

Business Intelligence (CIS 5270 )

Project-1: Tableau Project

*Topic: Call Center dataset analysis*

Team members: Nagender Chauhan & Lekha Ajit Kumar

## INTRODUCTION

We selected the Call Center dataset from the provided choices, and we plan to analyze the performance of call centers and channels. Analyzing the data based on customers is valuable and important for the company to know its customers. For instance, the firm can focus on cities that have less customer satisfaction and increase its performance. Knowing your customers, for example, their call reasons, average call durations, and their complaints, can give you some strategic ideas to improve the company and customer relations. Using the call center channels might give you investment ideas as to which channel you should improve, or stop the investment for that channel, or hire more talented staff for that channel. A consumer-based call center may help you improve complaint, response, and analytics rates, and solutions can evaluate call center staff and evaluate their customer service KPI indicators to multiple objectives. Because the data is provided in an interactive, intuitive, and visual manner, it is straightforward to see, absorb, evaluate, and share critical data, saving time and boosting cross-departmental cooperation.

In order to achieve our goals, we will be using various data analysis techniques, including exploratory data analysis, visualization, and statistical analysis. We will be looking for relationships between the different variables in our dataset and using data visualization tools to help us better understand our data. Our aim is to present the analysis by state and city using call center data, to demonstrate the performance of the call centers and the activity of the consumers, and to show the analysis of the call center channels.

Ultimately, our goal is to use the insights gained from the analysis to make data-driven decisions and improve the call center's performance. This project will provide valuable information that will help optimize operations and better serve the customers.

## DATA DESCRIPTION

Call Center	This column contains the location of the call centre. An example value: Baltimore/MD
Call Timestamp	It consists of the date when the call was made. An example value: 11/10/2020
Channel	The data set has 4 unique channels through which customers can reach out to the call centre. They are call centre, web, chatbot and email.
City	City name from which the call was issued An example value: Detroit
Customer Name	The name of the customer who raised an issue. An example value: Analise Gairdner
Id	Unique id for every call log An example value: QGK-72219678-w-102139-KY
Reason	Why did the customer call? This column has only 3 reasons for which a customer can call. They are Billing Questions, Payments and Service Outages.
Response Time	How long did it take to respond to the call? This column consists of their distinct values, they are Within SLA, Above SLA and Below SLA.
Sentiment	The feedback of the customer after the call. There are 6 distinct types of feedback the customer can give. They are as follows Very Negative, Negative, Neutral, Very Positive and Positive.
State	The state name from which the call was issued. An example value: Oregon
Avg Call Duration in Seconds	The Average call duration of a single call in seconds. An example value: 118
Avg Call Duration in Mins	The Average call duration of a single call in Minutes. An example value: 1:18
Average Customer Satisfaction	The average customer satisfaction rate is in the range of 1-10. An example value: 7
Call Duration	The call duration of a single call in Seconds. An example value: 120

Csat Score	Customer satisfaction score. The range of the score is between 1-10. An example value: 8
Sat rate	Total Satisfaction score. The range of the score is between 1-10. An example value: 6
Total Calls Received	It shows the Total Calls Received from a particular customer on a specific date. An example value: 1

## DATA CLEANING

1. We had no values for the column “Average customer Satisfaction” for a few of the rows (13% of the data set of about 33k rows) so we took an average of the remaining data for that column and added in the values via Excel formula.

### Pre-Data Cleaning

L1 X ✓ fx Average Customer Satisfaction

G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	son	Response Time Sentiment	State	Average Call Duration	Average Customer Satisfaction	Call Duration	Csat Score	sat rate	Total Calls Received				
2	ng Question	Within SLA	Neutral	Michigan	117	7	117	7	0.002125	1			
3	vice Outage	Within SLA	Very Positive	South Carolina	123		123			1			
4	ng Question	Above SLA	Negative	Florida	145		145			1			
5	ng Question	Within SLA	Very Negative	Oregon	112	1	112	1	0.0003036	1			
6	ments	Within SLA	Very Positive	Indiana	123		123			1			
7	ng Question	Within SLA	Neutral	Utah	125	5	125	5	0.0015179	1			
8	ng Question	Within SLA	Neutral	Texas	131	8	131	8	0.0024286	1			
9	ng Question	Below SLA	Positive	New York	137		137			1			
10	ng Question	Below SLA	Very Negative	Texas	137		137			1			
11	ng Question	Within SLA	Neutral	Ohio	112		112			1			
12	ng Question	Within SLA	Negative	Washington	135		135			1			
13	ng Question	Within SLA	Negative	West Virginia	118		118			1			
14	ng Question	Within SLA	Very Positive	Oregon	130		130			1			
15	ng Question	Within SLA	Positive	Massachusetts	120	7	120	7	0.002125	1			
16	ments	Below SLA	Very Negative	Maryland	122		122			1			
17	ng Question	Within SLA	Negative	New York	128	4	128	4	0.0012143	1			
18	ng Question	Above SLA	Negative	Alabama	136		136			1			
19	ng Question	Above SLA	Neutral	Kansas	137		137			1			
20	ng Question	Within SLA	Negative	Michigan	141		141			1			
21	ments	Within SLA	Very Negative	Michigan	109		109			1			
22	ng Question	Below SLA	Neutral	Indiana	135		135			1			
23	ng Question	Within SLA	Negative	California	127	5	127	5	0.0015179	1			
24	ng Question	Within SLA	Very Negative	California	120	2	120	2	0.0006071	1			
25	ng Question	Within SLA	Negative	Tennessee	143		143			1			
26	vice Outage	Below SLA	Negative	Tennessee	141		141			1			
27	ng Question	Below SLA	Positive	Texas	145		145			1			

### Post-Data Cleaning

L1	G	H	I	J	K	L	M	N	O	P
	Response Time	Sentiment	State	Average Call Duration	Average Customer Satisfaction	Call Duration	Csat Score	Rate	Total Calls Received	
1	son	Within SLA	Neutral	Michigan	117	7	117	7	0.002125	1
2	ng Question	Within SLA	Very Positive	South Carolina	123	6	123			1
3	vice Outage	Within SLA	Very Negative	Oregon	112	1	112	1	0.0003036	1
4	ng Question	Above SLA	Negative	Florida	145	6	145			1
5	ng Question	Within SLA	Very Positive	Indiana	123	6	123			1
6	ments	Within SLA	Neutral	Utah	125	5	125	5	0.0015179	1
7	ng Question	Within SLA	Neutral	Texas	131	8	131	8	0.0024286	1
8	ng Question	Below SLA	Positive	New York	137	6	137			1
9	ng Question	Below SLA	Very Negative	Texas	137	6	137			1
10	ng Question	Within SLA	Neutral	Ohio	112	6	112			1
11	ng Question	Within SLA	Negative	Washington	135	6	135			1
12	ng Question	Within SLA	Negative	West Virginia	118	6	118			1
13	ng Question	Within SLA	Very Positive	Oregon	130	6	130			1
14	ng Question	Within SLA	Positive	Massachusetts	120	7	120	7	0.002125	1
15	ng Question	Below SLA	Very Negative	Maryland	122	6	122			1
16	ments	Within SLA	Negative	New York	128	4	128	4	0.0012143	1
17	ng Question	Above SLA	Negative	Alabama	136	6	136			1
18	ng Question	Above SLA	Neutral	Kansas	137	6	137			1
19	ng Question	Within SLA	Negative	Michigan	141	6	141			1
20	ng Question	Within SLA	Very Negative	Michigan	109	6	109			1
21	ments	Below SLA	Neutral	Indiana	135	6	135			1
22	ng Question	Within SLA	Negative	California	127	5	127	5	0.0015179	1
23	ng Question	Within SLA	Very Negative	California	120	2	120	2	0.0006071	1
24	ng Question	Within SLA	Negative	Tennessee	143	6	143			1
25	ng Question	Below SLA	Negative	Tennessee	141	6	141			1
26	vice Outage	Below SLA	Positive	Texas	145	6	145			1
27	ng Question	Below SLA	Positive	Texas	145					

2. We have a column in our data set named “Channel” which specifies the mode of communication. This had a few “NULL” values (about 1% of the dataset). We took an average of each of the different modes of communication and assigned values based on the highest average in that category.

## Pre-Data Cleaning

C1	A	B	C	D	E	F	G	H	I	J	K
	Call Center	Call Timestamp	Channel	City	Customer Name	Id	Reason	Response Time	Sentiment	State	Average Call Duration
1	Call Center	29/10/20	Call-Center	Detroit	Analise Gairdne	DKK-570768	Billing Question	Within SLA	Neutral	Michigan	117
2	Los Angeles/CA	05/10/20	NULL	Spartanburg	Crichton Kilday	QGK-722196	Service Outage	Within SLA	Very Positive	South Carolina	123
3	Baltimore/MD	04/10/20	Call-Center	Gainesville	Averill Brundret	GVI-300259	Billing Question	Above SLA	Negative	Florida	145
4	Los Angeles/CA	17/10/20	Chatbot	Portland	Noreen Lafflina	ZJI-9680755	Billing Question	Within SLA	Very Negative	Oregon	112
5	Los Angeles/CA	17/10/20	NULL	Fort Wayne	Toma Van der B	DDU-694517	Payments	Within SLA	Very Positive	Indiana	123
6	Baltimore/MD	28/10/20	Call-Center	Salt Lake City	Kaylyn Emlen	JVI-972866	Billing Question	Within SLA	Neutral	Utah	125
7	Baltimore/MD	16/10/20	Chatbot	Tyler	Philippe Bowring	AZI-9505405	Billing Question	Within SLA	Neutral	Texas	131
8	Los Angeles/CA	21/10/20	Chatbot	New York City	Krysta de Tocqu	TWX-270075	Billing Question	Below SLA	Positive	New York	137
9	Baltimore/MD	03/10/20	Email	Dallas	Oran Lifsey	XNG-445991	Billing Question	Below SLA	Very Negative	Texas	137
10	Baltimore/MD	07/10/20	Chatbot	Cincinnati	Port Ingall	RLC-641082	Billing Question	Within SLA	Neutral	Ohio	112
11	Los Angeles/CA	09/10/20	Chatbot	Everett	Ella Cristoforo	RJF-0026392	Billing Question	Within SLA	Negative	Washington	135
12	Los Angeles/CA	11/10/20	Web	Huntington	Aubrey Surcomi	ZQN-328748	Billing Question	Within SLA	Negative	West Virginia	118
13	Baltimore/MD	02/10/20	NULL	Portland	Nicole Farewea	IDP-3514751	Billing Question	Within SLA	Very Positive	Oregon	130
14	Denver/CO	10/10/20	Chatbot	Springfield	Melesa Ricardot	DPT-564834	Billing Question	Within SLA	Positive	Massachusetts	120
15	Baltimore/MD	06/10/20	Call-Center	Hyattsville	Odell Catherseyed	ZOV-958613	Payments	Below SLA	Very Negative	Maryland	122
16	Denver/CO	18/10/20	Chatbot	New York City	Dani Stanfield	BEJ-6971144	Billing Question	Within SLA	Negative	New York	128
17	Baltimore/MD	11/10/20	Email	Huntsville	Margarette Jehn	DEC-837672	Billing Question	Above SLA	Negative	Alabama	136
18	Baltimore/MD	30/10/20	Call-Center	Wichita	Noni Greatrakes	XNY-041063	Billing Question	Above SLA	Neutral	Kansas	137
19	Baltimore/MD	26/10/20	Web	Lansing	Gerik Archell	GKH-065325	Billing Question	Within SLA	Negative	Michigan	141
20	Chicago/IL	11/10/20	Call-Center	Lansing	Tammie Bettins	DUJ-199778	Payments	Within SLA	Very Negative	Michigan	109
21	Baltimore/MD	12/10/20	Chatbot	Fort Wayne	Errol Follos	ADD-822192	Billing Question	Below SLA	Neutral	Indiana	135
22	Baltimore/MD	08/10/20	NULL	Hayward	Nanni Doy	YOB-404922	Billing Question	Within SLA	Negative	California	127
23	Chicago/IL	03/10/20	Chatbot	Santa Barbara	Sophie Kleinerm	GZD-504595	Billing Question	Within SLA	Very Negative	California	120
24	Baltimore/MD	28/10/20	Call-Center	Memphis	Timotheus Men	FQX-241188	Billing Question	Within SLA	Negative	Tennessee	143
25	Baltimore/MD	06/10/20	Chatbot	Murfreesboro	Allayne Lednor	SVH-328807	Service Outage	Below SLA	Negative	Tennessee	141
26	Denver/CO	22/10/20	Chatbot	Lubbock	Bethina Fazzoli	ISK-9496544	Billing Question	Below SLA	Positive	Texas	145

## Post-Data Cleaning

	A	B	C	D	E	F	G	H	I	J	K
	Call Center	Call Timestamp	Channel	City	Customer Name	Id	Reason	Response	Sentiment	State	Average Call Duration
1	Call Center	29/10/20	Call-Center	Detroit	Analise Gairdne	DKK-570768	Billing Question	Within SLA	Neutral	Michigan	117
2	Los Angeles/CA	29/10/20	Call-Center	Spartanburg	Crichton Kidsley	QGK-722196	Service Outage	Within SLA	Very Positive	South Carolina	123
3	Baltimore/MD	05/10/20	Call-Center	Gainesville	Averill Brundrett	GY-300259	Billing Question	Above SLA	Negative	Florida	145
4	Los Angeles/CA	04/10/20	Call-Center	Portland	Noreen Lafflina	ZJ-9680755	Billing Question	Within SLA	Very Negative	Oregon	112
5	Los Angeles/CA	17/10/20	Chatbot	Fort Wayne	Toma Van der B	DDU-694517	Payments	Within SLA	Very Positive	Indiana	123
6	Los Angeles/CA	17/10/20	Call-Center	Salt Lake City	Kaylyn Emlen	JV1-7972866	Billing Question	Within SLA	Neutral	Utah	125
7	Baltimore/MD	28/10/20	Call-Center	Tyler	Phillipe Bowring	AZI-9505405	Billing Question	Within SLA	Neutral	Texas	131
8	Baltimore/MD	16/10/20	Chatbot	New York City	Krysta de Tocou	TWX-270075	Billing Question	Below SLA	Positive	New York	137
9	Los Angeles/CA	21/10/20	Chatbot	Dallas	Oran Lifsey	XNG-445991	Billing Question	Below SLA	Very Negative	Texas	137
10	Baltimore/MD	03/10/20	Email	Cincinnati	Port Ingall	RLC-641082	Billing Question	Within SLA	Neutral	Ohio	112
11	Baltimore/MD	07/10/20	Chatbot	Everett	Ella Cristoforo	RJF-0026392	Billing Question	Within SLA	Negative	Washington	135
12	Los Angeles/CA	09/10/20	Chatbot	Huntington	Aubrey Surcom	ZQN-328748	Billing Question	Within SLA	Negative	West Virginia	118
13	Los Angeles/CA	11/10/20	Web	Portland	Nicolle Fareweia	JDP-351475	Billing Question	Within SLA	Very Positive	Oregon	130
14	Baltimore/MD	02/10/20	Call-Center	Springfield	Melesa Ricardot	DPT-564834	Billing Question	Within SLA	Positive	Massachusetts	120
15	Denver/CO	10/10/20	Chatbot	Hyattsville	Odell Cathesyed	ZOV-958613	Payments	Below SLA	Very Negative	Maryland	122
16	Baltimore/MD	06/10/20	Call-Center	Huntsville	Dani Stanfield	BEJ-6971144	Billing Question	Within SLA	Negative	New York	128
17	Denver/CO	18/10/20	Chatbot	Huntsville	Margarette Jeh	DEC-837672	Billing Question	Above SLA	Negative	Alabama	136
18	Baltimore/MD	11/10/20	Email	Wichita	Noni Greatrakes	XNY-041063	Billing Question	Above SLA	Neutral	Kansas	137
19	Baltimore/MD	30/10/20	Call-Center	Lansing	Gerik Archell	GKH-065325	Billing Question	Within SLA	Negative	Michigan	141
20	Baltimore/MD	26/10/20	Web	Lansing	Tammie Bettins	DJU-199778	Payments	Within SLA	Very Negative	Michigan	109
21	Chicago/IL	11/10/20	Call-Center	Fort Wayne	Errol Follos	ADD-822192	Billing Question	Below SLA	Neutral	Indiana	135
22	Baltimore/MD	12/10/20	Chatbot	Hayward	Nanni Doy	YOB-404922	Billing Question	Within SLA	Negative	California	127
23	Baltimore/MD	08/10/20	Call-Center	Santa Barbara	Sophie Kleinerm	GZD-504595	Billing Question	Within SLA	Very Negative	California	120
24	Chicago/IL	03/10/20	Chatbot	Memphis	Timotheus Men	FQX-241188	Billing Question	Within SLA	Negative	Tennessee	143
25	Baltimore/MD	28/10/20	Call-Center	Murfreesboro	Allayne Lednor	SVH-328807	Service Outage	Below SLA	Negative	Tennessee	141
26	Baltimore/MD	06/10/20	Chatbot	Lubbock	Bethina Fazzoli	ISK-9496544	Billing Question	Below SLA	Positive	Texas	145
27	Denver/CO	22/10/20	Chatbot	New York City	Stanwood Esley	WUU-907277	Billing Question	Within SLA	Very Negative	New York	119
28	Baltimore/MD	01/10/20	Call-Center	Oklahoma City	Anissa Kirnade	PKG-516912	Payments	Within SLA	Very Positive	Oklahoma	140
29	Chicago/IL	05/10/20	Call-Center	Omaha	Bradly Dinkin	YSU-893933	Billing Question	Below SLA	Very Positive	Nebraska	106
30	Chicago/IL	30/10/20	Chatbot								

3. The column “Call Timestamp” was in the incorrect format i.e., dd/mm/yy. We changed the format to mm/dd/yyyy as per the US format.

## Pre-Data Cleaning

	A	B	C	Formula Bar	
1	Call Center	Call Timestamp	Channel	City	Custom
2	Los Angeles/CA	29/10/20	Call-Center	Detroit	Analise
3	Baltimore/MD	05/10/20	Call-Center	Spartanburg	Crichton
4	Los Angeles/CA	04/10/20	Call-Center	Gainesville	Averill E
5	Los Angeles/CA	17/10/20	Chatbot	Portland	Noreen
6	Los Angeles/CA	17/10/20	Call-Center	Fort Wayne	Toma V
7	Baltimore/MD	28/10/20	Call-Center	Salt Lake City	Kaylyn I
8	Baltimore/MD	16/10/20	Chatbot	Tyler	Phillipe
9	Los Angeles/CA	21/10/20	Chatbot	New York City	Krysta C
10	Baltimore/MD	03/10/20	Email	Dallas	Oran L
11	Baltimore/MD	07/10/20	Chatbot	Cincinnati	Port Ing
12	Los Angeles/CA	09/10/20	Chatbot	Everett	Ella Cris
13	Los Angeles/CA	11/10/20	Web	Huntington	Aubrey
14	Baltimore/MD	02/10/20	Call-Center	Portland	Nicolle I
15	Denver/CO	10/10/20	Chatbot	Springfield	Melesa
16	Baltimore/MD	06/10/20	Call-Center	Hyattsville	Odell C
17	Denver/CO	18/10/20	Chatbot	New York City	Dani St
18	Baltimore/MD	11/10/20	Email	Huntsville	Margare
19	Baltimore/MD	30/10/20	Call-Center	Wichita	Noni Gr
20	Baltimore/MD	26/10/20	Web	Lansing	Gerik A
21	Chicago/IL	11/10/20	Call-Center	Lansing	Tammie
22	Baltimore/MD	12/10/20	Chatbot	Fort Wayne	Errol Fo
23	Baltimore/MD	08/10/20	Call-Center	Hayward	Nanni I
24	Chicago/IL	03/10/20	Chatbot	Santa Barbara	Sophie I
25	Baltimore/MD	28/10/20	Call-Center	Memphis	Timoth
26	Baltimore/MD	06/10/20	Chatbot	Murfreesboro	Allayne

## Post-Data Cleaning

	A	B	C	D	Custom
1	Call Center	Call Timestamp	Channel	City	Custom
2	Los Angeles/CA	10/29/2020	Call-Center	Detroit	Analise
3	Baltimore/MD	10/05/2020	Call-Center	Spartanburg	Crichton
4	Los Angeles/CA	10/04/2020	Call-Center	Gainesville	Averill I
5	Los Angeles/CA	10/17/2020	Chatbot	Portland	Noreer
6	Los Angeles/CA	10/17/2020	Call-Center	Fort Wayne	Toma V
7	Baltimore/MD	10/28/2020	Call-Center	Salt Lake City	Kaylyn I
8	Baltimore/MD	10/16/2020	Chatbot	Tyler	Phillipe
9	Los Angeles/CA	10/21/2020	Chatbot	New York City	Krysta C
10	Baltimore/MD	10/03/2020	Email	Dallas	Oran L
11	Baltimore/MD	10/07/2020	Chatbot	Cincinnati	Port Ing
12	Los Angeles/CA	10/09/2020	Chatbot	Everett	Ella Cris
13	Los Angeles/CA	10/11/2020	Web	Huntington	Aubrey
14	Baltimore/MD	10/02/2020	Call-Center	Portland	Nicolle I
15	Denver/CO	10/10/2020	Chatbot	Springfield	Melesa
16	Baltimore/MD	10/06/2020	Call-Center	Hyattsville	Odell C
17	Denver/CO	10/18/2020	Chatbot	New York City	Dani St
18	Baltimore/MD	10/11/2020	Email	Huntsville	Margare
19	Baltimore/MD	10/30/2020	Call-Center	Wichita	Noni Gr
20	Baltimore/MD	10/26/2020	Web	Lansing	Gerik A
21	Chicago/IL	10/11/2020	Call-Center	Lansing	Tammie
22	Baltimore/MD	10/12/2020	Chatbot	Fort Wayne	Errol Fo
23	Baltimore/MD	10/08/2020	Call-Center	Hayward	Nanni I
24	Chicago/IL	10/03/2020	Chatbot	Santa Barbara	Sophie I
25	Baltimore/MD	10/28/2020	Call-Center	Memphis	Timoth

4. We have a column called “sentiment” which describes the tone of the call and is denoted by the numbers from 1 to 5, where 1 represents very negative and 5 represents very positive and 3 represents normal. We assigned the numbers with their corresponding text for better readability

## Pre-Data Cleaning

G	H	I	J	K	
Reason	Response	Sentiment	State	Average Call Duration	Average Customer Satisfaction
Question	Within SLA	3	Michigan	117	
Outage	Within SLA	5	South Carolina	123	
Question	Above SLA	2	Florida	145	
Question	Within SLA	1	Oregon	112	
Comments	Within SLA	5	Indiana	123	
Question	Within SLA	3	Utah	125	
Question	Within SLA	3	Texas	131	
Question	Below SLA	4	New York	137	
Question	Below SLA	1	Texas	137	
Question	Within SLA	3	Ohio	112	
Question	Within SLA	2	Washington	135	
Question	Within SLA	2	West Virginia	118	
Question	Within SLA	5	Oregon	130	
Comments	Within SLA	4	Massachusetts	120	
Comments	Below SLA	1	Maryland	122	
Question	Within SLA	2	New York	128	
Question	Above SLA	2	Alabama	136	
Question	Above SLA	3	Kansas	137	
Comments	Within SLA	2	Michigan	141	
Comments	Within SLA	1	Michigan	109	
Question	Below SLA	3	Indiana	135	
Question	Within SLA	2	California	127	
Question	Within SLA	1	California	120	

## Post-Data Cleaning

G	H	I	J	K	
Reason	Response	Sentiment	State	Average Call Duration	Average Customer Satisfaction
Question	Within SLA	Neutral	Michigan	117	
Outage	Within SLA	Very Positive	South Carolina	123	
Question	Above SLA	Negative	Florida	145	
Question	Within SLA	Very Negative	Oregon	112	
Comments	Within SLA	Very Positive	Indiana	123	
Question	Within SLA	Neutral	Utah	125	
Question	Within SLA	Neutral	Texas	131	
Question	Below SLA	Positive	New York	137	
Question	Below SLA	Very Negative	Texas	137	
Question	Within SLA	Neutral	Ohio	112	
Question	Within SLA	Negative	Washington	135	
Question	Within SLA	Negative	West Virginia	118	
Question	Within SLA	Very Positive	Oregon	130	
Comments	Within SLA	Positive	Massachusetts	120	
Comments	Below SLA	Very Negative	Maryland	122	
Question	Within SLA	Negative	New York	128	
Question	Above SLA	Negative	Alabama	136	
Question	Above SLA	Neutral	Kansas	137	
Question	Within SLA	Negative	Michigan	141	
Comments	Within SLA	Very Negative	Michigan	109	
Question	Below SLA	Neutral	Indiana	135	
Question	Within SLA	Negative	California	127	
Question	Within SLA	Very Negative	California	120	
Question	Within SLA	Negative	Tennessee	143	
Outage	Below SLA	Negative	Tennessee	141	

5. The column “Average call Duration” had the data in number format. We added another column to represent the data in mm:ss format by using the Excel formula  
“=TEXT(K3/(24\*60\*60), "mm:ss")”

## Pre-Data Cleaning

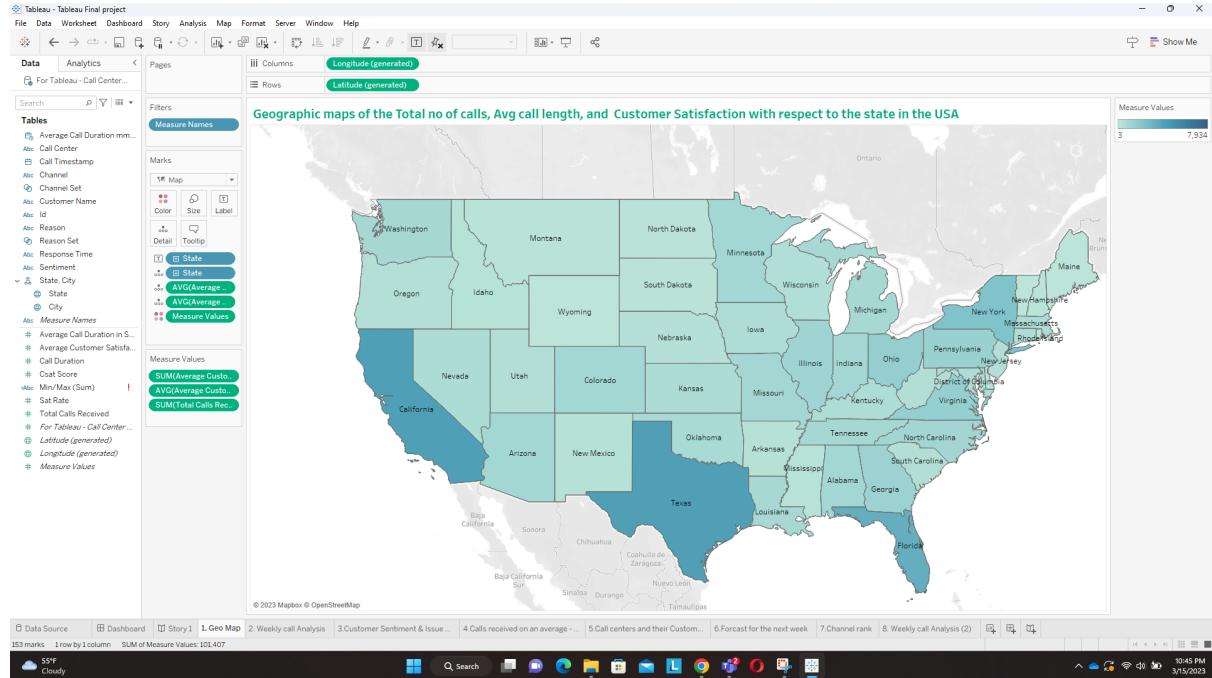
	I	J	K	L
	Sentiment	State	Average Call Duration	Average Customer Satisfaction
A	Neutral	Michigan	117	
A	Very Positive	South Carolina	123	
A	Negative	Florida	145	
A	Very Negative	Oregon	112	
A	Very Positive	Indiana	123	
A	Neutral	Utah	125	
A	Neutral	Texas	131	
A	Positive	New York	137	
A	Very Negative	Texas	137	
A	Neutral	Ohio	112	
A	Negative	Washington	135	
A	Negative	West Virginia	118	
A	Very Positive	Oregon	130	
A	Positive	Massachusetts	120	
A	Very Negative	Maryland	122	
A	Negative	New York	128	
A	Negative	Alabama	136	
A	Neutral	Kansas	137	
A	Negative	Michigan	141	
A	Very Negative	Michigan	109	
A	Neutral	Indiana	135	
A	Negative	California	127	

## Post-Data Cleaning

I	J	K	L	M
ment	State	Avg Call Duration in Seconds	Average Call Duration	Average Customer Satisfaction
ral	Michigan	117	01:57	
Positive	South Carolina	123	02:03	
itive	Florida	145	02:25	
Negative	Oregon	112	01:52	
Positive	Indiana	123	02:03	
ral	Utah	125	02:05	
ral	Texas	131	02:11	
ive	New York	137	02:17	
Negative	Texas	137	02:17	
ral	Ohio	112	01:52	
itive	Washington	135	02:15	
itive	West Virginia	118	01:58	
Positive	Oregon	130	02:10	
ive	Massachusetts	120	02:00	
Negative	Maryland	122	02:02	
itive	New York	128	02:08	
itive	Alabama	136	02:16	
ral	Kansas	137	02:17	
itive	Michigan	141	02:21	
Negative	Michigan	109	01:49	
ral	Indiana	135	02:15	
itive	California	127	02:07	
Negative	California	120	02:00	
itive	Tennessee	143	02:23	
itive	Tennessee	141	02:21	
ive	Texas	145	02:25	
Negative	New York	119	01:59	
Positive	Oklahoma	140	02:20	

## DATA VISUALIZATIONS

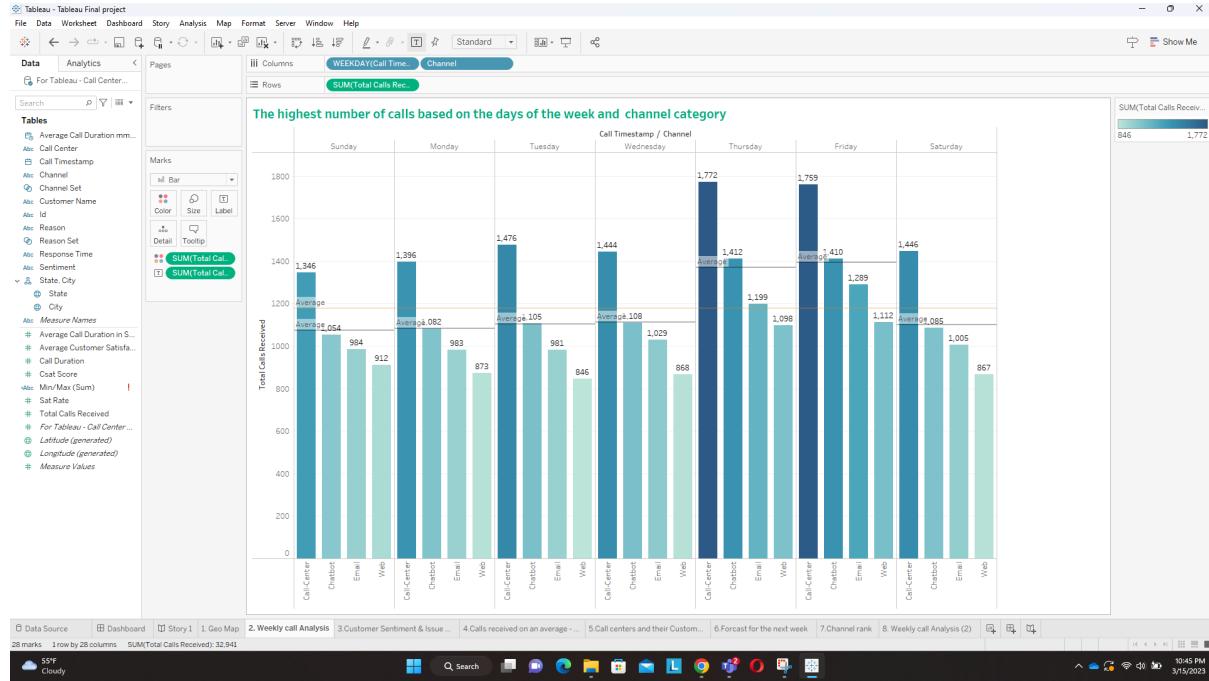
1. Geographic maps of the Total no of calls, Avg call length, and Customer Satisfaction with respect to the state in the USA.



From this above map, we can see which state is receiving the highest number of calls and what is the average duration of each call in seconds. We can also see the average customer satisfaction rate in each state. In the above screenshot, you can see Texas's records. The call centres in Texas receive 3572 calls per month, the call duration is approximately 125 seconds on an average. It also shows that the average customer satisfaction rate is 5.6. Also, It can be seen easily from the map that California has the largest amount of total calls. It can be useful to compare California and other states' data to deduct reasoning because the state with the highest population is California.

As we have columns like state and city, we have implemented Geo Maps in the above visualisation.

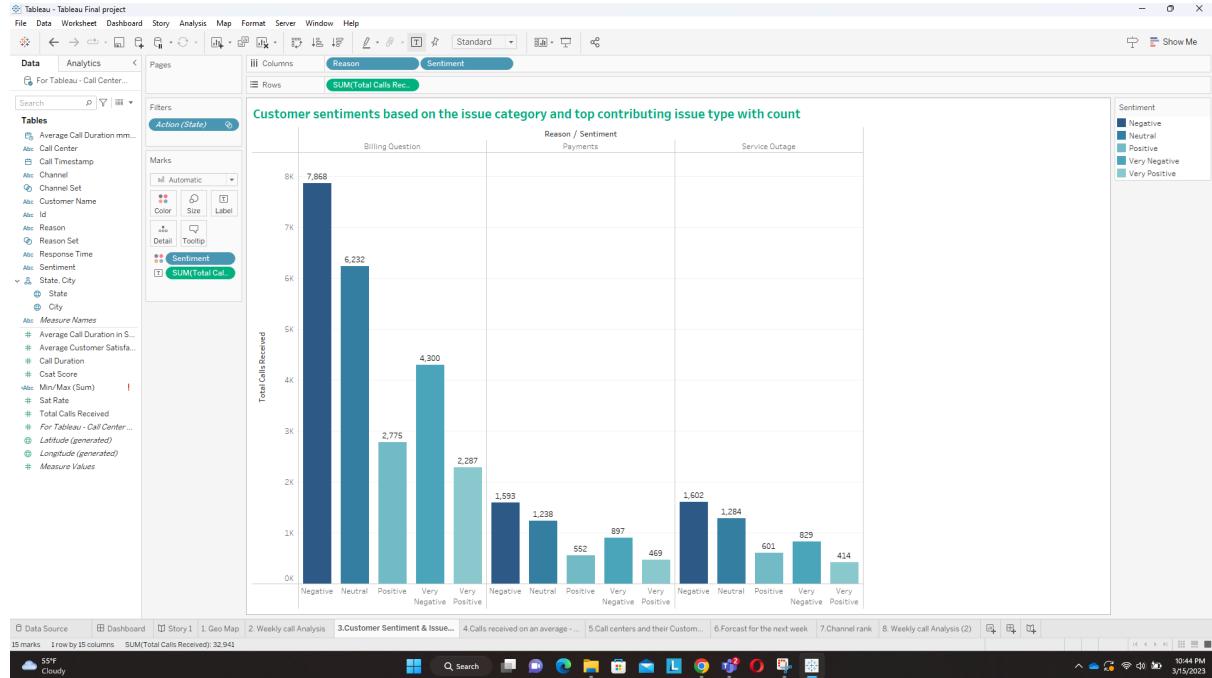
## 2. The highest number of calls based on the days of the week and channel category



From this above chart, we can see on which day of the week we receive the most number of complaints and for what particular issue they have raised an issue. We can see on Thursday and Friday we receive the maximum number of calls and complaints from all the channels. We also see the most preferred channel to raise an issue is the call centre followed by the chatbot, email and then the web. We also show the trend line and average reference line for each weekday and the total number of calls received during a month. On the line graph, the users can see that after Thursday, there was an increase in all sentiments, because the largest amount of calls was received on Thursday and Friday. It is effective to show dates with line graphs such as changing the total number of sentiments by weekdays.

In this above visualization, we use a bar chart with an average reference line to display the average calls received for each weekday and average calls received during a span of a month.

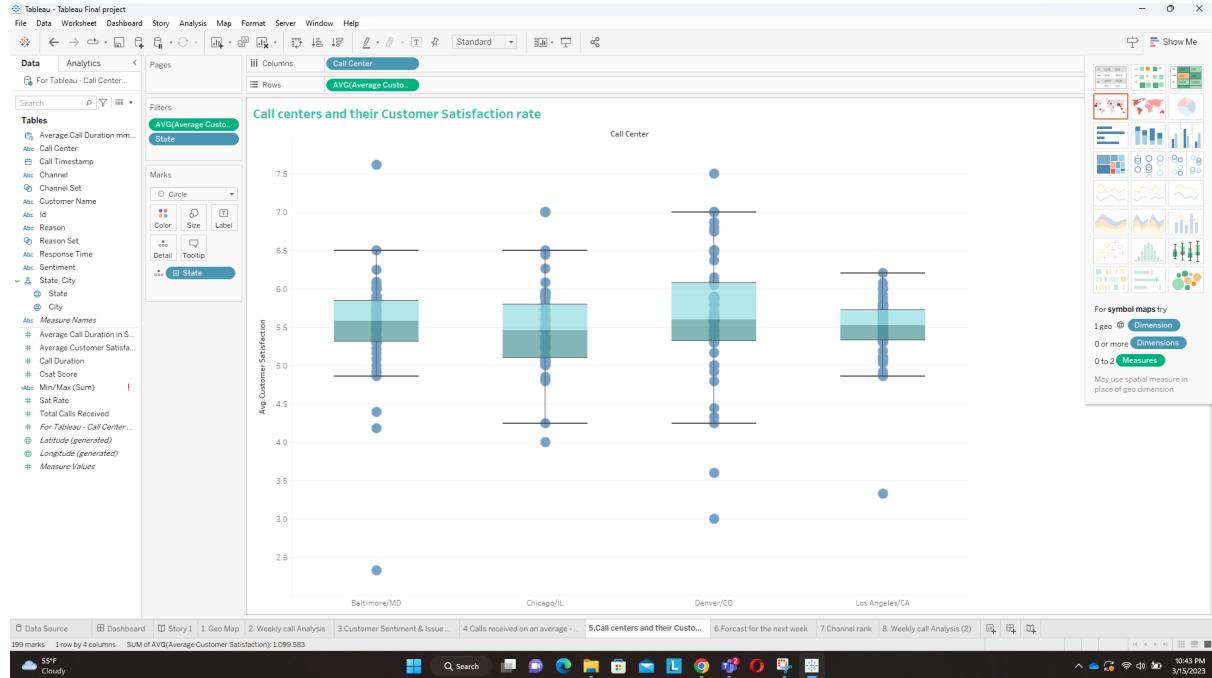
3. Customer sentiments based on the issue category and top contributing issue type with count In this cart we try to show.



In this visualization, we can see what is the most common type of issue for which a query is raised by the customer and what is the customer feedback of the customer after the call. We see that there are three main reasons namely Billing questions, Payments and service outages. Most of the queries are raised due to billing questions followed by Service outages and then payments. We can also see the customer feedback or response after the query was raised. Was it very negative, Negative, very positive, positive or neutral feedback? From this above visualization, we can see that most of the customers had negative feedback overall. Billing questions had the highest number of negative feedback. 71.2% of the total calls, that is, 23,462 customers, called the call centres due to a billing question.

In this chart, we use a simple bar chart and we filter it out for all the states to know the overall customer sentiment for each of the corresponding issues.

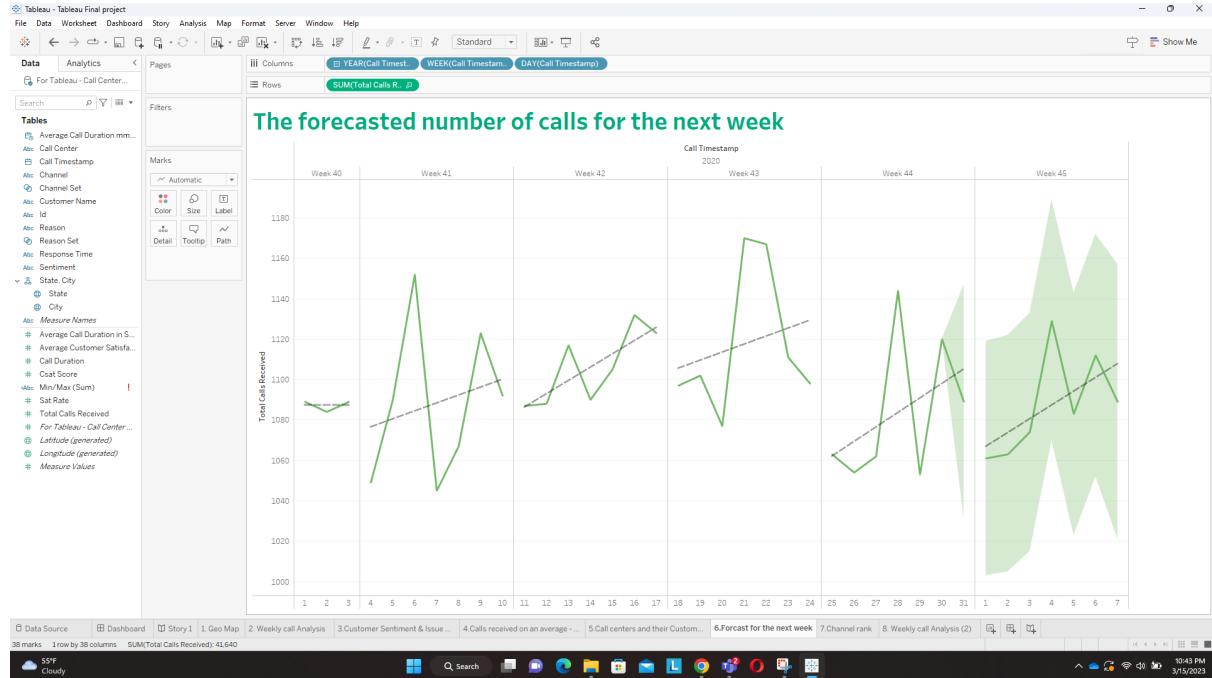
#### 4. Call centers and their customer satisfaction rate



There are a total of four call centres in the united states for this specific dataset. The four call centres are picking up calls from users who are located in the united states. These four call centres are located in Baltimore/MD, Chicago/IL, Denver/CO, and Los Angeles/CA. In this chart, we try to show the customer satisfaction rate for each of these call centres. It shows the customer satisfaction rate for calls it receives from each of the states. In this chart, we see that Denver has the highest customer satisfaction rate followed by Baltimore, Los Angeles and Chicago. We can also see specific customer satisfaction for the calls a particular call centre receives from a particular state. For example, the Denver call centre has a 7.5 customer satisfaction score for calls it receives from the state of Montana.

In this visualization, we use the block and whisker plot to see the average customer satisfaction rate for each of the call centres.

## 5. The forecasted number of calls for the next week.

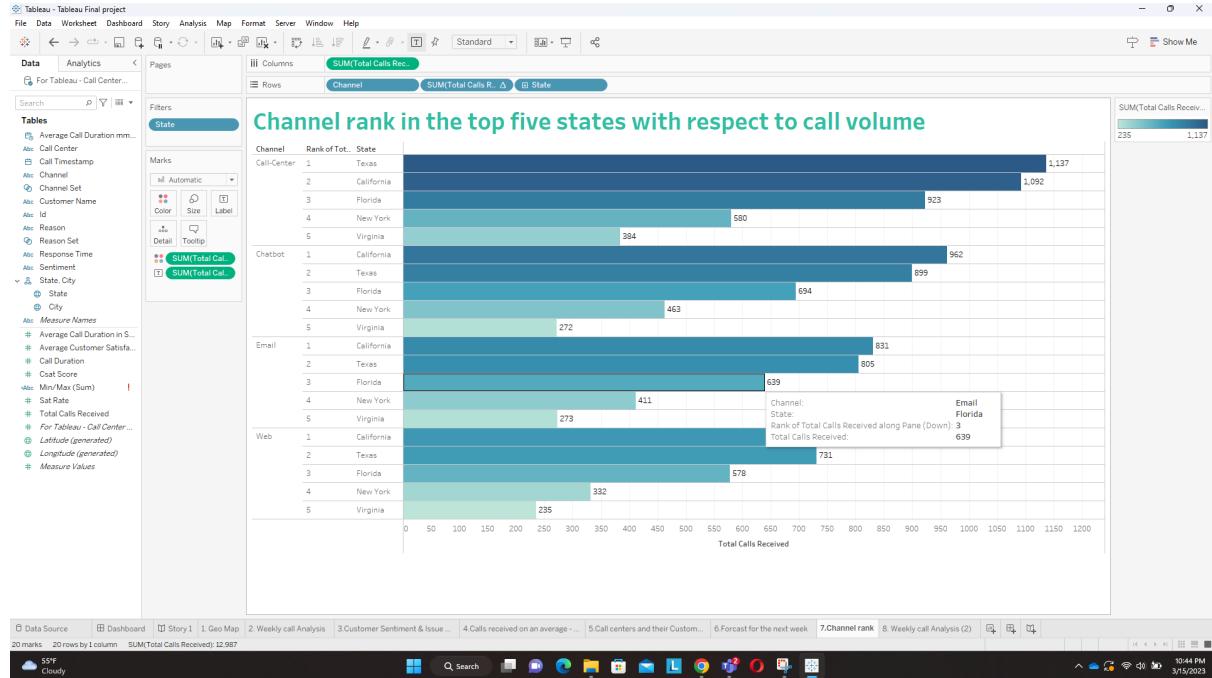


In this chart, we try to predict the number of calls the call centre might receive next week. We have October months data. We split the data into weeks, now it gives us an idea of how many calls we received during a particular week.

Now we have 4 weeks of data, and hence we can predict the following week's estimated number of calls. We have predicted how many calls we may receive in the first week of November. We also show the trend line and recalculated line.

In this cart, we have used the forecast trend lines concept and reference line to forecast next week's call volume. We have also used the concept of the date to split the data into 4 weeks.

6. Channel rank in the top five states with respect to call volume.

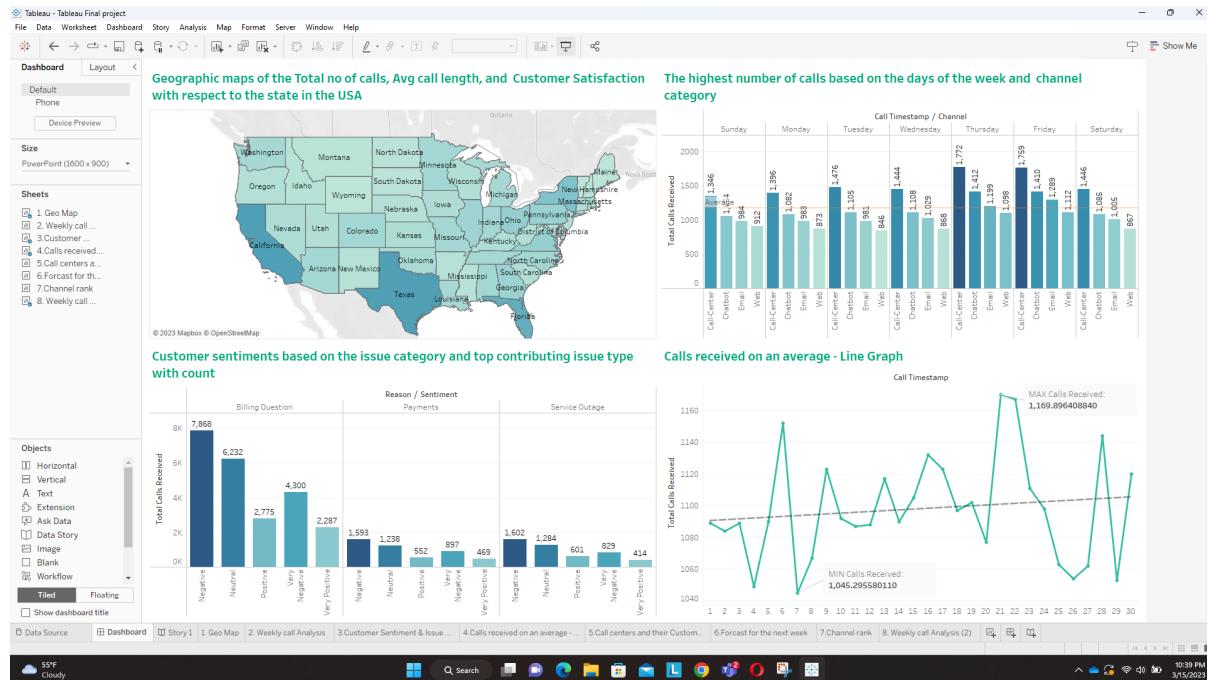


In this cart, we find out which state prefers what type of channel to address its issues. The four channels used by the company are Call-centres, chatbots, emails and the web to resolve customer issues. We rearranged the data using ranks, to find out only the top 5 states with the most call volume. We used ranks to filter out only the top 5 states, cause previously we had data for all the states which made the visualization very lengthy and difficult to read.

From this above visualization, we can see that people in Texas prefer solving their issues via a call center while users in California prefer solving their queries over the chatbot.

We have displayed the above findings using a simple bar diagram with ranks to find out the states and their preferred communication mode.

## DASHBOARD



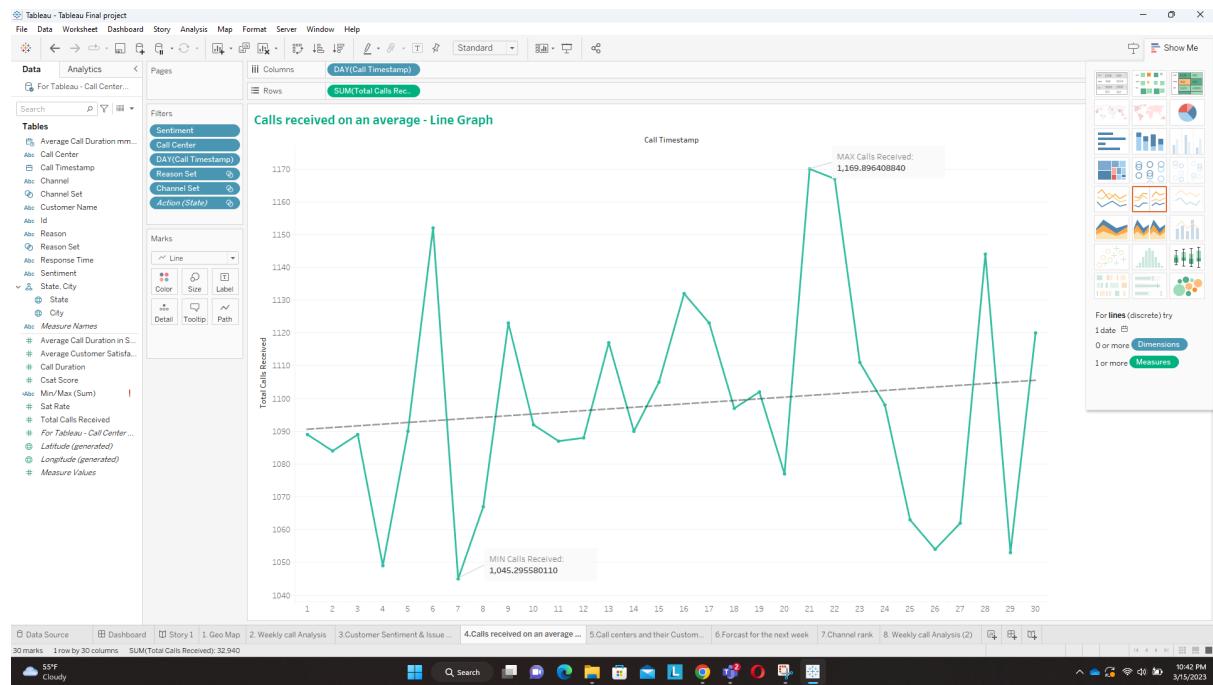
The dashboard aims to show some general specifications of the call center and States. It contains several visuals which all can be filtered by state to show the specifications. For instance, the States map is in accordance with the total number of calls received to show which state is the busiest in terms of calls, and it is always useful for the users to compare those states with others. It helps you to observe the data from a wide perspective of the states. You can select each state individually to know the total number of calls it receives on a particular day, what was the customer feedback for each of its call and using which channel most of the issues were sought out. We also show the customer satisfaction rate and average duration of the calls for each of the states. You can see a map distributed proportionally with shades of green based on the total number of calls. The darkest hue indicates the busiest state in terms of total calls. we chose the map to show this data because it is the proper way to show that geographical data and all states are distributed.

Knowing where your customers come from and how they reach you is often helpful and gives ideas about what to do with your business. Getting to know customers means you can always offer them better services or offers. To show the most popular cities and channels which are

used to reach the call centers, we chose a bar chart because helps to compare the different sets of data among different groups easily.

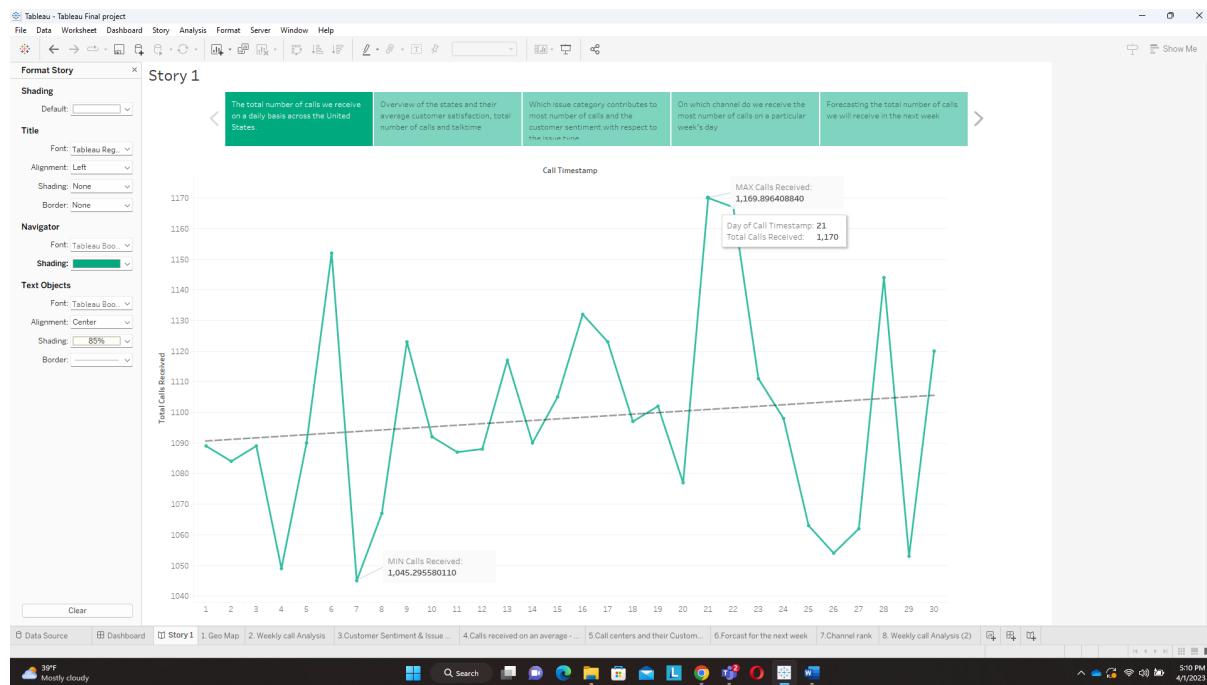
The chart we used especially for the dashboard is the following.

Calls received on an average - Line Graph. We used this chart so that when the user clicks on a particular state in the Geo-Map, he/she can see the total number of calls a state receives on a particular day. We also display the average trend line for the line graph.



## STORYBOARD

### 1. The total number of calls received on a daily basis across the United States.



One of the key metrics to measure the performance of a customer service center is the number of calls received daily. In this analysis, we will be presenting a line graph that illustrates the total number of calls received daily across the United States over the past 30 days.

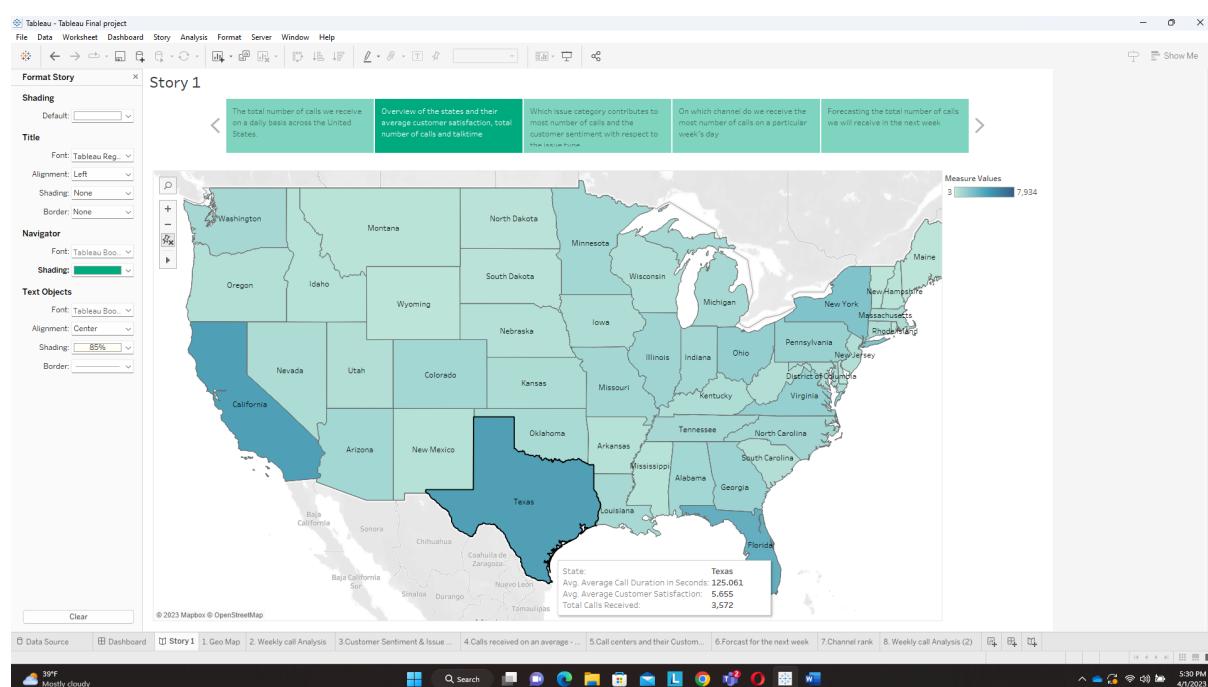
The line graph shows the total number of calls received daily across the United States over the past 30 days. The x-axis represents the number of days, while the y-axis represents the

total number of calls. Each point on the line represents the total number of calls received on that particular day.

The line graph shows that the total number of calls received daily across the United States has been fluctuating over the past 30 days. There is a clear pattern of spikes in the number of calls received on certain days, followed by a decrease in the number of calls on the following days. For example, on Day 7, the total number of calls received was 1,045, which was the least recorded, and on Day 21 the number of calls recorded was 1,169, which was the highest.

By monitoring and analyzing this data, we can identify any potential issues early on and take proactive steps to address them before they become larger problems. Overall, this line graph provides us with valuable insights into the volume of calls we are receiving, allowing us to make informed decisions that benefit both our customers and our business.

## 2. Overview of the states and their average customer satisfaction, the total number of calls, and talk time



After analyzing the daily call volume, we wanted to deep dive and see how each of the individual state's customer satisfaction, total no. of calls, and talk time was compared to the other states in the US.

The data visualization we used to present our findings was a Geo map that showed the average customer satisfaction, total number of calls, and talk time for each state.

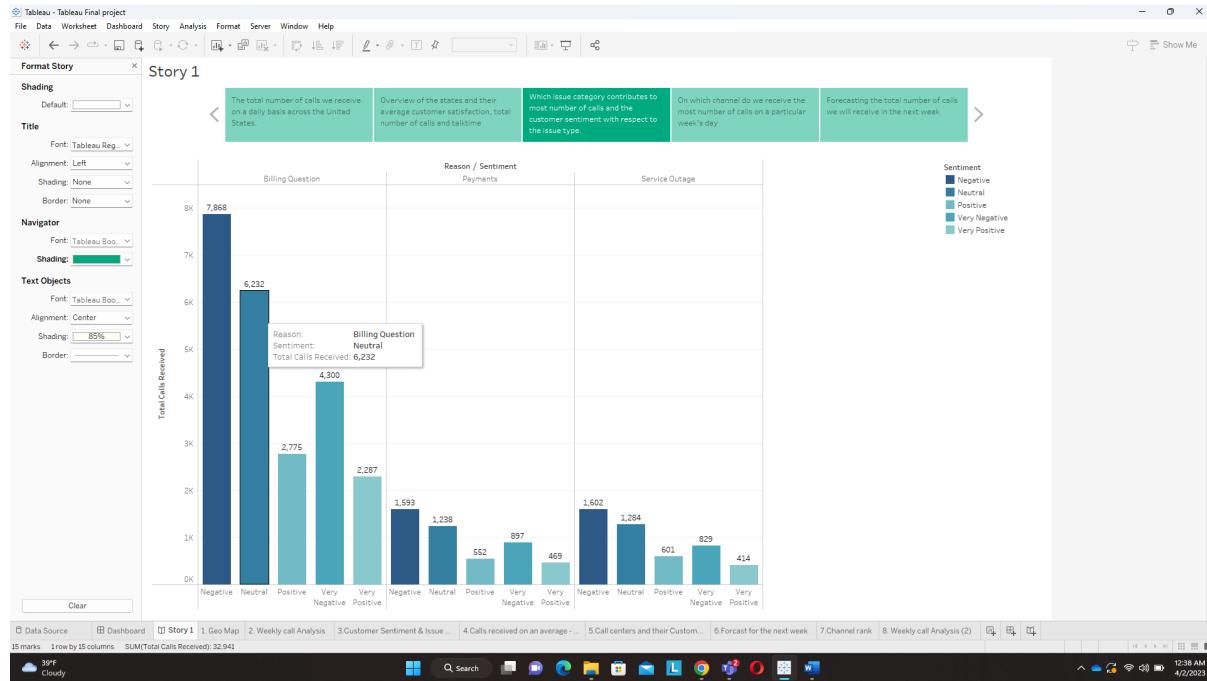
The map is color-coded, with shades of dark green representing higher call volumes satisfaction and shades of lighter green representing lower levels. Upon hovering over each state, info text provides us with the details of customer satisfaction, avg talk time and call volume.

Upon analyzing the map, we discovered that states with higher levels of customer satisfaction tended to have lower average talk times per call. In contrast, states with lower levels of customer satisfaction had higher average talk times per call. This suggests that customers in less satisfied states may be more likely to engage in longer conversations with customer service representatives.

Another interesting trend we noticed was that states with higher populations tended to have a higher total number of calls received. This is likely due to the fact that larger states have more customers who require customer service assistance.

Overall, the world map was a useful tool in visualizing and understanding customer data across different states. It allowed us to easily identify trends and patterns and helped us make informed decisions about how to improve customer satisfaction and optimize customer service processes.

3. Which issue category contributes to the most number of calls and the customer sentiment with respect to the issue type.



we have analyzed the data of the customer service calls to gain insights into which issue categories contribute to the most number of calls and the customer sentiment with respect to each issue type.

To understand this, we created a bar graph that represents the total number of calls received for each issue category and the corresponding customer sentiment for each issue type.

Upon analyzing the data, it was found that billing question issues were the most common reason for customers to call customer service. A total of 71.2% of calls received were for billing questions.

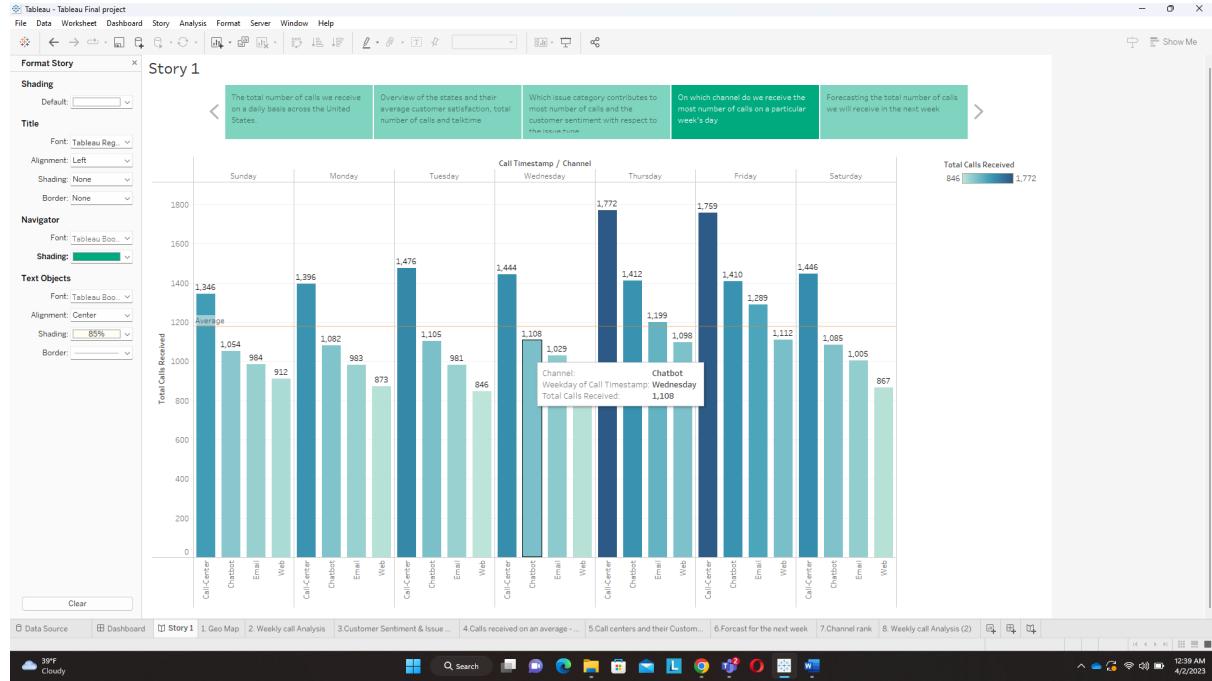
When it came to customer sentiment, it was found that customers who called in with billing-related issues had mixed sentiments. About 52.33% of customers' sentiment was very negative or negative with the resolution provided by the customer service representatives, while the remaining 47.77% of customers were positive and neutral.

The second most common reason for customers to call was related to payments, with 14.4% of calls received being for this issue category. Customers who called with payment inquiries had similar sentiment as that of the first issue category.

Another issue category was service outage which comprised the remaining 14.4% of calls received.

Overall, the data from the bar graph suggests that billing and payment issues are the biggest pain point for customers and that the customer service team should focus on improving customer satisfaction in this area.

4. On which channel do we receive the most number of calls on a particular week's day

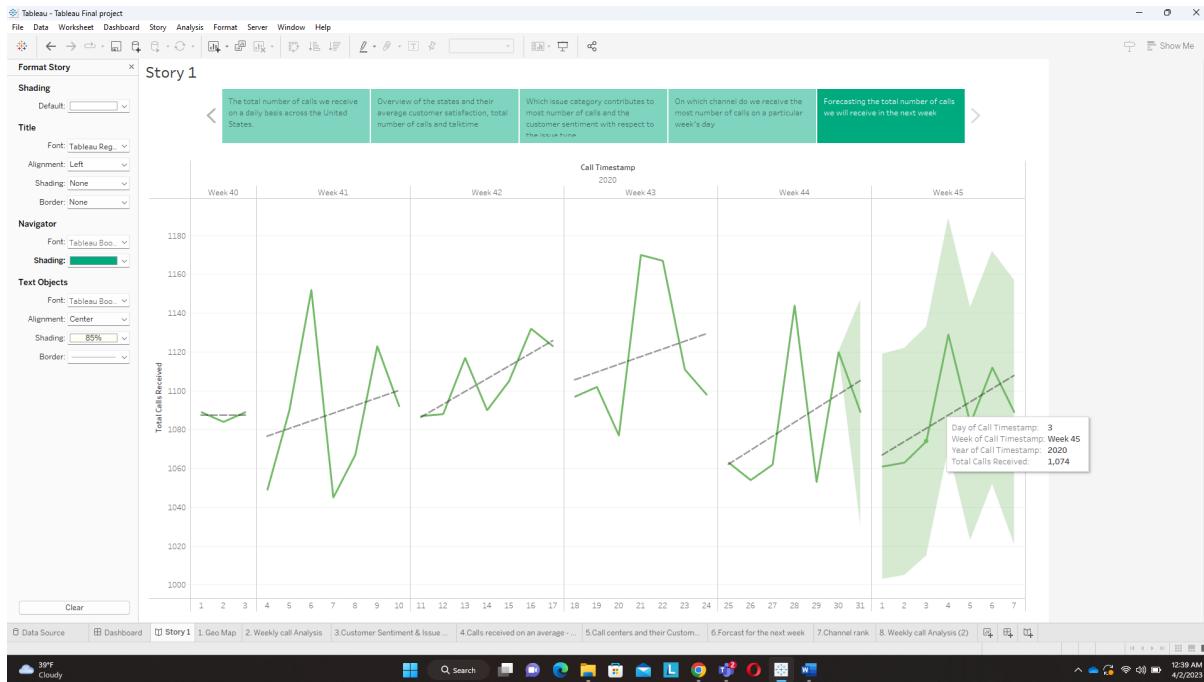


Looking at the graph, it is clear that the company receives the most number of calls on Thursdays & Fridays, with Thursday being the busiest day. The highest number of calls is received on the phone channel, followed by chat, email, and web channels. The company receives the least number of calls on Monday.

It is interesting to note that while the phone channel receives the most number of calls overall, there is a significant drop in the number of calls on Monday. On the other hand, the email channel sees a spike in the number of calls on Fridays.

Overall, this data highlights the importance of having adequate staffing and resources for the phone channel, particularly on weekdays. It also suggests that the company should consider dedicating more resources to the email channel on Fridays to manage the increased volume of calls.

## 5. Forecasting the total number of calls we will receive in the next week



it is essential to analyze and forecast the number of calls the company is expected to receive in the future. The following is a story on forecasting the total number of calls the company is likely to receive in the next week based on the previous month's trend.

Looking at the line graph, we can see that the company has received an increasing number of calls in the past month. The total number of calls received has steadily increased. Using this trend as a basis, we can forecast that the company will continue to receive an increasing number of calls in the next week. On the first day of the new month, we predict that the company will receive approximately 1,061 calls. This number is based on the average number of calls received on the first day of the previous month.

As the week progresses, we expect the number of calls to increase, peaking on Thursday with an estimated 1,129 calls. This is based on the trend observed in the previous month, which

saw an increasing number of calls throughout the week, peaking on Thursday and decreasing on Friday.

Based on this forecast, the company can prepare to handle an increasing number of calls in the upcoming week. They can schedule more staff members to handle the calls and ensure that they have sufficient resources to handle any potential increase in call volume.

It is important to note that while this forecast is based on the previous month's trend, there are several factors that can impact the actual number of calls received. These include external factors such as holidays, weather conditions, and marketing campaigns. Therefore, it is important to continually monitor the call volume and adjust the forecast accordingly.

## References

- (1) Brown, L., Gans, N., Mandelbaum, A. Sakov, A., Shen, H., Zeltyn, S., and Zhao, L. (2002),
- (2) "Statistical analysis of a telephone call center: A queueing-science perspective", Technical Report, University of Pennsylvania.
- (3) Brown, L., and Zhao, L. (2002) "A new test for the Poisson distribution", *Sankhya: The Indian Journal of Statistics, Series A*, 64, 611-625.
- (4) Gans, N., Koole, G., and Mandelbaum, A. (2002), "Telephone calls centers: a tutorial and literature review", Technical Report