

Data Preparation

PREDICTING CUSTOMER CHURN IN PYTHON



Mark Peterson

Senior Data Scientist, Alliance Data

Model assumptions

- Some assumptions that models make:
 - That the features are normally distributed
 - That the features are on the same scale

Data types

- Machine learning algorithms require numeric data types
 - Need to encode categorical variables as numeric

```
telco.dtypes
```

```
Account_Length      int64
Vmail_Message       int64
Day_Mins            float64
Eve_Mins            float64
Night_Mins          float64
Intl_Mins           float64
CustServ_Calls      int64
Churn               object
Intl_Plan           object
Vmail_Plan          object
Day_Calls           int64
Day_Charge          float64
Eve_Calls           int64
Eve_Charge          float64
Night_Calls         int64
Night_Charge        float64
Intl_Calls          int64
Intl_Charge         float64
State              object
Area_Code           int64
Phone              object
dtype: object
```

Encoding binary features

```
telco['Intl_Plan'].head()
```

```
0    no
1    no
2    no
3   yes
4   yes
Name: Intl_Plan, dtype: object
```

Encoding binary features

Option 1: `.replace()`

```
telco['Intl_Plan'].replace({'no':0 , 'yes':1})  
  
telco['Intl_Plan'].head()
```

```
0    0  
1    0  
2    0  
3    1  
4    1  
Name: Intl_Plan, dtype: int64
```

Option 2: `LabelEncoder()`

```
from sklearn.preprocessing import LabelEncoder  
  
LabelEncoder().fit_transform(telco["Intl_Plan"])  
  
telco['Intl_Plan'].head()
```

```
0    0  
1    0  
2    0  
3    1  
4    1  
Name: Intl_Plan, dtype: int64
```

Encoding state

```
telco[ 'State' ].head(4)
```

```
0    KS
1    OH
2    NJ
3    OH
Name: State, dtype: object
```

- Could assign a number to each state

```
0    0
1    1
2    2
3    1
Name: State, dtype: int64
```

- Bad idea
- Would make your model less effective

One hot encoding

State
KS
OH
NJ
OH

One hot encoding

State
KS
OH
NJ
OH



One hot encoding

State		State_KS	State_OH	State_NJ
KS	➔	1	0	0
OH		0	1	0
NJ		0	0	1
OH		0	1	0

Feature scaling

- Features should be on the same scale
- Rarely true of real-world data

Feature scaling

```
telco['Intl_Calls'].describe()
```

```
count    3333.000000
mean       4.479448
std        2.461214
min        0.000000
25%        3.000000
50%        4.000000
75%        6.000000
max       20.000000
Name: Intl_Calls, dtype: float64
```

```
telco['Night_Mins'].describe()
```

```
count    3333.000000
mean    200.872037
std     50.573847
min     23.200000
25%    167.000000
50%    201.200000
75%    235.300000
max    395.000000
Name: Night_Mins, dtype: float64
```

Standardization

- Centers the distribution around the mean
- Calculates the number of standard deviations away from the mean each point is

```
from sklearn.preprocessing import StandardScaler
```

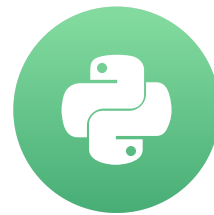
```
df = StandardScaler().fit_transform(df)
```

Let's practice!

PREDICTING CUSTOMER CHURN IN PYTHON

Feature selection and engineering

PREDICTING CUSTOMER CHURN IN PYTHON



Mark Peterson

Senior Data Scientist, Alliance Data

Dropping unnecessary features

- Unique identifiers
 - Phone numbers
 - Social security numbers
 - Account numbers
- `.drop()` method

```
telco.drop(['Soc_Sec', 'Tax_ID'], axis=1)
```


Dropping correlated features

- Highly correlated features can be dropped
- They provide no additional information to the model

```
telco.corr()
```

	Day_Mins	Eve_Mins	Night_Mins	Intl_Mins	CustServ_Calls	Day_Calls	Day_Charge	Eve_Calls	Eve_Charge	Night_Calls	Night_Charge	Intl_Calls	Intl_Charge
Day_Mins	1.000000	0.007043	0.004323	-0.010155	-0.013423	0.006750	1.000000	0.015769	0.007029	0.022972	0.004300	0.008033	-0.010092
Eve_Mins	0.007043	1.000000	-0.012584	-0.011035	-0.012985	-0.021451	0.007050	-0.011430	1.000000	0.007586	-0.012593	0.002541	-0.011067
Night_Mins	0.004323	-0.012584	1.000000	-0.015207	-0.009288	0.022938	0.004324	-0.002093	-0.012592	0.011204	0.999999	-0.012353	-0.015180
Intl_Mins	-0.010155	-0.011035	-0.015207	1.000000	-0.009640	0.021565	-0.010157	0.008703	-0.011043	-0.013605	-0.015214	0.032304	0.999993
CustServ_Calls	-0.013423	-0.012985	-0.009288	-0.009640	1.000000	-0.018942	-0.013427	0.002423	-0.012987	-0.012802	-0.009277	-0.017561	-0.009675
Day_Calls	0.006750	-0.021451	0.022938	0.021565	-0.018942	1.000000	0.006753	0.006462	-0.021449	-0.019557	0.022927	0.004574	0.021666
Day_Charge	1.000000	0.007050	0.004324	-0.010157	-0.013427	0.006753	1.000000	0.015769	0.007036	0.022972	0.004301	0.008032	-0.010094
Eve_Calls	0.015769	-0.011430	-0.002093	0.008703	0.002423	0.006462	0.015769	1.000000	-0.011423	0.007710	-0.002056	0.017434	0.008674
Eve_Charge	0.007029	1.000000	-0.012592	-0.011043	-0.012987	-0.021449	0.007036	-0.011423	1.000000	0.007596	-0.012601	0.002541	-0.011074
Night_Calls	0.022972	0.007586	0.011204	-0.013605	-0.012802	-0.019557	0.022972	0.007710	0.007596	1.000000	0.011188	0.000305	-0.013630
Night_Charge	0.004300	-0.012593	0.999999	-0.015214	-0.009277	0.022927	0.004301	-0.002056	-0.012601	0.011188	1.000000	-0.012329	-0.015186
Intl_Calls	0.008033	0.002541	-0.012353	0.032304	-0.017561	0.004574	0.008032	0.017434	0.002541	0.000305	-0.012329	1.000000	0.032372
Intl_Charge	-0.010092	-0.011067	-0.015180	0.999993	-0.009675	0.021666	-0.010094	0.008674	-0.011074	-0.013630	-0.015186	0.032372	1.000000

```
telco.corr()
```

	Day_Mins	Eve_Mins	Night_Mins	Intl_Mins	CustServ_Calls	Day_Calls	Day_Charge	Eve_Calls	Eve_Charge	Night_Calls	Night_Charge	Intl_Calls	Intl_Charge
Day_Mins	1.000000	0.007043	0.004323	-0.010155	-0.013423	0.006750	1.000000	0.015769	0.007029	0.022972	0.004300	0.008033	-0.010092
Eve_Mins	0.007043	1.000000	-0.012584	-0.011035	-0.012985	-0.021451	0.007050	-0.011430	1.000000	0.007586	-0.012593	0.002541	-0.011067
Night_Mins	0.004323	-0.012584	1.000000	-0.015207	-0.009288	0.022938	0.004324	-0.002093	-0.012592	0.011204	0.999999	-0.012353	-0.015180
Intl_Mins	-0.010155	-0.011035	-0.015207	1.000000	-0.009640	0.021565	-0.010157	0.008703	-0.011043	-0.013605	-0.015214	0.032304	0.999993
CustServ_Calls	-0.013423	-0.012985	-0.009288	-0.009640	1.000000	-0.018942	-0.013427	0.002423	-0.012987	-0.012802	-0.009277	-0.017561	-0.009675
Day_Calls	0.006750	-0.021451	0.022938	0.021565	-0.018942	1.000000	0.006753	0.006462	-0.021449	-0.019557	0.022927	0.004574	0.021666
Day_Charge	1.000000	0.007050	0.004324	-0.010157	-0.013427	0.006753	1.000000	0.015769	0.007036	0.022972	0.004301	0.008032	-0.010094
Eve_Calls	0.015769	-0.011430	-0.002093	0.008703	0.002423	0.006462	0.015769	1.000000	-0.011423	0.007710	-0.002056	0.017434	0.008674
Eve_Charge	0.007029	1.000000	-0.012592	-0.011043	-0.012987	-0.021449	0.007036	-0.011423	1.000000	0.007596	-0.012601	0.002541	-0.011074
Night_Calls	0.022972	0.007586	0.011204	-0.013605	-0.012802	-0.019557	0.022972	0.007710	0.007596	1.000000	0.011188	0.000305	-0.013630
Night_Charge	0.004300	-0.012593	0.999999	-0.015214	-0.009277	0.022927	0.004301	-0.002056	-0.012601	0.011188	1.000000	-0.012329	-0.015186
Intl_Calls	0.008033	0.002541	-0.012353	0.032304	-0.017561	0.004574	0.008032	0.017434	0.002541	0.000305	-0.012329	1.000000	0.032372
Intl_Charge	-0.010092	-0.011067	-0.015180	0.999993	-0.009675	0.021666	-0.010094	0.008674	-0.011074	-0.013630	-0.015186	0.032372	1.000000

```
telco.corr()
```

	Day_Mins	Eve_Mins	Night_Mins	Intl_Mins	CustServ_Calls	Day_Calls	Day_Charge	Eve_Calls	Eve_Charge	Night_Calls	Night_Charge	Intl_Calls	Intl_Charge
Day_Mins	1.000000	0.007043	0.004323	-0.010155	-0.013423	0.006750	1.000000	0.015769	0.007029	0.022972	0.004300	0.008033	-0.010092
Eve_Mins	0.007043	1.000000	-0.012584	-0.011035	-0.012985	-0.021451	0.007050	-0.011430	1.000000	0.007586	-0.012593	0.002541	-0.011067
Night_Mins	0.004323	-0.012584	1.000000	-0.015207	-0.009288	0.022938	0.004324	-0.002093	-0.012592	0.011204	0.999999	-0.012353	-0.015180
Intl_Mins	-0.010155	-0.011035	-0.015207	1.000000	-0.009640	0.021565	-0.010157	0.008703	-0.011043	-0.013605	-0.015214	0.032304	0.999993
CustServ_Calls	-0.013423	-0.012985	-0.009288	-0.009640	1.000000	-0.018942	-0.013427	0.002423	-0.012987	-0.012802	-0.009277	-0.017561	-0.009675
Day_Calls	0.006750	-0.021451	0.022938	0.021565	-0.018942	1.000000	0.006753	0.006462	-0.021449	-0.019557	0.022927	0.004574	0.021666
Day_Charge	1.000000	0.007050	0.004324	-0.010157	-0.013427	0.006753	1.000000	0.015769	0.007036	0.022972	0.004301	0.008032	-0.010094
Eve_Calls	0.015769	-0.011430	-0.002093	0.008703	0.002423	0.006462	0.015769	1.000000	-0.011423	0.007710	-0.002056	0.017434	0.008674
Eve_Charge	0.007029	1.000000	-0.012592	-0.011043	-0.012987	-0.021449	0.007036	-0.011423	1.000000	0.007596	-0.012601	0.002541	-0.011074
Night_Calls	0.022972	0.007586	0.011204	-0.013605	-0.012802	-0.019557	0.022972	0.007710	0.007596	1.000000	0.011188	0.000305	-0.013630
Night_Charge	0.004300	-0.012593	0.999999	-0.015214	-0.009277	0.022927	0.004301	-0.002056	-0.012601	0.011188	1.000000	-0.012329	-0.015186
Intl_Calls	0.008033	0.002541	-0.012353	0.032304	-0.017561	0.004574	0.008032	0.017434	0.002541	0.000305	-0.012329	1.000000	0.032372
Intl_Charge	-0.010092	-0.011067	-0.015180	0.999993	-0.009675	0.021666	-0.010094	0.008674	-0.011074	-0.013630	-0.015186	0.032372	1.000000


```
telco.corr()
```

	Day_Mins	Eve_Mins	Night_Mins	Intl_Mins	CustServ_Calls	Day_Calls	Day_Charge	Eve_Calls	Eve_Charge	Night_Calls	Night_Charge	Intl_Calls	Intl_Charge
Day_Mins	1.000000	0.007043	0.004323	-0.010155	-0.013423	0.006750	1.000000	0.015769	0.007029	0.022972	0.004300	0.008033	-0.010092
Eve_Mins	0.007043	1.000000	-0.012584	-0.011035	-0.012985	-0.021451	0.007050	-0.011430	1.000000	0.007586	-0.012593	0.002541	-0.011067
Night_Mins	0.004323	-0.012584	1.000000	-0.015207	-0.009288	0.022938	0.004324	-0.002093	-0.012592	0.011204	0.999999	-0.012353	-0.015180
Intl_Mins	-0.010155	-0.011035	-0.015207	1.000000	-0.009640	0.021565	-0.010157	0.008703	-0.011043	-0.013605	-0.015214	0.032304	0.999993
CustServ_Calls	-0.013423	-0.012985	-0.009288	-0.009640	1.000000	-0.018942	-0.013427	0.002423	-0.012987	-0.012802	-0.009277	-0.017561	-0.009675
Day_Calls	0.006750	-0.021451	0.022938	0.021565	-0.018942	1.000000	0.006753	0.006462	-0.021449	-0.019557	0.022927	0.004574	0.021666
Day_Charge	1.000000	0.007050	0.004324	-0.010157	-0.013427	0.006753	1.000000	0.015769	0.007036	0.022972	0.004301	0.008032	-0.010094
Eve_Calls	0.015769	-0.011430	-0.002093	0.008703	0.002423	0.006462	0.015769	1.000000	-0.011423	0.007710	-0.002056	0.017434	0.008674
Eve_Charge	0.007029	1.000000	-0.012592	-0.011043	-0.012987	-0.021449	0.007036	-0.011423	1.000000	0.007596	-0.012601	0.002541	-0.011074
Night_Calls	0.022972	0.007586	0.011204	-0.013605	-0.012802	-0.019557	0.022972	0.007710	0.007596	1.000000	0.011188	0.000305	-0.013630
Night_Charge	0.004300	-0.012593	0.999999	-0.015214	-0.009277	0.022927	0.004301	-0.002056	-0.012601	0.011188	1.000000	-0.012329	-0.015186
Intl_Calls	0.008033	0.002541	-0.012353	0.032304	-0.017561	0.004574	0.008032	0.017434	0.002541	0.000305	-0.012329	1.000000	0.032372
Intl_Charge	-0.010092	-0.011067	-0.015180	0.999993	-0.009675	0.021666	-0.010094	0.008674	-0.011074	-0.013630	-0.015186	0.032372	1.000000

```
telco.corr()
```

	Day_Mins	Eve_Mins	Night_Mins	Intl_Mins	CustServ_Calls	Day_Calls	Day_Charge	Eve_Calls	Eve_Charge	Night_Calls	Night_Charge	Intl_Calls	Intl_Charge
Day_Mins	1.000000	0.007043	0.004323	-0.010155	-0.013423	0.006750	1.000000	0.015769	0.007029	0.022972	0.004300	0.008033	-0.010092
Eve_Mins	0.007043	1.000000	-0.012584	-0.011035	-0.012985	-0.021451	0.007050	-0.011430	1.000000	0.007586	-0.012593	0.002541	-0.011067
Night_Mins	0.004323	-0.012584	1.000000	-0.015207	-0.009288	0.022938	0.004324	-0.002093	-0.012592	0.011204	0.999999	-0.012353	-0.015180
Intl_Mins	-0.010155	-0.011035	-0.015207	1.000000	-0.009640	0.021565	-0.010157	0.008703	-0.011043	-0.013605	-0.015214	0.032304	0.999993
CustServ_Calls	-0.013423	-0.012985	-0.009288	-0.009640	1.000000	-0.018942	-0.013427	0.002423	-0.012987	-0.012802	-0.009277	-0.017561	-0.009675
Day_Calls	0.006750	-0.021451	0.022938	0.021565	-0.018942	1.000000	0.006753	0.006462	-0.021449	-0.019557	0.022927	0.004574	0.021666
Day_Charge	1.000000	0.007050	0.004324	-0.010157	-0.013427	0.006753	1.000000	0.015769	0.007036	0.022972	0.004301	0.008032	-0.010094
Eve_Calls	0.015769	-0.011430	-0.002093	0.008703	0.002423	0.006462	0.015769	1.000000	-0.011423	0.007710	-0.002056	0.017434	0.008674
Eve_Charge	0.007029	1.000000	-0.012592	-0.011043	-0.012987	-0.021449	0.007036	-0.011423	1.000000	0.007596	-0.012601	0.002541	-0.011074
Night_Calls	0.022972	0.007586	0.011204	-0.013605	-0.012802	-0.019557	0.022972	0.007710	0.007596	1.000000	0.011188	0.000305	-0.013630
Night_Charge	0.004300	-0.012593	0.999999	-0.015214	-0.009277	0.022927	0.004301	-0.002056	-0.012601	0.011188	1.000000	-0.012329	-0.015186
Intl_Calls	0.008033	0.002541	-0.012353	0.032304	-0.017561	0.004574	0.008032	0.017434	0.002541	0.000305	-0.012329	1.000000	0.032372
Intl_Charge	-0.010092	-0.011067	-0.015180	0.999993	-0.009675	0.021666	-0.010094	0.008674	-0.011074	-0.013630	-0.015186	0.032372	1.000000

```
telco.corr()
```

	Day_Mins	Eve_Mins	Night_Mins	Intl_Mins	CustServ_Calls	Day_Calls	Day_Charge	Eve_Calls	Eve_Charge	Night_Calls	Night_Charge	Intl_Calls	Intl_Charge
Day_Mins	1.000000	0.007043	0.004323	-0.010155	-0.013423	0.006750	1.000000	0.015769	0.007029	0.022972	0.004300	0.008033	-0.010092
Eve_Mins	0.007043	1.000000	-0.012584	-0.011035	-0.012985	-0.021451	0.007050	-0.011430	1.000000	0.007586	-0.012593	0.002541	-0.011067
Night_Mins	0.004323	-0.012584	1.000000	-0.015207	-0.009288	0.022938	0.004324	-0.002093	-0.012592	0.011204	0.999999	-0.012353	-0.015180
Intl_Mins	-0.010155	-0.011035	-0.015207	1.000000	-0.009640	0.021565	-0.010157	0.008703	-0.011043	-0.013605	-0.015214	0.032304	0.999993
CustServ_Calls	-0.013423	-0.012985	-0.009288	-0.009640	1.000000	-0.018942	-0.013427	0.002423	-0.012987	-0.012802	-0.009277	-0.017561	-0.009675
Day_Calls	0.006750	-0.021451	0.022938	0.021565	-0.018942	1.000000	0.006753	0.006462	-0.021449	-0.019557	0.022927	0.004574	0.021666
Day_Charge	1.000000	0.007050	0.004324	-0.010157	-0.013427	0.006753	1.000000	0.015769	0.007036	0.022972	0.004301	0.008032	-0.010094
Eve_Calls	0.015769	-0.011430	-0.002093	0.008703	0.002423	0.006462	0.015769	1.000000	-0.011423	0.007710	-0.002056	0.017434	0.008674
Eve_Charge	0.007029	1.000000	-0.012592	-0.011043	-0.012987	-0.021449	0.007036	-0.011423	1.000000	0.007596	-0.012601	0.002541	-0.011074
Night_Calls	0.022972	0.007586	0.011204	-0.013605	-0.012802	-0.019557	0.022972	0.007710	0.007596	1.000000	0.011188	0.000305	-0.013630
Night_Charge	0.004300	-0.012593	0.999999	-0.015214	-0.009277	0.022927	0.004301	-0.002056	-0.012601	0.011188	1.000000	-0.012329	-0.015186
Intl_Calls	0.008033	0.002541	-0.012353	0.032304	-0.017561	0.004574	0.008032	0.017434	0.002541	0.000305	-0.012329	1.000000	0.032372
Intl_Charge	-0.010092	-0.011067	-0.015180	0.999993	-0.009675	0.021666	-0.010094	0.008674	-0.011074	-0.013630	-0.015186	0.032372	1.000000


```
telco.corr()
```

	Day_Mins	Eve_Mins	Night_Mins	Intl_Mins	CustServ_Calls	Day_Calls	Day_Charge	Eve_Calls	Eve_Charge	Night_Calls	Night_Charge	Intl_Calls	Intl_Charge
Day_Mins	1.000000	0.007043	0.004323	-0.010155	-0.013423	0.006750	1.000000	0.015769	0.007029	0.022972	0.004300	0.008033	-0.010092
Eve_Mins	0.007043	1.000000	-0.012584	-0.011035	-0.012985	-0.021451	0.007050	-0.011430	1.000000	0.007586	-0.012593	0.002541	-0.011067
Night_Mins	0.004323	-0.012584	1.000000	-0.015207	-0.009288	0.022938	0.004324	-0.002093	-0.012592	0.011204	0.999999	-0.012353	-0.015180
Intl_Mins	-0.010155	-0.011035	-0.015207	1.000000	-0.009640	0.021565	-0.010157	0.008703	-0.011043	-0.013605	-0.015214	0.032304	0.999993
CustServ_Calls	-0.013423	-0.012985	-0.009288	-0.009640	1.000000	-0.018942	-0.013427	0.002423	-0.012987	-0.012802	-0.009277	-0.017561	-0.009675
Day_Calls	0.006750	-0.021451	0.022938	0.021565	-0.018942	1.000000	0.006753	0.006462	-0.021449	-0.019557	0.022927	0.004574	0.021666
Day_Charge	1.000000	0.007050	0.004324	-0.010157	-0.013427	0.006753	1.000000	0.015769	0.007036	0.022972	0.004301	0.008032	-0.010094
Eve_Calls	0.015769	-0.011430	-0.002093	0.008703	0.002423	0.006462	0.015769	1.000000	-0.011423	0.007710	-0.002056	0.017434	0.008674
Eve_Charge	0.007029	1.000000	-0.012592	-0.011043	-0.012987	-0.021449	0.007036	-0.011423	1.000000	0.007596	-0.012601	0.002541	-0.011074
Night_Calls	0.022972	0.007586	0.011204	-0.013605	-0.012802	-0.019557	0.022972	0.007710	0.007596	1.000000	0.011188	0.000305	-0.013630
Night_Charge	0.004300	-0.012593	0.999999	-0.015214	-0.009277	0.022927	0.004301	-0.002056	-0.012601	0.011188	1.000000	-0.012329	-0.015186
Intl_Calls	0.008033	0.002541	-0.012353	0.032304	-0.017561	0.004574	0.008032	0.017434	0.002541	0.000305	-0.012329	1.000000	0.032372
Intl_Charge	-0.010092	-0.011067	-0.015180	0.999993	-0.009675	0.021666	-0.010094	0.008674	-0.011074	-0.013630	-0.015186	0.032372	1.000000


```
telco.corr()
```

	Day_Mins	Eve_Mins	Night_Mins	Intl_Mins	CustServ_Calls	Day_Calls	Day_Charge	Eve_Calls	Eve_Charge	Night_Calls	Night_Charge	Intl_Calls	Intl_Charge
Day_Mins	1.000000	0.007043	0.004323	-0.010155	-0.013423	0.006750	1.000000	0.015769	0.007029	0.022972	0.004300	0.008033	-0.010092
Eve_Mins	0.007043	1.000000	-0.012584	-0.011035	-0.012985	-0.021451	0.007050	-0.011430	1.000000	0.007586	-0.012593	0.002541	-0.011067
Night_Mins	0.004323	-0.012584	1.000000	-0.015207	-0.009288	0.022938	0.004324	-0.002093	-0.012592	0.011204	0.999999	-0.012353	-0.015180
Intl_Mins	-0.010155	-0.011035	-0.015207	1.000000	-0.009640	0.021565	-0.010157	0.008703	-0.011043	-0.013605	-0.015214	0.032304	0.999993
CustServ_Calls	-0.013423	-0.012985	-0.009288	-0.009640	1.000000	-0.018942	-0.013427	0.002423	-0.012987	-0.012802	-0.009277	-0.017561	-0.009675
Day_Calls	0.006750	-0.021451	0.022938	0.021565	-0.018942	1.000000	0.006753	0.006462	-0.021449	-0.019557	0.022927	0.004574	0.021666
Day_Charge	1.000000	0.007050	0.004324	-0.010157	-0.013427	0.006753	1.000000	0.015769	0.007036	0.022972	0.004301	0.008032	-0.010094
Eve_Calls	0.015769	-0.011430	-0.002093	0.008703	0.002423	0.006462	0.015769	1.000000	-0.011423	0.007710	-0.002056	0.017434	0.008674
Eve_Charge	0.007029	1.000000	-0.012592	-0.011043	-0.012987	-0.021449	0.007036	-0.011423	1.000000	0.007596	-0.012601	0.002541	-0.011074
Night_Calls	0.022972	0.007586	0.011204	-0.013605	-0.012802	-0.019557	0.022972	0.007710	0.007596	1.000000	0.011188	0.000305	-0.013630
Night_Charge	0.004300	-0.012593	0.999999	-0.015214	-0.009277	0.022927	0.004301	-0.002056	-0.012601	0.011188	1.000000	-0.012329	-0.015186
Intl_Calls	0.008033	0.002541	-0.012353	0.032304	-0.017561	0.004574	0.008032	0.017434	0.002541	0.000305	-0.012329	1.000000	0.032372
Intl_Charge	-0.010092	-0.011067	-0.015180	0.999993	-0.009675	0.021666	-0.010094	0.008674	-0.011074	-0.013630	-0.015186	0.032372	1.000000

```
telco.corr()
```

	Day_Mins	Eve_Mins	Night_Mins	Intl_Mins	CustServ_Calls	Day_Calls	Day_Charge	Eve_Calls	Eve_Charge	Night_Calls	Night_Charge	Intl_Calls	Intl_Charge
Day_Mins	1.000000	0.007043	0.004323	-0.010155	-0.013423	0.006750	1.000000	0.015769	0.007029	0.022972	0.004300	0.008033	-0.010092
Eve_Mins	0.007043	1.000000	-0.012584	-0.011035	-0.012985	-0.021451	0.007050	-0.011430	1.000000	0.007586	-0.012593	0.002541	-0.011067
Night_Mins	0.004323	-0.012584	1.000000	-0.015207	-0.009288	0.022938	0.004324	-0.002093	-0.012592	0.011204	0.999999	-0.012353	-0.015180
Intl_Mins	-0.010155	-0.011035	-0.015207	1.000000	-0.009640	0.021565	-0.010157	0.008703	-0.011043	-0.013605	-0.015214	0.032304	0.999993
CustServ_Calls	-0.013423	-0.012985	-0.009288	-0.009640	1.000000	-0.018942	-0.013427	0.002423	-0.012987	-0.012802	-0.009277	-0.017561	-0.009675
Day_Calls	0.006750	-0.021451	0.022938	0.021565	-0.018942	1.000000	0.006753	0.006462	-0.021449	-0.019557	0.022927	0.004574	0.021666
Day_Charge	1.000000	0.007050	0.004324	-0.010157	-0.013427	0.006753	1.000000	0.015769	0.007036	0.022972	0.004301	0.008032	-0.010094
Eve_Calls	0.015769	-0.011430	-0.002093	0.008703	0.002423	0.006462	0.015769	1.000000	-0.011423	0.007710	-0.002056	0.017434	0.008674
Eve_Charge	0.007029	1.000000	-0.012592	-0.011043	-0.012987	-0.021449	0.007036	-0.011423	1.000000	0.007596	-0.012601	0.002541	-0.011074
Night_Calls	0.022972	0.007586	0.011204	-0.013605	-0.012802	-0.019557	0.022972	0.007710	0.007596	1.000000	0.011188	0.000305	-0.013630
Night_Charge	0.004300	-0.012593	0.999999	-0.015214	-0.009277	0.022927	0.004301	-0.002056	-0.012601	0.011188	1.000000	-0.012329	-0.015186
Intl_Calls	0.008033	0.002541	-0.012353	0.032304	-0.017561	0.004574	0.008032	0.017434	0.002541	0.000305	-0.012329	1.000000	0.032372
Intl_Charge	-0.010092	-0.011067	-0.015180	0.999993	-0.009675	0.021666	-0.010094	0.008674	-0.011074	-0.013630	-0.015186	0.032372	1.000000

Feature engineering

- Creating new features to help improve model performance
- Should consult with business and subject matter experts

Examples of feature engineering

- Total Minutes: Sum of `Day_Mins` , `Eve_Mins` , `Night_Mins` , `Intl_Mins`
- Ratio between Minutes and Charge

```
telco['Day_Cost'] = telco['Day_Mins'] / telco['Day_Charge']
```

Let's practice!

PREDICTING CUSTOMER CHURN IN PYTHON