|  |  |  |
| --- | --- | --- |
| utn | Universidad Tecnológica Nacional  Facultad Regional Buenos Aires  Sintaxis y Semántica del Lenguaje.  Curso: K2054  Turno: Noche |  |

Trabajo Práctico Nro 1

Web Scraping Financiero

|  |  |
| --- | --- |
| **Nombre y Apellido** | **Legajo** |
| Juan Martín Torres | 1595829 |
| Facundo Mamani | 1674602 |
| Natalí Huamaní Cataño | 1531876 |

|  |  |
| --- | --- |
| **FECHA DE PRESENTACIÓN:** | 5/09/2019 |
| **FECHA DE DEVOLUCIÓN:** |  |
| **CALIFICACIÓN** |  |
| **FIRMA PROFESOR** |  |

## Informe:

### Consideraciones generales:

Para la implementación del TP se descargó previamente el HTML de la tabla de BOLSAR y se alojó en la raíz del proyecto en el archivo “lideres\_bcba.html”.

### Modelado:

Se crea la siguiente estructura para almacenar las acciones

typedef struct{

char nombre[255];

char vto[255];

int cantidadNominal;

double precioCompra;

double precioVenta;

int cantNominal2;

double ultimo;

double variacion;

double apertura;

double maximo;

double minimo;

double cierreAnt;

int volNominal;

int montoOperado;

int cantOperada;

char horaCotizacion[255];

} Accion;

Donde luego se define una variable acciones que va a consistir en un array de 20 posiciones del tipo de dato Accion.

### Parseo del archivo:

Para el parseo del archivo se implementó el método obtenerHTMLDeArchivo donde se realiza una lectura linea a linea del archivo buscando las líneas donde se encuentran las filas de la tabla y se ignora la fila de la tabla.

Por cada fila de cada tabla se realiza un parseo para obtener cada campo de la fila moviendo el cursor tantos caracteres como corresponda para ignorar los tags <td> y <tr>, luego estos campos se parsean como array de chars, entero o double según corresponda.

### Generación reporte CSV:

Se implementó un método para generar un reporte en formato CSV mediante el cual se recorre la estructura donde se almacenan las acciones y se generará un archivo donde la primera línea es el título de las columnas del reporte separadas con el carácter separador “;” y luego el resto de las filas se corresponden a cada acción donde también cada columna está separada con el mismo carácter.

Este archivo generado por este reporte se almacena en la raíz del proyecto.

### Generación reporte HTML:

Se implementó un método para generar un reporte en formato HTML donde en el cúal se genera un archivo con la extensión .html donde hay tanto un inicio de archivo como fin de archivo estáticos con la siguiente estructura

<table border = \"1\">

<tr><th>Especie</th><th>Variacion</th></tr>

*CONTENIDO DINÁMICO*

</table>

Donde este contenido dinámico es cada fila de las acciones y en caso de que se cumpla la condición de que el precio de compra y venta sean menores al precio de apertura se agrega en cada tag <tr> el siguiente atributo “style=color:red;” que define que el color del texto de esta fila sea color rojo, caso contrario no se agrega nada y el color de la fuente es el negro por defecto.