



Patrones Secuenciales

Minería de Patrones:

Consiste en describir patrones interesantes, útiles e inesperados en una base de datos.

Tipos de Patrones:

- ✓ Itemsets frecuentes.
- ✓ Asociaciones.
- ✓ Sub grafos.
- ✓ Reglas secuenciales.
- ✓ Patrones periódicos.



La tarea de minería de patrones secuenciales se especializa en analizar datos secuenciales y descubrir patrones secuenciales. Esto se puede traducir a encontrar sub-secuencias interesantes dentro de un conjunto de secuencias.



Cada una de las secuencias representa elementos comprados por un cliente en un tiempo determinado.

SI D	SECUENCIA
1	((A, B), (C), (F, G), (G), (E))
2	((A, D), (C), (B), (A, B, E, F))
3	((A), (B), (F, G), (E))
4	((B), (F, G))



Patrones secuenciales se utiliza para encontrar subsecuencias que aparecen con frecuencia en una base de datos (subsecuencias de patrones frecuentes).

Esto puede ayudarnos a entender el comportamiento de clientes en cuanto a decisiones de mercado.



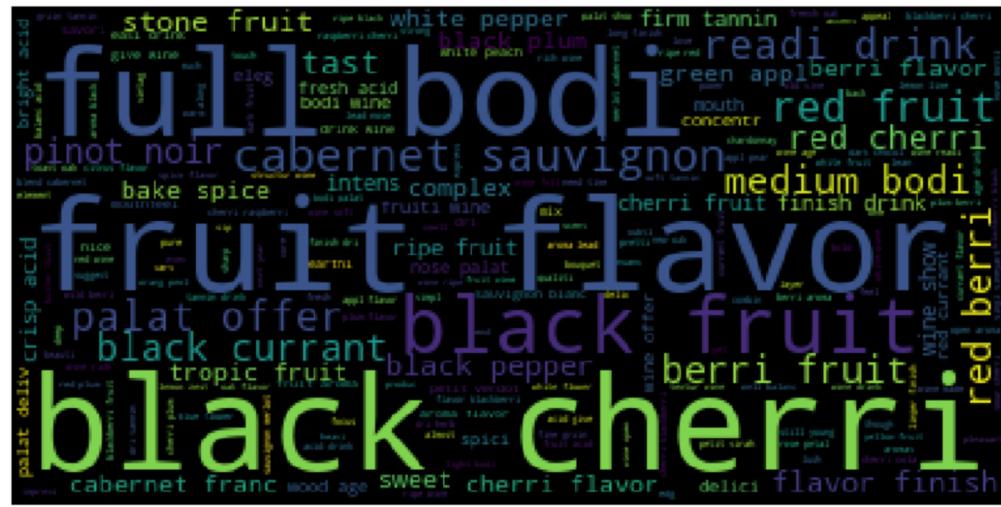
Para hacer patrones secuenciales el usuario debe de dar una base de datos secuencial y especificar lo que se conoce como **umbral de soporte mínimo**

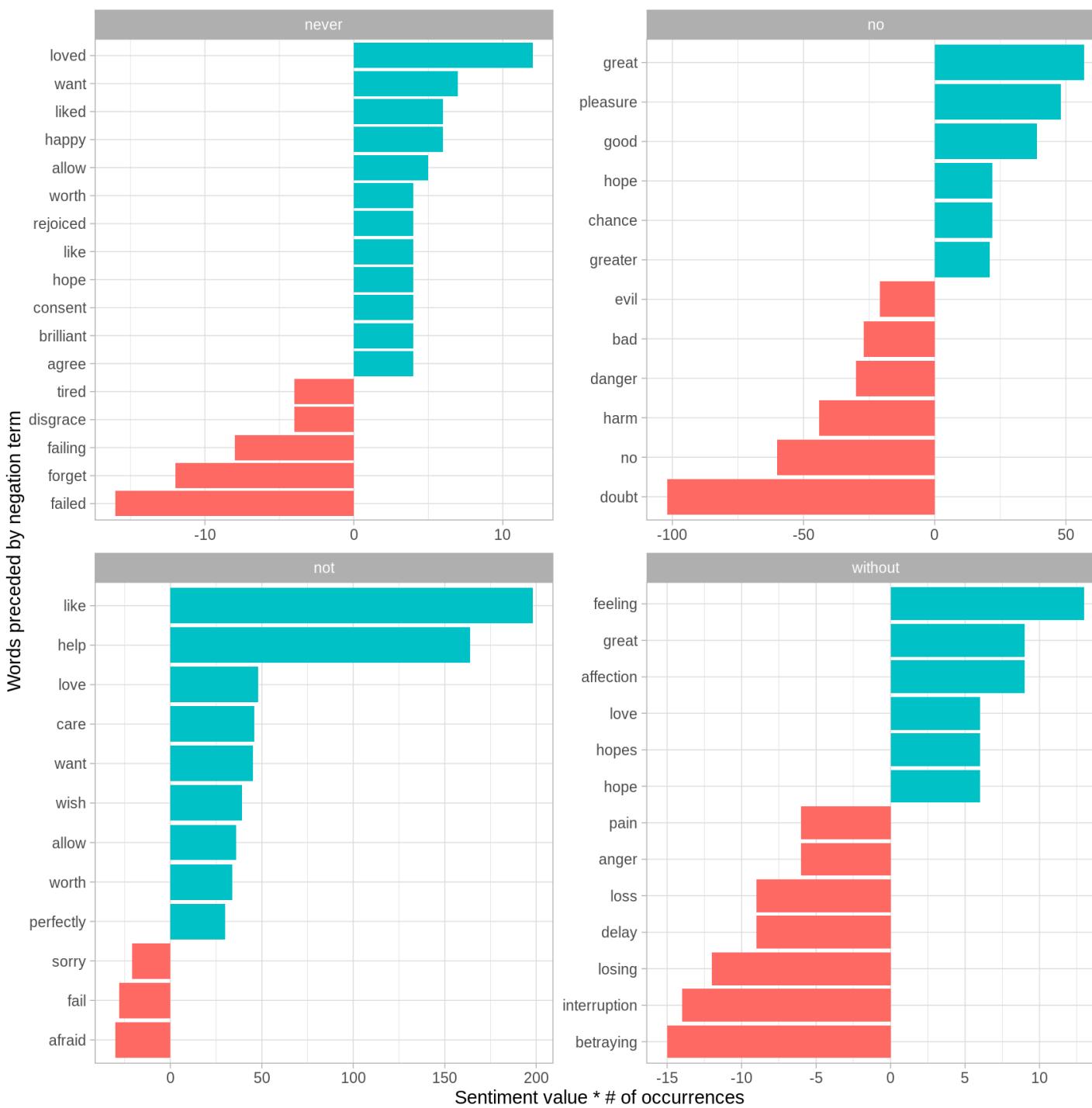
En el ejemplo podemos tomar un valor de soporte mínimo como 2. Estos patrones son interesantes ya que representan alguno de los comportamientos comunes de los clientes.

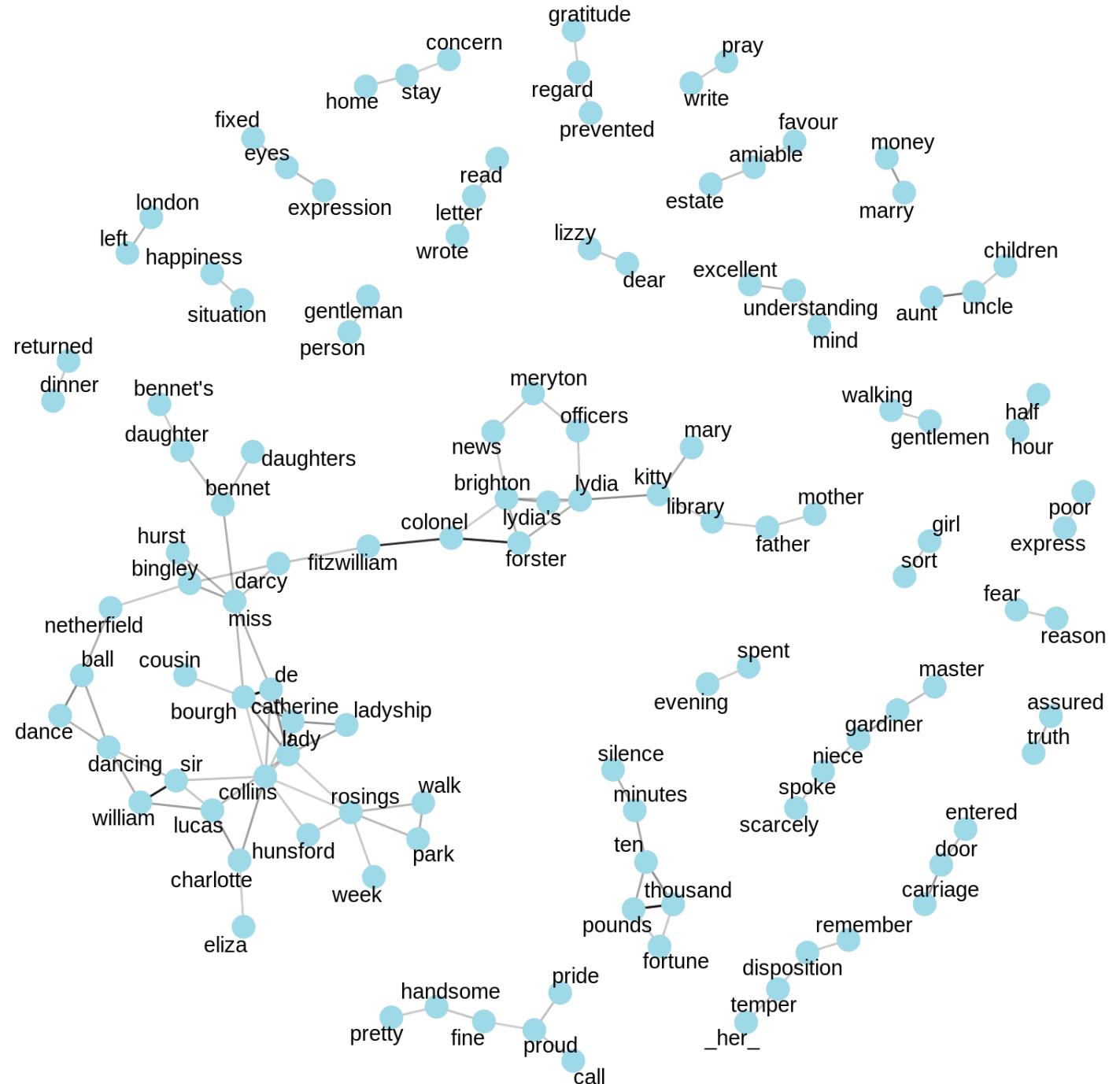
SI D	SECUENCIA
1	((A, B), (C), (F, G), (G), (E))
2	((A, D), (C), (B), (A, B, E, F))
3	((A), (B), (F, G), (E))
4	((B), (F, G))

PATRON	SOPORTE
((A))	3
((A), (G))	2
((A), (G), (E))	2
((A), (F))	3
((A), (F), (E))	2
((A), (C))	2
((A), (C), (F))	2
.	.
.	.
.	.

Otro ejemplo de la aplicación de patrones secuenciales puede presentarse en análisis de textos. Un conjunto de frases de un texto pueden ser vistas como la base de datos de secuencias, y el objetivo de patrón secuencial es encontrar las palabras más utilizadas en el texto.







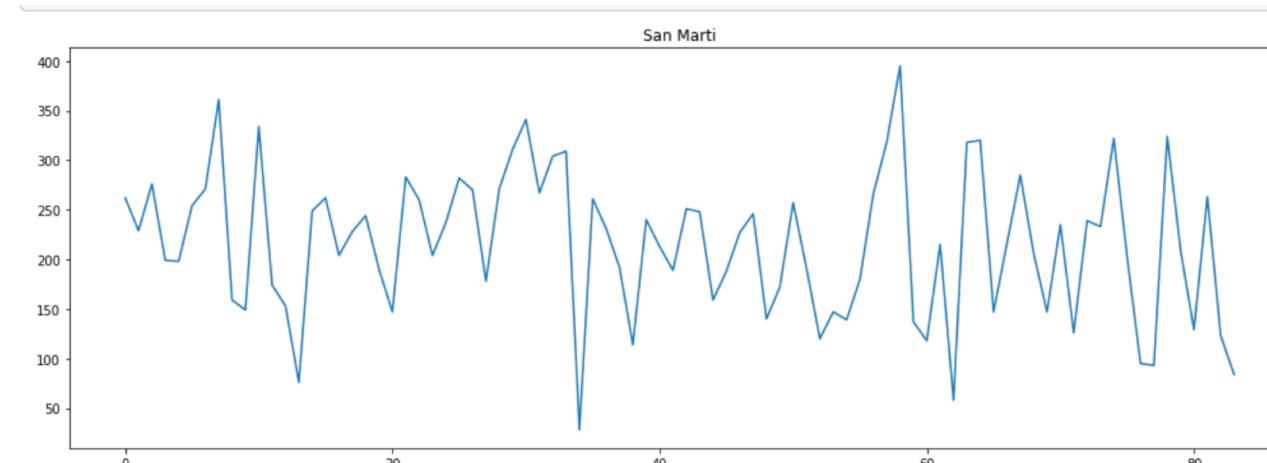


Además de secuencias, esta técnica puede ser aplicada en series de tiempo., siempre y cuando se realice una discretización de datos como pre procesamiento.

```
In [93]: d = df.drop(["Id", "Neig", "Street"], axis=1)  
d.head(5)
```

Out[93]:

	District	Weekday	Month	Day	acc	p_day	m_injuries	s_injuries	Victims	Vehicles	Longitude	Latitude
0	1	1	1	13	8	Morning	2	0	2	2	2.125624	41.340045
1	1	1	2	1	13	Morning	2	0	2	2	2.120452	41.339426
2	1	1	3	8	21	Afternoon	5	0	5	2	2.167356	41.360885
3	1	1	4	21	2	Night	1	0	1	2	2.124529	41.337668
4	2	2	5	25	14	Afternoon	1	0	1	3	2.185272	41.416365

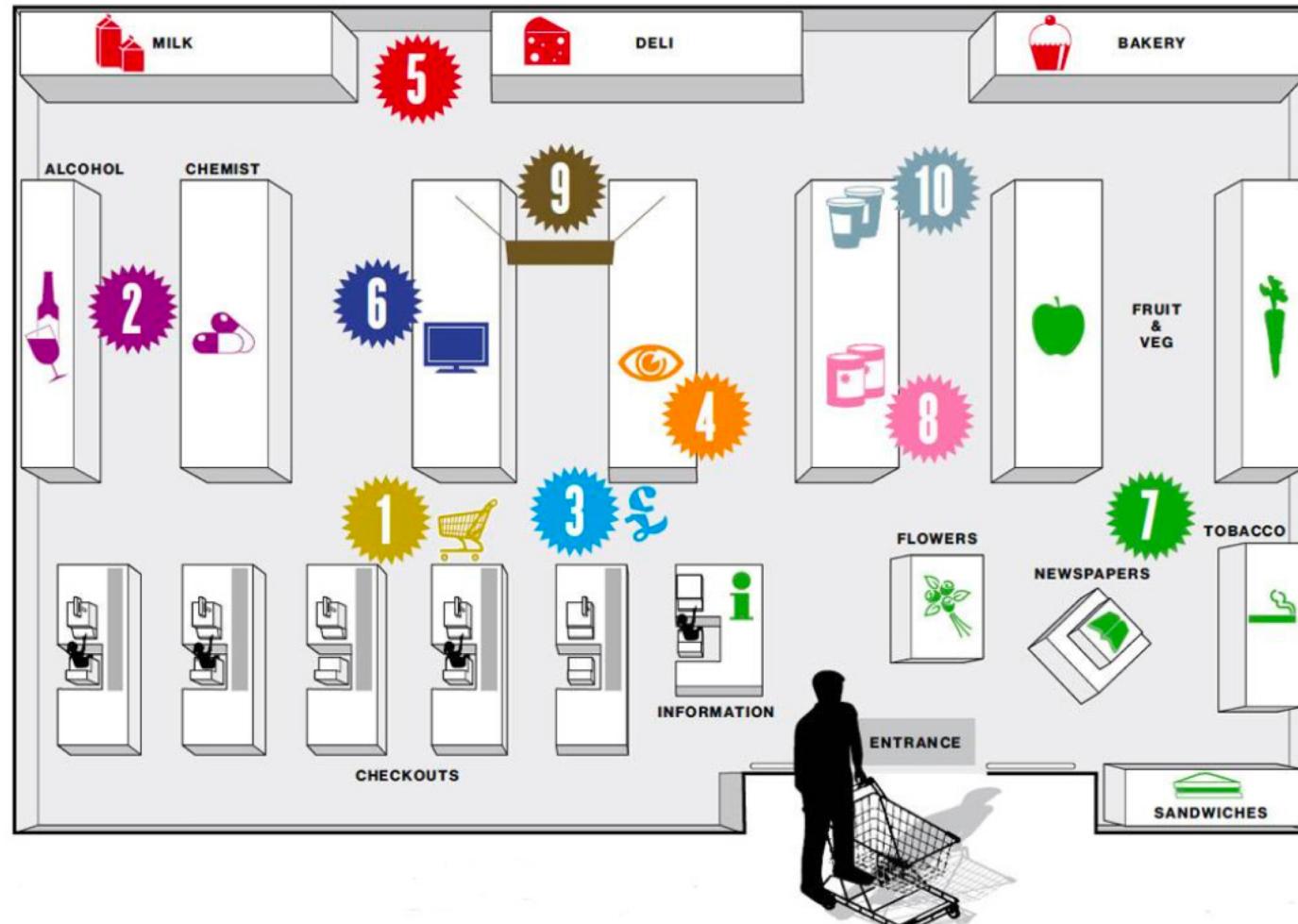


Existen varios algoritmos para resolver problemas de patrones secuenciales. Algunos de los más mencionados en investigaciones son:

- PrefixSpan
- Spade.
- SPAM
- GSP



Minería de Patrones Periódicos.



La minería de patrones periódicos es una tarea importante ya que de manera periódica pueden aparecer diferentes tipos de datos, y conviene entenderlos para tomar decisiones estratégicas.

• Análisis Biológico Secuencial.

Compara y analiza las secuencias biológicas, lo cual resulta crucial en el análisis bio informático.

Estas secuencias pueden estar formadas de nucleoides o amino ácidos.

A) CCAGTTGCA
CATAGGGCC
AGCCATTAA
GGGGATGCT
CCAGCGGCC
GTAGTCCAG
CAAGACCGG
TTCGAGGCC
GTCGTGGCC
CCAGTTGCA
...

B) CCAGAGGCC
C) 034350001:A
013021000:T
320515621:G
631111266:C

