1. A partir del fichero access.log que se corresponde con un log del servidor web Apache.
   a. Descubre el n.º de líneas del fichero.
   b. Revisa la primera línea y compara su estructura con la siguiente tabla.

%h	**%l**	**%u**	**%t**	**"%r"**	**%>s**	**%b**	**"%{Referer}i"**	**"%{User-agent}i"**
IP cliente	-	usuario	fecha petición	Request	Status	tamaño respuesta sin cabeceras	Referer	UserAgent

   c. Ver IPs de los clientes que hacen las solicitudes, sin duplicados.
   d. Ver IPs de clientes y el nº de peticiones que hacen ordenadas de mayor a menor. Mostrar únicamente los 15 primeros.
   e. Ver peticiones del equipo más activo.
   f. Ver peticiones del equipo más activo; pero en vez de mostrar toda la línea, que se vea exclusivamente la fecha, la *Request* y el *Status* (código de respuesta).
   g. Ver las peticiones realizadas por equipos de la red 192 (IPs del tipo 192.x.x.x) --> investiga como indicar el inicio de línea en una expresión regular.
   h. Ver códigos de respuesta (*Status*) → nº de veces y ordenados de mayor a menor.
   i. Ver líneas asociadas al código de respuesta 404.
   j. Ver los User-agent y cuántas peticiones se ejecutan con cada uno de ellos ordenados de mayor a menor.
   k. Buscar las peticiones realizadas por el agente Twitterbot/1.0.
   l. Ver las peticiones donde en la Request aparece la palabra *dhcp.*
   m.
2. Las técnicas usadas para analizar logs se pueden emplear para otro tipo de archivos. Por ejemplo, el fichero 24022007.txt.zip contiene información de flujos de red:
   a. Descomprime el fichero 24022007.txt.zip.
   b. Descubre el n.º de líneas del fichero.
   c. Revisa las primeras 5 líneas y analiza su estructura.
   d. Guarda en fichero de nombre ip.txt un listado con las direcciones IP origen de flujos udp sin duplicados.
   e. Guarda en fichero de nombre ip_priv.txt un listado con las direcciones IP origen de flujos udp que sean de clase A privadas (direcciones IP del tipo 10.x.x.x) sin duplicados.
   f. Guarda en fichero de nombre ip_pub.txt un listado con las direcciones IP origen de flujos udp que NO sean de clase A privadas (direcciones IP del tipo 10.x.x.x) sin duplicados.
   g. Ver IPs de origen y el nº de peticiones tcp que hacen ordenadas de mayor a menor. Mostrar únicamente los 15 primeros.
3. Las técnicas usadas para analizar logs se pueden emplear para la salida de otros comandos. Por ejemplo, tshark es la versión en línea de comandos de Wireshark y permite analizar ficheros de captura. Usando el fichero lab-01.pcap del boletín de filtros de visualización:
   a. Ejecuta el siguiente comando para ver las Estadísticas de Endpoints de IPv4:

```
tshark -q -n -r lab-01.pcap -z endpoints,ip
```

      - `-q` → no muestra los paquetes de la captura
      - `-n` → no hace resolución DNS

- `-r lab-01.pcap` → lee el fichero lab-01.pcap
- `-z endpoints,ip` → muestra las estadísticas de Endpoints IPv4

b. Ejecuta el siguiente comando

```
tshark -q -n -r lab-01.pcap -Y http.host -E header=y -T fields -e ip.src -e http.host
```

- `-Y http.host` → filtro de visualización para quedarse con los paquetes http que tengan la cabecera Host.
- `-E header=y` → muestra un encabezado con el nombre de los campos que se indiquen a continuación
- `-T fields -e ip.src -e http.host` → muestra el valor de los campos IP origen y la cabecera Host de los mensajes http.

c. La salida del comando anterior se puede procesar como si fuese un fichero. Trata de adivinar que se obtendrá de la ejecución de un comando como el anterior pero con algunos cambios:

```
tshark -q -n -r lab-01.pcap -Y http.host -T fields -e ip.src -e http.host | sort | uniq -c | sort -nr
```

- se quita `-E header=y`
- se añade `| sort | uniq -c | sort -nr`