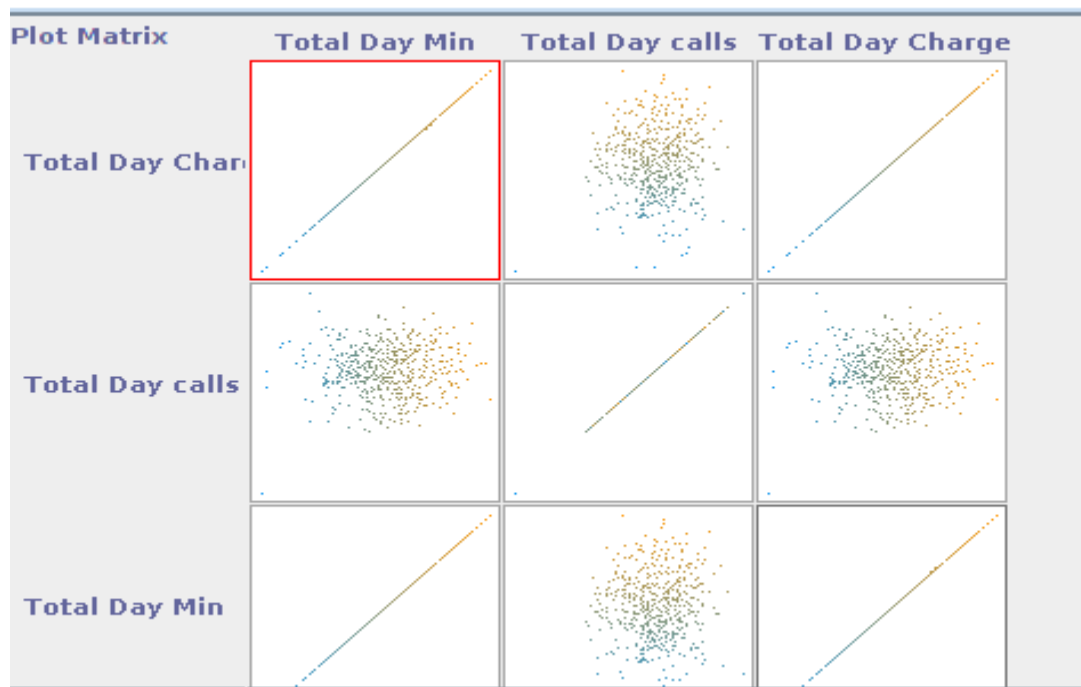
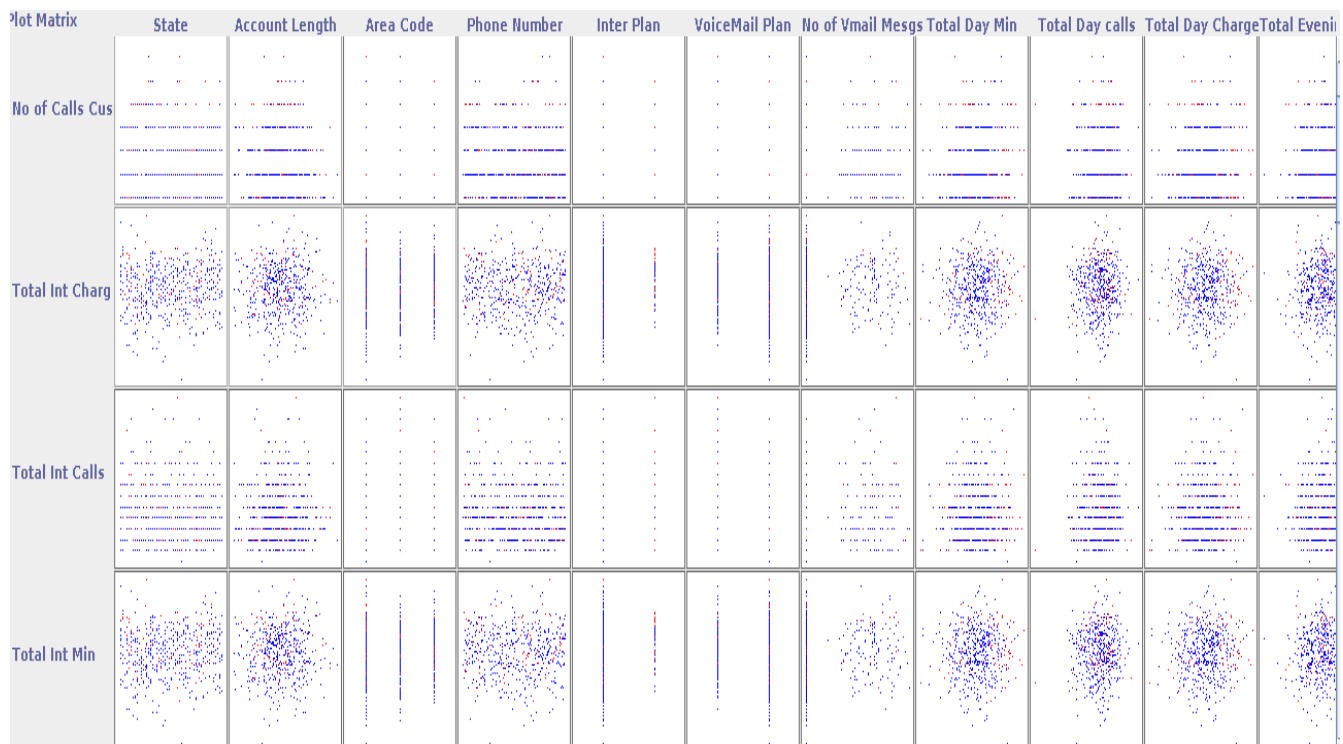


1. Posibles variables correlacionadas.

Utilizando weka para identificar posibles variables correlacionadas tenemos:



tenemos variables que son dependientes o que se puede obtener una a partir de la otra esto a simple vista estas son:

Total int charge – Total int min
Total nighth charge – Total nighth min
Total evening charge – Total evening min
Total day charge – Total day

Para verificar esto debemos hallar el coeficiente de correlacion.

```
Attributes:  2
             Total Day Min
             Total Day Charge
Test mode:10-fold cross-validation

=== Classifier model (full training set) ===

Linear Regression Model

Total Day Charge =

      0.17   * Total Day Min +
      0.0006

Time taken to build model: 0.21 seconds

=== Cross-validation ===
=== Summary ===

Correlation coefficient          1
Mean absolute error             0.0025
Root mean squared error         0.0029
Relative absolute error         0.0337 %
Root relative squared error     0.0309 %
Total Number of Instances      3333
```

```

Attributes: 2
            Total Int Min
            Total Int Charge
Test mode:10-fold cross-validation

=== Classifier model (full training set) ===

Linear Regression Model

Total Int Charge =

            0.27 * Total Int Min +
            0.0006

Time taken to build model: 0.01 seconds

=== Cross-validation ===
=== Summary ===

Correlation coefficient          1
Mean absolute error             0.0025
Root mean squared error         0.0029
Relative absolute error         0.1338 %
Attributes: 2
            Total Night Minutes
            Total Night Charge
Test mode:10-fold cross-validation

=== Classifier model (full training set) ===

Linear Regression Model

Total Night Charge =

            0.045 * Total Night Minutes +
            -0.0001

Time taken to build model: 0.03 seconds

=== Cross-validation ===
=== Summary ===

Correlation coefficient          1
Mean absolute error             0.0025
Root mean squared error         0.0029
Relative absolute error         0.1358 %
Root relative squared error     0.1253 %
Total Number of Instances       3333

```

```

Attributes:  2
              Total Evening Min
              Total Evening Charge
Test mode:10-fold cross-validation

=== Classifier model (full training set) ===

Linear Regression Model

Total Evening Charge =

      0.085 * Total Evening Min +
      0.0003

Time taken to build model: 0.03 seconds

=== Cross-validation ===
=== Summary ===

Correlation coefficient          1
Mean absolute error             0.0025
Root mean squared error         0.0029
Relative absolute error         0.0727 %
Root relative squared error     0.067 %
Total Number of Instances      3333

```

Mediante esto podemos corroborar que el coeficiente de correlacion es 1, osea una es totalmente obtenible a partir de la otra.

2. Se pueden eliminar variables.

Si se pueden eliminar variables:

State, Area Code, Numero Telefono son irrelevantes para soluciones al problema.

Total int charge

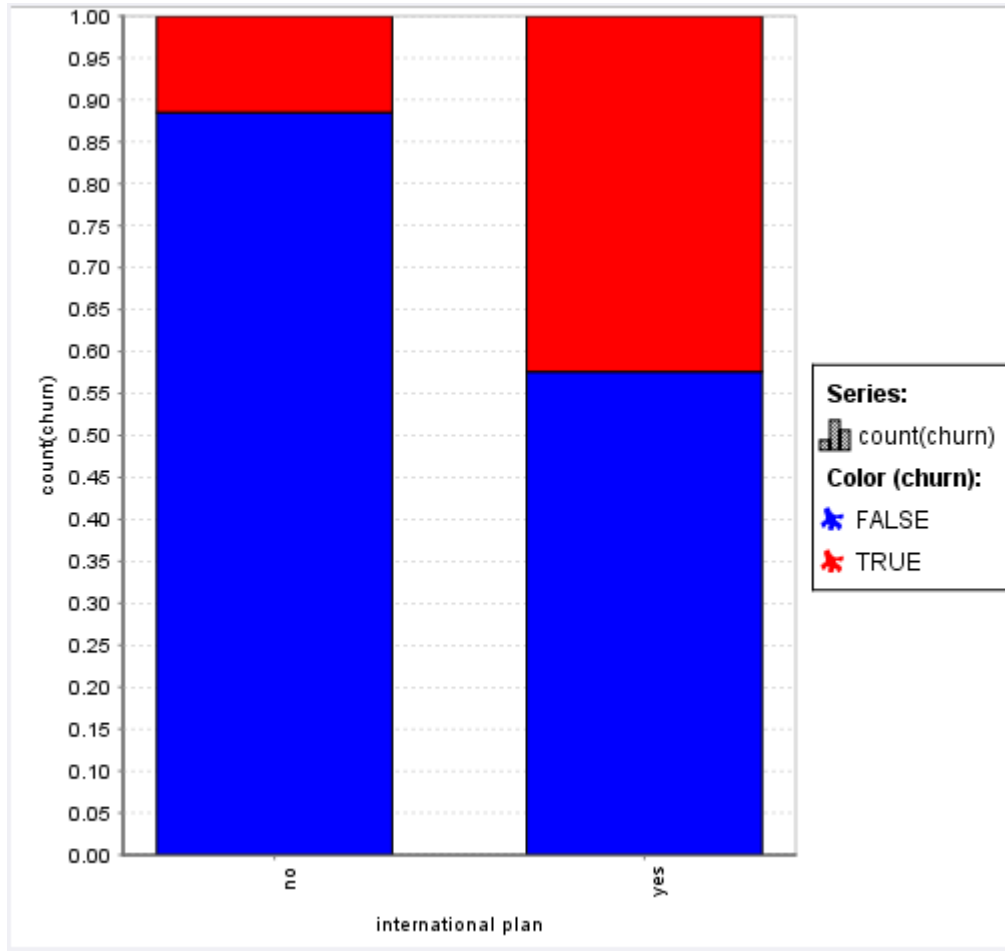
Total nighth charge

Total evening charge

Total day charge

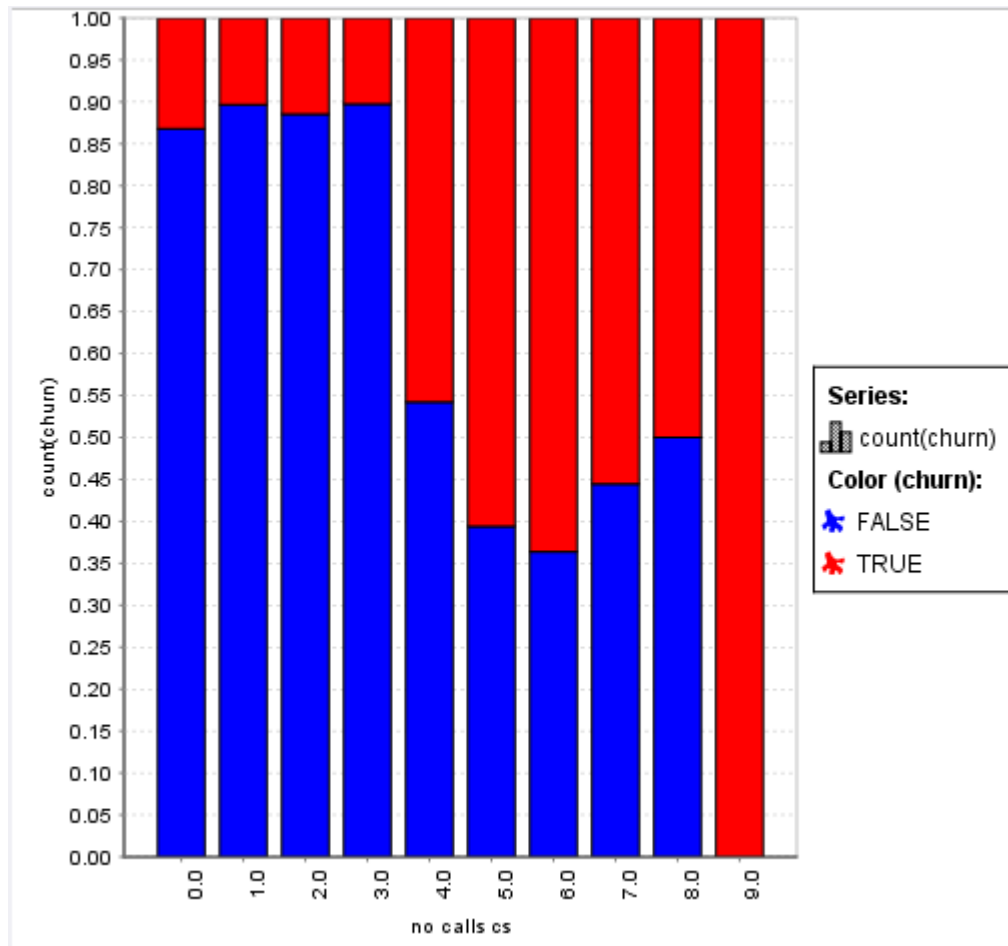
Pueden ser eliminadas por ser correlacionadas de otras.

3.



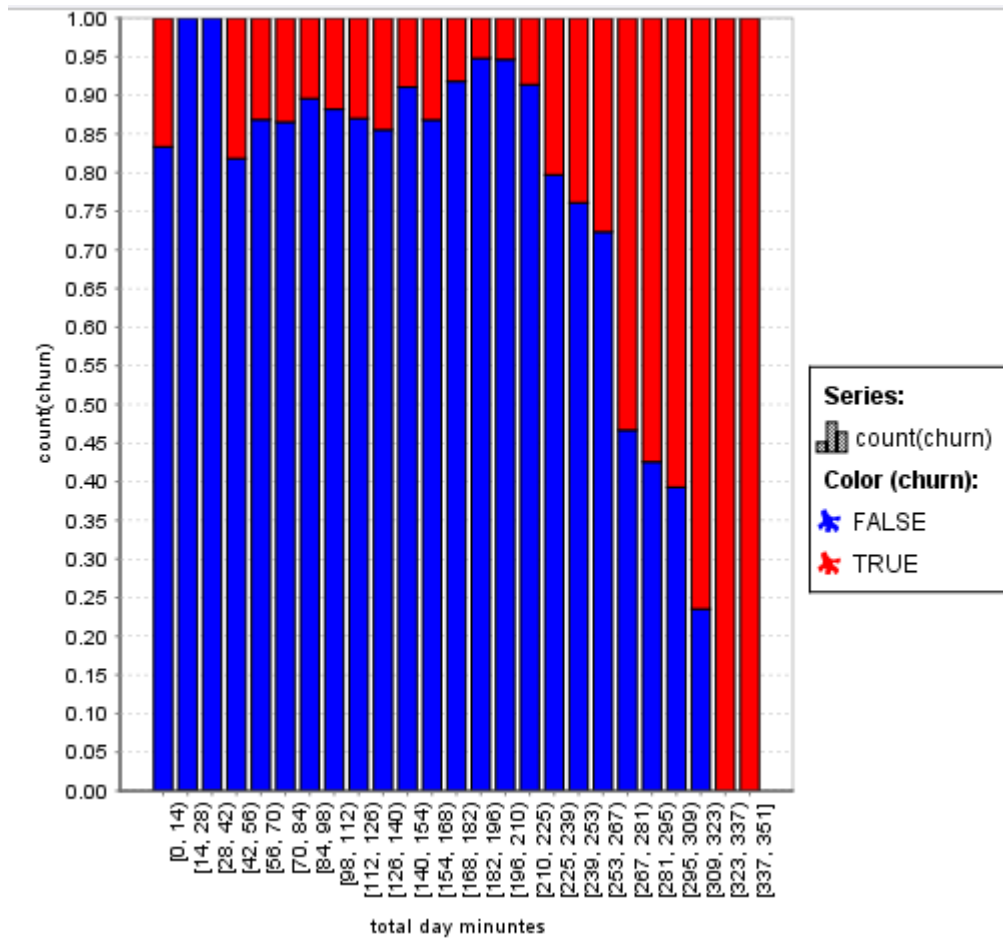
Los clientes que no poseen un plan internacional tienen mayor tendencia a permanecer en la empresa.

4.



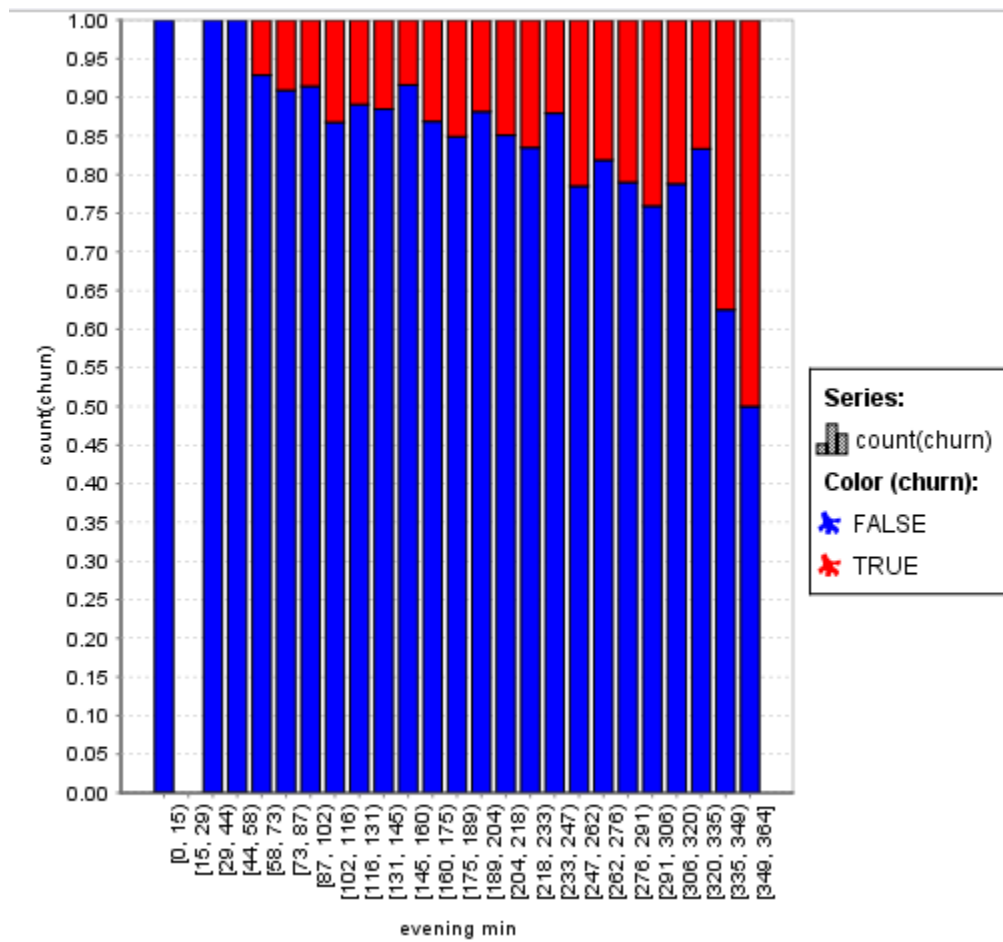
Las personas que poco llaman al servicio al cliente tienden a permanecer en la empresa.

5.



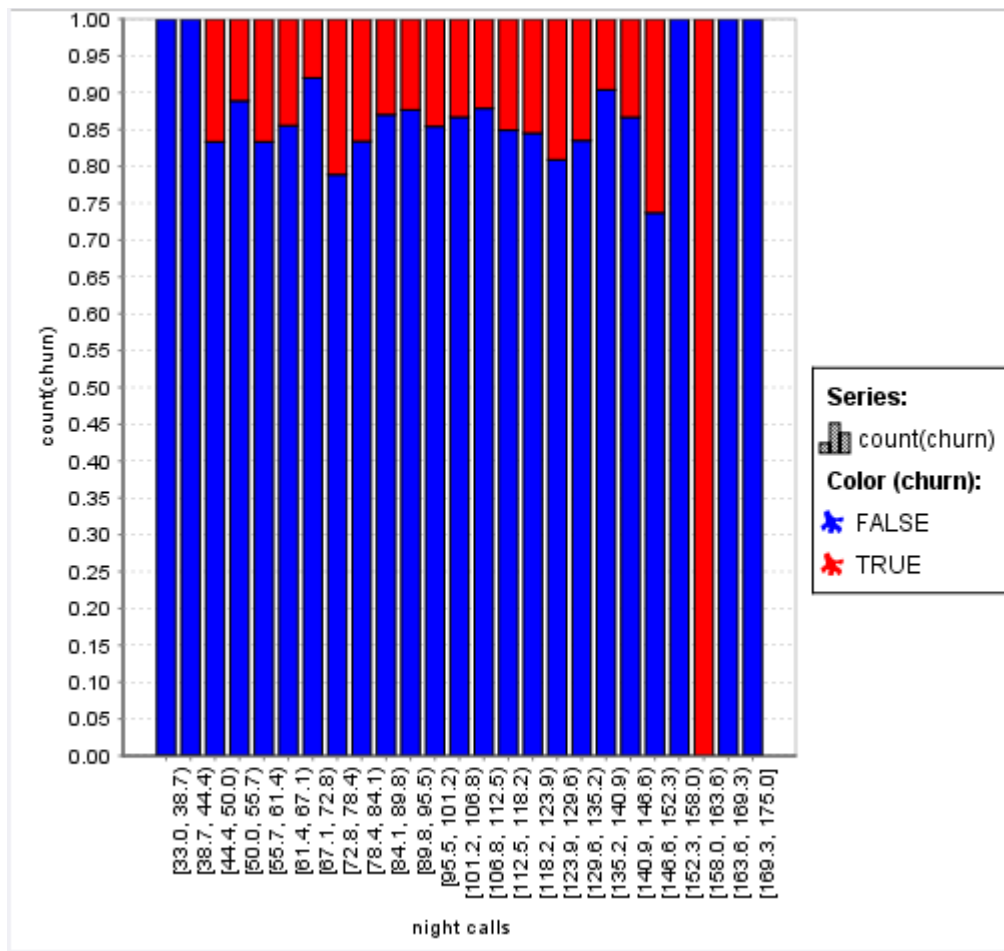
Las personas que mas consumen minutos en el dia tienden a abandonar la empresa.

6.

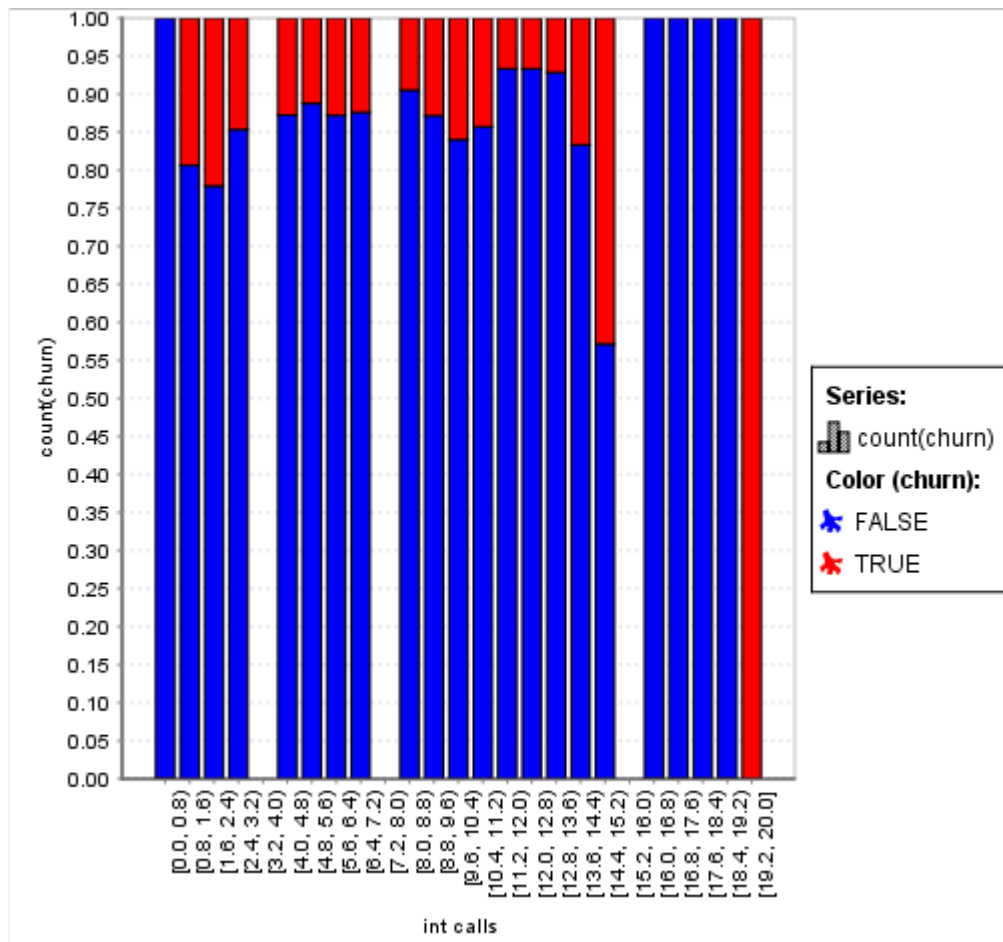


Los clientes que consumen pocos minutos en la tarde no abandonan la empresa.

7.



las personas que mas llamadas realizan en la noche no abandonan lqa empresa.



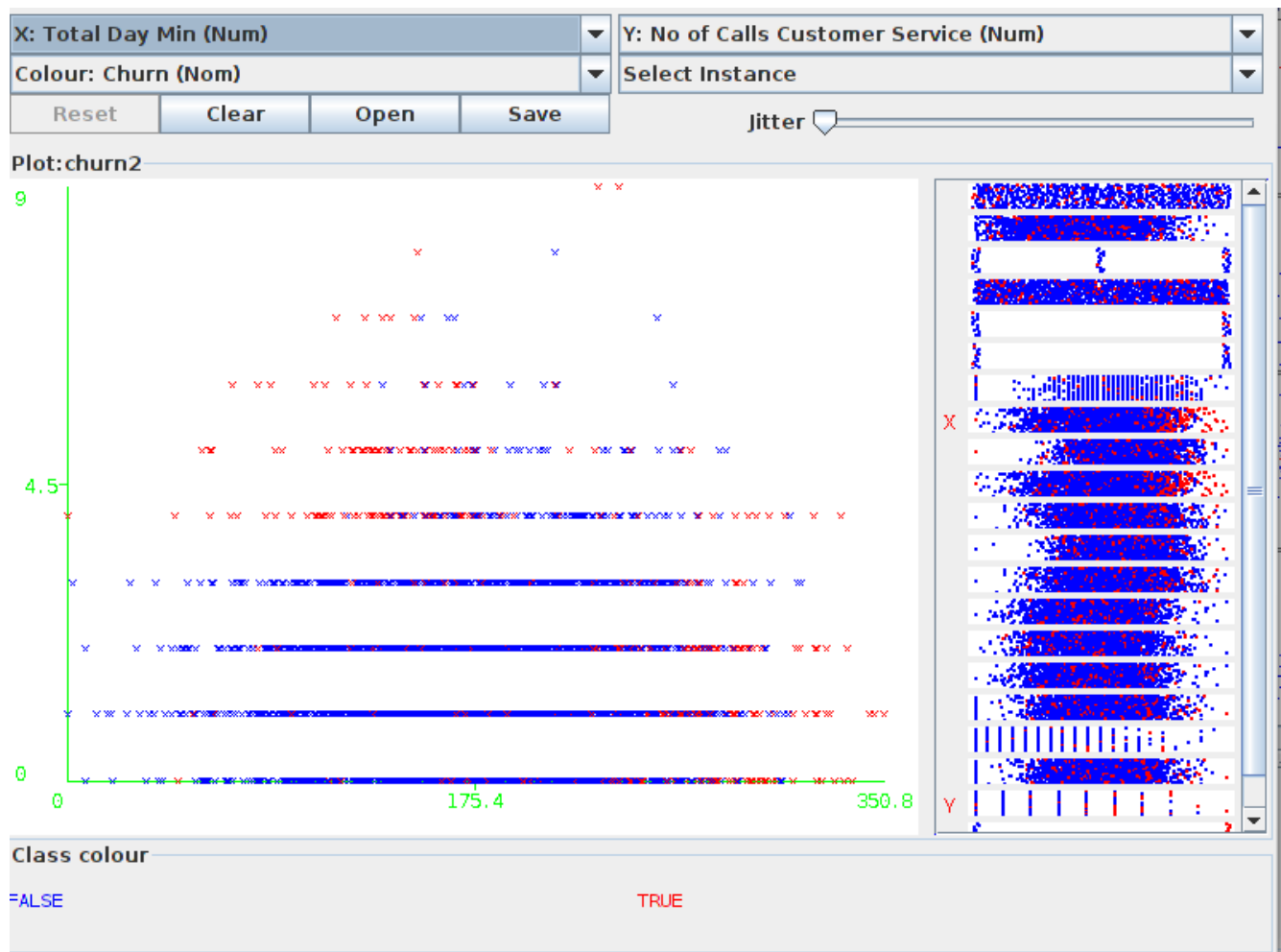
Las personas que mas realizan llamadas internacionales abandonan la empresa.
 Las personas que no realizan llamadas internacionales no abandonan la empresa.

8.

Variable	Decision
State	Anomala : no se puede hallar una distribucion normal
Code area	Anomala
Phone number	No ofrece informacion para el modelo.
International plan	Retenida : ofrece informacion para la solucion
Voicemail plan	Retenida
Number of voice mail messages	Retenida aunque no ofrece informacion acorde o bien distribuida
Today day call	Retenida
Total day min	Retenida
Total day charge	Eliminada : por correlacion
Total evening min	Retenida
Total evening Charge	Eliminada
Total evening calls	Retenida
Total nighth min	Retenida
Total nighth calls	Retenida
Total nighth charge	Eliminada
Total international mins	Retenida
Total international calls	Retenida
Total international charge	Eliminada
Calls cs	Retenida

1.1.1.

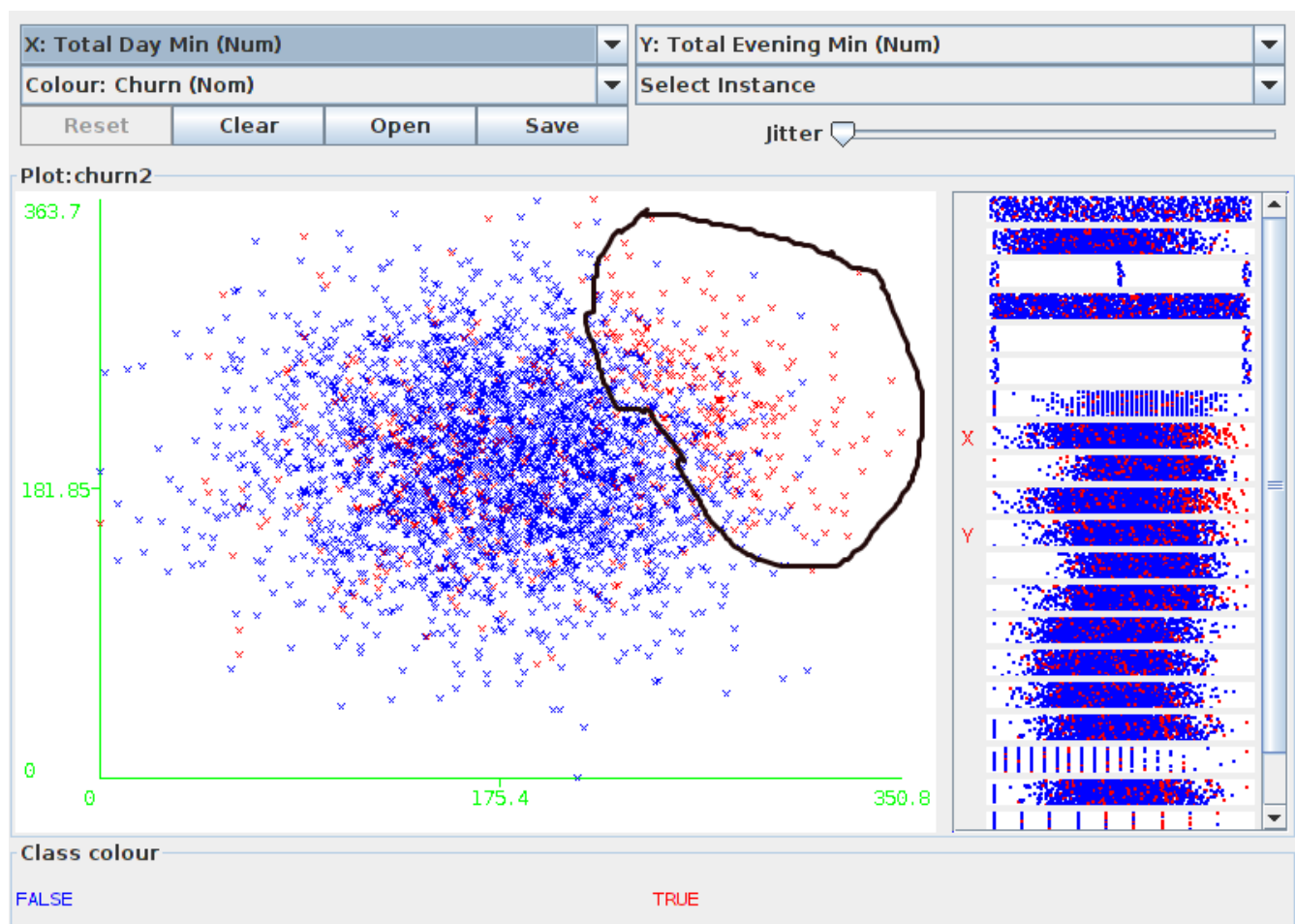
1.



Se puede obtener zonas de churn a partir de las 4 llamadas al servicio al cliente con una curva exponencial cuadrada.

Y a partir de los 245 minutos.

2.



la relacion es 1/3 churn por cliente.