

**110.5.**

Wykonaj zestawienie gmin miejskich o nazwach zaczynających się na „J”, podając nazwy powiatów i województw, w których są one położone.

**Zadanie 111.****Wiązka zadań *Malware***

*Malware Domain List* to niekomercyjny projekt społecznościowy, w którym tworzona jest lista adresów stron internetowych i dokumentów stwarzających zagrożenie: trojanów, exploitów itp. Listę na bieżąco uzupełniają profesjonaliści bezpieczeństwa w sieci.

W trzech plikach tekstowych przedstawiono dane zaczerpnięte z tej listy. Dane w wierszach oddzielone są pojedynczymi znakami tabulacji. W każdym pliku pierwszy wiersz jest wierszem nagłówkowym.

W pliku `malware.txt` znajdują się pozycje z tej listy, wybrane z okresu od stycznia 2014 do stycznia 2015 roku włącznie. Podano: datę rejestracji zagrożenia (`data`), adres IP komputera udostępniającego zagrożenie (`IP`), opis zagrożenia (`opis`), numer ASN sieci, do której ten komputer należy (`ASN`), ścieżkę dostępu do szkodliwej strony lub do zasobu (`URL`).

**Przykład**

<code>data</code>	<code>IP</code>	<code>opis</code>	<code>ASN</code>	<code>URL</code>
2014-12-17	62.76.74.228	Trojan.Downloader	51408	my-screenshot.net/
2014-12-04	31.41.218.232	CryptoLocker	42655	mysda24.com/f/pacchetto_38.zip
2014-11-25	89.218.31.11	Script.exploit	9198	zakonodatelstvo.kz/russ.html

W pliku `asn.txt` znajdują się informacje o sieciach komputerowych, zawierające m.in. informacje o numerze ASN (*Autonomic System Number*) — identyfikatorze sieci, wykorzystywanym w konfiguracji routerów. W pliku podano dla każdej sieci: numer ASN (`ASN`), internetowy identyfikator kraju (`ID_kraju`), nazwę organizacji regionalnej przydzielającej adres ASN (`region`) oraz nazwę firmy zarządzającej siecią (`siec`).

**Przykład**

<code>ASN</code>	<code>ID_kraju</code>	<code>region</code>	<code>siec</code>
1267	it	ripence	ASN-INFOSTRADA WIND Telecomunicazioni S.p.A.
2514	jp	apnic	INFOSPHERE NTT PC Communications, Inc
2914	us	arin	NTT-COMMUNICATIONS-2914 - NTT America, Inc

Regionalne organizacje przydzielające adresy ASN obejmują:

- `apnic` (*Asia Pacific Network Information Centre*) — rejon Azji i Pacyfiku,
- `arin` (*American Registry for Internet Numbers*) — rejon Ameryki Północnej,
- `lacnic` (*Latin-American and Caribbean*) — rejon Ameryki Łacińskiej i Wysp Karaibskich,
- `ripence` (*Réseaux IP Européens*) — rejon Europy, Bliskiego Wschodu i centralnej Azji,
- `afrinic` — rejon Afryki.

W pliku `kraje.txt` podano nazwy krajów (`kraj`) oraz ich 2-literowe identyfikatory internetowe (`ID_kraju`).

**Przykład**

kraj	ID_kraju
Australia	au
France	fr
Hungary	hu

Rozwiąż poniższe zadania, wykorzystując dostępne narzędzia informatyczne. Wyniki zamieść w pliku tekstowym o nazwie `malware_wynik.txt`. Do oceny oddaj plik tekstowy zawierający wyniki oraz plik zawierający realizację komputerową Twojego rozwiązania.

**111.1.**

Znajdź te pozycje złośliwego oprogramowania, których celem jest **phishing**, czyli wyłudzenie informacji od użytkownika (w polu: *opis* zawierają łańcuch znaków: *phish* lub *Phish*). Podaj listę zawierającą dla każdej pozycji:

- nazwę kraju, z którego pochodzi komputer udostępniający zagrożenie,
- opis zagrożenia (*opis*),
- pełną ścieżkę dostępu do szkodliwego pliku (URL).

**111.2.**

Znajdź pięć sieci, z których komputery udostępniły najwięcej pozycji złośliwego oprogramowania. Podaj zestawienie zawierające dla każdej takiej sieci: nazwę sieci, nazwę kraju, w którym znajduje się ta sieć, liczbę stron lub dokumentów zawierających złośliwe oprogramowanie oraz liczbę różnych adresów IP, z których to oprogramowanie udostępniono.

**111.3.**

Dla każdego wpisu na listę *malware* określ domenę, z której udostępniono szkodliwe oprogramowanie. Nazwę domeny stanowią znaki pola URL liczone kolejno od lewej aż do pierwszego wystąpienia znaku „/”, bez tego znaku. W każdym polu URL danych znak „/” występuje przynajmniej jeden raz.

- Podaj liczbę domen, z których pochodzi szkodliwe oprogramowanie.
- Serwer DNS można skonfigurować tak, aby odpowiadał kilkoma adresami IP dla jednej domeny. Wyszukaj wśród domen te pozycje, którym odpowiada więcej niż jeden adres IP. Podaj nazwy tych domen i odpowiadające im liczby różnych adresów IP.

**111.4.**

Sporządź w postaci tabeli zestawienie, w którym podasz liczbę zarejestrowanych pozycji złośliwego oprogramowania **w każdym miesiącu roku 2014**, w podziale na poszczególne regiony: apnic, arin, lacnic, ripenc, afrinic. W zestawieniu nie uwzględniaj danych ze stycznia 2015 r.

**111.5.**

Złośliwe oprogramowanie może być ukryte w plikach graficznych i uaktywniać się podczas wyświetlania obrazu, wykorzystując luki w programach odtwarzających obraz.

Znajdź wszystkie wpisane na listę *malware* pozycje informujące o złośliwym kodzie ukrytym w obrazach zapisanych w formatach: jpg i png (pole *URL* kończy się: „.jpg” lub „.png”).

Sporządź w postaci tabeli zestawienie, w którym dla każdego, w którym znajdował się komputer udostępniający szkodliwy kod w plikach graficznych, podasz liczby tych plików, z podziałem na format jpg i png.

### Zadanie 112.

#### Wiązka zadań *Kod EAN*

Kod EAN (*European Article Number*) jest kodem kreskowym powszechnie stosowanym na opakowaniach towarów będących przedmiotem handlu. Czytnik kodu przetwarza go na ciąg cyfr z zakresu 0..9.

Dla uproszczenia zadania (choć rzeczywistość wykracza poza te reguły) przyjęto, że:

- kody wszystkich towarów są 13-cyfrowe,
- pierwsze 3 cyfry kodu są oznaczeniem kraju, w którym zarejestrowany jest producent,
- następne 4 cyfry identyfikują oddział firmy producenta,
- kolejne 5 cyfr jest identyfikatorem towaru,
- ostatnia cyfra jest cyfrą kontrolną.

W pliku `towary.txt` zebrano dane o wybranych słodyczach, napojach, deserach i używkach dostępnych na polskim rynku (nieco zmodyfikowane na potrzeby zadania). Kody EAN towarów są unikatowe. Towary o tej samej nazwie, lecz w opakowaniach o różnych masach, posiadają różne kody EAN. Dla każdego towaru podano: nazwę (*nazwa*), jego rodzaj (*rodzaj*), masę towaru w gramach (*masa*), podatek VAT (*VAT*) oraz kod EAN (*EAN*). Każdy towar opisany jest w osobnym wierszu. Przykład:

<i>nazwa</i>	<i>rodzaj</i>	<i>masa</i>	<i>VAT</i>	<i>EAN</i>
TIKI TAKI	czekolada	100	0,23	5909102008735
KOPIEC KRETA	ciasto	410	0,05	5909437089003
NIC NAC'S	przekaska	40	0,08	4018077798818

W pliku `kraje.txt` znajdują się nazwy krajów (*kraj*), w których zarejestrowani są producenci towarów, oraz przydzielone tym krajom 3-cyfrowe kody (*kod\_kraju*), będące pierwszymi trzema cyframi kodów EAN. Większość krajów posiada jeden kod, ale są kraje, którym przydzielono więcej kodów. Przykład:

<i>kod_kraju</i>	<i>kraj</i>
304	Francja
338	Francja
380	Bulgaria

W pliku `producenci.txt` znajdują się informacje o oddziałach firm producentów towarów. Dla każdego oddziału podano: jego 4-cyfrowy kod stosowany w kodzie EAN (*kod\_oddzialu*), 3-cyfrowy kod kraju, w którym oddział został zarejestrowany (*kod\_kraju*), nazwę firmy (*nazwa\_firmy*) oraz lokalizację jej oddziału — nazwę miasta, ewentualnie z nazwą dzielnicy (*lokalizacja*). Niektóre firmy mają wiele oddziałów, które mogą być zarejestrowane w różnych krajach. Przykład: