Case Study

"Privacy Protection"

Abderrahmane Hammia

Kaouther Nouni

2023

Problematic:

In this case study we try to act like attackers and try to locate a person in a dataset based on some information that was collected after a survey

> ➤ This person is a <u>woman</u>, born between <u>1957</u> and <u>1960</u>, she is an <u>architect</u> and <u>at the time of the survey</u> she was <u>working independently</u> in the same field. She <u>lives in Barcelona,</u> which is also her <u>place of birth</u> and has either <u>two</u> or <u>three</u> children.

Code:

We used python to apply filters to track down the targeted woman in the database, here's how our code works:

1- Read database file: we read the database file as it is structured into array of database rows as string, without removing spaces because spaces are important, and they indicate that the value of a specific attribute is not filled.
2- We will introduce an array "filter_positions" with positions of attributes we want to filter, and another array "filter" of values we want to find.
3- We will start going database rows, splitting the string based on the number of characters in each attribute, and get the attribute values we want with "filter_positions" from the string and compare it to our filter if it does match, we keep this record, otherwise we discard it.
4- We write the remaining records in a file and display them.

Results and discussions:

1. First Case:

In this case we added the filter TAMUNI that indicates the population of the city and the professional situation because she works independently in the times of the survey, in addition of all the required attributes to be able to locate her better.

Attribute	Variable	Position	Length	Required Value	Code
Civil statue	A_10	38	1	Married	2
Autonomous Community	CCAA	1:2	2	Catalunya	09
Gender	SEXO	3	1	Woman	6
Year of Birth	A2B	6:9	3	1957:1960	1957- 1958 1960
Place of birth	A_3A	13	1	Yes	1
Community of Birth	A_3B	14:15	2	Barcelona - Catalunya-	09

Economic activity (last week)	A_6_1	23	1	Worked in the last week	1
Economic activity (last 12 months)	A_5	21:22	2	Works	01
Professional situation	A_9	37	1	Architect	2
Population of the city	TAMUNI	310	1	Population of the city	>500 000
Cohabitation	B12C	42	1	1 or two children	1

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Comunida	SEXO	Año de na	Nacido en	Si es en Es	Nivel de estudios	Actividad económic	Actividad o	Ocupación	Situación profesional	Estado civil	Convivenc	Tamaño del muni	Personas con las que ha convivido en lo
2	Cataluña	Mujer	1958	España	Cataluña	Estudios universitarios	Trabajando	Si	Técnicos y profesionales científicos	Asalariado/a fijo (a sueldo,	Casado/a	Si	Más de 500000 h	Cónyuge/Pareja y otro
3	Cataluña	Mujer	1958	España	Cataluña	Estudios universitarios	Trabajando	Si	Técnicos y profesionales científicos	Asalariado/a fijo (a sueldo,	Casado/a	Si	Más de 500000 h	Cónyuge/Pareja y otro
4														
5														

The two individuals found in the record above match the basic requirements as: Gender, Autonomous Community, Year of Birth, works as an architect which falls under the classification of "Técnicos y profesionales científicos e intelectuales". also lives in Barcelona, so the city has more than 500000 inhabitants.

However, the two found records both don't work independently.

2. Second case:

Attribute	Variable	Position	Length	Value	Code
Civil statue	A_10	38	1	Married	2
Autonomous Community	CCAA	1:2	2	Catalunya	09
Gender	SEXO	3	1	Woman	6
Year of Birth	A_2B	6:9	3	1957:1960	1957:1960
Place of birth	A_3A	13	1	Yes	1
Community of Birth	A_3B	14:15	2	Barcelona - Catalunya-	09
Economic activity (last week)	A_6_1	23	1	Worked in the last week	1

Economic activity (last 12 months)	A_5	21:22	2	Works	01
Professional Situation	A_9	37	1	Independent	2
Cohabitation	B12C	42	1	1 or two children	1

In this case, we found one record that aligns with the given background, has a university level of studies (architect). However, the city where this individual declared they live in has only 50000 inhabitants (which is obviously not Barcelona's population.) but it can be in the Province of Barcelona yet not in the city itself.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	Comunidad	SEXO	Año de na	Nacido en	Si es en Es	Nivel de estudios	Actividad eco	Actividad eco	Ocupación	Situación profesional	Estado civ	Convivenc	Tamaño del municipio	Personas con las que ha convivido en los últim			
2	Cataluña	Mujer	1960	España	Cataluña	Estudios universitarios	Trabajando	Trabajando	Técnicos y profesionales de apoy	Empresario/a sin asalari	Casado/a	Si	De 10001 a 50000 hab	Cónyuge/Pareja y otro			
3																	

It is not possible to locate one single woman with the exact given background, it can be that the data was very well anonymized (altered per example), further details of answers in the sensitive question may help in re-identifying her.

Since we know that the woman lives in the city of Barcelona, so the first case is highly stronger to be true than the second case, therefore should be one of the 2 results found.

We put the code in the following github repository: https://github.com/Abdo99ab/reidentification_challenge.