

Analyzing customer churn (CASE STUDY)

1.1 Background

The problem on this case study is customer churn. We'll be using a fictitious churn dataset from a Telecom provider called Databel where I am hired as a consultant, and my task is to analyze why customers are churning, or in other words, leaving Databel. This case study helps to understand why customers are churning at the rate they are, and how to reduce churn.

1.2 Churn Definition

what is churn exactly? A good definition is the one from Investopedia: "The churn rate, also known as the rate of attrition or customer churn, is the rate at which customers stop doing business with an entity." Churn reminds me of the leaking bucket dilemma. You can add additional water to the bucket (or new clients in this example), but your overall revenue will not improve if existing customers leave. Because retaining customers is simpler than attracting new ones, reducing churn is a top aim for many businesses.

1.2.1 Churn Calculation

The simplified formula for churn is to divide customers lost by the total number of customers. If we have a total of 100 customers in a certain period, and 10 end up leaving, we have a churn rate of 10%. There are multiple methods to calculate churn, and depending on the industry. A traditional e-commerce platform might consider a certain customer a churner if he or she hasn't made a purchase in the last 12 months.

1.3 Data Background

The Databel dataset consists of 29 different columns and has one row per customer. The dataset contains numerous dimensions, the first one being Customer_id.

The Customer_id is a unique ID that identifies an individual customer.

The second column is called Churn Label, and it indicates if a customer churned with "Yes" and "No" labels. The dataset



The dataset contains more than just dimensions. The Total Charges column, for example, takes the sum of all monthly charges billed to a customer.

2. Exploratory data analysis (EDA)

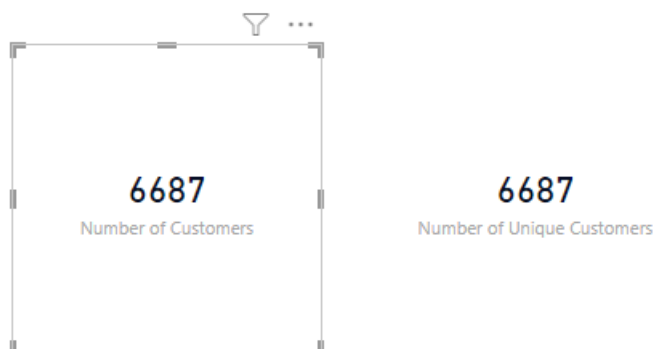
The first step in any analysis is doing a data check. We'll be creating two measures to check if the count of customer ids is equal to the count of unique customer ids. This check is particularly important, because in case there are duplicate rows I might double-count costs later.

Data check time

The first step in any analysis is doing a data check. We'll be creating two measures to check if the count of customer ids is equal to the count of unique customer ids. This check is particularly important, because in case there are duplicate rows I might double-count costs later.

```
1 Number of Customers = count('Databel - Data'[Customer ID])
```

```
1 Number of Unique Customers = DISTINCTCOUNT('Databel - Data'[Customer ID])
```



The count of unique customer match the count of customers

Calculating Churn

Before deep-diving into the analysis, there is a churn label column that indicates "Yes" or "No." Because working with this column is difficult.

- convert it to a binomial column 1 if a customer churned and 0 in case customer didnt churned.

```
1 Churned = if('Databel - Data'[Churn Label]="Yes",1,0)
```

- Create a measure for the number of churned customers

```
1 Number of Churned Customers = sum('Databel - Data'[Churned])
```

- Calculate the churn rate

```
1 Churn Rate = [Number of Churned Customers]/[Number of Customers]
```

26.86 %

Churn Rate

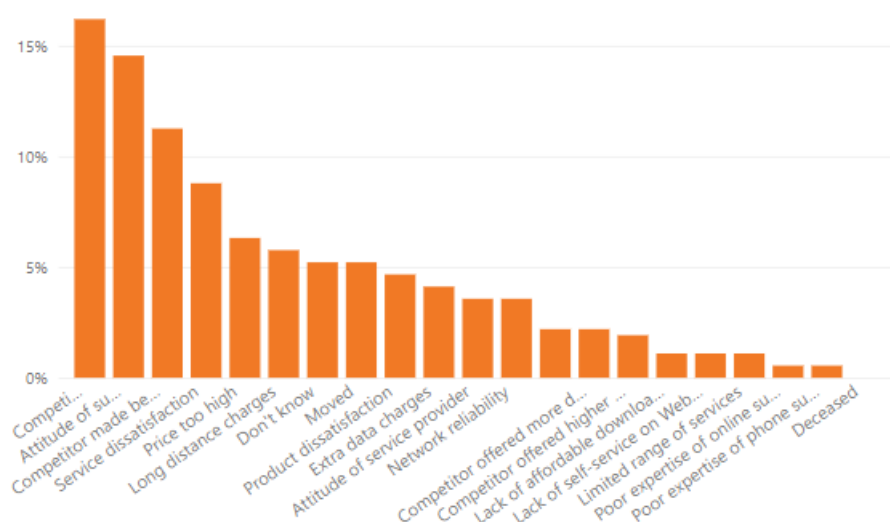
The total churn rate for "Databel" is 26.86%, which seems fairly high.

Investigating Churn reasons

The next step is to investigate the different reasons why customers churned

Using Number of Customers and Churn Reason in a bar Chart in a descending order

Churn Reasons



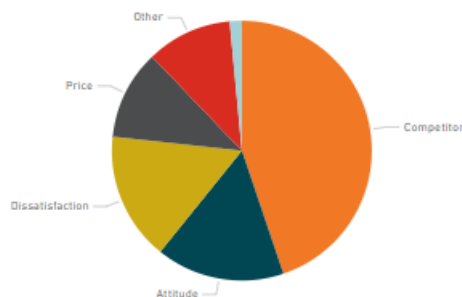
The top three reasons are:

- Competitor made better offer (1st)
- Competitor had better devices (2nd)
- Attitude of support person (3rd)

Digging deeper into churn categories

Churn Reasons are grouped together in the Churn Category column. The "Extra data charges", "Price too high" and other price related reasons are grouped together in

Churn by category



the "Price" category. We'll be displaying all churn categories in one clear visualization

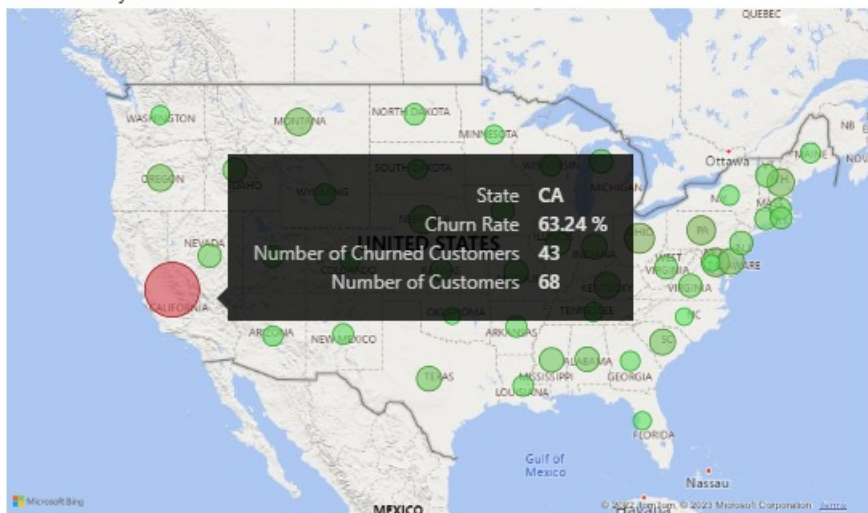
Almost half of all the customers churning are related to the competitor category

Using Maps

We want to check if competitors have run strong promotions in specific states, and Databel wants to know if this has had an impact on their customers. Our task is to

create a map that will allow me to look at the churn rate by state.

Churn Rate by State



California has a tremendous churn rate of 63.24 %, which We will investigate further as we move forward.

We already found some key insights:

- the average churn rate is around 27%, and that the main reason why customers churn is related to competitors. This could raise questions such as "Is Databel competitive enough?"
- the churn rate in California is abnormally high at 63.24%, We don't have a clear explanation yet for the relatively high churn rate.

3. Investigating Churn Patterns

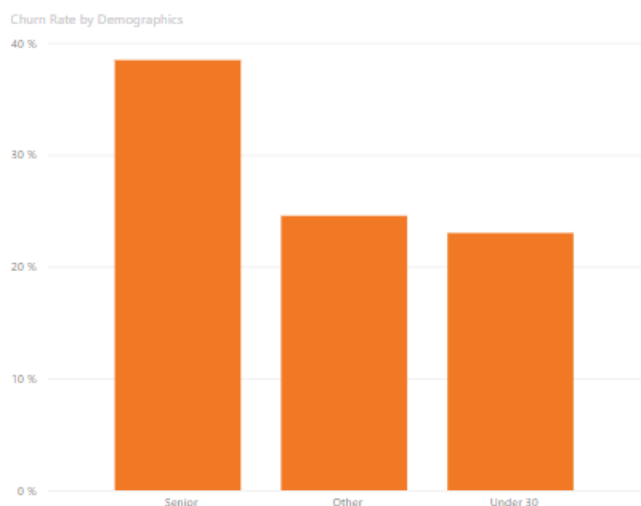
This time we'll be analyzing the demographics of Databel

Analyzing Demographics

In a new column called Demographics, we'll use IF() method to create a column with three categories:

- "Senior"
- "Under 30"
- "Other" for the chart below we can see The churn rate for senior citizens is around 10% higher than average.

```
1 Demographics = if('Databel - Data'[Senior]="Yes", "Senior", If('Databel - Data'[Under 30]="Yes", "Under 30", "Other"))
```



Age groups

from the previous visual we found a great insight that senior citizens churn more often. This suggests that it might be a good idea to analyze the customer age in general. We'll will create different age bins and make a combo chart visualizing the number of customers per bracket and their respective churn rates.

Groups



Name *

Age (bins)

Field

Age

Group type

Bin

Bin type

Size of bins

Min value

19

Max value

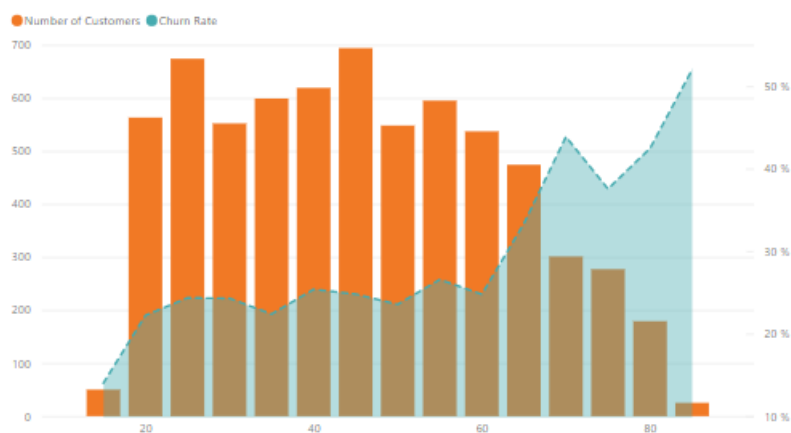
85

Binning splits numeric or date/time data into equally sized groups. Enter bin size.

Bin size *

5

Number of Customers and Churn Rate by Age (bins)



We can clearly see that:

- The churn rate has an increasing trend through the age brackets. -The churn rate for customers aged more than 80 is exactly 52%.

We have positive correlation between age and churn rate.

Inspecting Groups

Databel offers group contracts to customers from the same household. The advantage for the customer is a discounted rate, while it's a great way for Databel to grow its customer base. Our task is to determine whether consumers who belong to a group have a lower phone bill and whether this has an effect on churn.

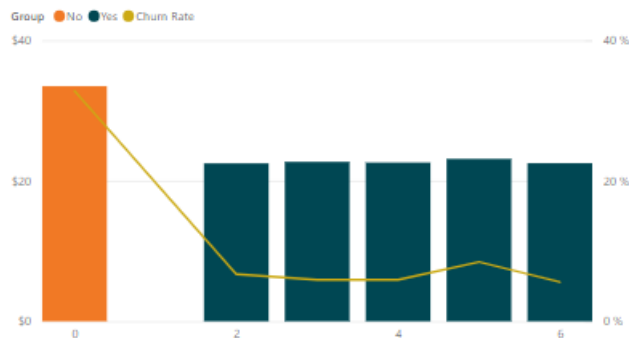
Average of Monthly Charge by Number of Customers in Group



It appears the Monthly Charge is significantly lower for people who are in a group of 2 or more.

Adding churn rate

Average of Monthly Charge and Churn Rate by Number of Customers in Group and Group



The group that has the lowest churn rate is group 6

Multiple field investigation

We'll be using the function SWITCH() which allows us to create a new column by assigning new results to the values in a column. This way we can easily create new groups for categorical fields. There are three different contract types: "One Year", "Two Year", and "Month-to-Month". It would be a good idea to gather the values of yearly contracts into one. In this way, we can observe the difference between the customers who have only yearly contracts and those having monthly contracts.

```
1 Contract Category = SWITCH('Databel - Data'[Contract Type], "One Year", "Yearly", "Two Year", "Yearly", "Monthly")
```

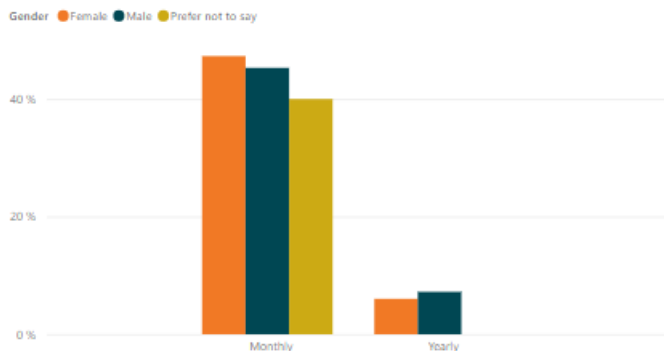
Monthly
46.29 %
Churn Rate

Yearly
6.62 %
Churn Rate

There's a significant difference: monthly contract customers churn more than yearly contract customers;

We'll need to investigate whether gender influences the churn rate. We'll be creating a clustered column chart to see how customers differ in terms of churn rate by looking at their contract categories and gender

Churn Rate by Contract Category and Gender



The female churn rate on a monthly basis is 47.31%

Unlimited plan

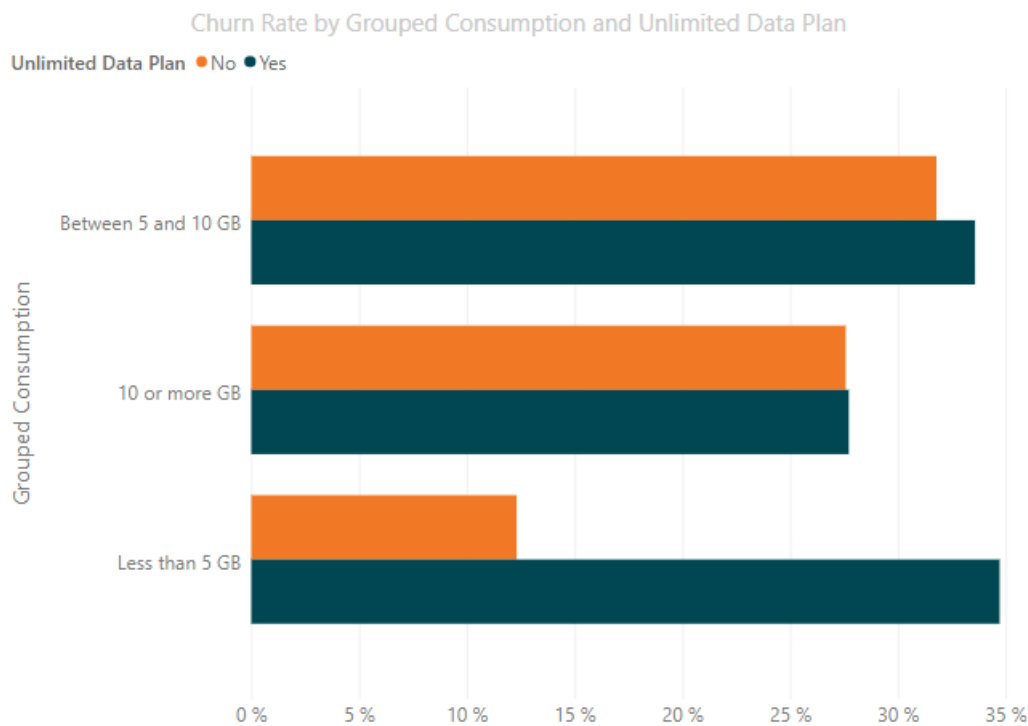
Databel has a hypothesis where people who are not on an unlimited data plan are more likely to churn. we are about to investigate how the Unlimited Data Plan influences the churn rate.

Churn Rate	Number of Customers	Unlimited Data
16.10 %	2193	No
32.11 %	4494	Yes
26.86 %	6687	

Unexpected It appears that customers who are on an unlimited plan are more likely to churn.

To see whether it's related to the amount of mobile data (GB) used a new column called (Grouped Consumption) will be created that categorizes the average monthly GB download into the following groups:

- Less than 5 GB. [Group 1]
- Between 5 and 10 GB. [Group 2]
- 10 or more GB. [Group 3]



There's a difference in churn rate by categories which evident by the chart.

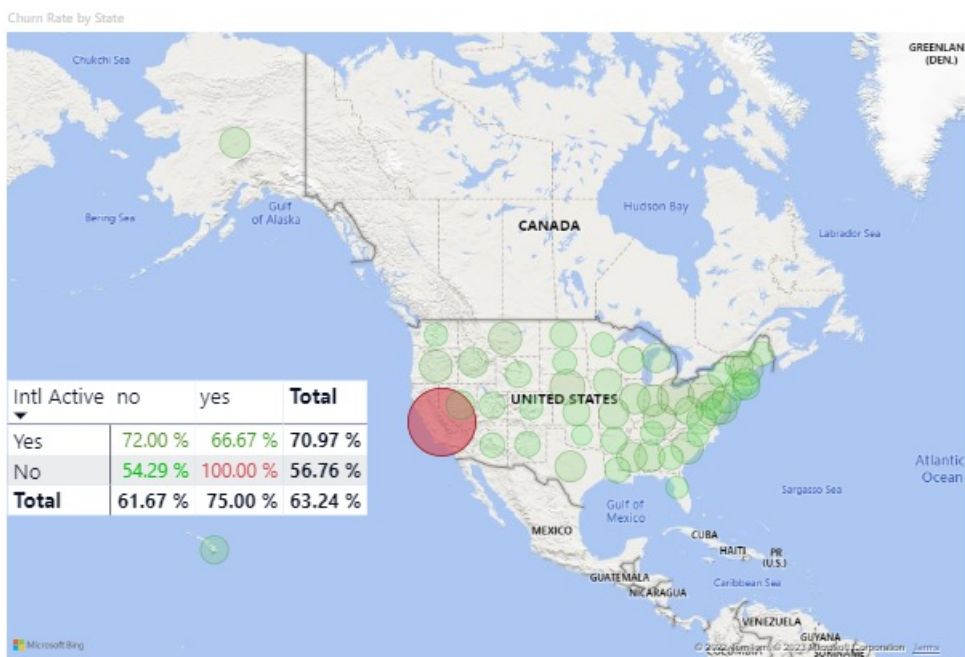
International Calls

The analysis requirement given by Databel includes a request to analyze the international activity of customers and its relationship to churn. They are curious about the behavior of customers who call internationally, and if paying for an international plan influences their loyalty.

A matrix below showing the churn rate by "Intl Plan" and "Intl Active"

Intl Active	no	yes	Total
No	20.01 %	71.19 %	22.21 %
Yes	40.34 %	7.59 %	34.31 %
Total	27.07 %	24.88 %	26.86 %

Databel wants to focus on the budget on a state-by-state basis for the promotion of the international plan.



This is important information that can help databel make a good decision while allocating the budget for the new promotions.

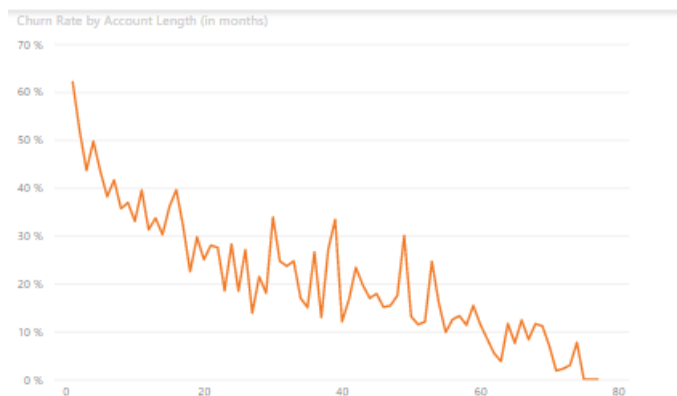
we discovered by this analysis that customers who pay for an international plan but don't call internationally is sky-high.

We recommend that Databel contact customers who have an international plan but have not made any international calls and suggest that

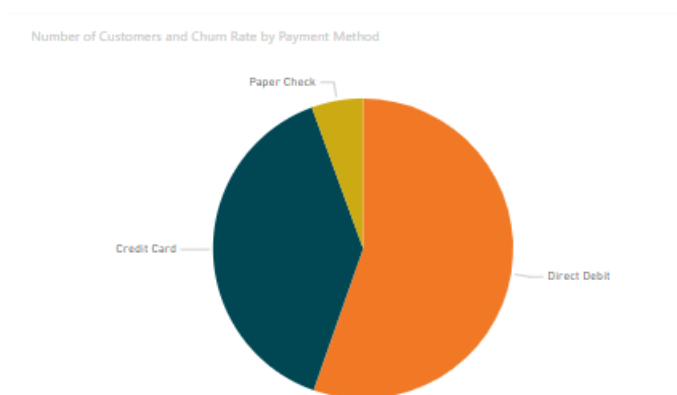
they downgrade their plan.

Contract Type

Databel also aims to enhance customer service after certain complaints. our anslysis is to job entails looking at three key aspects of customers: payment methods, contract types, and the length of time a client has been a customer.



It seems the churn rate does decrease over time; right now we want to investigate how this decrease behave through the different types of contracts. Exploring the different payment methods, we can see that there is one type that is much smaller than the others.



The number of customers who pay with paper checks is 371, accounting for only 5.5% of all customers. The pattern appears to be ambiguous solely due to the small number of cases.

4. Dashboards

in this section we will convert our analysis into dashboard-style pages and arrange them into stories.

4.1. Overview Page

The overview dashboard contain the main insights from our analysis. Contains three main KPIs: Number of customers, number of churned customers, and finally the churn rate and to explain how the different churn categories, types of contracts, reasons, and states can affect Churn.

<https://user-images.githubusercontent.com/128379502/226742144-5c58aa39-8213-4129-aa7b-e1f67a5b9314.mov>

By looking at the state with the highest churn rate. California had the highest churn rate of 79 % for clients on month-to-month contracts.

4.2. Age Groups

this report, will explain information on the age buckets and groups.

<https://user-images.githubusercontent.com/128379502/235319311-d6ae64d7-78a2-43c1-8f2c-57a33b2ec57a.mov>

The churn rate for consumers who are not in a group plan, that belong to the age group 50, and who have an account length of 12 months or less is 50.67 %. We start to observe peaks in the churn rate after this age. Fortunately, switching customers to a one- or two-year contract can help reduce churn.

4.3. Payment and Contract

Until now we hadn't looked into any customer service calls. So We'll design a chart that show a relation between contract and payment type. lets create a new measure that calculates the average customer service calls per customer

```
1 Avg Customer Service Calls = SUM('Databel - Data'[Customer Service Calls]) / [Number of Customers]
```

To validate our charts, lets ask questions like: What is the average number of customer support calls for monthly contract customers who pay via direct debit?

<https://user-images.githubusercontent.com/128379502/235319324-fb44523c-e1ed-4dfa-b14f-d52f78fe4574.mov>

1.47 calls per customer is high! Databel should definitely investigate what's going on here. Perhaps there was an issue with the payment method that needed to be investigated. These customers' with 53.90% churn rate is pretty high! It is now obvious that those that call customer service frequently also churn more than the average. They are monthly contract customers that prefer direct debit as a payment option in our situation.

4.4. Extra Charges

The final report would include the data insights as well as any relevant charges, and answer a questions of what are the average additional data charges for customers who are not on an unlimited data plan and have spent 10 or more gigabytes?.

<https://user-images.githubusercontent.com/128379502/235319329-c6023650-8406-4a7e-a9c4-04d850b84dcb.mov>

The customers pay extra for their additional data usage because they are not on an unlimited data plan.

4.5 More Insights

Its a good idea to look at the data through different lens. We will check the characteristics of each states regarding the customer service calls and churn label. To answer "Which state has the highest churn rate yet the lowest number of customer support calls?" And how many overall customer service calls does that state have?

<https://user-images.githubusercontent.com/128379502/235319345-91966ffa-6b53-43fd-a7e5-3c48f2123e42.mov>

Even though we observe that use of customer service is higher for churners , California is one of a kind. the customer there have highest churn rate. However, the number of customer service calls in the same state is the lowest.

Thank you for taking time to read our analysis.
