

Clase **11**

# Explotación de datos textuales



El pack Sphinx dedicado a la enseñanza de métodos y técnicas para encuestas

# Objetivos

## Objetivos metodológicos

- Comprensión del valor de la recopilación de datos textuales en un estudio
- Ser capaz de explotar los principales resultados de un análisis textual elemental

## Objetivos Sphinx

- Saber activar e interpretar las funciones de análisis de texto de Sphinx

# Tabla de contenidos

1. El valor de la recopilación de datos textuales en un estudio
2. Ejemplos de protocolos de recopilación de textos
3. Explotación de las respuestas textuales
  - 3.1 Análisis léxico
  - 3.2 Codificación
  - 3.3 Extracción de los verbatim

# 1. El valor de la recopilación de datos textuales en un estudio

- Las preguntas abiertas permiten registrar respuestas más **espontáneas**.
- Permiten cambiar el formato de respuesta y así asegurar un protocolo más **variado**.
- Si el encuestado está involucrado y/o es voluntario, se puede recopilar información muy **rica**, a veces inesperada, por ejemplo, sobre los sentimientos, la imagen o la motivación. De esta manera, se introduce un poco de "**cualitativo**" en un enfoque de cuestionario que es generalmente cuantitativo.
- Los verbatim (textos creados por los encuestados) pueden "**dar vida**" a resultados que a veces son demasiado numéricos.

# 2. Ejemplos de protocolos de recopilación de textos

- **Comentarios/sugerencias al final del cuestionario:** *Finalmente, ¿tiene algún comentario, crítica o sugerencia sobre esto o cualquier otro tema?*
- **Otro, precise:** *¿Qué otras novedades podrían interesarle?*
- **Completar una frase:** *Para usted, dar regalos de Navidad es...*
- **3 palabras:** *Diga tres palabras (o tres adjetivos) para calificar al Real Madrid.*
- **Asociación de ideas:** *Si digo Real Madrid, piensa... Si el Real Madrid fuera una canción, sería...*
- **Relanzar una pregunta cerrada:** *Está insatisfecho con..., por favor, díganos por qué.*
- **Narración:**
  - *La última vez que hizo esto, o compró eso... ¿cómo fue? ¿Cómo se sintió?*
  - *Por ejemplo: ¿recuerda la última película que vio que le gustara? En pocas palabras, ¿por qué le gustó esa película?*



## Consejos:

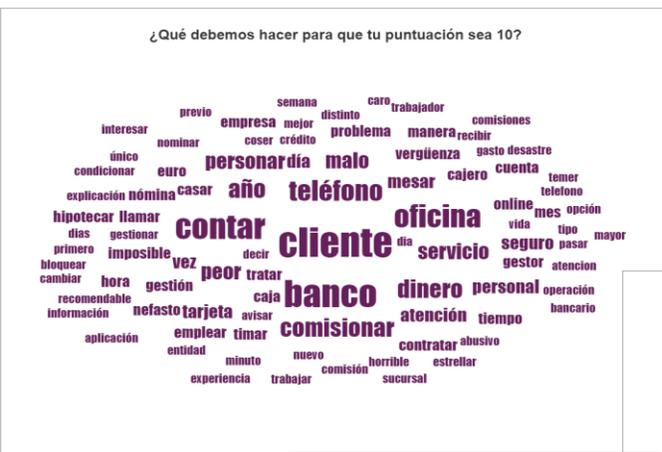
- Planifique el número de líneas correspondiente al volumen de respuesta previsto
- Si es necesario, para una pregunta abierta particularmente importante, no dude en definir una respuesta obligatoria
- Las preguntas abiertas no representan más del 10-15% de la encuesta

# 3. Explotación de las respuestas textuales

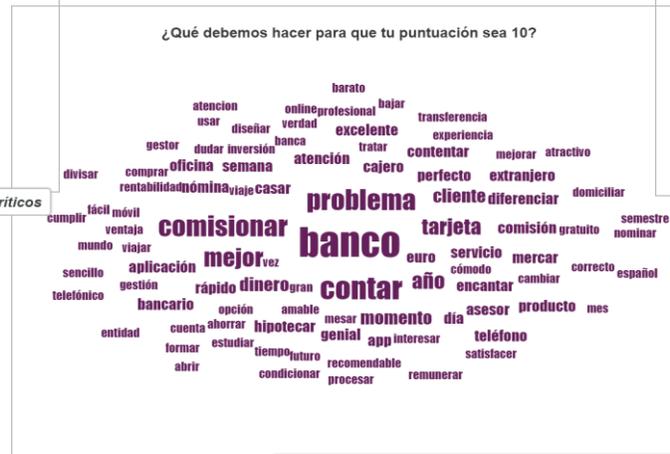
- Para analizar los verbatim, hay básicamente tres enfoques:
  - **El resumen intelectual:** a partir de una lectura casi exhaustiva del texto literal y sobre la base de su conocimiento del tema, el analista hace una síntesis e identifica los n temas principales del texto. Tradición de análisis cualitativo.
  - **La codificación:** lectura de la literalidad y codificación de cada respuesta según una tabla de codificación (o libro de códigos) predefinida por el analista. Análisis del contenido.
  - **Análisis léxico:** las palabras y expresiones citadas se encuentran en todo el corpus y de ellas se deducen los contenidos temáticos.
- Para el segundo y tercer enfoque, se propondrán tablas y/o gráficos para apoyar las conclusiones.
- En los tres casos, los resultados se complementarán con **verbatim significativos** escogidos para ilustrar cada categoría temática.

# 3.1 Análisis léxico

- El análisis léxico tiene por objeto restablecer las palabras o expresiones más frecuentemente citadas para captar rápidamente el contenido en todo el corpus o por categorías de encuestados.
- Las listas de palabras clave pueden presentarse como una lista, tabla o nube de etiquetas (*tag cloud*).



Información de la muestra : Clases\_NPS =Críticos



Información de la muestra : Clases\_NPS =Promotores



Información de la muestra : Clases\_NPS =Promotores

En el ejemplo, podemos ver que las palabras clave citadas por los usuarios varían según el número de estrellas del banco.

# 3.2 Codificación

- El análisis de contenido mediante la codificación "manual" consiste en la lectura de todos los verbatim pertinentes y en codificarlos según una tabla de codificación más o menos elaborada, que ha sido predefinida por el analista, en función de sus preguntas de investigación.
- Los programas informáticos suelen ofrecer un entorno específico en el que se elige la variable que se va a codificar, en el que se define la lista de codificación. Aquí se consultan los verbatim y se codifican uno tras otro.
- La variable así constituida está entonces disponible para realizar un análisis descriptivo o cruzado.

CODIFICACIÓN MANUAL - Variable analizada: ¿Qué debemos hacer para que tu puntuación sea 10? (COMENTARIOS\_ES)

“ No está mal, es un clásico que siempre está ahí y nunca falla

57 - Hombre

11 Observación no codificada

Mostrar los extractos

servicio	Online	Responsabilidad social	Comentarios
<input type="checkbox"/> Calidad 10 obs. 8 obs. No está mal, es un clásico que siempre está ahí y nunca falla	<input type="checkbox"/> Conectividad 1 obs.	<input type="checkbox"/> Ética	<input type="checkbox"/> Positivos
<input type="checkbox"/> Rapidez 3 obs.	<input type="checkbox"/> Usabilidad 1 obs.	<input type="checkbox"/> Programas de fomento	<input type="checkbox"/> Negativos
<input type="checkbox"/> Satisfactorio	<input type="checkbox"/> Descarga de la app	<input type="checkbox"/> Imagen social	+ Añadir un tema
+ Añadir un tema	+ Añadir un tema	<input type="checkbox"/> Colaboraciones	



# Conceptos clave

- **Análisis de datos textuales:** conjunto de métodos dedicados al tratamiento de textos no estructurados destinados, en particular, a la comprensión de su contenido temático o semántico.
- **Análisis de contenido:** técnica que consiste en leer una respuesta textual (o una foto, por ejemplo) y luego codificar el contenido según una lista temática predefinida (o libro de códigos).
- **Nube de palabras:** (o *tag cloud*) representación visual en forma de nube, de las palabras o expresiones más citadas dentro de un texto que estamos estudiando. El tamaño de las palabras generalmente indica su frecuencia.

# Ejercicios

## **Exercice 11a : adjectifs pour la chaine favorite**

A partir de l'enquête « Télévision », analysez les résultats de la question 9 qui enregistre les adjectifs associés par les téléspectateurs à leur chaine favorite. Quels sont les 5 adjectifs qui sont globalement les plus cités ? Observe-t-on des différences en fonction de la chaine regardée le plus souvent ?

## **Exercice 11b : verbatim culture**

Editez les verbatim contenant des mots liés à la « culture ». Quelles sont les chaines les plus fréquemment concernées ?

# Más recursos

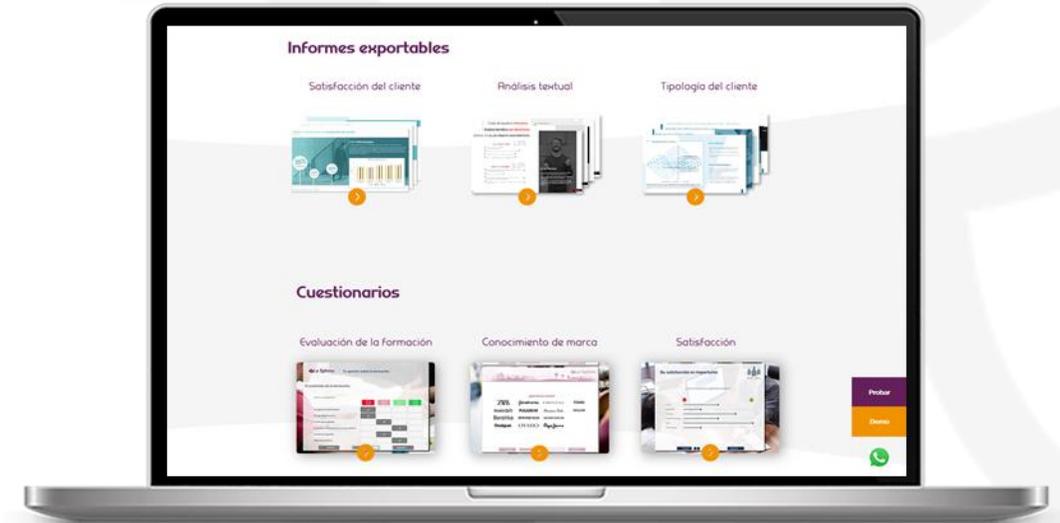
## Tutoriales

[www.lesphinx.es/videos](http://www.lesphinx.es/videos)



## Ejemplos

[www.lesphinx.es/ejemplos](http://www.lesphinx.es/ejemplos)



# Descargar una licencia

[www.lesphinx.es/probar-gratis](http://www.lesphinx.es/probar-gratis)

