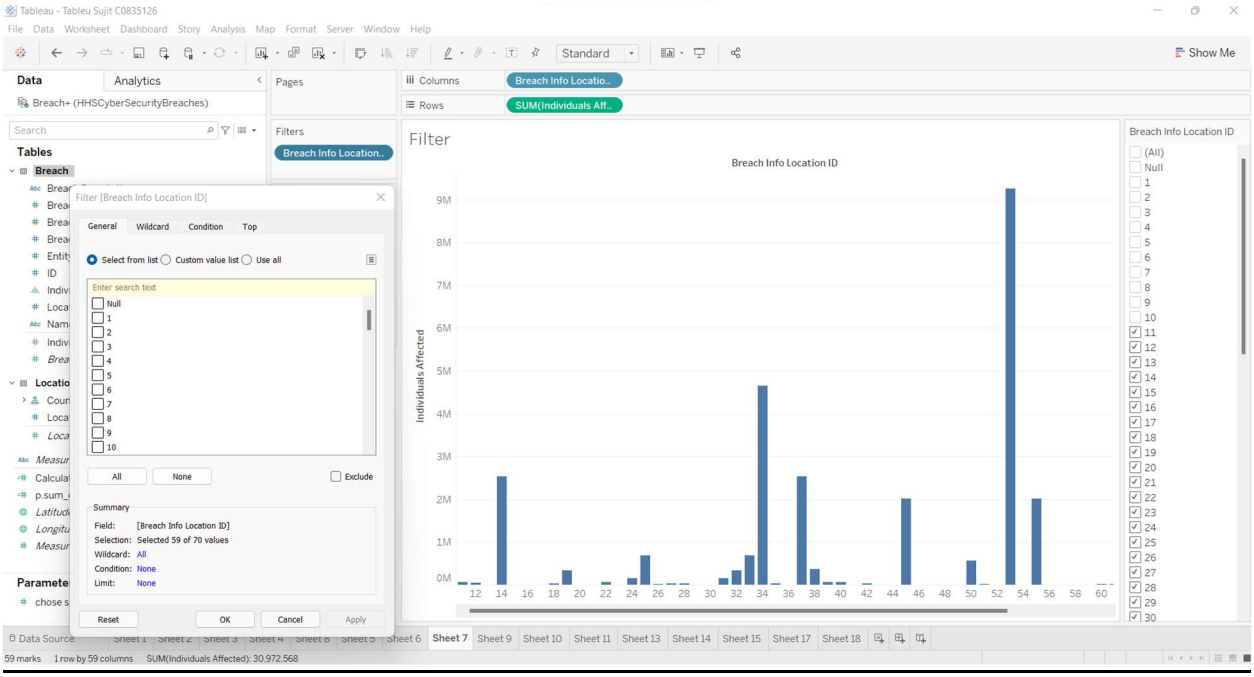


6. Sort - Ascending , Descending

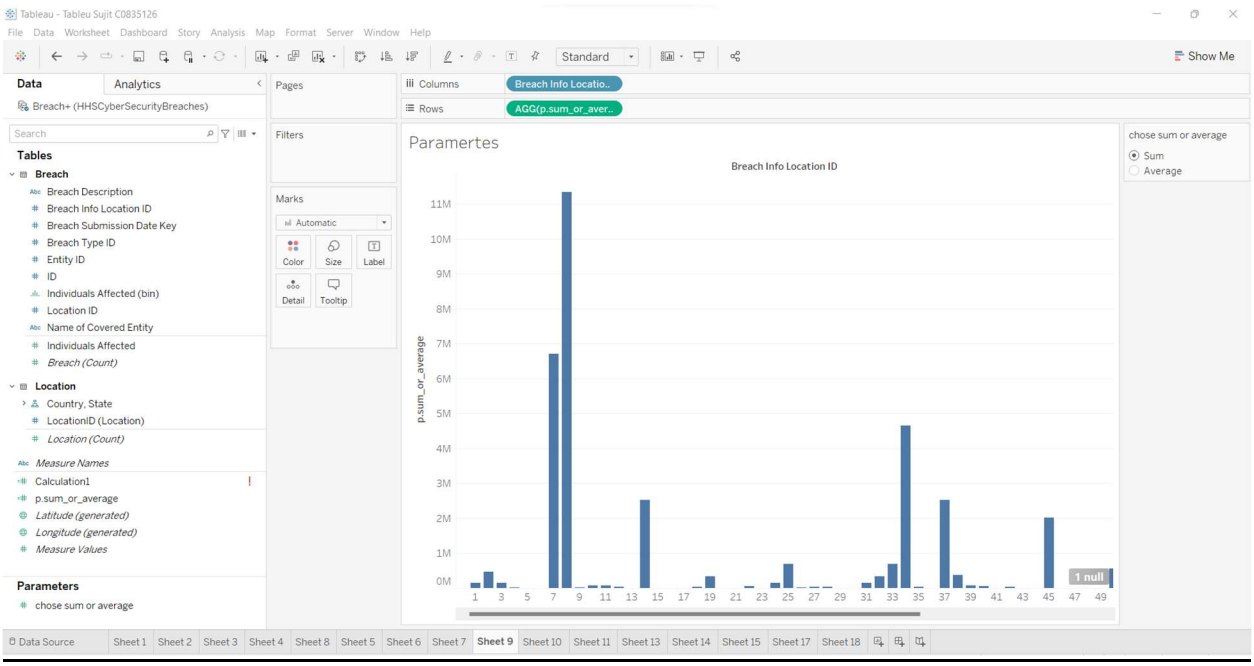




7. Creating a selection filter/Applying filter

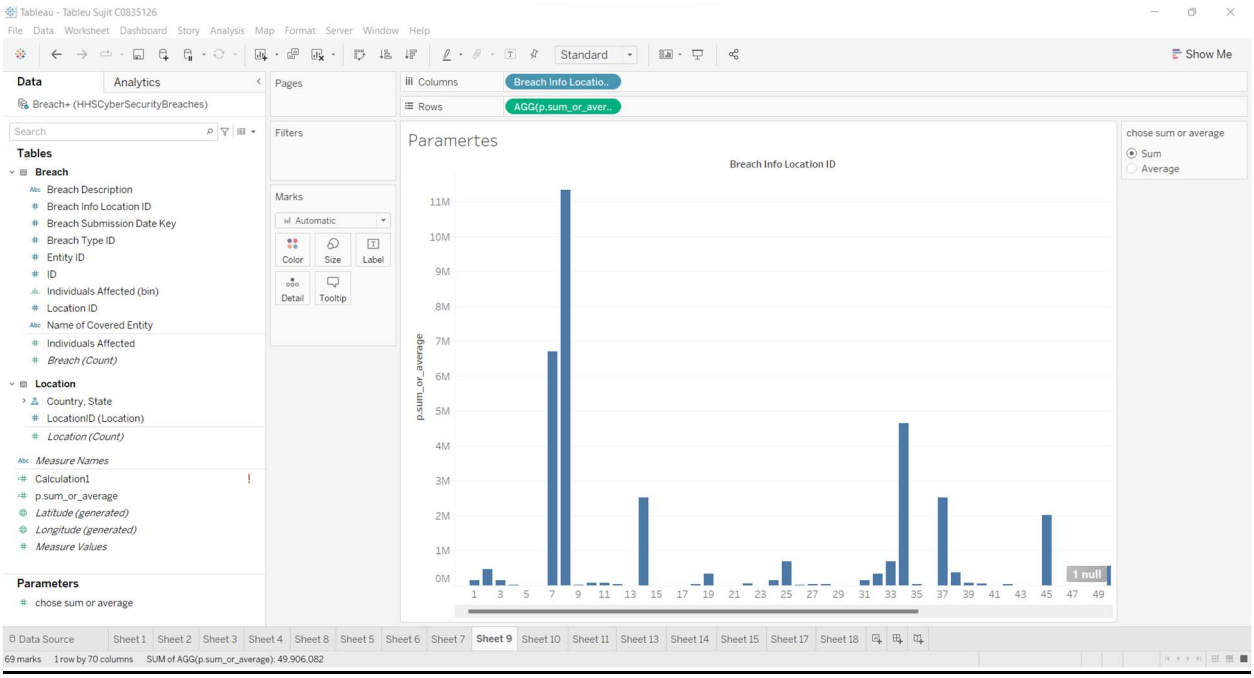


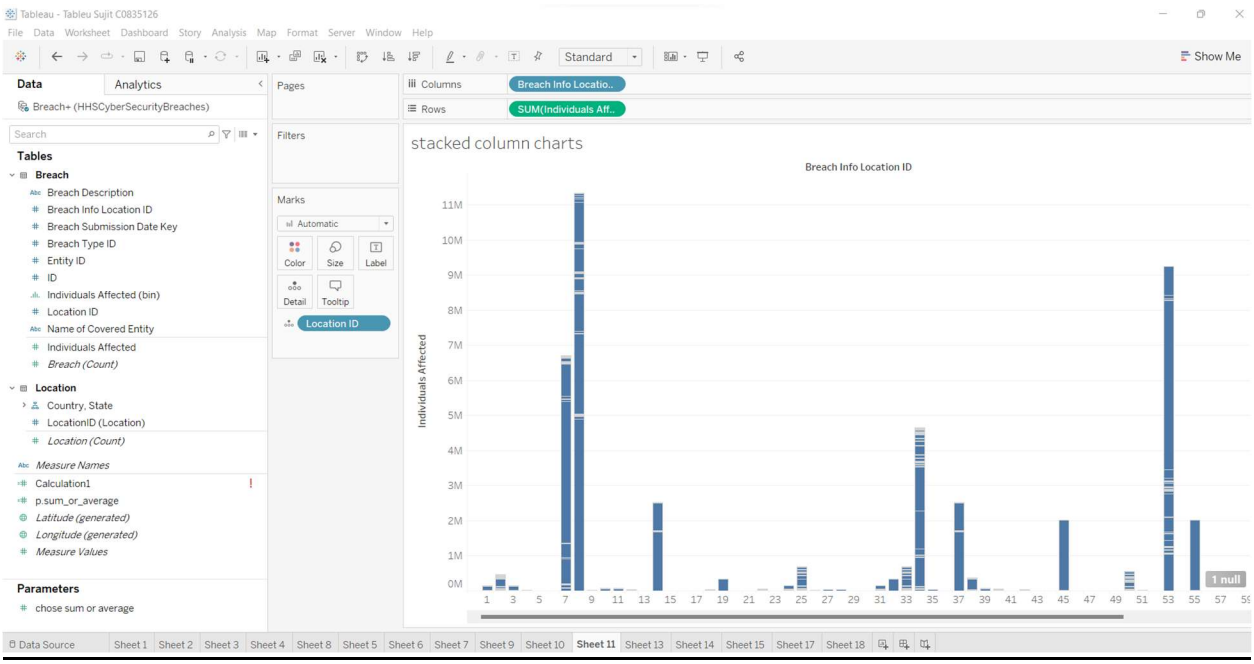
8. Editing and Deleting parameters



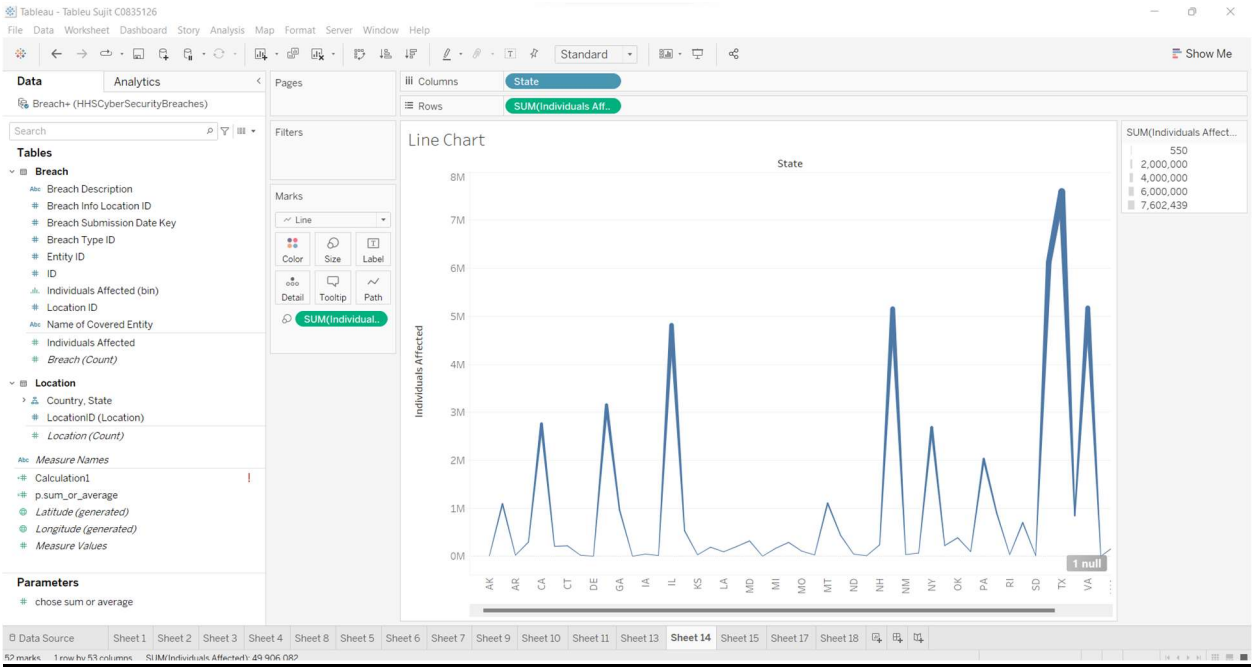
9. Create Basic Visualizations

a. Column and stacked column charts

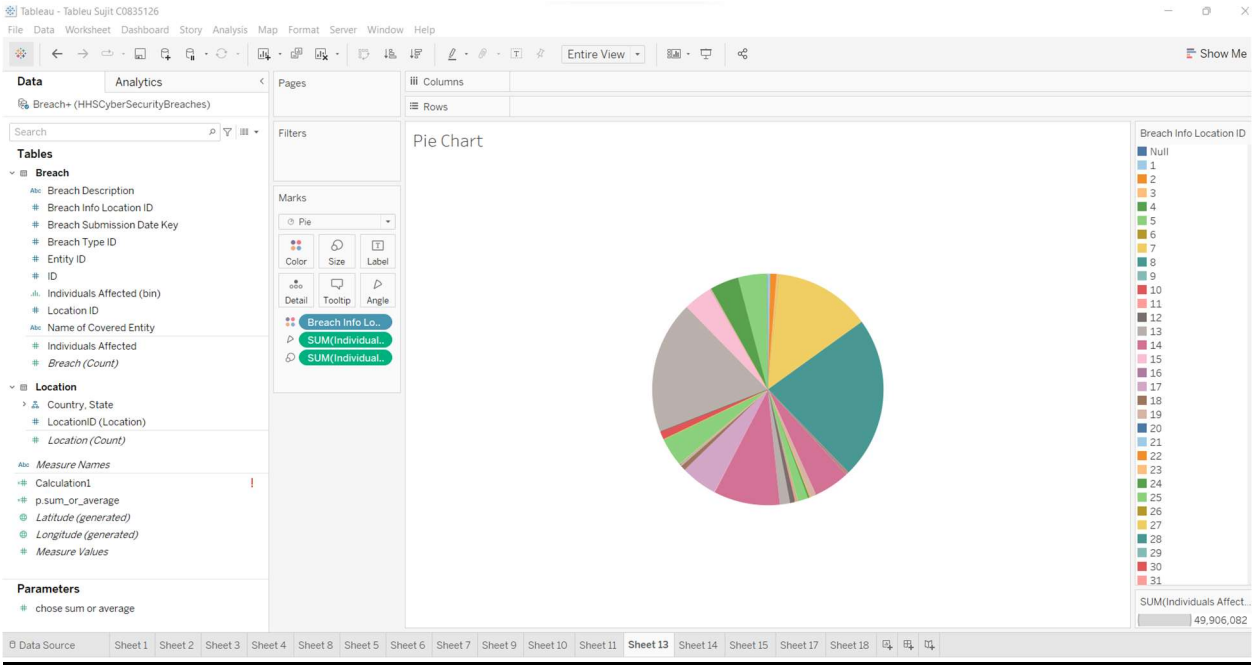




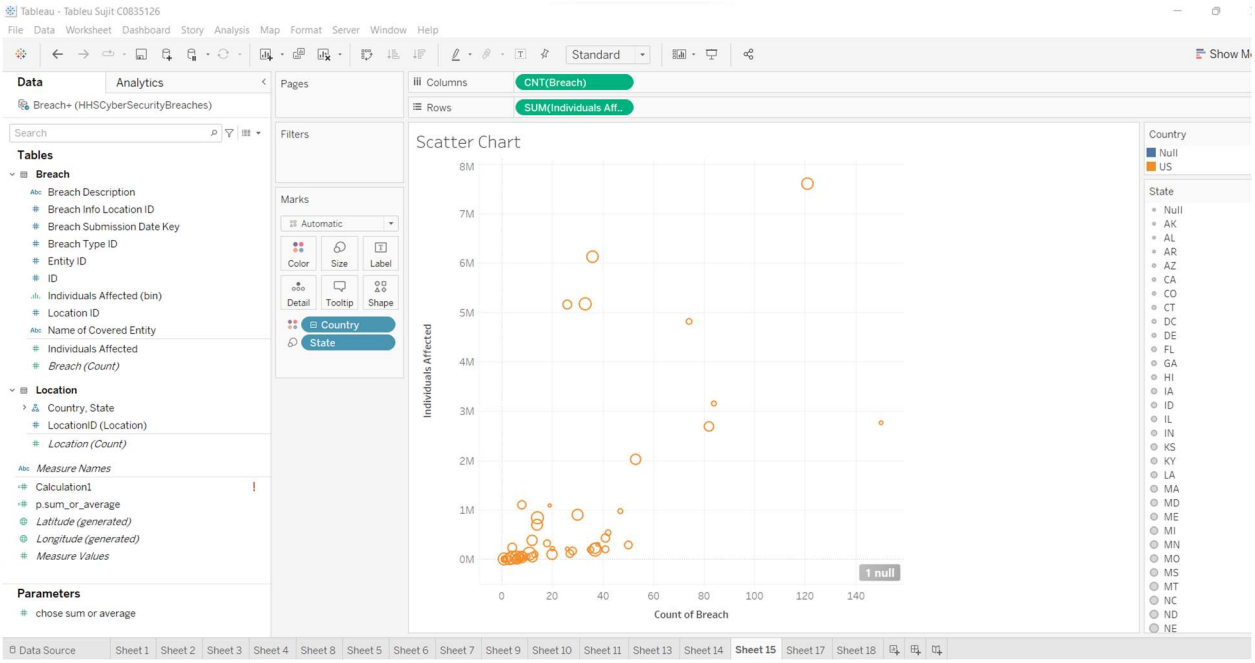
b. Line Chart



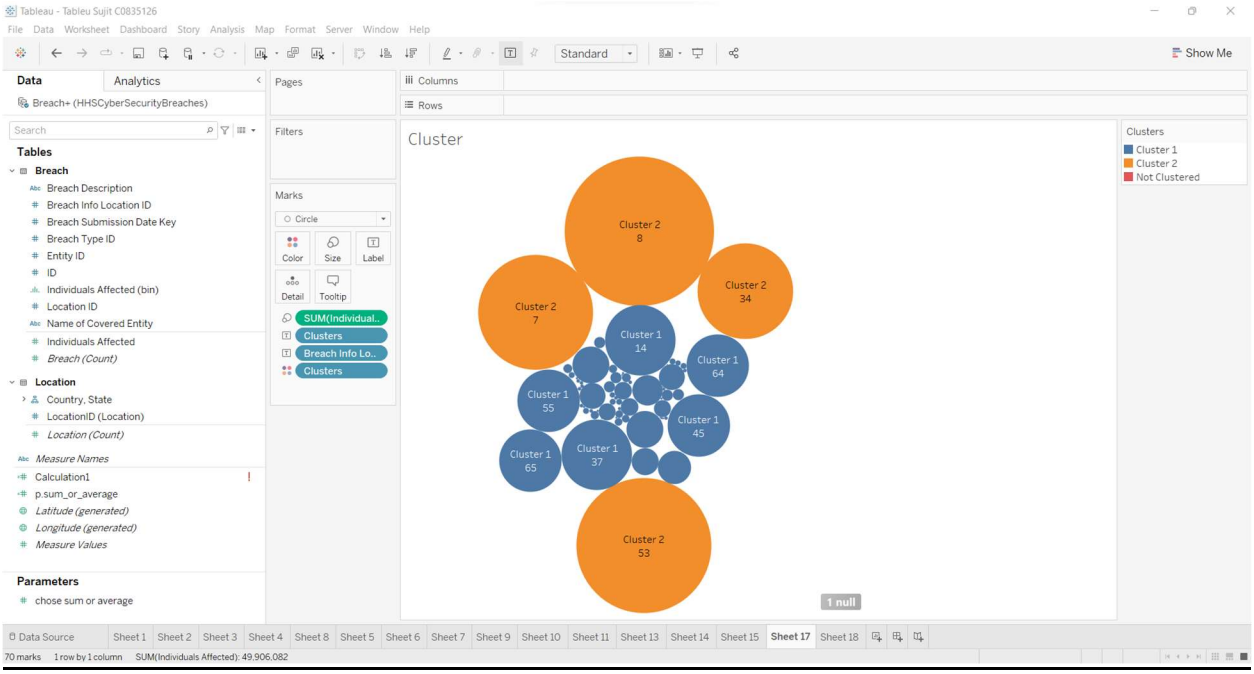
c. Pie Chart



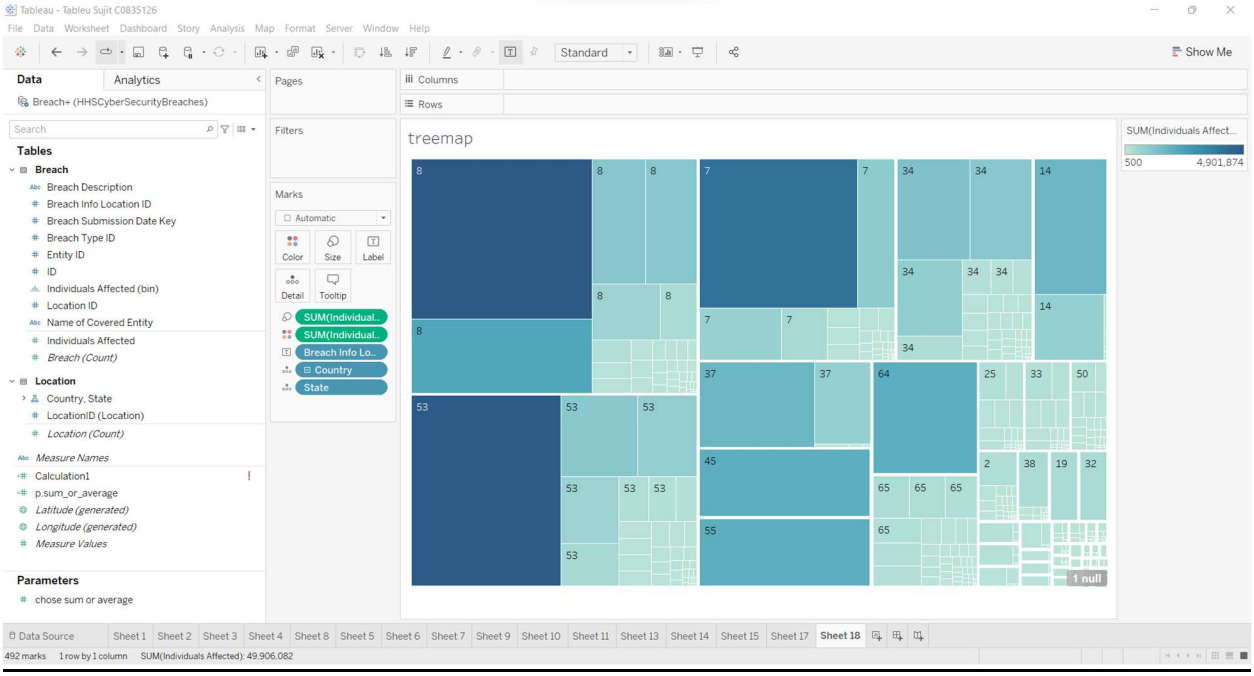
d. Scatter Chart



e. Data Cluster



f. Tree Map



10. Dashboard

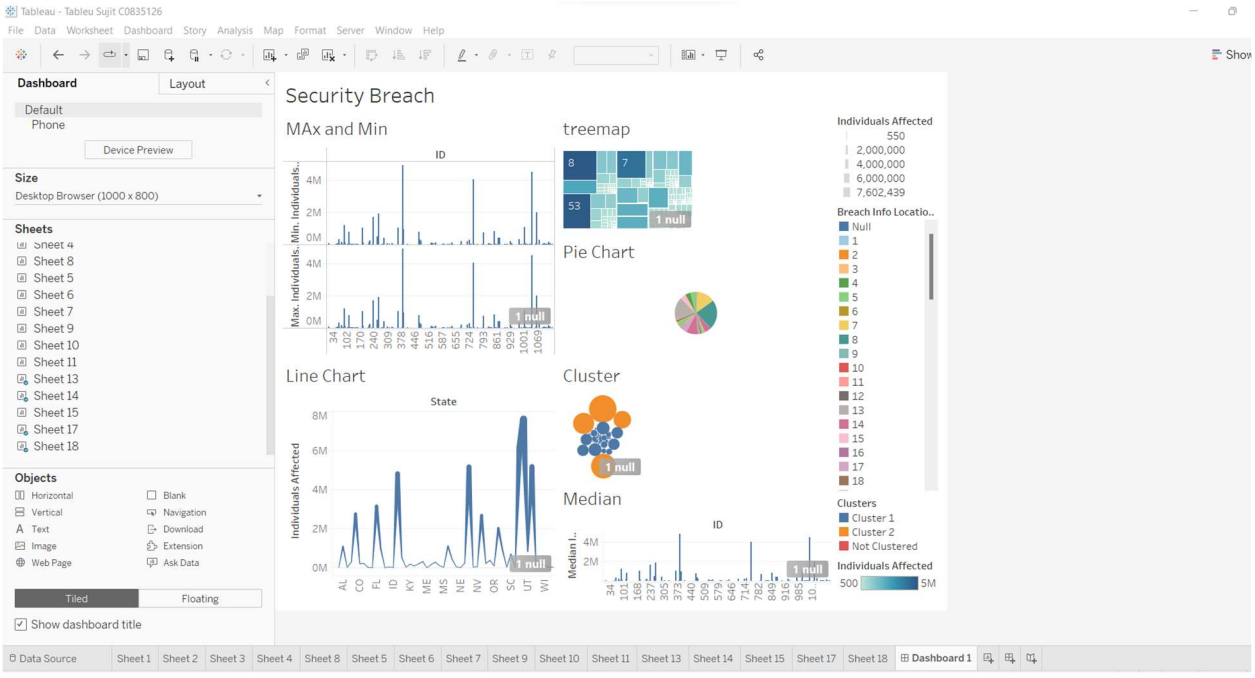


Tableau - Tableau Sujit C0835126
File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Dashboard
Layout
Default
Phone
Device Preview
Size
Desktop Browser (1000 x 800)
Sheets
Sheet 8
Sheet 5
Sheet 6
Sheet 7
Sheet 9
Sheet 10
Sheet 11
Sheet 13
Sheet 14
Sheet 15
Sheet 17
Sheet 18
Objects
Horizontal
Vertical
Text
Image
Web Page
Blank
Navigation
Download
Extension
Ask Data
Tiled
Floating
Show dashboard title
Security Breach
MAx and Min
ID
Max. individuals.. Min. individuals..
1 null
treemap
1 null
Pie Chart
Line Chart
State
Individuals Affected
1 null
Cluster
1 null
Median
ID
Median I..
1 null
Individuals Affected
550
2,000,000
4,000,000
6,000,000
7,602,439
Breach Info Locatio..
Null
Clusters
Cluster 1
Cluster 2
Not Clustered
Individuals Affected
500
5M
Data Source
Sheet 1
Sheet 2
Sheet 3
Sheet 4
Sheet 8
Sheet 5
Sheet 6
Sheet 7
Sheet 9
Sheet 10
Sheet 11
Sheet 13
Sheet 14
Sheet 15
Sheet 17
Sheet 18
Dashboard 1