CustomerChurn

June 21, 2019

0.1 Projeto com Feedback 4

0.2 Customer Churn em uma Operadora de Telecom

0.3 MARCIO DE LIMA

0.4

0.5 DICIONARIO DE DADOS - data/projeto4_telecom_treino.csv

Unnamed: 0 int64 state object account_length int64 area_code object international_plan object voice_mail_plan object number_vmail_messages int64 total_day_minutes float64 total_day_calls int64 total_day_charge float64 total_eve_minutes float64 total_eve_calls int64 total_eve_charge float64 total_night_minutes float64 total_night_calls int64 total_night_charge float64 total_intl_minutes float64 total_intl_calls int64 total_intl_charge float64 number_customer_service_calls int64 churn object => VARIAVEL TARGET

0.6

```
In [1]: # Importando as bibliotecas
        import pandas as pd
        import numpy as np
        import matplotlib.pyplot as plt
        import seaborn as sns
        import warnings

        %matplotlib inline
        warnings.filterwarnings("ignore")

In [2]: # Importando o arquivo
        df = pd.read_csv('data/projeto4_telecom_treino.csv')

In [3]: df.head(10)

Out[3]:    Unnamed: 0 state  account_length      area_code international_plan  \
        0           1    KS             128  area_code_415                 no
        1           2    OH             107  area_code_415                 no
        2           3    NJ             137  area_code_415                 no
        3           4    OH              84  area_code_408                yes
        4           5    OK              75  area_code_415                yes
```

```
5          6    AL         118  area_code_510               yes
6          7    MA         121  area_code_510                no
7          8    MO         147  area_code_415               yes
8          9    LA         117  area_code_408                no
9         10    WV         141  area_code_415               yes

   voice_mail_plan  number_vmail_messages  total_day_minutes  total_day_calls  \
0              yes                     25              265.1              110
1              yes                     26              161.6              123
2               no                      0              243.4              114
3               no                      0              299.4               71
4               no                      0              166.7              113
5               no                      0              223.4               98
6              yes                     24              218.2               88
7               no                      0              157.0               79
8               no                      0              184.5               97
9              yes                     37              258.6               84

   total_day_charge  ...  total_eve_calls  total_eve_charge  \
0             45.07  ...               99             16.78
1             27.47  ...              103             16.62
2             41.38  ...              110             10.30
3             50.90  ...               88              5.26
4             28.34  ...              122             12.61
5             37.98  ...              101             18.75
6             37.09  ...              108             29.62
7             26.69  ...               94              8.76
8             31.37  ...               80             29.89
9             43.96  ...              111             18.87

   total_night_minutes  total_night_calls  total_night_charge  \
0                244.7                 91               11.01
1                254.4                103               11.45
2                162.6                104                7.32
3                196.9                 89                8.86
4                186.9                121                8.41
5                203.9                118                9.18
6                212.6                118                9.57
7                211.8                 96                9.53
8                215.8                 90                9.71
9                326.4                 97               14.69

   total_intl_minutes  total_intl_calls  total_intl_charge  \
0                10.0                 3               2.70
1                13.7                 3               3.70
2                12.2                 5               3.29
3                 6.6                 7               1.78
4                10.1                 3               2.73
```

```
5              6.3             6           1.70
6              7.5             7           2.03
7              7.1             6           1.92
8              8.7             4           2.35
9             11.2             5           3.02

   number_customer_service_calls  churn
0                              1     no
1                              1     no
2                              0     no
3                              2     no
4                              3     no
5                              0     no
6                              3     no
7                              0     no
8                              1     no
9                              0     no

[10 rows x 21 columns]
```

In [4]: df.dtypes

Out[4]: Unnamed: 0   int64
        state   object
        account_length   int64
        area_code   object
        international_plan   object
        voice_mail_plan   object
        number_vmail_messages   int64
        total_day_minutes   float64
        total_day_calls   int64
        total_day_charge   float64
        total_eve_minutes   float64
        total_eve_calls   int64
        total_eve_charge   float64
        total_night_minutes   float64
        total_night_calls   int64
        total_night_charge   float64
        total_intl_minutes   float64
        total_intl_calls   int64
        total_intl_charge   float64
        number_customer_service_calls   int64
        churn   object
        dtype: object

In [5]: #Dados estatisticos - analise descritiva
        df.describe()

Out[5]:   Unnamed: 0 account_length number_vmail_messages total_day_minutes \
        count 3333.00000   3333.000000   3333.000000   3333.000000
```

|       |            |            |            |            |
|-------|------------|------------|------------|------------|
| mean  | 1667.00000 | 101.064806 | 8.099010   | 179.775098 |
| std   | 962.29855  | 39.822106  | 13.688365  | 54.467389  |
| min   | 1.00000    | 1.000000   | 0.000000   | 0.000000   |
| 25%   | 834.00000  | 74.000000  | 0.000000   | 143.700000 |
| 50%   | 1667.00000 | 101.000000 | 0.000000   | 179.400000 |
| 75%   | 2500.00000 | 127.000000 | 20.000000  | 216.400000 |
| max   | 3333.00000 | 243.000000 | 51.000000  | 350.800000 |

|       | total_day_calls | total_day_charge | total_eve_minutes | total_eve_calls \ |
|-------|-----------------|------------------|-------------------|-------------------|
| count | 3333.000000     | 3333.000000      | 3333.000000       | 3333.000000       |
| mean  | 100.435644      | 30.562307        | 200.980348        | 100.114311        |
| std   | 20.069084       | 9.259435         | 50.713844         | 19.922625         |
| min   | 0.000000        | 0.000000         | 0.000000          | 0.000000          |
| 25%   | 87.000000       | 24.430000        | 166.600000        | 87.000000         |
| 50%   | 101.000000      | 30.500000        | 201.400000        | 100.000000        |
| 75%   | 114.000000      | 36.790000        | 235.300000        | 114.000000        |
| max   | 165.000000      | 59.640000        | 363.700000        | 170.000000        |

|       | total_eve_charge | total_night_minutes | total_night_calls \ |
|-------|------------------|---------------------|---------------------|
| count | 3333.000000      | 3333.000000         | 3333.000000         |
| mean  | 17.083540        | 200.872037          | 100.107711          |
| std   | 4.310668         | 50.573847           | 19.568609           |
| min   | 0.000000         | 23.200000           | 33.000000           |
| 25%   | 14.160000        | 167.000000          | 87.000000           |
| 50%   | 17.120000        | 201.200000          | 100.000000          |
| 75%   | 20.000000        | 235.300000          | 113.000000          |
| max   | 30.910000        | 395.000000          | 175.000000          |

|       | total_night_charge | total_intl_minutes | total_intl_calls \ |
|-------|--------------------|--------------------|--------------------|
| count | 3333.000000        | 3333.000000        | 3333.000000        |
| mean  | 9.039325           | 10.237294          | 4.479448           |
| std   | 2.275873           | 2.791840           | 2.461214           |
| min   | 1.040000           | 0.000000           | 0.000000           |
| 25%   | 7.520000           | 8.500000           | 3.000000           |
| 50%   | 9.050000           | 10.300000          | 4.000000           |
| 75%   | 10.590000          | 12.100000          | 6.000000           |
| max   | 17.770000          | 20.000000          | 20.000000          |

|       | total_intl_charge | number_customer_service_calls |
|-------|-------------------|-------------------------------|
| count | 3333.000000       | 3333.000000                   |
| mean  | 2.764581          | 1.562856                      |
| std   | 0.753773          | 1.315491                      |
| min   | 0.000000          | 0.000000                      |
| 25%   | 2.300000          | 1.000000                      |
| 50%   | 2.780000          | 1.000000                      |
| 75%   | 3.270000          | 2.000000                      |
| max   | 5.400000          | 9.000000                      |

```
In [6]: #Tratamento dos dados NA - Missing
        df.fillna(0)

Out[6]:         Unnamed: 0 state  account_length       area_code international_plan  \
        0                1    KS             128   area_code_415                no
        1                2    OH             107   area_code_415                no
        2                3    NJ             137   area_code_415                no
        3                4    OH              84   area_code_408               yes
        4                5    OK              75   area_code_415               yes
        5                6    AL             118   area_code_510               yes
        6                7    MA             121   area_code_510                no
        7                8    MO             147   area_code_415               yes
        8                9    LA             117   area_code_408                no
        9               10    WV             141   area_code_415               yes
        10              11    IN              65   area_code_415                no
        11              12    RI              74   area_code_415                no
        12              13    IA             168   area_code_408                no
        13              14    MT              95   area_code_510                no
        14              15    IA              62   area_code_415                no
        15              16    NY             161   area_code_415                no
        16              17    ID              85   area_code_408                no
        17              18    VT              93   area_code_510                no
        18              19    VA              76   area_code_510                no
        19              20    TX              73   area_code_415                no
        20              21    FL             147   area_code_415                no
        21              22    CO              77   area_code_408                no
        22              23    AZ             130   area_code_415                no
        23              24    SC             111   area_code_415                no
        24              25    VA             132   area_code_510                no
        25              26    NE             174   area_code_415                no
        26              27    WY              57   area_code_408                no
        27              28    MT              54   area_code_408                no
        28              29    MO              20   area_code_415                no
        29              30    HI              49   area_code_510                no
        ...            ...   ...             ...             ...               ...
        3303          3304    WI             114   area_code_415                no
        3304          3305    IL              71   area_code_510               yes
        3305          3306    IN              58   area_code_415                no
        3306          3307    AL             106   area_code_408                no
        3307          3308    OK             172   area_code_408                no
        3308          3309    IA              45   area_code_415                no
        3309          3310    VT             100   area_code_408               yes
        3310          3311    NY              94   area_code_415                no
        3311          3312    LA             128   area_code_415                no
        3312          3313    SC             181   area_code_408                no
        3313          3314    ID             127   area_code_408                no
        3314          3315    MO              89   area_code_415                no
        3315          3316    ME             149   area_code_415                no
```

5

3316	3317	MS	103	area_code_510	no
3317	3318	SD	163	area_code_415	yes
3318	3319	OK	52	area_code_415	no
3319	3320	WY	89	area_code_415	no
3320	3321	GA	122	area_code_510	yes
3321	3322	VT	60	area_code_415	no
3322	3323	MD	62	area_code_408	no
3323	3324	IN	117	area_code_415	no
3324	3325	WV	159	area_code_415	no
3325	3326	OH	78	area_code_408	no
3326	3327	OH	96	area_code_415	no
3327	3328	SC	79	area_code_415	no
3328	3329	AZ	192	area_code_415	no
3329	3330	WV	68	area_code_415	no
3330	3331	RI	28	area_code_510	no
3331	3332	CT	184	area_code_510	yes
3332	3333	TN	74	area_code_415	no

	voice_mail_plan	number_vmail_messages	total_day_minutes \
0	yes	25	265.1
1	yes	26	161.6
2	no	0	243.4
3	no	0	299.4
4	no	0	166.7
5	no	0	223.4
6	yes	24	218.2
7	no	0	157.0
8	no	0	184.5
9	yes	37	258.6
10	no	0	129.1
11	no	0	187.7
12	no	0	128.8
13	no	0	156.6
14	no	0	120.7
15	no	0	332.9
16	yes	27	196.4
17	no	0	190.7
18	yes	33	189.7
19	no	0	224.4
20	no	0	155.1
21	no	0	62.4
22	no	0	183.0
23	no	0	110.4
24	no	0	81.1
25	no	0	124.3
26	yes	39	213.0
27	no	0	134.3
28	no	0	190.0

```
29              no                0                119.3
...             ...              ...                 ...
3303            yes               26               137.1
3304            no                0                186.1
3305            yes               22               224.1
3306            yes               29                83.6
3307            no                0                203.9
3308            no                0                211.3
3309            no                0                219.4
3310            no                0                190.4
3311            no                0                147.7
3312            no                0                229.9
3313            no                0                102.8
3314            no                0                178.7
3315            yes               18               148.5
3316            yes               29               164.1
3317            no                0                197.2
3318            no                0                124.9
3319            no                0                115.4
3320            no                0                140.0
3321            no                0                193.9
3322            no                0                321.1
3323            no                0                118.4
3324            no                0                169.8
3325            no                0                193.4
3326            no                0                106.6
3327            no                0                134.7
3328            yes               36               156.2
3329            no                0                231.1
3330            no                0                180.8
3331            no                0                213.8
3332            yes               25               234.4

      total_day_calls  total_day_charge  ...  total_eve_calls  \
0             110             45.07      ...             99
1             123             27.47      ...            103
2             114             41.38      ...            110
3              71             50.90      ...             88
4             113             28.34      ...            122
5              98             37.98      ...            101
6              88             37.09      ...            108
7              79             26.69      ...             94
8              97             31.37      ...             80
9              84             43.96      ...            111
10            137             21.95      ...             83
11            127             31.91      ...            148
12             96             21.90      ...             71
13             88             26.62      ...             75
```

14	70	20.52	...	76
15	67	56.59	...	97
16	139	33.39	...	90
17	114	32.42	...	111
18	66	32.25	...	65
19	90	38.15	...	88
20	117	26.37	...	93
21	89	10.61	...	121
22	112	31.11	...	99
23	103	18.77	...	102
24	86	13.79	...	72
25	76	21.13	...	112
26	115	36.21	...	112
27	73	22.83	...	100
28	109	32.30	...	84
29	117	20.28	...	109
...	...	...	...	...
3303	88	23.31	...	125
3304	114	31.64	...	140
3305	127	38.10	...	85
3306	131	14.21	...	131
3307	109	34.66	...	123
3308	87	35.92	...	97
3309	112	37.30	...	102
3310	91	32.37	...	107
3311	94	25.11	...	83
3312	130	39.08	...	93
3313	128	17.48	...	95
3314	81	30.38	...	74
3315	106	25.25	...	106
3316	111	27.90	...	96
3317	90	33.52	...	113
3318	131	21.23	...	118
3319	99	19.62	...	115
3320	101	23.80	...	77
3321	118	32.96	...	110
3322	105	54.59	...	122
3323	126	20.13	...	97
3324	114	28.87	...	105
3325	99	32.88	...	88
3326	128	18.12	...	87
3327	98	22.90	...	68
3328	77	26.55	...	126
3329	57	39.29	...	55
3330	109	30.74	...	58
3331	105	36.35	...	84
3332	113	39.85	...	82

```
     total_eve_charge  total_night_minutes  total_night_calls  \
0               16.78                244.7                 91
1               16.62                254.4                103
2               10.30                162.6                104
3                5.26                196.9                 89
4               12.61                186.9                121
5               18.75                203.9                118
6               29.62                212.6                118
7                8.76                211.8                 96
8               29.89                215.8                 90
9               18.87                326.4                 97
10              19.42                208.8                111
11              13.89                196.0                 94
12               8.92                141.1                128
13              21.05                192.3                115
14              26.11                203.0                 99
15              27.01                160.6                128
16              23.88                 89.3                 75
17              18.55                129.6                121
18              18.09                165.7                108
19              13.56                192.8                 74
20              20.37                208.8                133
21              14.44                209.6                 64
22               6.20                181.8                 78
23              11.67                189.6                105
24              20.84                237.0                115
25              23.55                250.7                115
26              16.24                182.7                115
27              13.22                102.1                 68
28              21.95                181.5                102
29              18.28                178.7                 90
...               ...                  ...                ...
3303            13.23                247.6                 94
3304            16.88                206.5                 80
3305            20.30                174.2                 86
3306            17.33                229.5                 73
3307            19.89                160.7                 65
3308            14.08                265.9                 72
3309            19.18                255.3                 95
3310             7.82                224.8                108
3311            24.08                188.3                124
3312            12.27                262.4                110
3313            12.21                191.4                 97
3314            19.86                131.9                120
3315             9.73                178.3                 98
3316            18.62                220.3                108
3317            16.02                211.1                 94
3318            25.54                192.5                106
```

```
3319             17.84              280.9                112
3320             16.69              120.1                133
3321              7.23              210.1                134
3322             22.57              180.5                 72
3323             21.19              227.0                 56
3324             16.80              193.7                 82
3325              9.94              243.3                109
3326             24.21              178.9                 92
3327             16.12              221.4                128
3328             18.32              279.1                 83
3329             13.04              191.3                123
3330             24.55              191.9                 91
3331             13.57              139.2                137
3332             22.60              241.4                 77

      total_night_charge  total_intl_minutes  total_intl_calls  \
0                  11.01                10.0                 3
1                  11.45                13.7                 3
2                   7.32                12.2                 5
3                   8.86                 6.6                 7
4                   8.41                10.1                 3
5                   9.18                 6.3                 6
6                   9.57                 7.5                 7
7                   9.53                 7.1                 6
8                   9.71                 8.7                 4
9                  14.69                11.2                 5
10                  9.40                12.7                 6
11                  8.82                 9.1                 5
12                  6.35                11.2                 2
13                  8.65                12.3                 5
14                  9.14                13.1                 6
15                  7.23                 5.4                 9
16                  4.02                13.8                 4
17                  5.83                 8.1                 3
18                  7.46                10.0                 5
19                  8.68                13.0                 2
20                  9.40                10.6                 4
21                  9.43                 5.7                 6
22                  8.18                 9.5                19
23                  8.53                 7.7                 6
24                 10.67                10.3                 2
25                 11.28                15.5                 5
26                  8.22                 9.5                 3
27                  4.59                14.7                 4
28                  8.17                 6.3                 6
29                  8.04                11.1                 1
...                  ...                 ...               ...
3303               11.14                11.5                 7
```

```
3304            9.29              13.8                  5
3305            7.84              11.5                  7
3306           10.33               8.1                  3
3307            7.23              17.8                  4
3308           11.97              13.3                  6
3309           11.49              12.0                  4
3310           10.12              13.6                 17
3311            8.47               6.9                  5
3312           11.81              14.2                  4
3313            8.61              10.0                  5
3314            5.94               9.1                  4
3315            8.02               6.5                  4
3316            9.91              12.3                  9
3317            9.50               7.8                  8
3318            8.66              11.6                  4
3319           12.64              15.9                  6
3320            5.40               9.7                  4
3321            9.45              13.2                  8
3322            8.12              11.5                  2
3323           10.22              13.6                  3
3324            8.72              11.6                  4
3325           10.95               9.3                  4
3326            8.05              14.9                  7
3327            9.96              11.8                  5
3328           12.56               9.9                  6
3329            8.61               9.6                  4
3330            8.64              14.1                  6
3331            6.26               5.0                 10
3332           10.86              13.7                  4

     total_intl_charge  number_customer_service_calls  churn
0                 2.70                              1     no
1                 3.70                              1     no
2                 3.29                              0     no
3                 1.78                              2     no
4                 2.73                              3     no
5                 1.70                              0     no
6                 2.03                              3     no
7                 1.92                              0     no
8                 2.35                              1     no
9                 3.02                              0     no
10                3.43                              4    yes
11                2.46                              0     no
12                3.02                              1     no
13                3.32                              3     no
14                3.54                              4     no
15                1.46                              4    yes
16                3.73                              1     no
```

```
17           2.19                      3     no
18           2.70                      1     no
19           3.51                      1     no
20           2.86                      0     no
21           1.54                      5     yes
22           2.57                      0     no
23           2.08                      2     no
24           2.78                      0     no
25           4.19                      3     no
26           2.57                      0     no
27           3.97                      3     no
28           1.70                      0     no
29           3.00                      1     no
...          ...                      ...   ...
3303         3.11                      2     no
3304         3.73                      4     yes
3305         3.11                      2     no
3306         2.19                      1     no
3307         4.81                      4     no
3308         3.59                      1     no
3309         3.24                      4     no
3310         3.67                      2     no
3311         1.86                      2     no
3312         3.83                      2     no
3313         2.70                      1     no
3314         2.46                      1     no
3315         1.76                      0     no
3316         3.32                      0     no
3317         2.11                      1     no
3318         3.13                      2     no
3319         4.29                      3     no
3320         2.62                      4     yes
3321         3.56                      3     no
3322         3.11                      4     yes
3323         3.67                      5     yes
3324         3.13                      1     no
3325         2.51                      2     no
3326         4.02                      1     no
3327         3.19                      2     no
3328         2.67                      2     no
3329         2.59                      3     no
3330         3.81                      2     no
3331         1.35                      2     no
3332         3.70                      0     no

[3333 rows x 21 columns]
```

In [7]: *#Tratamento e Transformação das colunas do DataFrame*

```
         # remoção da coluna Unnamed: 0
         df = df.drop('Unnamed: 0', axis=1)

         #Colunas international_plan, voice_mail_plan, churn,
         df['churn'] = df['churn'].map(lambda x: 1 if x == 'yes' else 0 )
         df['voice_mail_plan'] = df['voice_mail_plan'].map(lambda x: 1 if x == 'yes' else 0 )
         df['international_plan'] = df['international_plan'].map(lambda x: 1 if x == 'yes' else

         # Coluna area_code
         df['area_code'] =  df['area_code'].map(lambda x: x.replace('area_code_',''))

In [21]: #Transformando o tipo das colunas
         df['area_code'] = df['area_code'].apply(pd.to_numeric, downcast='integer')
         df['voice_mail_plan'] = df['voice_mail_plan'].apply(pd.to_numeric, downcast='integer')
         df['international_plan'] = df['international_plan'].apply(pd.to_numeric, downcast='in
         df['churn'] = df['churn'].apply(pd.to_numeric, downcast='integer')

         df.dtypes

Out[21]: state                          object
         account_length                  int64
         area_code                       int64
         international_plan              int64
         voice_mail_plan                int64
         number_vmail_messages          int64
         total_day_minutes             float64
         total_day_calls                int64
         total_day_charge              float64
         total_eve_minutes             float64
         total_eve_calls                int64
         total_eve_charge              float64
         total_night_minutes           float64
         total_night_calls              int64
         total_night_charge            float64
         total_intl_minutes            float64
         total_intl_calls               int64
         total_intl_charge             float64
         number_customer_service_calls  int64
         churn                          int64
         dtype: object

In [47]: df.shape

Out[47]: (3333, 20)

In [45]: #Gerando gráficos para analise das variaveis

         #Histogramas
```

```
df.plot(kind = 'hist', subplots = True, layout = (4,5), sharex = False, figsize=(30,10
plt.show()
```


In [43]: #Gráfico de Densidade
```
df.plot(kind = 'density', subplots = True, layout = (4,5), sharex = False, figsize=(30
plt.show()
```


In [42]: #Box-Plots
```
df.plot(kind = 'box', subplots = True, layout = (4,5), sharex = False, sharey = False
plt.show()
```

```
In [46]: #Dados de Correlação entre as variáveis
         df.corr()
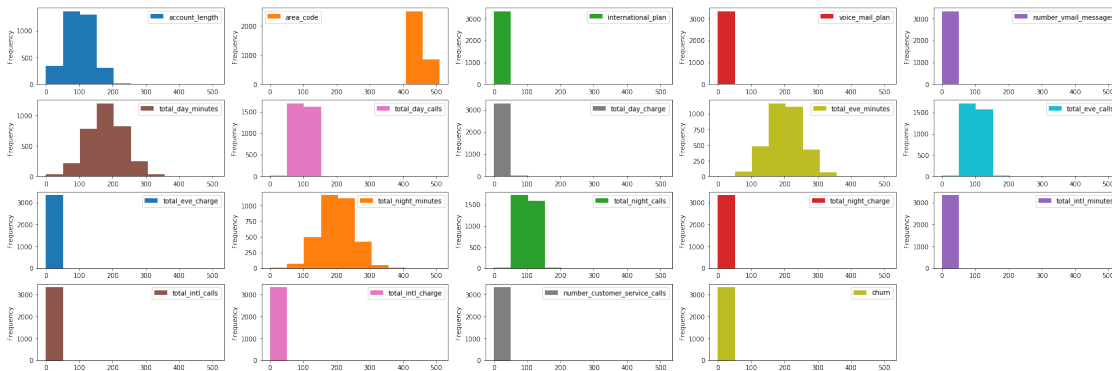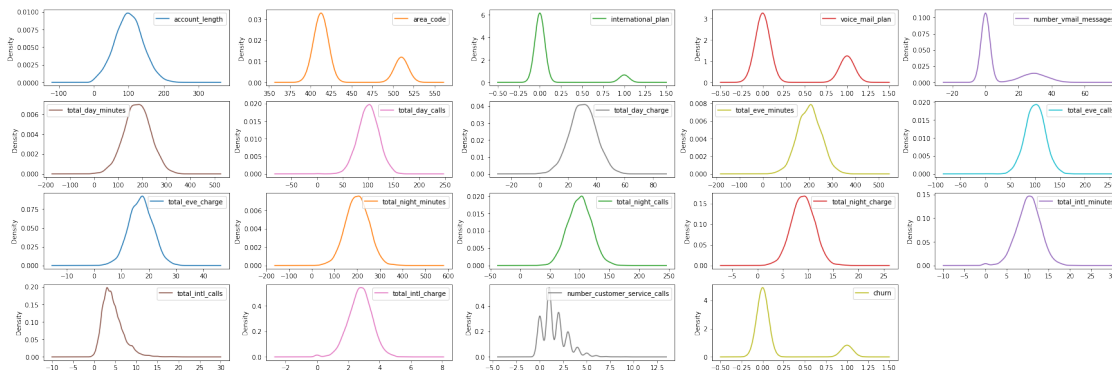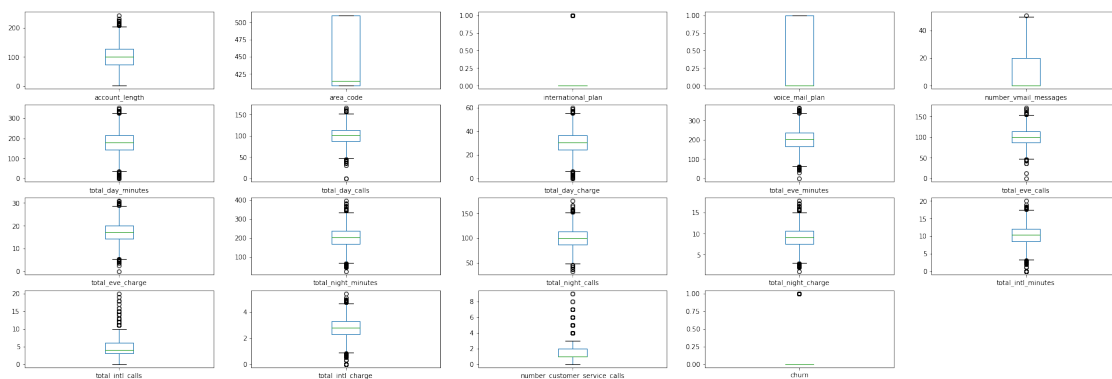
Out[46]:                                 account_length  area_code  international_plan  \
         account_length                        1.000000  -0.012463           0.024735
         area_code                            -0.012463   1.000000           0.048551
         international_plan                     0.024735   0.048551           1.000000
         voice_mail_plan                       0.002918  -0.000747           0.006006
         number_vmail_messages                -0.004628  -0.001994           0.008745
         total_day_minutes                     0.006216  -0.008264           0.049396
         total_day_calls                       0.038470  -0.009646           0.003755
         total_day_charge                      0.006214  -0.008264           0.049398
         total_eve_minutes                    -0.006757   0.003580           0.019100
         total_eve_calls                       0.019260  -0.011886           0.006114
         total_eve_charge                     -0.006745   0.003607           0.019106
         total_night_minutes                  -0.008955  -0.005825          -0.028905
         total_night_calls                    -0.013176   0.016522           0.012451
         total_night_charge                   -0.008960  -0.005845          -0.028913
         total_intl_minutes                    0.009514  -0.018288           0.045871
         total_intl_calls                      0.020661  -0.024179           0.017366
         total_intl_charge                     0.009546  -0.018395           0.045780
         number_customer_service_calls        -0.003796   0.027572          -0.024522
         churn                                 0.016541   0.006174           0.259852

                                         voice_mail_plan  number_vmail_messages  \
         account_length                         0.002918              -0.004628
         area_code                             -0.000747              -0.001994
         international_plan                      0.006006               0.008745
         voice_mail_plan                        1.000000               0.956927
         number_vmail_messages                  0.956927               1.000000
         total_day_minutes                     -0.001684               0.000778
         total_day_calls                       -0.011086              -0.009548
         total_day_charge                      -0.001686               0.000776
         total_eve_minutes                      0.021545               0.017562
         total_eve_calls                       -0.006444              -0.005864
         total_eve_charge                       0.021559               0.017578
         total_night_minutes                    0.006079               0.007681
         total_night_calls                      0.015553               0.007123
         total_night_charge                     0.006064               0.007663
         total_intl_minutes                    -0.001318               0.002856
         total_intl_calls                       0.007618               0.013957
         total_intl_charge                     -0.001276               0.002884
         number_customer_service_calls         -0.017824              -0.013263
         churn                                 -0.102148              -0.089728
```

15

	total_day_minutes	total_day_calls \
account_length	0.006216	0.038470
area_code	-0.008264	-0.009646
international_plan	0.049396	0.003755
voice_mail_plan	-0.001684	-0.011086
number_vmail_messages	0.000778	-0.009548
total_day_minutes	1.000000	0.006750
total_day_calls	0.006750	1.000000
total_day_charge	1.000000	0.006753
total_eve_minutes	0.007043	-0.021451
total_eve_calls	0.015769	0.006462
total_eve_charge	0.007029	-0.021449
total_night_minutes	0.004323	0.022938
total_night_calls	0.022972	-0.019557
total_night_charge	0.004300	0.022927
total_intl_minutes	-0.010155	0.021565
total_intl_calls	0.008033	0.004574
total_intl_charge	-0.010092	0.021666
number_customer_service_calls	-0.013423	-0.018942
churn	0.205151	0.018459

	total_day_charge	total_eve_minutes \
account_length	0.006214	-0.006757
area_code	-0.008264	0.003580
international_plan	0.049398	0.019100
voice_mail_plan	-0.001686	0.021545
number_vmail_messages	0.000776	0.017562
total_day_minutes	1.000000	0.007043
total_day_calls	0.006753	-0.021451
total_day_charge	1.000000	0.007050
total_eve_minutes	0.007050	1.000000
total_eve_calls	0.015769	-0.011430
total_eve_charge	0.007036	1.000000
total_night_minutes	0.004324	-0.012584
total_night_calls	0.022972	0.007586
total_night_charge	0.004301	-0.012593
total_intl_minutes	-0.010157	-0.011035
total_intl_calls	0.008032	0.002541
total_intl_charge	-0.010094	-0.011067
number_customer_service_calls	-0.013427	-0.012985
churn	0.205151	0.092796

	total_eve_calls	total_eve_charge \
account_length	0.019260	-0.006745
area_code	-0.011886	0.003607
international_plan	0.006114	0.019106
voice_mail_plan	-0.006444	0.021559
number_vmail_messages	-0.005864	0.017578

```
total_day_minutes                  0.015769           0.007029
total_day_calls                    0.006462          -0.021449
total_day_charge                   0.015769           0.007036
total_eve_minutes                 -0.011430           1.000000
total_eve_calls                    1.000000          -0.011423
total_eve_charge                  -0.011423           1.000000
total_night_minutes               -0.002093          -0.012592
total_night_calls                  0.007710           0.007596
total_night_charge                -0.002056          -0.012601
total_intl_minutes                 0.008703          -0.011043
total_intl_calls                   0.017434           0.002541
total_intl_charge                  0.008674          -0.011074
number_customer_service_calls      0.002423          -0.012987
churn                              0.009233           0.092786

                              total_night_minutes  total_night_calls  \
account_length                          -0.008955          -0.013176
area_code                               -0.005825           0.016522
international_plan                       -0.028905           0.012451
voice_mail_plan                          0.006079           0.015553
number_vmail_messages                    0.007681           0.007123
total_day_minutes                        0.004323           0.022972
total_day_calls                          0.022938          -0.019557
total_day_charge                         0.004324           0.022972
total_eve_minutes                       -0.012584           0.007586
total_eve_calls                         -0.002093           0.007710
total_eve_charge                        -0.012592           0.007596
total_night_minutes                      1.000000           0.011204
total_night_calls                        0.011204           1.000000
total_night_charge                       0.999999           0.011188
total_intl_minutes                      -0.015207          -0.013605
total_intl_calls                        -0.012353           0.000305
total_intl_charge                       -0.015180          -0.013630
number_customer_service_calls           -0.009288          -0.012802
churn                                    0.035493           0.006141

                              total_night_charge  total_intl_minutes  \
account_length                         -0.008960            0.009514
area_code                              -0.005845           -0.018288
international_plan                      -0.028913            0.045871
voice_mail_plan                         0.006064           -0.001318
number_vmail_messages                   0.007663            0.002856
total_day_minutes                       0.004300           -0.010155
total_day_calls                         0.022927            0.021565
total_day_charge                        0.004301           -0.010157
total_eve_minutes                      -0.012593           -0.011035
total_eve_calls                        -0.002056            0.008703
total_eve_charge                       -0.012601           -0.011043
```

total_night_minutes	0.999999	-0.015207
total_night_calls	0.011188	-0.013605
total_night_charge	1.000000	-0.015214
total_intl_minutes	-0.015214	1.000000
total_intl_calls	-0.012329	0.032304
total_intl_charge	-0.015186	0.999993
number_customer_service_calls	-0.009277	-0.009640
churn	0.035496	0.068239

	total_intl_calls	total_intl_charge \
account_length	0.020661	0.009546
area_code	-0.024179	-0.018395
international_plan	0.017366	0.045780
voice_mail_plan	0.007618	-0.001276
number_vmail_messages	0.013957	0.002884
total_day_minutes	0.008033	-0.010092
total_day_calls	0.004574	0.021666
total_day_charge	0.008032	-0.010094
total_eve_minutes	0.002541	-0.011067
total_eve_calls	0.017434	0.008674
total_eve_charge	0.002541	-0.011074
total_night_minutes	-0.012353	-0.015180
total_night_calls	0.000305	-0.013630
total_night_charge	-0.012329	-0.015186
total_intl_minutes	0.032304	0.999993
total_intl_calls	1.000000	0.032372
total_intl_charge	0.032372	1.000000
number_customer_service_calls	-0.017561	-0.009675
churn	-0.052844	0.068259

	number_customer_service_calls	churn
account_length	-0.003796	0.016541
area_code	0.027572	0.006174
international_plan	-0.024522	0.259852
voice_mail_plan	-0.017824	-0.102148
number_vmail_messages	-0.013263	-0.089728
total_day_minutes	-0.013423	0.205151
total_day_calls	-0.018942	0.018459
total_day_charge	-0.013427	0.205151
total_eve_minutes	-0.012985	0.092796
total_eve_calls	0.002423	0.009233
total_eve_charge	-0.012987	0.092786
total_night_minutes	-0.009288	0.035493
total_night_calls	-0.012802	0.006141
total_night_charge	-0.009277	0.035496
total_intl_minutes	-0.009640	0.068239
total_intl_calls	-0.017561	-0.052844
total_intl_charge	-0.009675	0.068259

```
        number_customer_service_calls                         1.000000  0.208750
        churn                                                  0.208750  1.000000
```

In [23]: *# Distribuição da varíavel TARGET*
         df.groupby('churn').size()

Out[23]: churn
         0   2850
         1   483
         dtype: int64

In [91]: *# Separando os dados*
         array = df.values
         colunas = df.columns[1:19]

         # Separando o array em componentes de input (X) e output (Y)
         X = array[:,1:19]
         Y = array[:,19]

In [75]: *# Normalizando e padronizando os dados do dataframe*

         from **sklearn.preprocessing** **import** MinMaxScaler
         from **sklearn.preprocessing** **import** StandardScaler

         # Gerando a nova escala (normalizando os dados)
         scaler = MinMaxScaler(feature_range = (0, 1))
         rescaledX = scaler.fit_transform(X)

         #Gerando padronização dos valores
         scalerP = StandardScaler().fit(rescaledX)
         standardX = scalerP.transform(rescaledX)

In [107]: *#Montando novo DataFrame com State, Churn e as colunas normalizadas na escala e Padr*
          df_new = pd.DataFrame (data=standardX, columns=colunas)
          df_new['state'] = df['state']
          df_new['churn'] = df['churn']

          df_new.head(10)

          array1 = df_new.values
          X1 = array1[:,0:18]
          Y1 = array1[:,19]
          Y1 = Y1.astype(int)

In [108]: *#Gerando dados de Treino e de Teste para os modelos*
          from **sklearn.model_selection** **import** train_test_split

          seed = 1313
          X_treino, X_teste, y_treino, y_teste = train_test_split(X1, Y1, test_size = 0.30, ra
```

```
In [109]: y_treino

Out[109]: array([0, 0, 0, ..., 0, 0, 0])

In [110]: # Feature selection com Random Forest
          from sklearn.ensemble import RandomForestClassifier
          from sklearn.feature_selection import SelectFromModel

          clf = RandomForestClassifier(random_state=seed)
          selector = clf.fit(X_treino, y_treino)
          fs = SelectFromModel(selector, prefit=True)

          X_treino_new = fs.transform(X_treino)
          X_teste_new = fs.transform(X_teste)

          print(X_treino_new.shape, X_teste_new.shape)

(2333, 6) (1000, 6)


In [111]: #Criando o Modelo - Versao 1
          from sklearn.linear_model import LogisticRegression

          modelo = LogisticRegression()

In [112]: # Treinamento do modelo
          modelo.fit(X_treino_new, y_treino)

Out[112]: LogisticRegression(C=1.0, class_weight=None, dual=False, fit_intercept=True,
                    intercept_scaling=1, max_iter=100, multi_class='warn',
                    n_jobs=None, penalty='l2', random_state=None, solver='warn',
                    tol=0.0001, verbose=0, warm_start=False)

In [113]: #Resultado do Modelo - Versao 1
          resultado = modelo.score(X_teste_new, y_teste)
          print("Acurácia nos Dados de Teste: %.3f%%" % (resultado * 100.0))

Acurácia nos Dados de Teste: 84.200%


In [119]: #Criando o modelo 2 - XGBClassifier
          #!pip install xgboost
          from sklearn.metrics import accuracy_score
          from xgboost import XGBClassifier

          modelo2 = XGBClassifier(n_estimators=110, nthread=-1, seed=seed)

          # Treinando o modelo
          modelo2.fit(X_treino_new, y_treino, eval_metric="auc", verbose = False)
```

```
Out[119]: XGBClassifier(base_score=0.5, booster='gbtree', colsample_bylevel=1,
                colsample_bynode=1, colsample_bytree=1, gamma=0, learning_rate=0.1,
                max_delta_step=0, max_depth=3, min_child_weight=1, missing=None,
                n_estimators=110, n_jobs=1, nthread=-1, objective='binary:logistic',
                random_state=0, reg_alpha=0, reg_lambda=1, scale_pos_weight=1,
                seed=1313, silent=None, subsample=1, verbosity=1)

In [120]: # Fazendo previsões
          y_pred = modelo2.predict(X_teste_new)
          previsoes = [round(value) for value in y_pred]

In [121]: #Resultado do Modelo - Versao 2
          resultado2 = accuracy_score(y_teste, previsoes)
          print("Acuracia do Modelo 2: %.3f" % (resultado2 * 100.0))

Acuracia do Modelo 2: 90.000
```

0.7 FIM

0.8 OBRIGADO