

## ÍNDICE

INTRODUCCIÓN .....	Pág.	11
1. LAS RESPUESTAS ABIERTAS, UN MATERIAL PARA EL ESTUDIO ESTADÍSTICO: UNIDADES Y TABLAS LÉXICAS .....		15
1.1. Las encuestas por cuestionario y el cuestionamiento abierto .....		15
1.2. Presentación del ejemplo <i>Identidad</i> .....		15
1.3. Estadística y variable textual .....		17
1.4. Recogida de las respuestas libres .....		18
1.5. Preproceso del corpus y segmentación del corpus en unidades léxicas .....		19
1.5.1. <i>Corpus</i> .....		19
1.5.2. <i>Preproceso del corpus</i> .....		19
1.5.3. <i>Unidad léxica: palabra o lema</i> .....		20
1.5.4. <i>Interés de la lematización</i> .....		22
1.5.5. <i>Stoplist</i> .....		23
1.5.6. <i>Unidad léxica compuesta: el segmento repetido</i> .....		23
1.5.7. <i>"Stematización" y/o reagrupación de sinónimos</i> .....		24
1.5.8. <i>Umbral de frecuencia</i> .....		24
1.6. Codificación del texto y tablas léxicas .....		25
1.6.1. <i>Tabla léxica individuos × palabras</i> .....		25
1.6.2. <i>Tabla léxica agregada palabras × categorías</i> .....		25
2. PRINCIPIOS BÁSICOS DEL SOFTWARE SPAD7 .....		27
2.1. Introducción .....		27
2.2. Ejemplo de proyecto SPAD .....		27
2.2.1. <i>El archivo BASE de SPAD</i> .....		27
2.2.2. <i>Construcción de una cadena</i> .....		27
2.2.3. <i>Apertura de una cadena favorita</i> .....		39
2.2.4. <i>Creación de una cadena favorita</i> .....		40

2.3.	Importación y exportación de archivos .....	42
2.4.	Conexiones .....	42
2.5.	Familias de métodos .....	43
3.	GLOSARIOS Y CONCORDANCIAS .....	49
3.1.	Introducción .....	49
3.2.	Preproceso del corpus <i>Identidad</i> .....	49
3.2.1.	<i>Selección de la pregunta abierta, de los individuos y de los delimitadores</i> .....	50
3.2.2.	<i>Selección del vocabulario</i> .....	53
3.2.3.	<i>Modificación del vocabulario</i> .....	54
3.3.	Glosario de palabras y segmentos .....	57
3.3.1.	<i>Glosario de palabras</i> .....	57
3.3.2.	<i>Glosarios de segmentos</i> .....	58
3.3.3.	<i>Temas más frecuentes</i> .....	60
3.4.	Concordancias de palabras .....	60
3.5.	Medidas estilísticas .....	62
3.5.1.	<i>Distribución de las frecuencias</i> .....	62
3.5.2.	<i>Longitud de las respuestas abiertas</i> .....	63
3.5.3.	<i>Riqueza del vocabulario</i> .....	63
4.	COMPARACIÓN DE CATEGORÍAS DE RESPUESTAS. ANÁLISIS DE CORRESPONDENCIAS DE UNA TABLA LÉXICA AGREGADA .....	65
4.1.	Introducción .....	65
4.2.	Construcción de la tabla léxica agregada .....	65
4.2.1.	<i>Breve resumen cuantitativo sobre las ocho categorías de respuestas</i> .....	67
4.2.2.	<i>Tabla léxica agregada</i> .....	67
4.2.3.	<i>Visualización de la tabla léxica agregada</i> .....	68
4.3.	Análisis de correspondencias de una tabla léxica agregada .....	69
4.3.1.	<i>Análisis de correspondencias y análisis textual</i> .....	69
4.3.2.	<i>Análisis de correspondencias: formulación matemática</i> .....	70
4.3.2.1.	Transformación de la tabla de efectivos en tabla de proporciones .....	70
4.3.2.2.	Tablas de perfiles-fila y de perfiles-columna .....	70

4.3.2.3.	Enfoque geométrico .....	71
4.3.2.4.	Inercia de las nubes de perfiles-fila y de la nube de perfiles-columna .....	72
4.3.2.5.	Visualización de la estructura de las nubes de perfiles-fila y de perfiles-columna .....	72
4.3.2.6.	Ejes principales y vectores de coordenadas principales .....	72
4.3.2.7.	Formulas de transición y elementos suplementarios .....	74
4.3.2.8.	Análisis de correspondencias de una tabla léxica .....	74
4.4.	Pequeño ejemplo .....	74
4.5.	Análisis de correspondencias de la tabla léxica agregada por Sexo $\times$ Edad .....	78
4.5.1.	<i>Inercia de los factores</i> .....	78
4.5.1.1.	Inercia y valores propios .....	78
4.5.1.2.	Número de factores a conservar e interpretar .....	79
4.5.2.	<i>Interpretación de los factores</i> .....	80
4.5.2.1.	Contribuciones de las filas y columnas ...	81
4.5.2.2.	Interpretación de un factor .....	81
4.5.2.3.	Interpretación de un plano .....	81
4.5.2.4.	Primer plano principal .....	86
4.5.2.5.	Breve síntesis .....	87
5.	COMPARACIÓN DE CATEGORÍAS DE RESPUESTAS. SELECCIÓN DE PALABRAS Y RESPUESTAS CARACTERÍSTICAS .....	89
5.1.	Introducción .....	89
5.2.	Cadena de análisis .....	89
5.3.	Palabras características .....	90
5.3.1.	<i>Palabras características de las ocho categorías de Edad_Sexo</i> .....	90
5.4.	Respuestas características .....	94
5.4.1.	<i>Selección a partir de las palabras características</i> ....	94
5.4.2.	<i>Selección a partir de la distancia del <math>\chi^2</math> entre perfiles léxicos</i> .....	94
5.4.3.	<i>Lectura global de los resultados</i> .....	96
6.	ANÁLISIS DIRECTO DE LAS RESPUESTAS .....	97
6.1.	Introducción .....	97

6.2.	Análisis de correspondencias de la tabla léxica .....	98
6.2.1.	<i>Procedimiento SPAD para el análisis directo del corpus</i> Identidad .....	98
6.2.2.	<i>Valores propios</i> .....	100
6.2.2.1.	Conjuntos sintagmáticos de palabras ....	101
6.2.3.	<i>Identificación de los grupos sintagmáticos</i> .....	102
6.2.3.1.	Nube y contribución de los individuos ...	102
6.2.3.2.	Nube y contribución de las palabras ....	103
6.2.3.3.	Identificación de los grupos sintagmáticos	103
6.2.3.4.	Detección de los grupos sintagmáticos asociados a los dos primeros ejes en el corpus <b>Identidad</b> .....	105
6.2.3.5.	Los segmentos como elementos suplementarios .....	105
6.2.4.	<i>Categorías como elementos suplementarios</i> .....	108
6.2.4.1.	Características sociodemográficas, una ayuda a la interpretación .....	108
6.2.4.2.	Valores-test .....	108
6.2.4.3.	Variables continuas suplementarias.....	111
6.3.	Conclusión y perspectivas .....	111
CONCLUSIÓN	.....	113
REFERENCIAS BIBLIOGRÁFICAS	.....	115