

Asignatura: Tipología y ciclo de vida de los datos

Ejercicio: Práctica 2

Alumno: Alejandro Belchí Hernández

Practica 2

ENLACE GITHUB: https://github.com/abelchih/practica_2.wiki.git

1. Descripción del dataset

<https://www.kaggle.com/philmohun/cryptocurrency-financial-data>

En el momento actual, donde la palabra criptomoneda se han puesto de moda, se presenta un dataset con las diferentes criptomonedas que se encuentran en el mercado actual (a la fecha de actualización del dataset), recolectado en CoinMarketCap.com, donde se nos muestra los movimientos en los mercados de cada una de ellas en un periodo de tiempo, indicando la fecha de apertura, precio de apertura, precio máximo, precio de cierre, volumen de transacciones realizadas en dicha fecha y la capitalización bursátil.

Con este dataset se pretende dar una primera aproximación a las cotizaciones de las diferentes criptomonedas y así entrar a analizar y actuar dentro de uno de los principales mercados bursátiles actuales.

2. Selección de datos de interés a analizar

En un primer momento nos quedamos con todos los datos obtenidos, que son:

- Currency name (e.g. bitcoin): nombre criptomoneda
- Date: fecha de la operación
- Open: valor de salida
- High: valor más alto
- Low: valor más bajo
- Close: Valor de cierre
- Volume: volumen de transacciones
- Marketcap: capitalización bursátil

Y hemos comprobado si los datos tienen algún caso vacío

```
> sapply(crypto, function(x) (sum(is.na(x))))  
Currency      Date      Open      High      Low      Close      Volume Market.Cap  
      0          0          0          0          0          0          0          0
```

Como podemos observar, no hay ningún campo vacío.

En principio no tenemos outlier, si hay datos que salen de la gráfica en cotizaciones se debe de averiguar el motivo pero no tratar dicho valor como outlier.

Asignatura: Tipología y ciclo de vida de los datos

Ejercicio: Práctica 2

Alumno: Alejandro Belchí Hernández

3. Análisis de datos

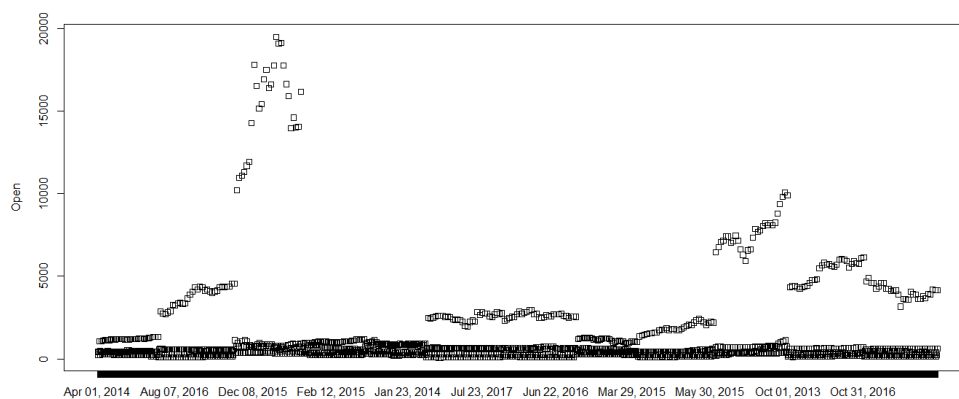
```
> summary(cripto)
      Currency      Date      Open      High      Low
bitcoin   : 1705   Dec 24, 2017: 200   Min.   : 0.000   Min.   : 0.000   Min.   : 0.000
litecoin  : 1705   Dec 25, 2017: 200   1st Qu.: 0.005   1st Qu.: 0.005   1st Qu.: 0.004
namecoin  : 1705   Dec 26, 2017: 200   Median : 0.093   Median : 0.103   Median : 0.085
peercoin  : 1704   Dec 27, 2017: 200   Mean    : 28.663   Mean    : 30.209   Mean    : 27.313
feathercoin: 1700 Dec 21, 2017: 199   3rd Qu.: 1.150   3rd Qu.: 1.270   3rd Qu.: 1.060
ripple    : 1607   Dec 22, 2017: 199   Max.    :19475.800   Max.    :20089.000   Max.    :18974.100
(Other)   :87130   (Other)   :96058
      Close      Volume      Market.Cap
Min.   : 0.000   -      : 1390   -      : 3629
1st Qu.: 0.005   1      : 435   451,600 : 197
Median : 0.094   0      : 381   6,951,590: 122
Mean    : 28.862   2      : 253   251,600 : 74
3rd Qu.: 1.160   4      : 211   1,451,600: 50
Max.    :19497.400 3      : 206   1,451,590: 41
      (Other):94380   (Other) :93143
.
```

En esta primera aproximación vemos algunas de las criptomonedas, que máximos hay en los valores de apertura, cierre, máximo y mínimos.

4. Representación de los resultados

Ahora mostraremos algunos datos mas detallados (como tenemos una gran cantidad de datos, para visualizar las estadísticas vamos a centrarnos en una única moneda, Bitcoin, por ser esta la más conocida)

Apertura de bitcoin

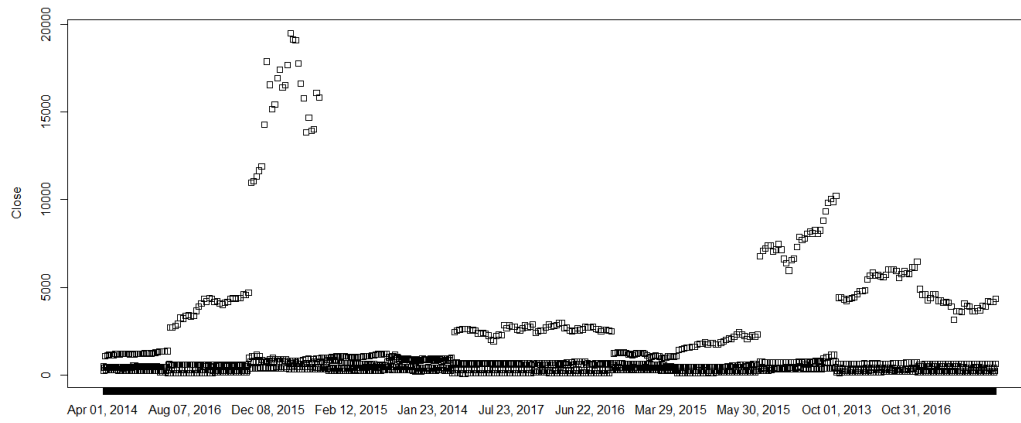


Asignatura: Tipología y ciclo de vida de los datos

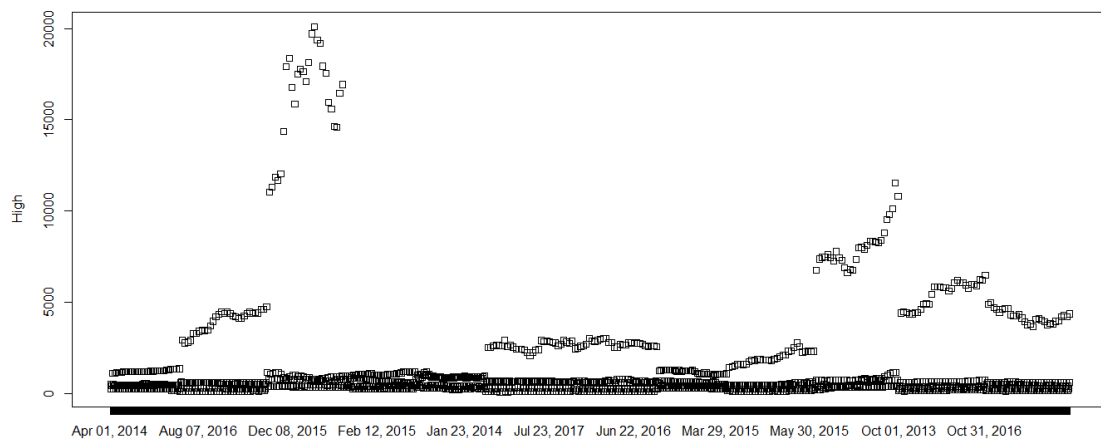
Ejercicio: Práctica 2

Alumno: Alejandro Belchí Hernández

Cierre de bitcoin



Precio más alto

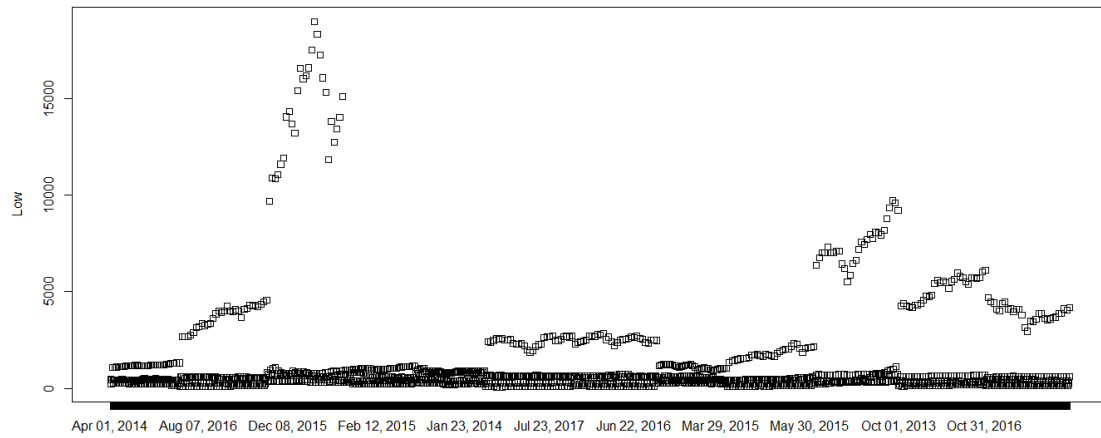


Precio mas bajo

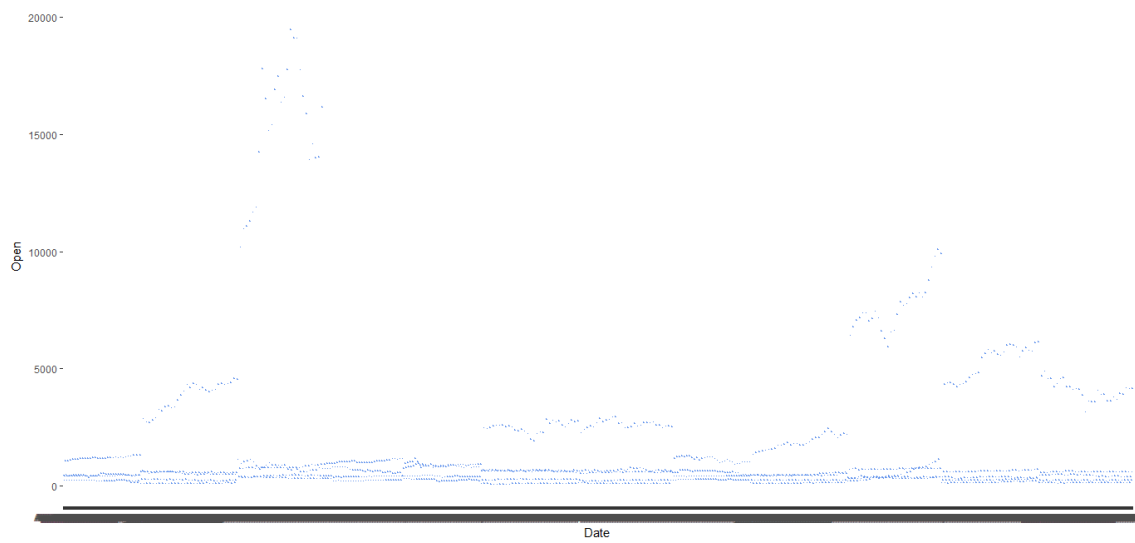
Asignatura: Tipología y ciclo de vida de los datos

Ejercicio: Práctica 2

Alumno: Alejandro Belchí Hernández



Vemos el comportamiento del precio de apertura del bitcoin

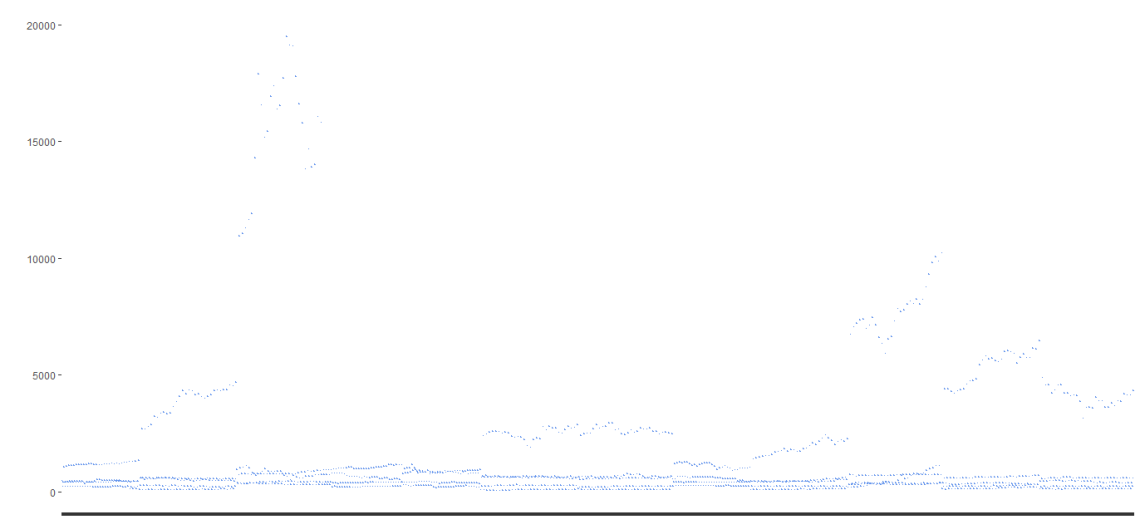


De cierre

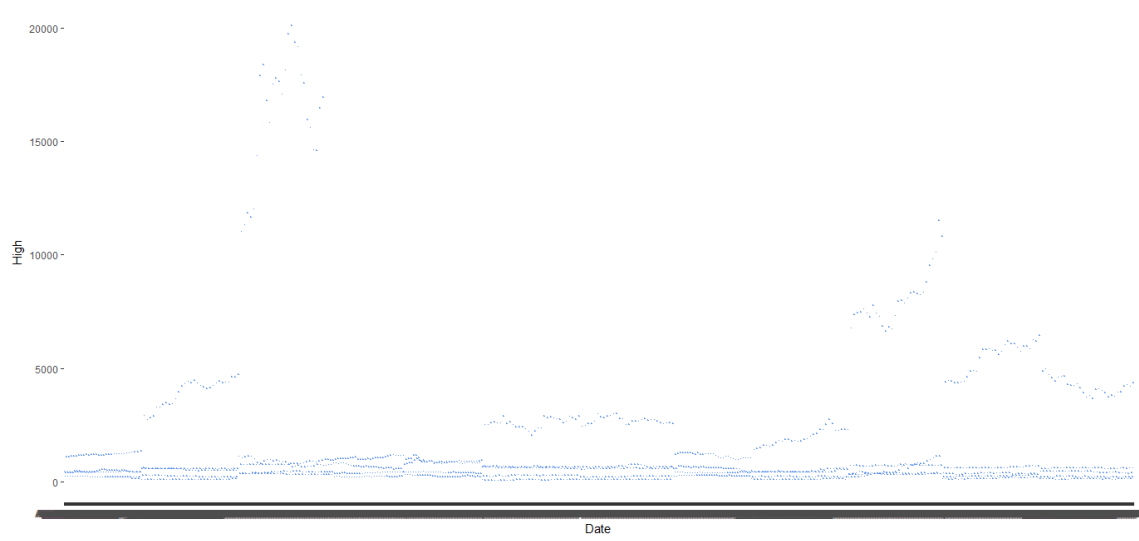
Asignatura: Tipología y ciclo de vida de los datos

Ejercicio: Práctica 2

Alumno: Alejandro Belchí Hernández



Valor más alto

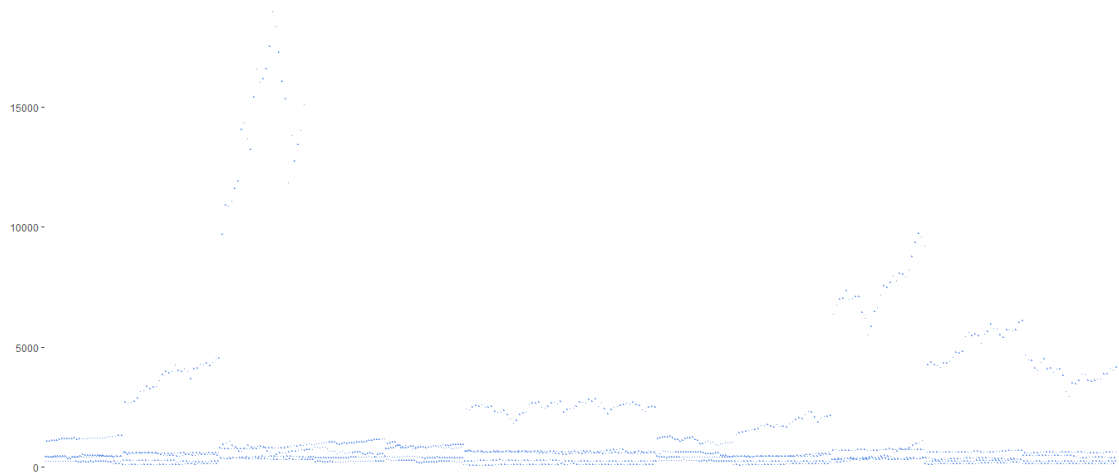


Asignatura: Tipología y ciclo de vida de los datos

Ejercicio: Práctica 2

Alumno: Alejandro Belchí Hernández

Valor más bajo



5. Resolución del problema

El dataset representado no concurría sobre ningún problema, simplemente era una primera aproximación a las cotizaciones de las criptomonedas.

Se han listado una serie de criptomonedas y los precios de cotización, esto también aporta la el presupuesto que tendríamos que tener para obtener una cantidad de criptomonedas y además nos permite analizar el comportamiento de las mismas respecto a las noticias ocurridas en un determinado día. Se pueden marcar los puntos donde se han producido alguna noticia importante, ataques hackers a las criptomonedas u otras empresas, cambio de presidencia en países como Estado Unidos, etc...