Práctica Data Driven Security

Grupo G: Yaneth Gonzalez, Miriam Jiménez, Anabel Vilchez y Alejandro Moreno

2018-06-03

1. Introducción

El propósito de la práctica es evidenciar las posibles vulnerabilidades industriales que pueden darse en diferentes países. La pregunta que intentamos responder es si **existe algun país más propenso a tener vulnerabilidades industriales**. Para este análisis se utilizaran diferentes fuentes de datos que serán tratados con R.

La fase más importante del análisis es la búsqueda en Shodan (http://www.shodan.io) de todos los dispositivos industriales públicos en Internet los cuales los cruzaremos con las vulnerabilidades que tengan este tipo de modelos de dispositivos industriales públicos en la Common Vulnerabilities and Exposures (https://cve.mitre.org). Para la extración de datos en Shodan utilizaremos el package publico en Github de hrbrmstr (https://github.com/hrbrmstr/shodan)

Shodan es un motor de búsquedas que ofrece información sobre equipos conectados a Internet ya sean routers, servidores, cámaras de videovigilancia, maquinaria industrial o cualquier otro tipo de dispositivo que se encuentre conectado. Permite buscar información masiva de los dispositivos y hacer un filtrado según los intereses que tengamos. Para la realización de esta práctica, vamos a hacer una búsqueda por fabricante y producto de dispositivos industriales

Los repositorios utilizados para los cálculos y descarga de las fuentes de datos se pueden encontrar en el siguiente reposiorio de Github de esta práctica. En este repositorio se encuentra: el script en R para descargar los datos de Shodan, el script para los cálculos intermedios de análisis y los packages dedicados para la práctica.

- https://github.com/amperis/upc-master-datadriven (https://github.com/amperis/upc-master-datadriven)

En el siguiente repositorio se encuentra el informe en R de la documentación de la práctica:

- https://github.com/amperis/upc-master-datadriven-informe (https://github.com/amperis/upc-master-datadriven-informe)

2. Descripción de los modelos de datos

Las fuentes de datos que utilizaremos en la práctica serán las siguientes:

1. Fuente de datos con dispositivos industriales según fabricantes.
2. Fuente de datos Shodan de dispositivos industriales accesibles desde Internet.
3. Fuente de datos de CVE.

A continuación se describen las características de cada una de las fuentes de datos.

2.1 Fuente de datos con dispositivos industriales según fabricantes

Esta fuente de datos es un archivo CSV generado manualmente a través de los diferentes dispositivos industriales más comunes actualmente. Nos hemos basado en diferente documentacion técnica disponible en Internet así como en la documentación impartida en el máster. Un ejemplo de esta información se puede encontrar en la siguiente URL:

- https://cybersecuritylaboratory.wordpress.com/2016/10/05/utilizando-shodan-para-encontrar-sistemas-de-control-industrial (https://cybersecuritylaboratory.wordpress.com/2016/10/05/utilizando-shodan-para-encontrar-sistemas-de-control-industrial)

Esta lista debe actualizarse a medida que se conozcan más fabricantes. El encabezado de dicha tabla es el siguiente:

Campo	Descripción
fabricante	Campo de tipo string que contiene el nombre del fabricante del dispositivo industrial.
producto	Campo de tipo string que contiene el modelo del producto industrial.
cadena_busqueda	String de búsqueda en Shodan.

A continuación se muestran los 15 primeros registros de este CSV:

Tabla 1 - Lista de dispositivos industriales

fabricante	producto	cadena_busqueda
Philips	Philips Hue Lighing	http.title:“hue personal wireless lighting”
Schneider Electric	PowerLogic ECC	Schneider Electric ECC21
Schneider Electric	PowerLogic EGX	Schneider Electric EGX100MG
Schneider Electric	PowerLogic ION	Power Measurement Ltd ION8650

Schneider Electric	PowerLogic ION	Meter ION
Schneider Electric	PowerLogic PM	Schneider Electric PM820SD
Schneider Electric	PowerLogic PM	PowerLogic PM800
Schneider Electric	PowerLogic PM	PowerLogic PM800
Schneider Electric	PowerLogic PM1	Schneider Electric PM870SD
General Electric	Proficy	ProficyPortal
Reliance	Reliance 4	Reliance 4 Control Server
Elster EnergyICT	RTU	EnergyICT RTU
ABB	RTU500	ABB RTU560
Siemens	Scalance S	Scalance S
Siemens	Simatic HMI	SIMATIC HMI

Actualmente se disponen de un total de 85 dispositivos industriales.

La lista completa de dispositivos industriales (https://github.com/amperis/upc-master-datadriven/blob/master/R/SHODAN/FUENTESDATOS/ORIG/dispositivos_industriales.csv) puede econtrarse en el repositoriode GitHub de la practica.

2.2 Fuente de datos Shodan de dispositivos industriales accesibles desde Internet

Esta fuente de datos es un archivo de datos CVE obtenido a través de un proceso interativo programado en R desde el cual por cada dispositivo industrial obtenido del punto anterior es buscado en Shodan. Esta búsqueda nos devuelve diferente información del dispositivo industrial encontrado y público en Internet. La extración de datos de Shodan pueden encontrarse en el script de R de la practica (https://github.com/amperis/upc-master-datadriven/tree/master/R/SHODAN).

Las cabeceras obtenidas de la tabla son las siguientes:

Campo	Descripción
v_fabricante	String con el nombre del fabricante del dispositovo industrial.
v_producto	String con el modelo del producto del dispositivo industrial.
v_cadena_busqueda	String con la cadena de búsqueda Shodan que se ha utilizado para encontrar este dispositivo.
aux.matches.ip_str	String con la IP pública del dispositivo en formato ICANN de 32 bits.
aux.matches.isp	String con el nombre comercial del ISP (Internet Service Provider) de dicha dirección IP.
aux.matches.location.city	String de la ciudad donde está ubicada fisicamente dicha IP.
aux.matches.location.longitude	String con la coordenada de longitud geográfica.
aux.matches.location.latitude	String con la coordenada de latitud geográfica.
aux.matches.location.country_name	String con el país.
aux.matches.location.country_code3	String con el país en formato ISO-3166.1.
aux.matches.os.au	String con el identificador del Sistema Autónomo al que pertenece la IP.

A continuación se muestran los 10 primeros regisotros de este CSV de consultas Shodan:

Tabla 2 - Listado de dispositivos industriales encontrados en Shodan

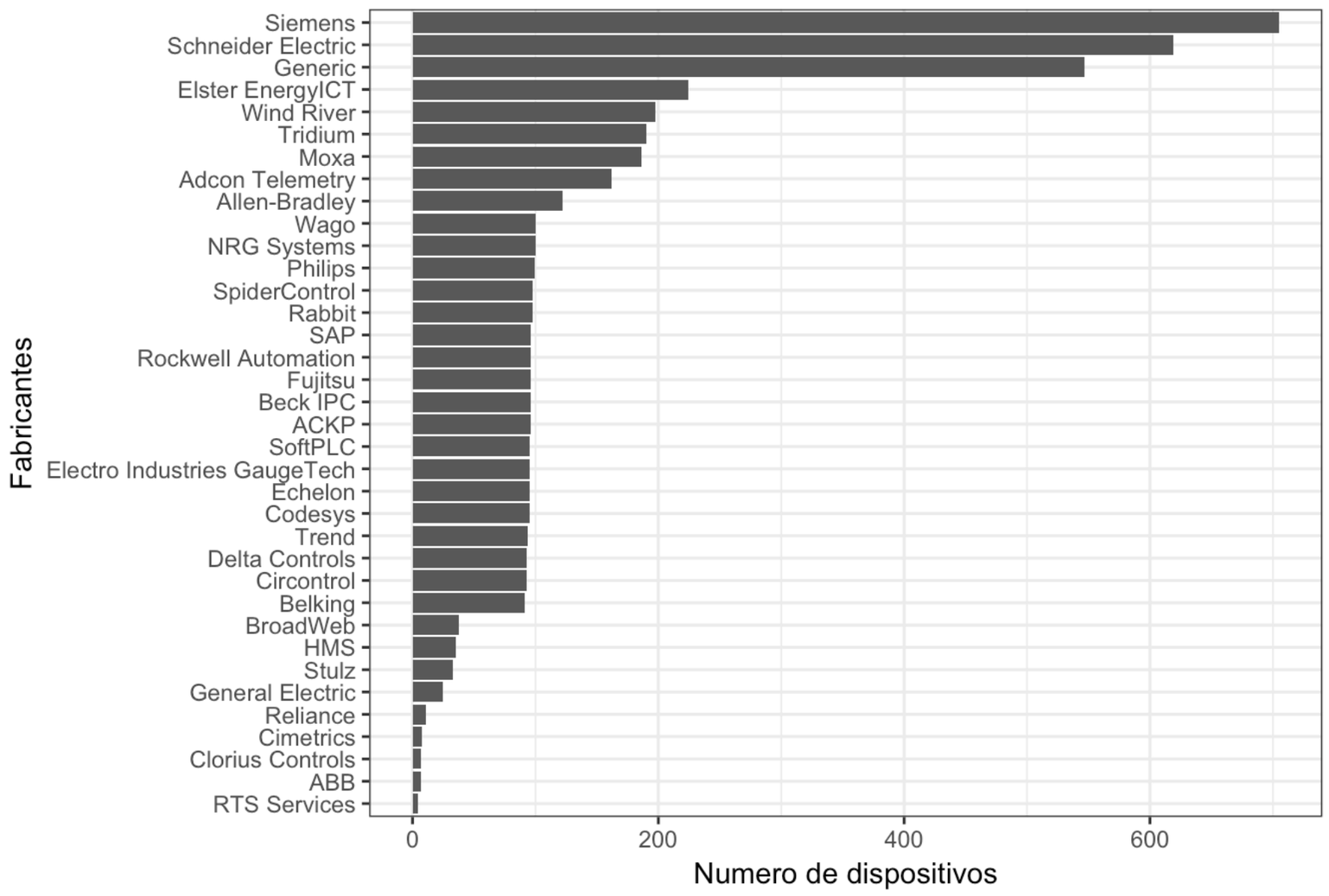
v_fabricante	v_producto	v_cadena_busqueda	aux.matches.ip_str	aux.matches.isp	aux.matches.location.city	aux.matches.location.long
Philips	Philips Hue Lighing	http.title:hue personal wireless lighting	82.197.175.26	Init7 (Switzerland) Ltd.	Zurich	8
Philips	Philips Hue Lighing	http.title:hue personal wireless lighting	188.2.21.180	Serbia BroadBand-Srpske Kablovske mreze d.o.o.	Novi Sad	19
Philips	Philips Hue Lighing	http.title:hue personal wireless lighting	213.65.61.145	Telia Company	Falun	15
Philips	Philips Hue Lighing	http.title:hue personal wireless lighting	145.14.221.218	Glattwerk AG	Wallisellen	8
Philips	Philips Hue Lighing	http.title:hue personal wireless lighting	99.154.9.16	AT&T U-verse	NA	-97
Philips	Philips Hue Lighing	http.title:hue personal wireless lighting	66.74.181.239	Time Warner Cable	Thousand Oaks	-118
Philips	Philips Hue Lighing	http.title:hue personal wireless lighting	160.3.37.217	Cable One	Pampa	-100

Philips	Philips Hue Lighing	http.title:hue personal wireless lighting	87.100.253.26	DNA Oyj	Helsinki	24
Philips	Philips Hue Lighing	http.title:hue personal wireless lighting	72.182.192.32	Time Warner Cable	Irving	-96
Philips	Philips Hue Lighing	http.title:hue personal wireless lighting	82.169.170.245	KPN	Breda	4

Actualmente se disponen de un total de 4847 dispositivos industriales.

La siguiente figura muestra el reparto de dispositivos industriales encontrados. Como puede verse la mayoria de dispositivos industriales econtrados y publicos en Internet son Siemens y Schneider Electric.

Figura 1 - Reparto de dispositivos industriales encontrados



2.3 Fuente de datos de CVE

Para la extración de datos de la Common Vulnerabilities and Exposures (https://cve.mitre.org) hemos utilizado el packager publico en Github de la r-net.tools (https://github.com/r-net-tools/security.datasets/raw/master/net.security/sysdata.rda). Esta fuente de datos contiene informacion de vulnerabilidades así como parches o soluciones aportadas por fabricantes y desarrolladores.

A continuación se describen los campos de esta tabla:

Campo	Descripción
cve.id	Identificador del CVE
affects	Versiones afectas
problem.type	Tipo de problema
references	Informacion extra sobre el CVE
description	Breve descripción de la vulnerabilidad o exposición de seguridad.
vulnerable.configuration	
cvss3.vector	Vector de ataque utilizado
cvss3.score.exploit	Puntuacion del exploit
cvss3.score.impact	Puntuacion del impacto
cvss2.vector	Vector de ataque utilizado
published.date	Fecha de publicación del CVE.
last.modified	Fecha de modificación del CVE.

A continuación se muestran los 3 primeros registros de CVE:

Tabla 3 - Listado de vulnerabilidads CVE

cve.id	affects	problem.type	references
CVE-1999-0001	[{“vendor_name”:“bsdi”,“product”: {“product_data”: [{“product_name”:“bsd_os”,“version”: {“version_data”: [{“version_value”:“3.1”}]}}]}}, {“vendor_name”:“freebsd”,“product”: {“product_data”:	[“CWE-20”]	[{“url”:“http://www.openbsd.org/errata23.html#tcpfix (http://www.openbsd.org/errata23.html#tcpfix)”,“name”:“http://www.openbsd.org/errata2 (http://www.openbsd.org/errata23.html#tcpfix)”,“refsource”:“CONFIRM”}, {“url”:“http://www.osvdb.org/5707 (http://www.osvdb.org/5707)”,“name”:“5707”,“refsource

	[{“product_name”:“freebsd”,“version”: {“version_data”:[{“version_value”:“1.0”}, {“version_value”:“1.1”}, {“version_value”:“1.1.5.1”}, {“version_value”:“1.2”}, {“version_value”:“2.0”}, {“version_value”:“2.0.1”}, {“version_value”:“2.0.5”}, {“version_value”:“2.1.5”}, {“version_value”:“2.1.6”}, {“version_value”:“2.1.6.1”}, {“version_value”:“2.1.7”}, {“version_value”:“2.1.7.1”}, {“version_value”:“2.2”}, {“version_value”:“2.2.2”}, {“version_value”:“2.2.3”}, {“version_value”:“2.2.4”}, {“version_value”:“2.2.5”}, {“version_value”:“2.2.6”}, {“version_value”:“2.2.8”}, {“version_value”:“3.0”}]}}]}}, {“vendor_name”:“openbsd”,“product”: {“product_data”: [{“product_name”:“openbsd”,“version”: {“version_data”:[{“version_value”:“2.3”}, {“version_value”:“2.4”}]}}]}}]		
CVE-1999-0002	[{“vendor_name”:“bsdi”,“product”: {“product_data”: [{“product_name”:“bsd_os”,“version”: {“version_data”: [{“version_value”:“1.1”}]}}]}}, {“vendor_name”:“caldera”,“product”: {“product_data”: [{“product_name”:“openlinux”,“version”: {“version_data”: [{“version_value”:“1.2”}]}}]}}, {“vendor_name”:“redhat”,“product”: {“product_data”: [{“product_name”:“linux”,“version”: {“version_data”:[{“version_value”:“2.0”}, {“version_value”:“2.1”}, {“version_value”:“3.0.3”}, {“version_value”:“4.0”}, {“version_value”:“4.1”}, {“version_value”:“4.2”}, {“version_value”:“5.0”}, {“version_value”:“5.1”}]}}]}}]	[“CWE-119”]	[{“url”:“ftp://patches.sgi.com/support/free/security/advisories/19981006-01-I (ftp://patches.sgi.com/support/free/security/advisories/19981006-01-I)”,“name”:“19981006-I”,“refsource”:“SGI”},{“url”:“http://www.ciac.org/ciac/bulletins/j-006.shtml (http://www.ciac.org/ciac/bulletins/j-006.shtml)”,“name”:“J-006”,“refsource”:“CIAC”}, {“url”:“http://www.securityfocus.com/bid/121 (http://www.securityfocus.com/bid/121)”,“name”:“121”,“refsource”:“BID”}]
CVE-1999-0003	[{“vendor_name”:“tritreal”,“product”: {“product_data”: [{“product_name”:“ted_cde”,“version”: {“version_data”: [{“version_value”:“4.3”}]}}]}}, {“vendor_name”:“hp”,“product”: {“product_data”:[{“product_name”:“hp-ux”,“version”:{“version_data”: [{“version_value”:“10.01”}, {“version_value”:“10.02”}, {“version_value”:“10.03”}, {“version_value”:“11.00”}]}}]}}, {“vendor_name”:“ibm”,“product”: {“product_data”: [{“product_name”:“aix”,“version”: {“version_data”:[{“version_value”:“4.1”}, {“version_value”:“4.1.1”}, {“version_value”:“4.1.2”}, {“version_value”:“4.1.3”}, {“version_value”:“4.1.4”}, {“version_value”:“4.1.5”}, {“version_value”:“4.2”}, {“version_value”:“4.2.1”}, {“version_value”:“4.3”}]}}]}}, {“vendor_name”:“sgi”,“product”: {“product_data”: [{“product_name”:“irix”,“version”: {“version_data”:[{“version_value”:“5.2”}, {“version_value”:“5.3”}, {“version_value”:“6.0”},	[“NVD-CWE-Other”]	[{“url”:“ftp://patches.sgi.com/support/free/security/advisories/19981101-01-A (ftp://patches.sgi.com/support/free/security/advisories/19981101-01-A)”,“name”:“19981101-A”,“refsource”:“SGI”},{“url”:“ftp://patches.sgi.com/support/free/security/advisories/19981101-01-PX (ftp://patches.sgi.com/support/free/security/advisories/19981101-01-PX)”,“name”:“19981101-PX”,“refsource”:“SGI”},{“url”:“http://www.securityfocus.com/bid/122 (http://www.securityfocus.com/bid/122)”,“name”:“122”,“refsource”:“BID”}]

{“version_value”:“6.1”},
{“version_value”:“6.2”},
{“version_value”:“6.3”},
{“version_value”:“6.4”}]}}]}},
{“vendor_name”:“sun”,“product”:
{“product_data”:
[{“product_name”:“solaris”,“version”:
{“version_data”:[{“version_value”:“1.1”},
{“version_value”:“1.2”},
{“version_value”:“2.0”},
{“version_value”:“2.1”},
{“version_value”:“2.2”},
{“version_value”:“2.3”},
{“version_value”:“2.4”},
{“version_value”:“2.5”},
{“version_value”:“2.5.1”},
{“version_value”:“2.6”}]}}]}}]

Actualmente se encuentran un total de 106717 vulnerabilidades.

3. Objetivos del análisis de datos

Los objetivos de esta práctica son:

- Conocer y mantenernos informados sobre los países que poseen riesgos asociados a sus dispositivos industriales, por vulnerabilidades detectadas y reportadas.
- Diferenciar visualmente entre los países que poseen más dispositivos instalados versus los países que poseen más dispositivos con vulnerabilidad.
- Detectar cuál es el fabricante con mayor cantidad de vulnerabilidades reportadas en CVE.

La metodología utilizada para los cálculos que se presentarán a continuación es el siguiente:

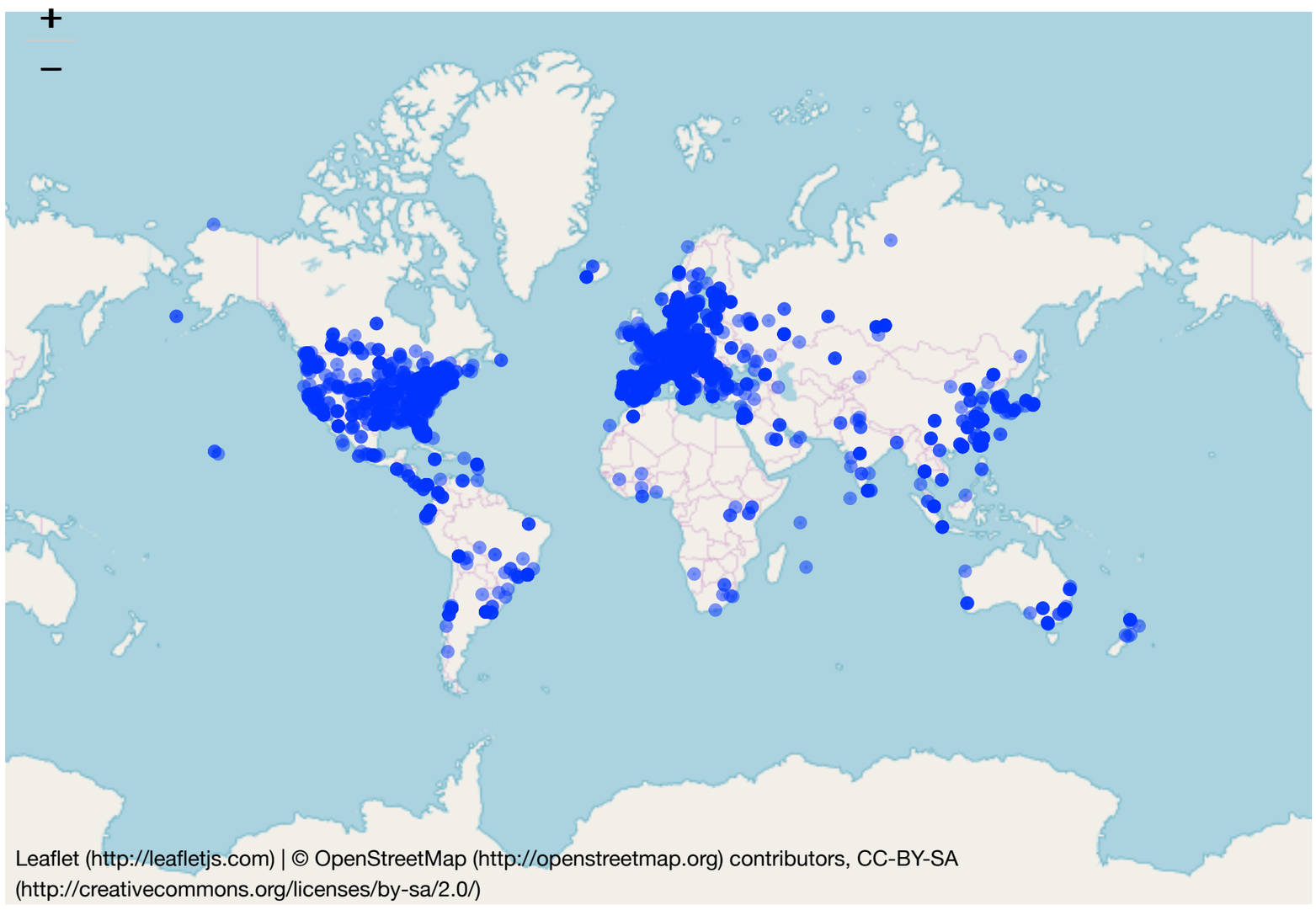
1. Dado un listado CSV de dispositivos industriales visto en el punto 2.1 de esta práctica se realiza un proceso interativo de busqueda en Shodan.
2. El resultado de la busqueda en Shodan es un nuevo listado CSV de dispositivos industriales publicados en Internet tal como puede verse en el punto 2.2 de la practica.
3. Con el listado de CSV de Shodan se buscan las vulnerabilidades CVE que contengan cada uno de estos dispositivos en cada uno de sus paises. Esta información se sumariza para obtener cuales son los paises que más vulnerabilidades tienen.

4. Análisis de datos

A continuación se analizan los países que más dispositivos industriales públicos en Internet se han encontrado a través de Shodan a nivel mundial.

Utilizamos un mapa de los puntos GPS de dichos dispositivos industriales. Aparentemente los paises que más dispositivos industriales utilizandon son EEUU y Europa:

Figura 2 - Posicion GPS de los dispositivos industriales

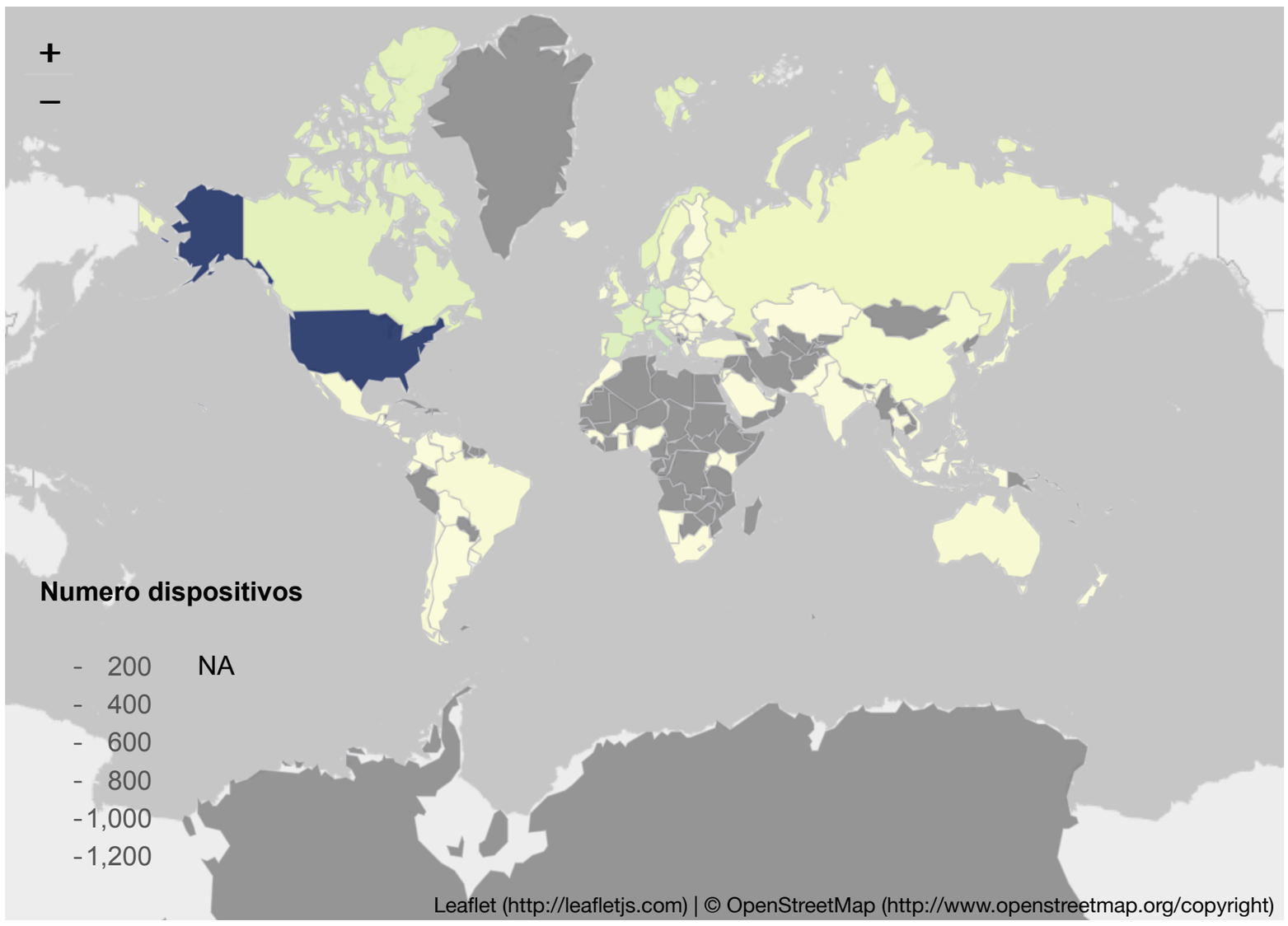


Se analiza a través de un mapa de calor dicha densidad de puntos para estar seguros. Revisando los datos CSV aparece una casuística que explica la diferencia de los gráficos, la cuál es que muchos dispositivos tienen asociada la misma latitud y longitud, por lo que los puntos se sobreponen una encima de otro tantas veces como se encuentren. Esto es así porque dichos dispositivos industriales pertenecen al mismo ISP

(Internet Service Provider)

Figura 3 - Mapa de calor de los dispostivos industriales

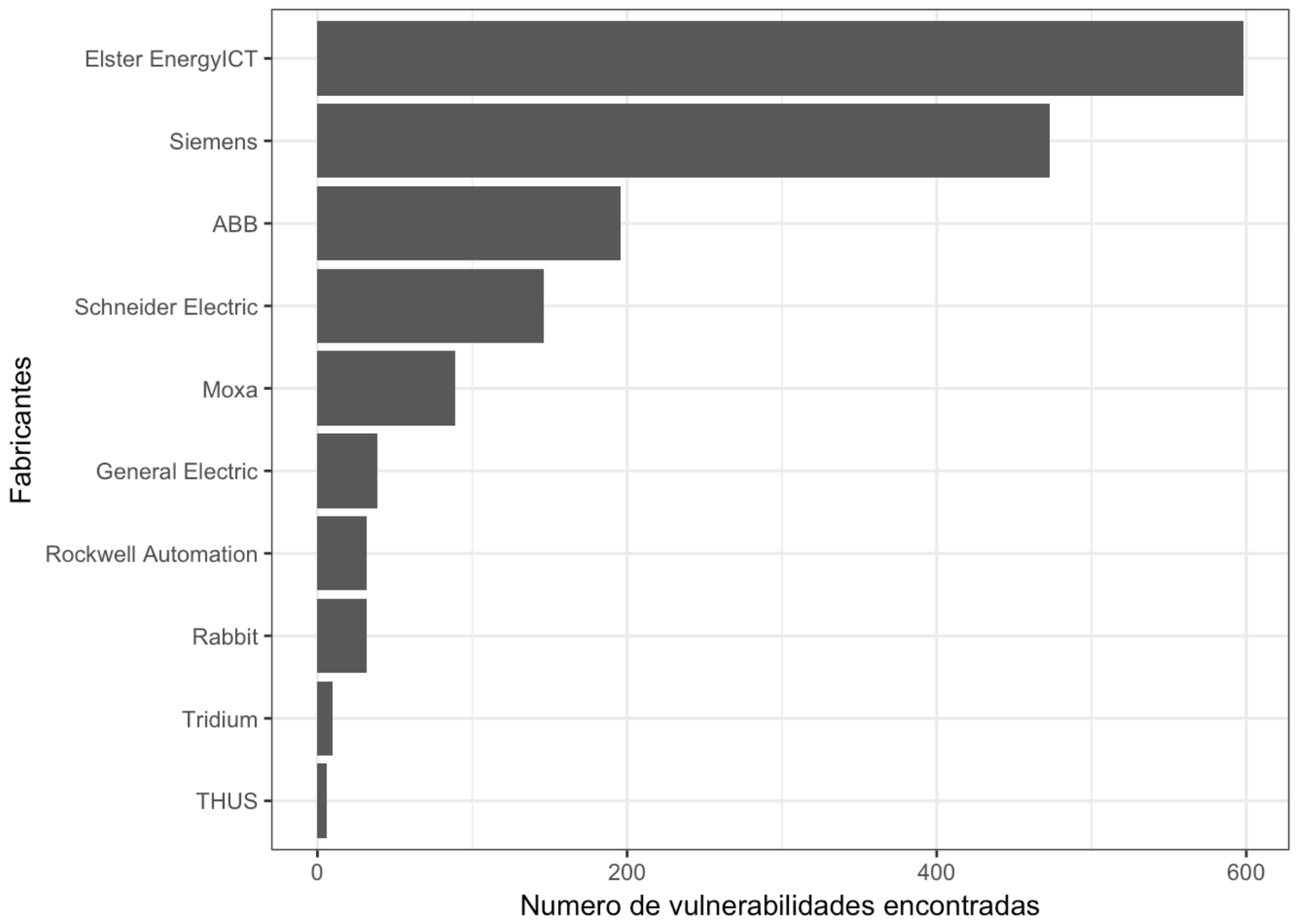
```
#> Joining, by = "id"
```



Analizamos cuales son los 10 fabricantes que tienen asociados más vulnerabilidades, para obtener esta relación se ha tomado la relación de los dispositivos resultados en la fuente Shodan por fabricante y producto y se ha realizado una sumarizacion por cada una de las coincidencias obtenidas en la fuente de vulnerabilidades.

Como conclusión podemos afirmar que el fabricante que más dispositivos tiene con vulnerabilidades públicas en CVE es Elster Energy.

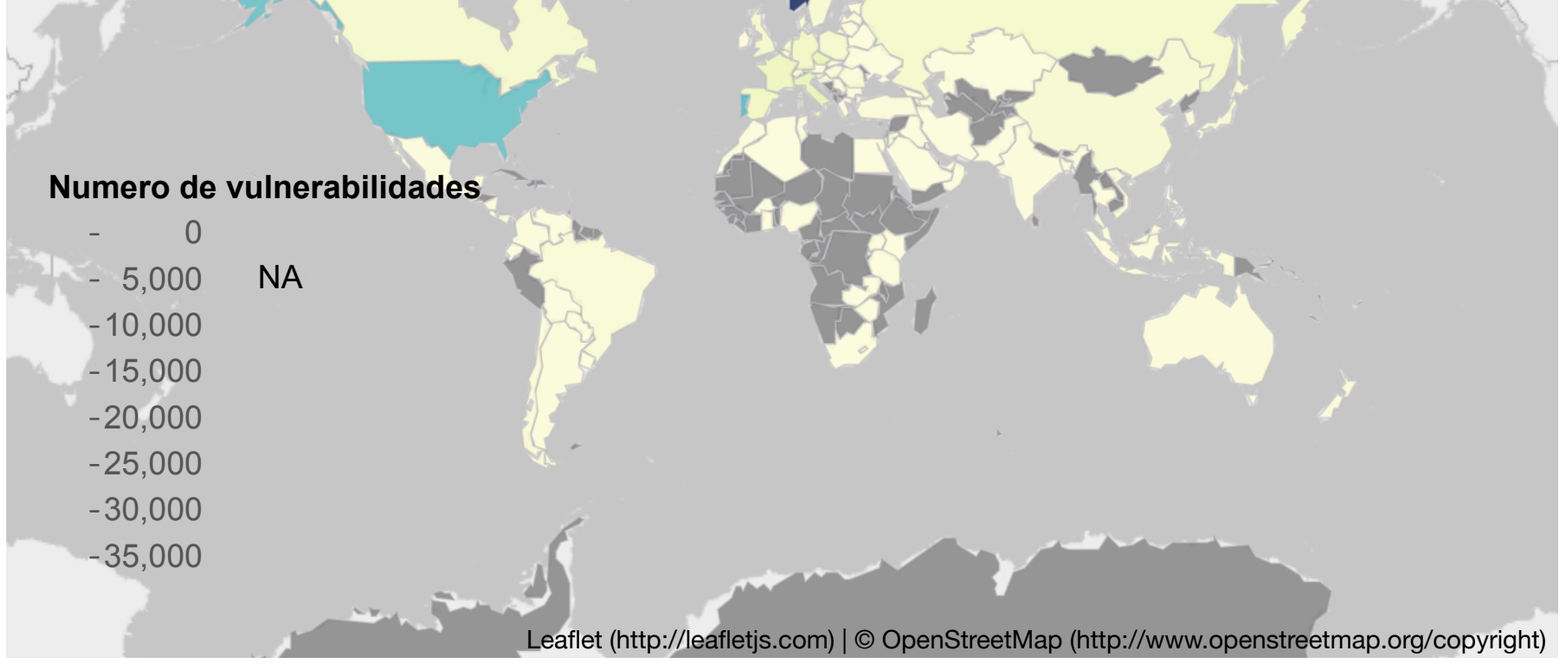
Figura 4 - Vulnerabilidades por fabricante



El siguiente paso es visualizar los paises que más vulneratilidades tienen en función de todos los dispostivos industriles encontrados. Se analiza en un mapa de calor los paises con más vulnerabilidades.

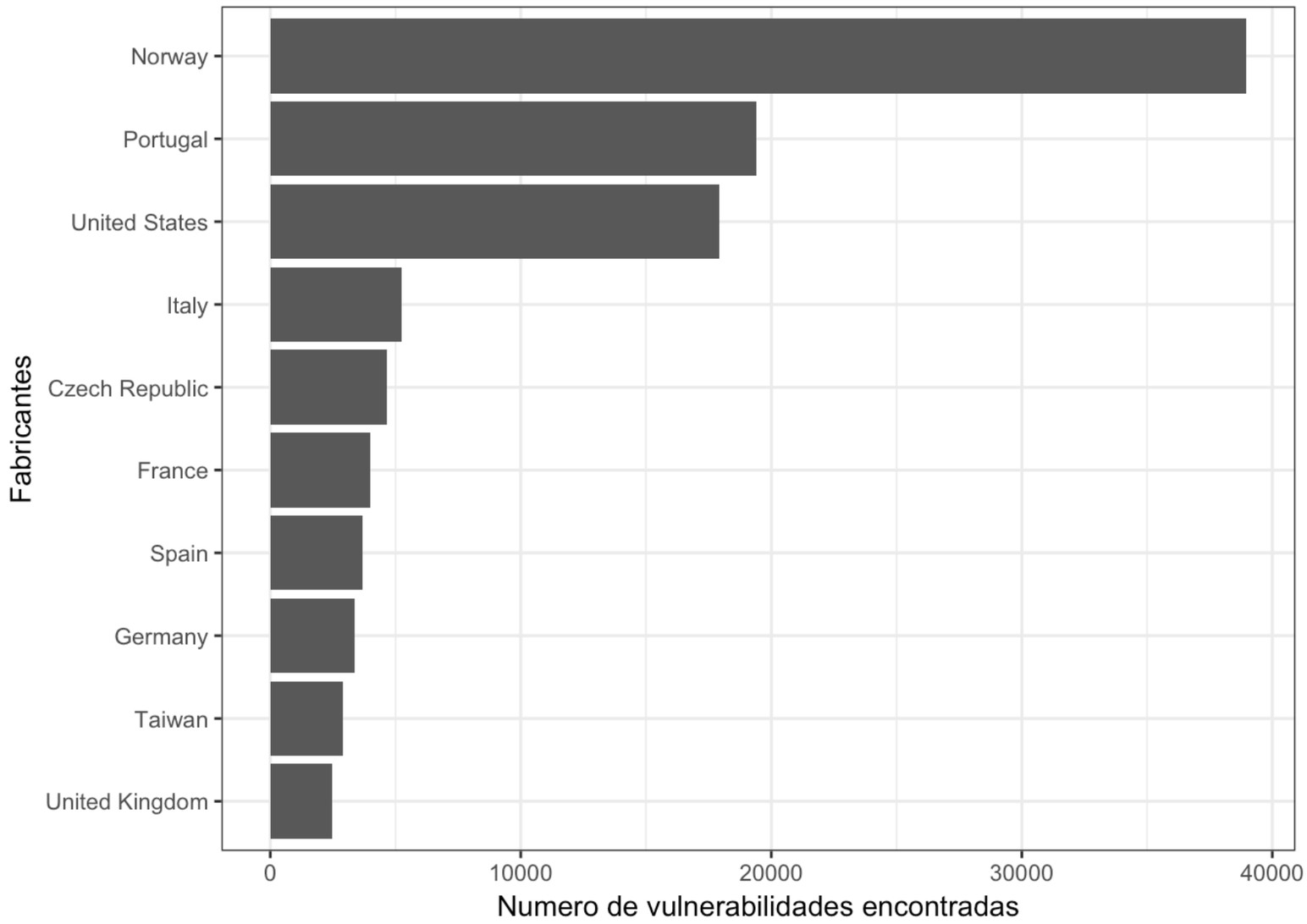
Figura 5 - Vulnerabilidades por paises

```
#> Joining, by = "id"
```



Los 5 paises con más vulnerabilidas industriales son:

Figura 6 - Top 10 de paises con más vulnerabilidades



Comparando los resultados reflejados en la Figura 3 y en e Figura 5, se observa como existe una diferencia entre la primera y la segunda porque no se ven reflejados la mayoría de las ubicaciones de las vulnerabilidades con el país que refleja que tiene más dispositivos vulnerables. Es decir, a pesar que EEUU tiene más dispositivos industriales es Noruega el que tiene más dispositivos vulnerables. Noruega tiene menos dispositivos industriales pero tiene mmuchos del fabricante Elster EneryICT. Tal como hemos dicho, este fabricante es el primero en vulnerabilidades publicadas en el CVE.

5. Conclusiones

Del analisis anteriormente realizado podemos concluir qué:

1. Estados Unidos tiene la mayor cantidad de dispositivos industriales publicos en Internet pero su nivel de vulnerabilidad es no es tan critico como otros paises.
2. Noruega tiene más dispositivos vulnerables asociados a un mismo fabricante lo cual hace que sea más vulnerable que paisis como EEUU con muchos más dispostivos industriales.
3. Los fabricantes Elster EnergyICT y Siemens son los fabricantes con más vulnerabilidades debido principalmente a que son son los fabricantes con mayor catalogo de dispostivos industriales.

6. Listado de tablas y figuras

7. Referencias utilizadas

- Hacer y documentar un paquete de R en 20 minutos (https://mauriciogtec.github.io/rGallery/entries/tutoriales/crear_paquetes/crear_paquete.html)

- R Markdown Cheat Sheet (https://www.rstudio.com/wp-content/uploads/2015/02/rmarkdown-cheatsheet.pdf)
- Shodan API (https://developer.shodan.io)
- CVE Mitre (https://cve.mitre.org)
- Leaflet Maps (https://rstudio.github.io/leaflet/)