

ÍNDICE

PRÓLOGO	IX
CAPÍTULO 1. EL LENGUAJE DE PROGRAMACIÓN PYTHON	1
Introducción	1
Programas y lenguajes	1
¿Por qué Python?	2
Cómo aprender Python	2
El entorno Jupyter Notebooks	3
Google Colaboratory	3
Subir ficheros de datos a Google Colaboratory	4
Instalación a través de Anaconda	5
¿Los Jupyter Notebooks son programas en Python?	6
Tipos simples en Python	6
Tipos numéricos	7
Tipo lógico	8
Variables	9
Secuencias	10
Concatenación	11
Acceso por posición	12
Otras operaciones para secuencias	13
Operaciones específicas para strings	14
Operaciones específicas para listas	18
Instrucciones condicionales	19
Iterando	21
Listas intensionales	24
Diccionarios	25
Funciones de usuario	27
Uso de bibliotecas Python	30
Importar bibliotecas	30

Instalar bibliotecas	32
CAPÍTULO 2. FICHEROS	35
Introducción	35
Ficheros de texto	35
Hojas de cálculo	39
CSV	49
JSON	55
CAPÍTULO 3. PALABRAS Y ELEMENTOS BÁSICOS	59
Introducción	59
Extracción de unidades léxicas con spaCy	60
Tipos de símbolos	61
Palabras vacías	65
Nubes de palabras	68
Modelos pre-entrenados	70
¿Qué es un modelo en spaCy?	70
Instalación y carga de los modelos	71
Los componentes del modelo	72
División de un texto en oraciones	73
Lematización	77
Extracción de lemas con spaCy	77
Aplicación: contando palabras	78
Extracción de raíces	80
CAPÍTULO 4. LOS ELEMENTOS DE LA ORACIÓN	83
Introducción	83
Análisis morfológico de palabras	84
Análisis sintáctico de oraciones	89
Detección de entidades con nombre	99
Otras bibliotecas para la detección de entidades con nombre	101
Referencias	102
CAPÍTULO 5. ANÁLISIS SEMÁNTICO	103
Introducción	103
Análisis de sentimiento	104
Semejanza entre palabras y entre frases	110
Modelado de temas	116
Análisis Semántico Latente	118
Asignación Latente de Dirichlet	121
Asignando tema a nuevos textos	124

Mejorando el modelado de temas	126
CAPÍTULO 6. ANÁLISIS DE TEXTOS MEDIANTE MÉTRICAS	129
Introducción	129
Estadística descriptiva	132
Facilidad de lectura	134
Distancia de dependencia	137
Proporción de categorías gramaticales	140
Coherencia de textos	141
Visualización de los análisis	143
Creación de tablas en pandas	145
Diagrama de barras de la longitud de las frases	148
Diagrama de barras de varias medidas	150
Histograma de longitudes de frase	153
Diagrama de franjas sobre la distribución de longitudes de frase	156
Diagrama de caja de longitudes de frase	157
Diagrama de barras apiladas de proporciones de categorías gramaticales	159
Referencias	161
ÍNDICE ANALÍTICO	165