Telecommunication Company

June 23, 2021

1 Data Analysis with Python

Challenge:

You work for a telecommunication company and you have clients for several different services, including the main ones: internet and telephone.

The problem is that, looking at the customer history of the last few years, you noticed that the company has a churn of more than 26% of customers.

This represents a loss of millions for the company.

What does the company need to do to resolve this?

Database: https://www.kaggle.com/radmirzosimov/telecom-users-dataset

- Step 1: Importing the database
- Step 2: Visualize the database

- Step 3: Changing and improving the database (null values, numbers as text)
- Step 4: Exploratory data analysis (analyzing data sets to summarize their main characteristics)
- Step 5: Conclusions and Ideas

[76]:
```
import pandas as pd
import numpy as np
```

[77]:
```
# Step 1
table = pd.read_csv("telecom_users.csv")
```

[78]:
```
# Step 2
display(table)
```

```
      Unnamed: 0  customerID  gender  SeniorCitizen Partner Dependents  \
0           1869  7010-BRBUU    Male              0     Yes        Yes
1           4528  9688-YGXVR  Female              0      No         No
2           6344  9286-DOJGF  Female              1     Yes         No
3           6739  6994-KERXL    Male              0      No         No
4            432  2181-UAESM    Male              0      No         No
…            …         …     …              …      …        …
5981        3772  0684-AOSIH    Male              0     Yes         No
5982        5191  5982-PSMKW  Female              0     Yes        Yes
```

```
5983        5226  8044-BGWPI    Male              0     Yes        Yes
5984        5390  7450-NWRTR    Male              1      No         No
5985         860  4795-UXVCJ    Male              0      No         No

      tenure PhoneService MultipleLines InternetService  …  \
0         72          Yes           Yes             No  …
1         44          Yes            No    Fiber optic  …
2         38          Yes           Yes    Fiber optic  …
3          4          Yes            No            DSL  …
4          2          Yes            No            DSL  …
…      …           …           …             …  …
5981       1          Yes            No    Fiber optic  …
5982      23          Yes           Yes            DSL  …
5983      12          Yes            No             No  …
5984      12          Yes           Yes    Fiber optic  …
5985      26          Yes            No             No  …

         DeviceProtection          TechSupport          StreamingTV  \
0     No internet service  No internet service  No internet service
1                     Yes                   No                  Yes
2                      No                   No                   No
3                      No                   No                   No
4                     Yes                   No                   No
…                   …                   …                   …
5981                   No                   No                  Yes
5982                  Yes                  Yes                  Yes
5983  No internet service  No internet service  No internet service
5984                  Yes                   No                  Yes
5985  No internet service  No internet service  No internet service

          StreamingMovies        Contract PaperlessBilling  \
0     No internet service        Two year               No
1                      No  Month-to-month              Yes
2                      No  Month-to-month              Yes
3                     Yes  Month-to-month              Yes
4                      No  Month-to-month               No
…                   …             …              …
5981                  Yes  Month-to-month              Yes
5982                  Yes        Two year              Yes
5983  No internet service  Month-to-month              Yes
5984                  Yes  Month-to-month              Yes
5985  No internet service        One year               No

                 PaymentMethod MonthlyCharges TotalCharges  Churn
0       Credit card (automatic)          24.10      1734.65     No
1       Credit card (automatic)          88.15       3973.2     No
2     Bank transfer (automatic)          74.95      2869.85    Yes
3               Electronic check          55.90        238.5     No
```

```
4              Electronic check          53.45        119.5     No
…                          …            …          …     …
5981           Electronic check          95.00           95    Yes
5982    Credit card (automatic)          91.10       2198.3     No
5983           Electronic check          21.15       306.05     No
5984           Electronic check          99.45      1200.15    Yes
5985    Credit card (automatic)          19.80        457.3     No

[5986 rows x 22 columns]
```

[79]:
```
# Step 3
# column Unnamed is not necessary -> delete it (axis=1(column) axis=0(row))
table = table.drop("Unnamed: 0", axis=1)
display(table)
```

```
      customerID  gender  SeniorCitizen Partner Dependents  tenure  \
0     7010-BRBUU    Male              0     Yes        Yes      72
1     9688-YGXVR  Female              0      No         No      44
2     9286-DOJGF  Female              1     Yes         No      38
3     6994-KERXL    Male              0      No         No       4
4     2181-UAESM    Male              0      No         No       2
…          …     …              …      …        …     …
5981  0684-AOSIH    Male              0     Yes         No       1
5982  5982-PSMKW  Female              0     Yes        Yes      23
5983  8044-BGWPI    Male              0     Yes        Yes      12
5984  7450-NWRTR    Male              1      No         No      12
5985  4795-UXVCJ    Male              0      No         No      26

     PhoneService MultipleLines InternetService       OnlineSecurity  …  \
0             Yes           Yes              No  No internet service  …
1             Yes            No     Fiber optic                  No  …
2             Yes           Yes     Fiber optic                  No  …
3             Yes            No             DSL                  No  …
4             Yes            No             DSL                 Yes  …
…           …           …              …                   …  …
5981          Yes            No     Fiber optic                 Yes  …
5982          Yes           Yes             DSL                 Yes  …
5983          Yes            No              No  No internet service  …
5984          Yes           Yes     Fiber optic                  No  …
5985          Yes            No              No  No internet service  …

        DeviceProtection          TechSupport          StreamingTV  \
0     No internet service  No internet service  No internet service
1                     Yes                  No                  Yes
2                      No                  No                   No
3                      No                  No                   No
4                     Yes                  No                   No
…                     …                    …                    …
```

```
5981                  No                  No                  Yes
5982                 Yes                 Yes                  Yes
5983  No internet service  No internet service  No internet service
5984                 Yes                  No                  Yes
5985  No internet service  No internet service  No internet service

          StreamingMovies        Contract PaperlessBilling  \
0     No internet service        Two year               No
1                      No  Month-to-month              Yes
2                      No  Month-to-month              Yes
3                     Yes  Month-to-month              Yes
4                      No  Month-to-month               No
…                     …              …                …
5981                  Yes  Month-to-month              Yes
5982                  Yes        Two year              Yes
5983  No internet service  Month-to-month              Yes
5984                  Yes  Month-to-month              Yes
5985  No internet service        One year               No

                 PaymentMethod MonthlyCharges  TotalCharges Churn
0       Credit card (automatic)          24.10       1734.65    No
1       Credit card (automatic)          88.15        3973.2    No
2     Bank transfer (automatic)          74.95       2869.85   Yes
3              Electronic check          55.90         238.5    No
4              Electronic check          53.45         119.5    No
…                          …              …            …   …
5981           Electronic check          95.00            95   Yes
5982    Credit card (automatic)          91.10        2198.3    No
5983           Electronic check          21.15        306.05    No
5984           Electronic check          99.45       1200.15   Yes
5985    Credit card (automatic)          19.80         457.3    No

[5986 rows x 21 columns]
```

[80]:
```
# column TotalCharges as number
table["TotalCharges"] = pd.to_numeric(table["TotalCharges"], errors="coerce")
```

[81]:
```
# nul value (NaN)
# delete first null columns and then null rows (how delete a null column?)
# table = table.dropna(how="all", axis=1)
# how=all (when all of them are empty)
# how delete a null row?
# table = table.dropna(how="any", axis=0)
# how=any (when there is any null value)
```

[82]:
```
table.info()
```

```
<class 'pandas.core.frame.DataFrame'>
```

```
RangeIndex: 5986 entries, 0 to 5985
Data columns (total 21 columns):
 #   Column            Non-Null Count  Dtype
---  ------            --------------  -----
 0   customerID        5986 non-null   object
 1   gender            5986 non-null   object
 2   SeniorCitizen     5986 non-null   int64
 3   Partner           5986 non-null   object
 4   Dependents        5986 non-null   object
 5   tenure            5986 non-null   int64
 6   PhoneService      5986 non-null   object
 7   MultipleLines     5986 non-null   object
 8   InternetService   5986 non-null   object
 9   OnlineSecurity    5986 non-null   object
 10  OnlineBackup      5986 non-null   object
 11  DeviceProtection  5986 non-null   object
 12  TechSupport       5986 non-null   object
 13  StreamingTV       5986 non-null   object
 14  StreamingMovies   5986 non-null   object
 15  Contract          5986 non-null   object
 16  PaperlessBilling  5986 non-null   object
 17  PaymentMethod     5986 non-null   object
 18  MonthlyCharges    5986 non-null   float64
 19  TotalCharges      5976 non-null   float64
 20  Churn             5986 non-null   object
dtypes: float64(2), int64(2), object(17)
memory usage: 982.2+ KB
```

Feature details

- Customers who left within the last month – the column is called Churn
- Services that each customer has signed up for – phone, multiple lines, internet, online security, online backup, device protection, tech support, and streaming TV and movies
- Customer account information – how long they've been a customer, contract, payment method, paperless billing, monthly charges, and total charges
- Demographic info about customers – gender, age range, and if they have partners and dependents

[83]:
```
# Step 4
display(table["Churn"].value_counts())
display(table["Churn"].value_counts(normalize=2)*100)
```

```
No     4399
Yes    1587
Name: Churn, dtype: int64
```

```
No     73.488139
Yes    26.511861
Name: Churn, dtype: float64
```

[84]:
```
#Bring the table to another file
# table.to_excel("File_name.xlsx")
```

[85]:
```
# Step 5
import plotly.express as px
```

[87]:
```
graph = px.histogram(table, x="tenure", color="Churn")
graph.show()
#tenure=months as client
#if I wanted to do per row instead of per column= for column in table.index:
```

[90]:
```
graph = px.histogram(table, x="Dependents", color="Churn")
graph.show()
```

[91]:
```
graph = px.histogram(table, x="Contract", color="Churn")
graph.show()
```

[92]:
```
graph = px.histogram(table, x="PaymentMethod", color="Churn")
graph.show()
```

for column in table: graph = px.histogram(table, x=column, color="Churn") graph.show()

Conclusions and Ideas:

- Clients who have bigger families tend to cancel less.
- When Tenure is low, there is A LOT OF cancellations.
  - Perhaps there is a problem of customer retention in the first months.
  - Idea: We could create a program to encourage customers staying with us longer.
  - Perhaps the customer acquisition is not efficient (We may be bringing the wrong clients).
  - Perhaps the after-sales/customer experiences is not good.
- Fiber Optic -> much higher cancellation rate than other services (triple cancellations compared to DSL, even with fewer customers) We may have a problem with that service.
- People who don't have the extra services (Technical Support, Device Protection, Online Security) tend to cancel a lot more than usual.
  - Big opportunity: we have to create a program to encourage people to get new services.
- Monthly contracts tend to be MUCH more canceled than annual contracts. Almost 90% of cancellations are on monthly contracts.
  - Idea: We need to encourage clients to change to yearly contracts. How? We may give them discount.
- Payment methods:
  - Automatic methods have a lower churn rate
  - The Electronic check is horrible, if you can, let's avoid it. Idea: we could offer benefits if the person chooses an automatic method (automatic debit, credit card).