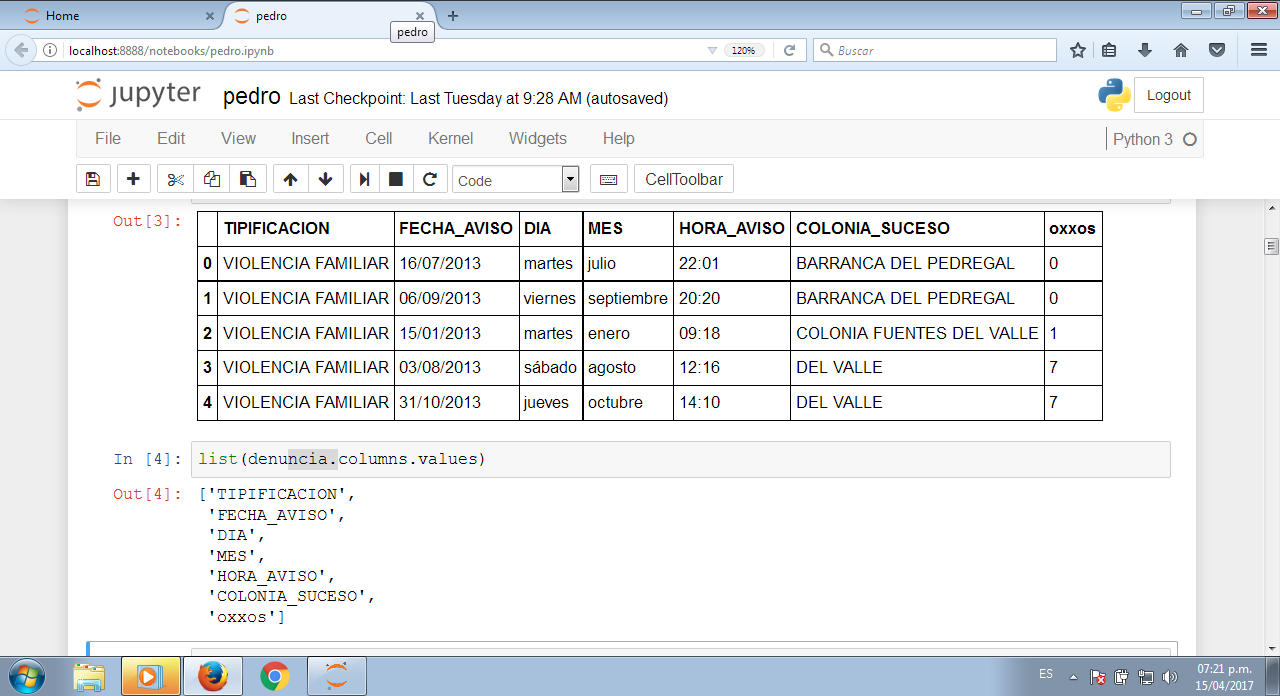
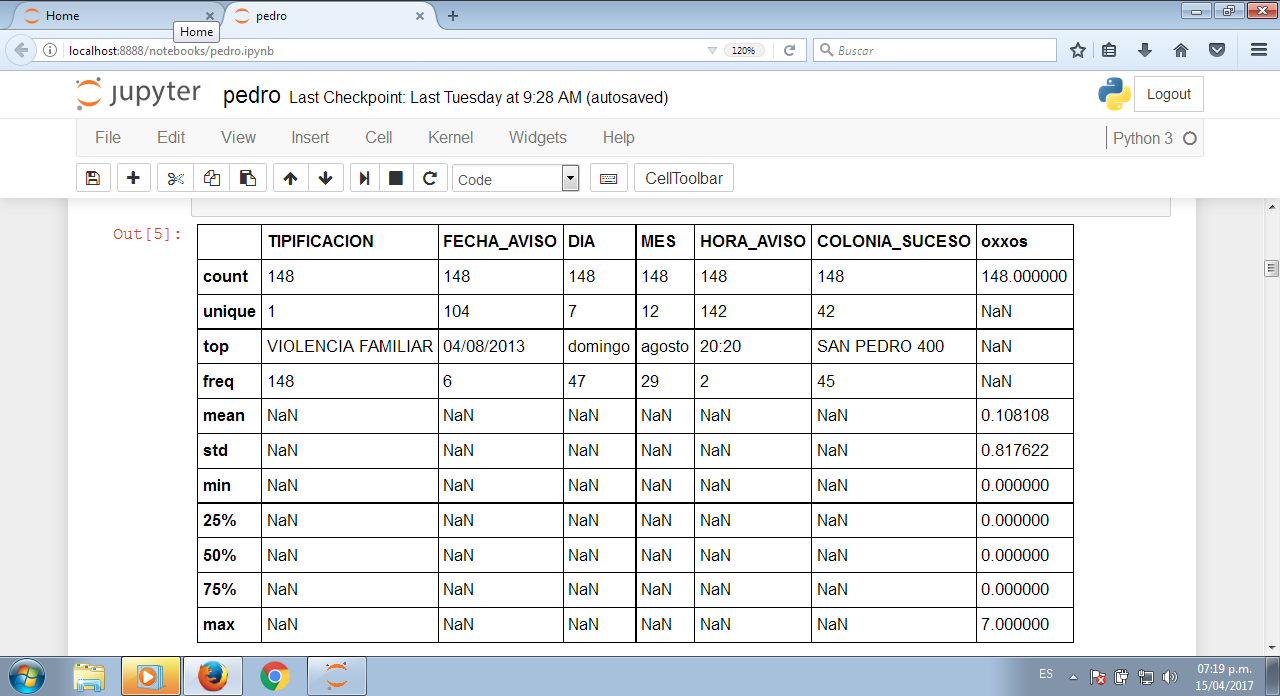
Trabajos sobre hallazgos en las bases de datos de san Pedro Nuevo León.

Primero cargaremos la base datos.

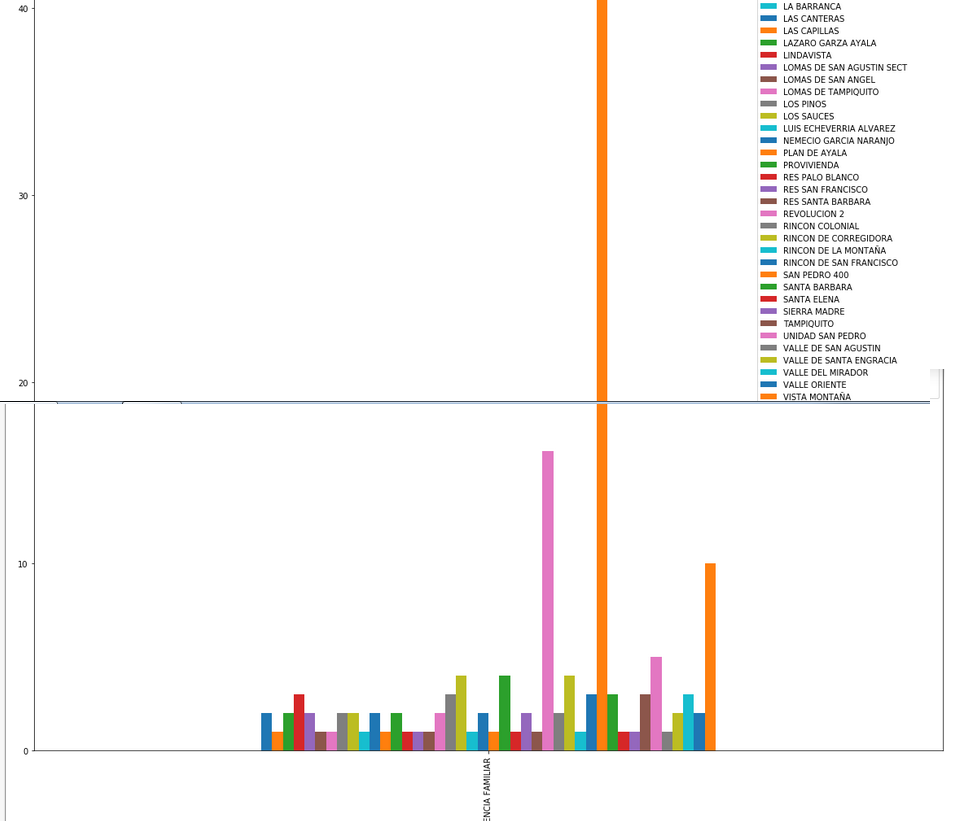
A la base de datos se creó una nueva variable llamada oxxos, dónde está el número de oxxos por colonia, para ello se clasifico cada oxxo en su colonia de procedencia.

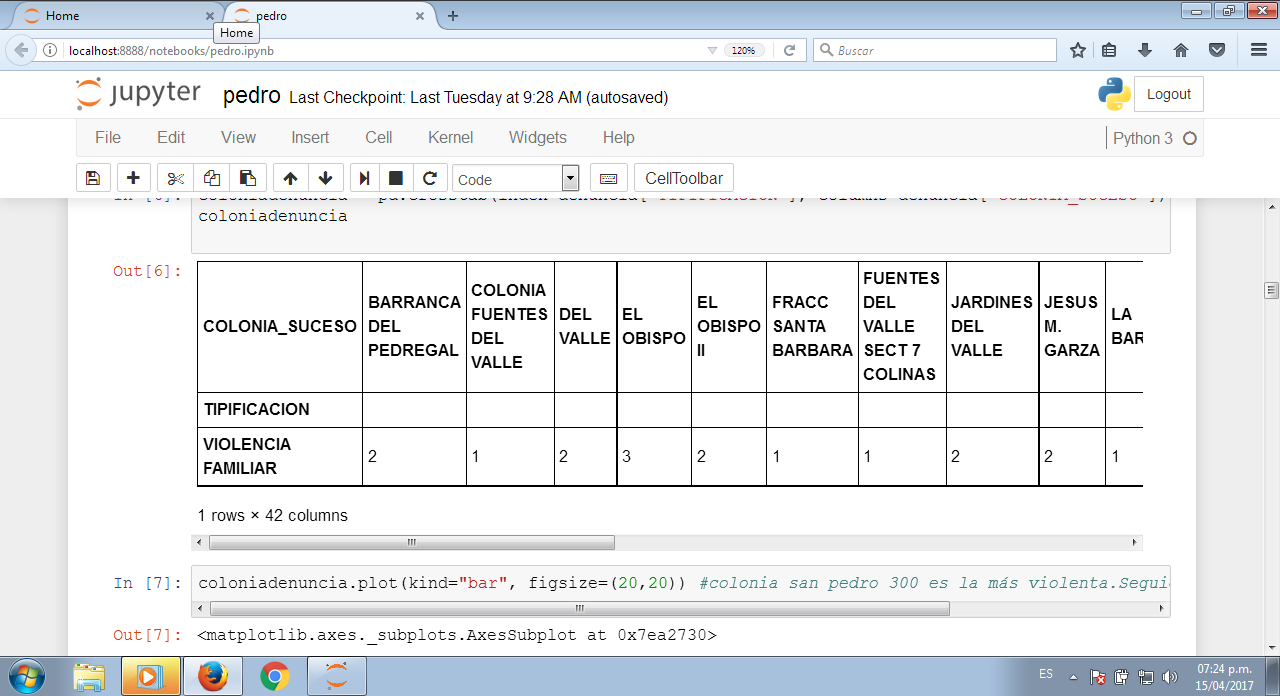
  
Utilizando la función describe podemos hallar el número de elementos de cada tipo.

Dando como información los valores top.

Los valores top son,

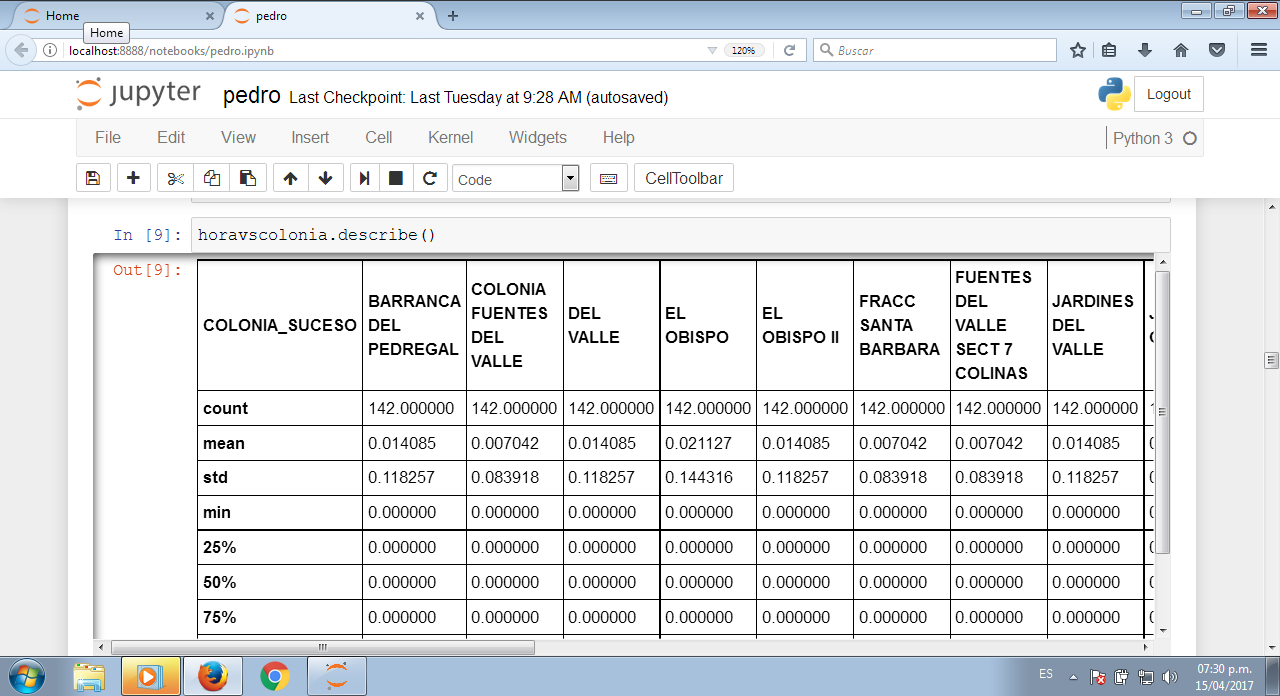
Fecha de aviso 04/08/2013, Domingo, agosto, hora de aviso 20:20, san pedro 400.

Para saber cual es la colonia de mayor denuncia crearé una nueva variable llamada “colonia suceso” que contiene la tipifica ción y la colonia.



*Paradójicamente el lugar con más oxxos(7) que sería una zona económicamente activa es una zona también que cuenta con más denuncias. San pedro 400, es la zona con más denuncias*. Seguido de revolución 2.

Para estudiar las horas y colonias de más denuncias se hace la misma función describe.



Sin embargo por ser “strings” no emite ningún valor, así que tendremos que hacer un cambio de variables.

Para ello quedan de la siguiente manera las variable.

TIPIFICACION

1 VIOLENCIA FAMILIAR

FECHA DE AVISO

....

DIA

1LUNES

2MARTES

3MIERCOLES

4JUEVES

5VIERNES

6SABADO

7DOMINGO

MES

1ENERO

2FEBRERO

COLONIA\_SUCESO

BARRANCA DEL PEDREGAL 1

COLONIA FUENTES DEL VALLE 2

DEL VALLE 3

EL OBISPO 4

EL OBISPO II 5

FRACC SANTA BARBARA 6

FUENTES DEL VALLE SECT 7 COLINAS 7

JARDINES DEL VALLE 8

JESUS M. GARZA 9

LA BARRANCA 10

LAS CANTERAS 11

LAS CAPILLAS 12

LAZARO GARZA AYALA 13

LINDAVISTA 14

LOMAS DE SAN AGUSTIN SECT 15

LOMAS DE SAN ANGEL 16

LOMAS DE TAMPIQUITO 17

LOS PINOS 18

LOS SAUCES 19

LUIS ECHEVERRIA ALVAREZ 20

NEMECIO GARCIA NARANJO 21

PLAN DE AYALA 22

PROVIVIENDA 23

RES PALO BLANCO 24

RES SAN FRANCISCO 25

RES SANTA BARBARA 26

REVOLUCION 2 27

RINCON COLONIAL 28

RINCON DE CORREGIDORA 29

RINCON DE LA MONTAÑA 30

RINCON DE SAN FRANCISCO 31

SAN PEDRO 400 32

SANTA BARBARA 33

SANTA ELENA 34

SIERRA MADRE 35

TAMPIQUITO 36

UNIDAD SAN PEDRO 37

VALLE DE SAN AGUSTIN 38

VALLE DE SANTA ENGRACIA 39

VALLE DEL MIRADOR 40

VALLE ORIENTE 41

VISTA MONTAÑA 42

OXXO

1

2

3

3MARZO

4ABRIL

5MAYO

6JUNIO

7JULIO

8AGOSTO

9SEPTIEMBRE

10OCTUBRE

11NOVIEMBRE

12DICIEMBRE

HORA AVISO

.....

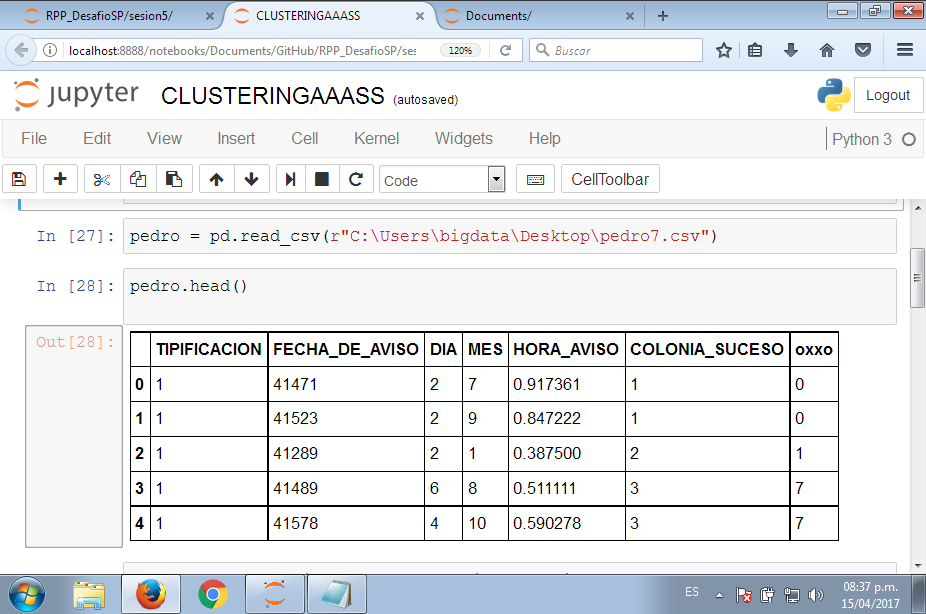
De las fechas y horas de aviso.

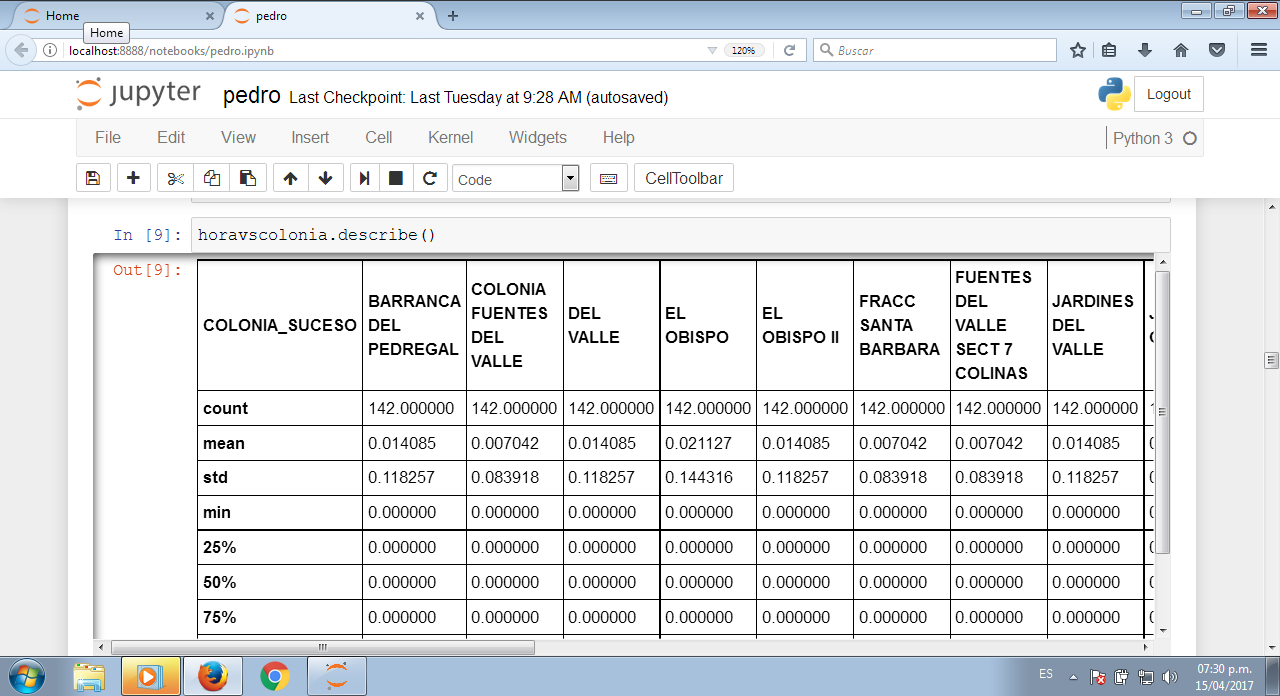
Microsoft Office Excel almacena las fechas como números secuenciales denominados valores de serie. Por ejemplo, en Microsoft Office Excel para Windows, la fecha 1 de enero de 1900 es el número de serie 1, y la fecha 1 de enero de 2008 es el número de serie 39448, porque es 39.448 días después del 1 de enero de 1900.

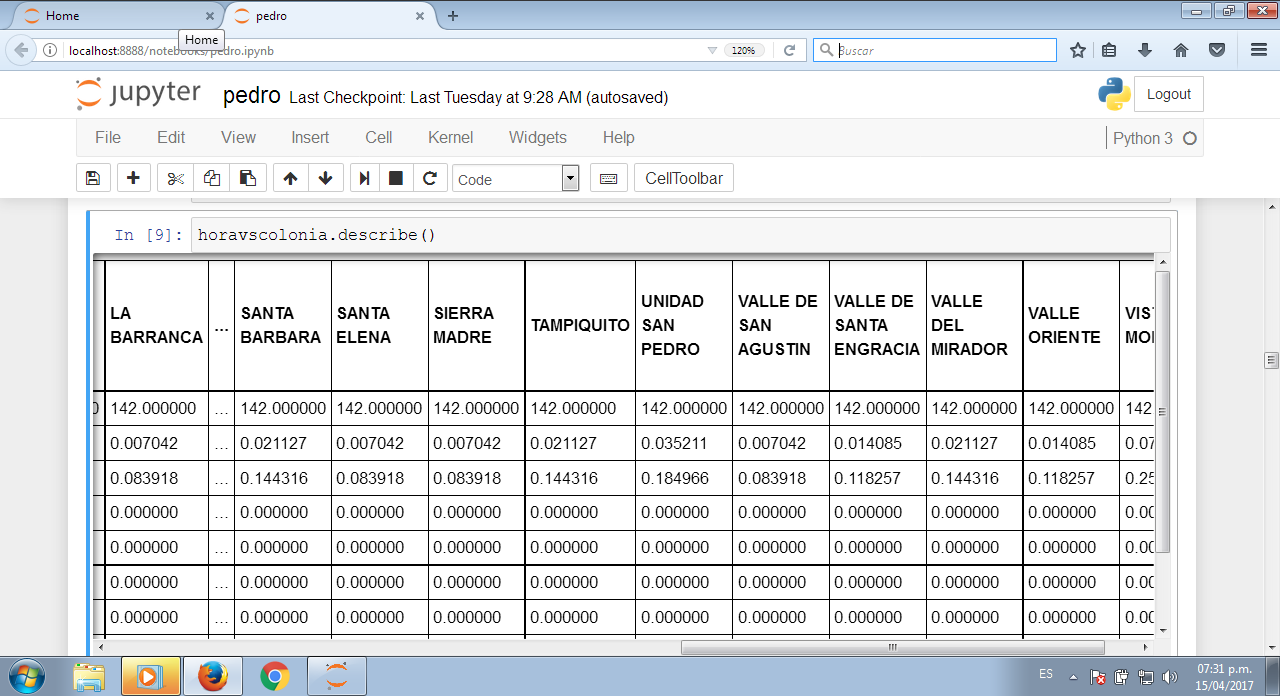
Office Excel almacena las horas como fracciones decimales, ya que la hora se considera una porción del día. El número decimal es un valor entre 0 (cero) y 0,99999999, y se corresponde con los momentos del día entre las 0:00:00 horas (12:00:00 a.m.) y las 23:59:59 horas (11:59:59 p.m.).

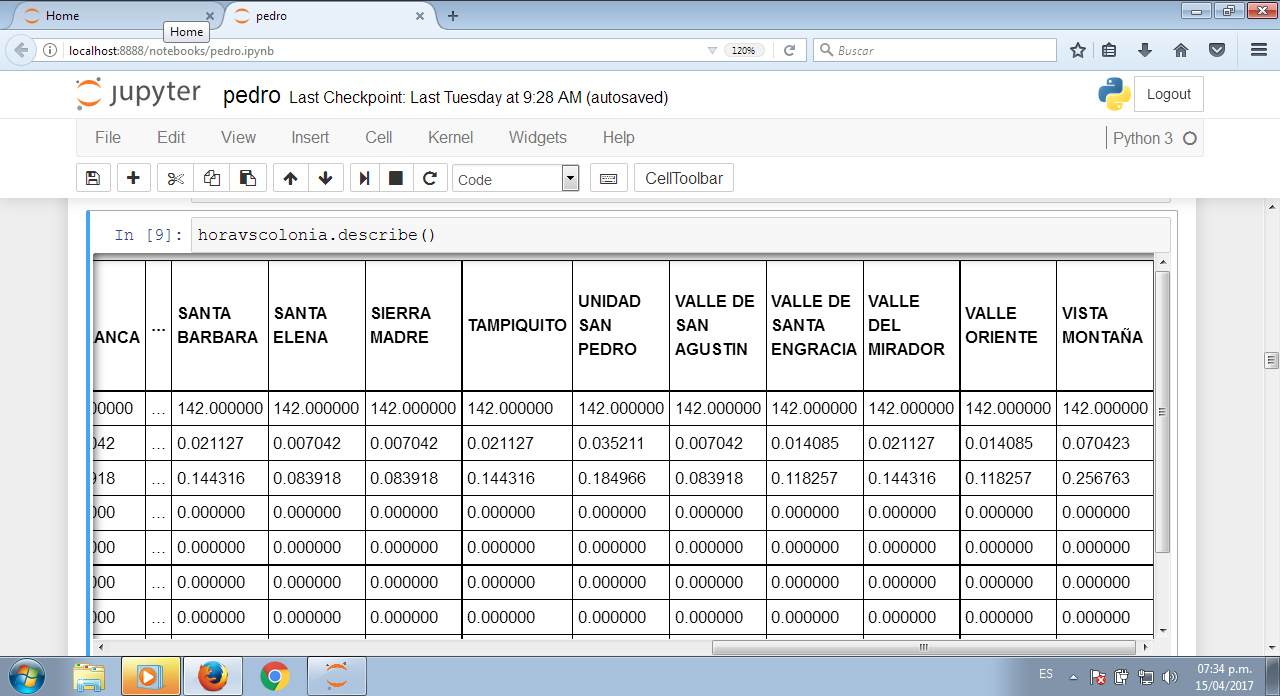
Las fechas y las horas son valores y, por lo tanto, pueden sumarse, restarse e incluirse en otros cálculos. Puede ver una fecha como un valor de serie y una hora como una fracción decimal, cambiando el formato de la celda que contenga la fecha o la hora a formato General.

Una vez hecho esta conversión podemos hacer un análisis con describe y realizar el clustering y regresión lineal.







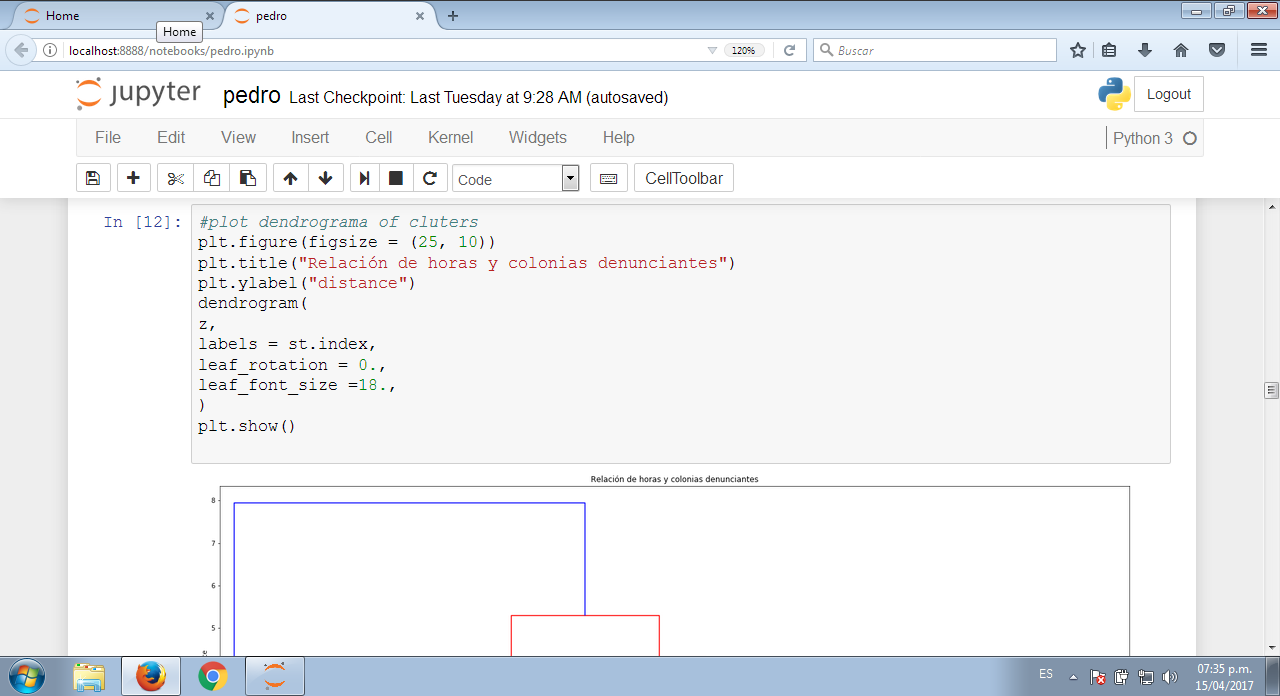


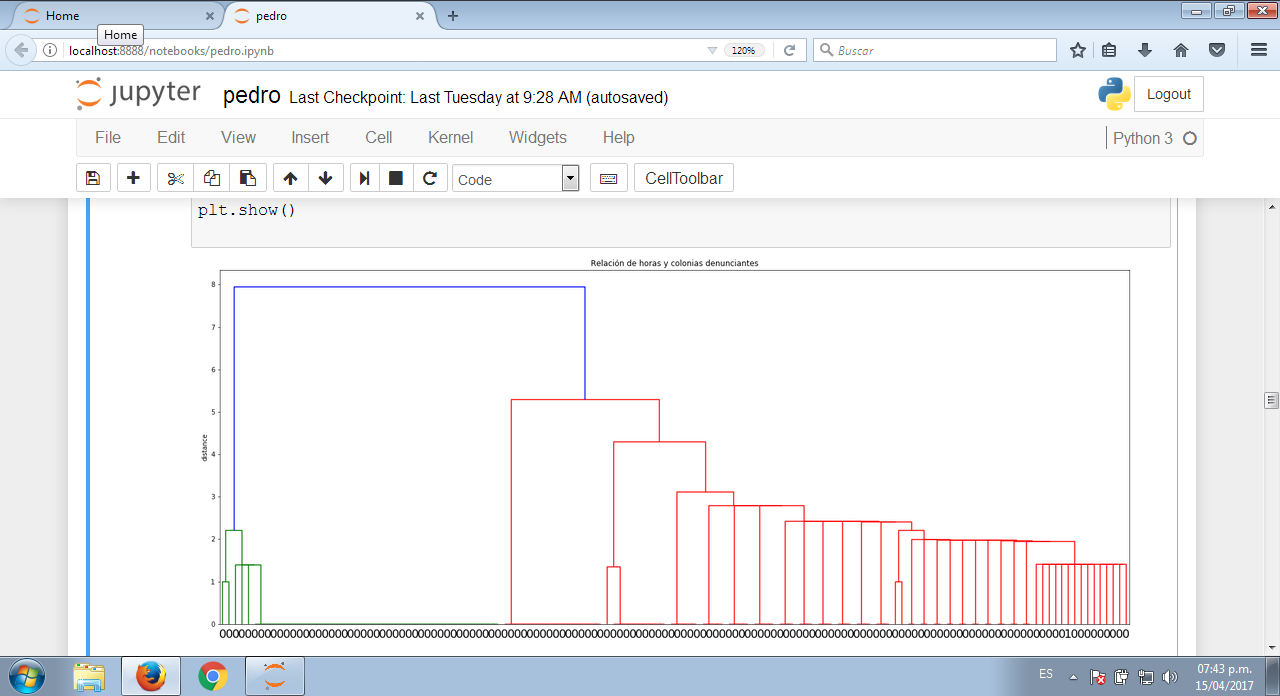
**Clustering**.

El clustering es una técnica que nos permite agrupar variables.

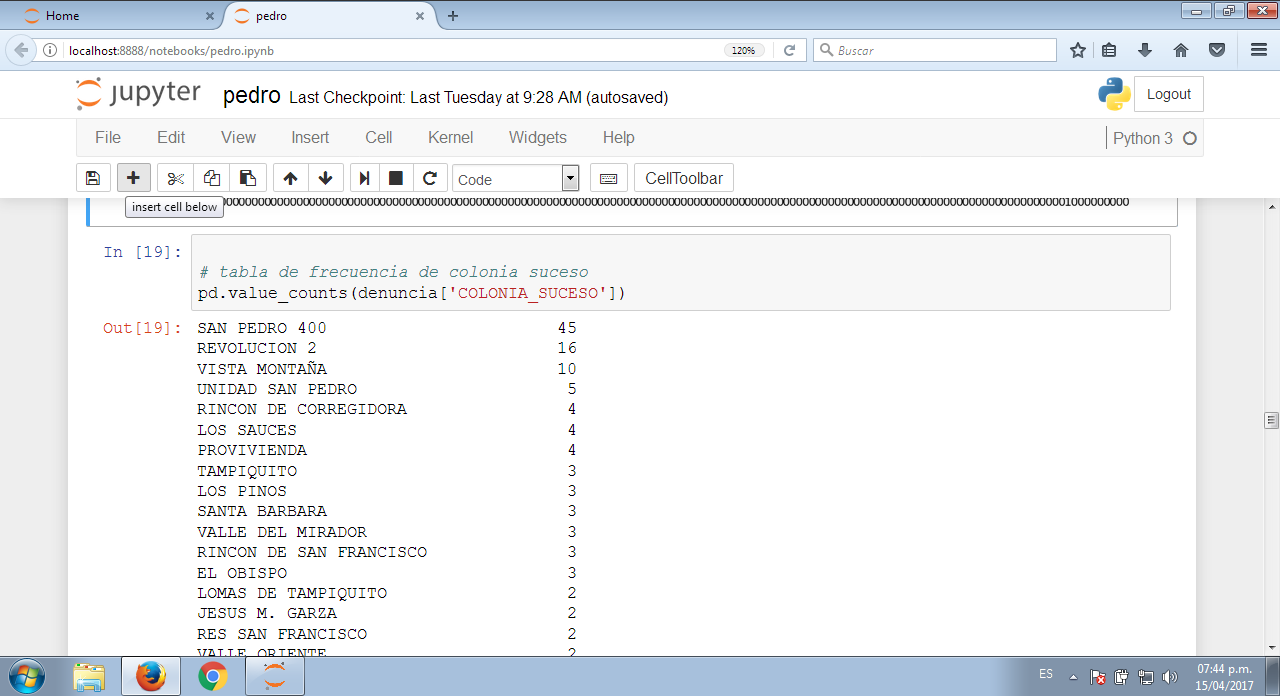
Hay dos tipos de clustering, el de jerarqueias y el de centroides.

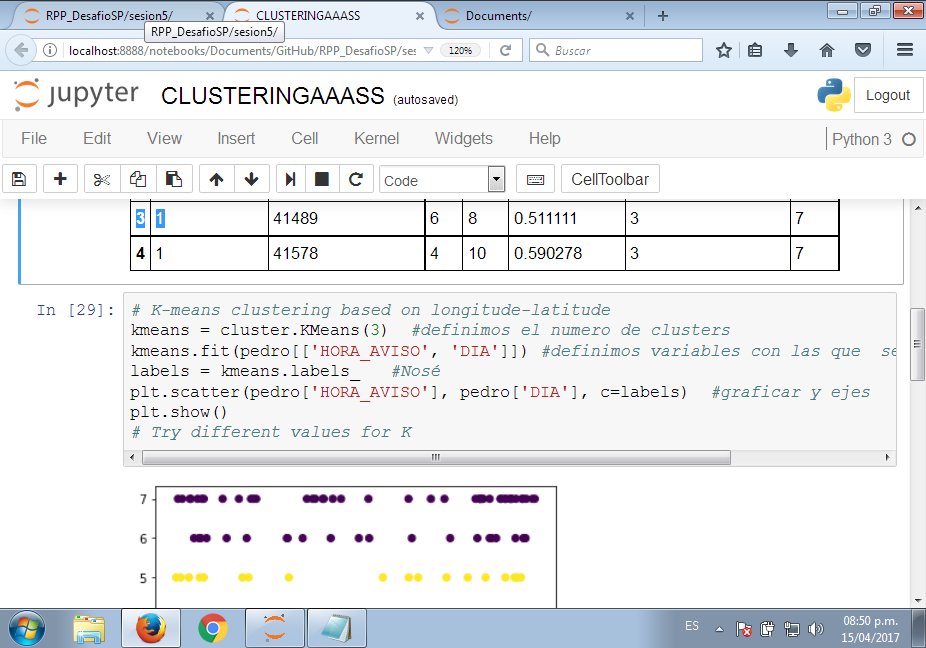
Haciendo los dos clustering el de jerarquía queda de esta manera, siendo erróneo, y no ayudando en nuestro proceso.

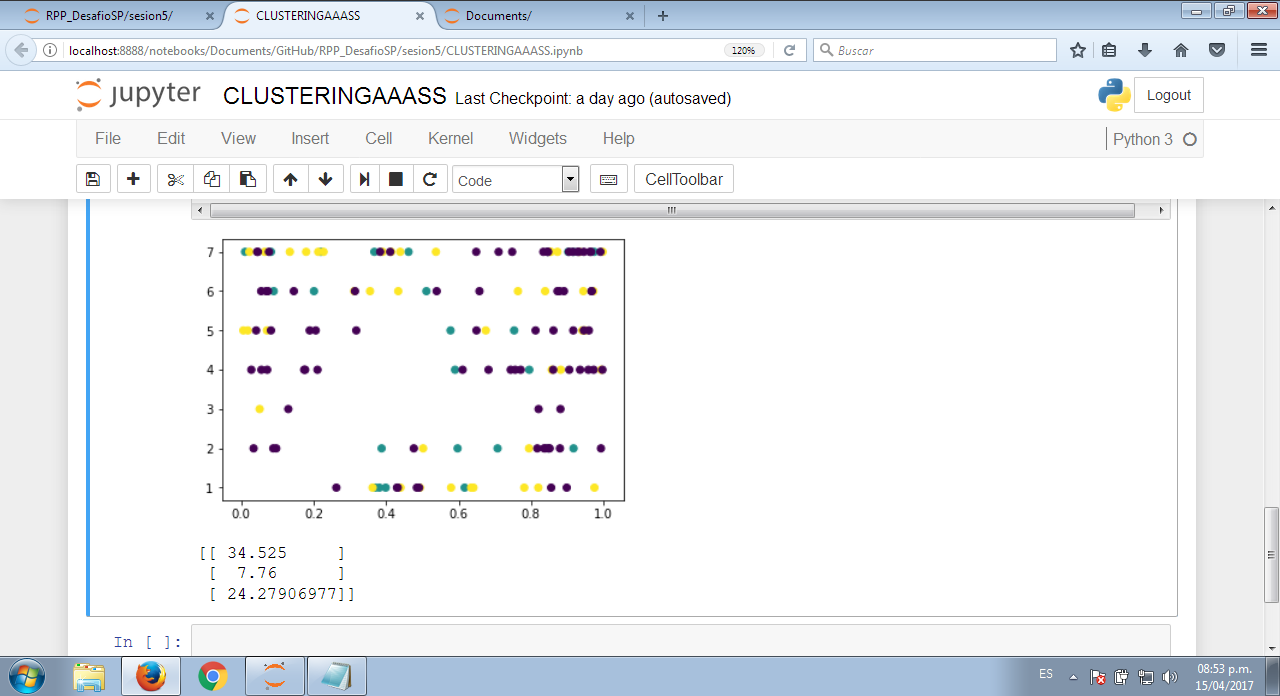




Por su parte el clustering con centroides queda de la siguiente manera.

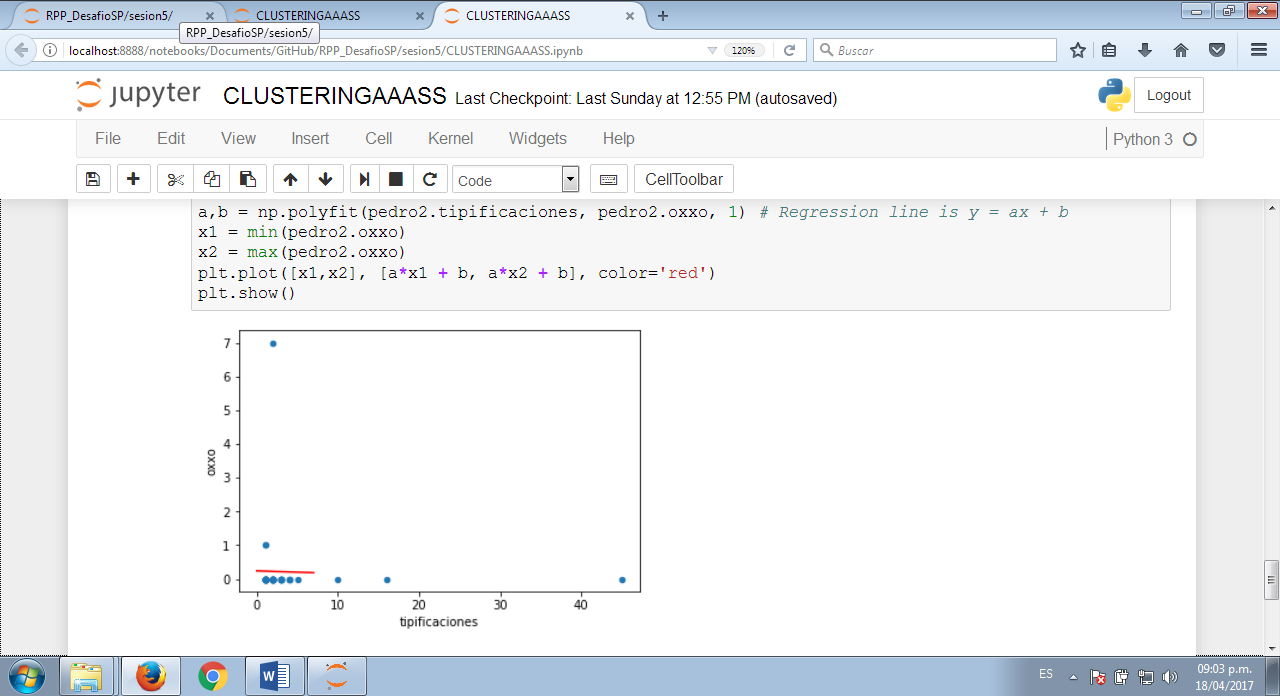


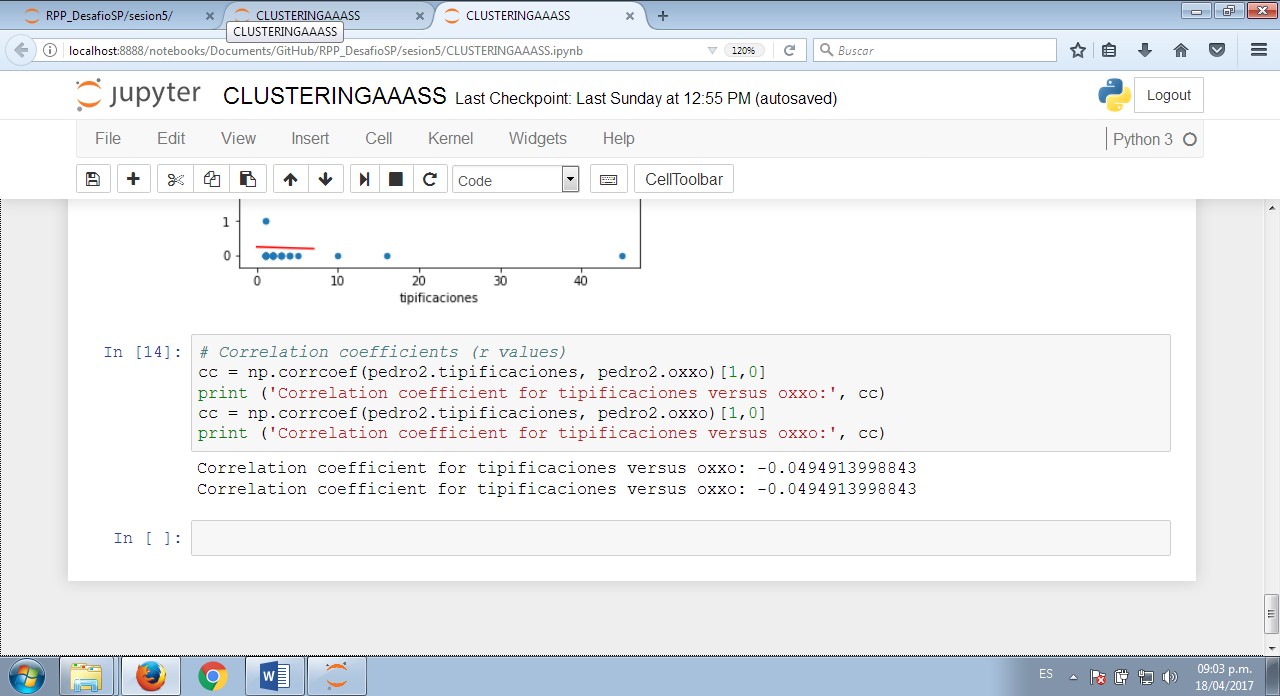




Obteniendo los centroides de los tres cluster que definimos.

**Regresión lineal.**





La pendiente de las variables son negativas y el valor de la correlación no supera el 0.30, por lo tanto no hay relación entre el desarrollo económico y el número de oxxos, puesto que la colonía con más oxxos también es san pedro 400, san pedro 400, revolución y vista montaña.

Recomendaciones:

Cerca de las 8 a 10 pm tener algunas patrullar cerca de san pedro 400.

También en los meses de agosto tener una disposición más amplia de patrullas rondando san pedro 400, revolución y vista montaña.