

Procesamiento de datos del corpus PRESEEA-Valencia

generación de informes automatizados con R, Shiny y ChatGPT

Adrián Cabedo Nebot

Universitat de València

2025-10-22

El proyecto ECOS-C/N

Estudio de los condicionantes sociales del español actual en el centro y norte de España: *Nuevas identidades, nuevos retos, nuevas soluciones (ECOS-C/N)*

- Proyecto de investigación **PID2023-148371NB-C42**
del **Ministerio de Ciencia, Innovación y Universidades**
(Cabedo Nebot, A. e Illamola Gómez, C., coords., © 2024)
- **Objetivo:** estudiar las variaciones lingüísticas contemporáneas en el español de España para anticipar futuros cambios.
- **Metodología:** combinación de lingüística de corpus y análisis sociolingüístico.
- **Vinculación:** conectado con el proyecto **PRESEEA**.
- **Equipo:** más de 24 investigadores de distintas regiones de España y Baleares.
- **Enfoque:** metodología unificada y colaborativa.

<https://ecoscn.org>

Diapositiva 1

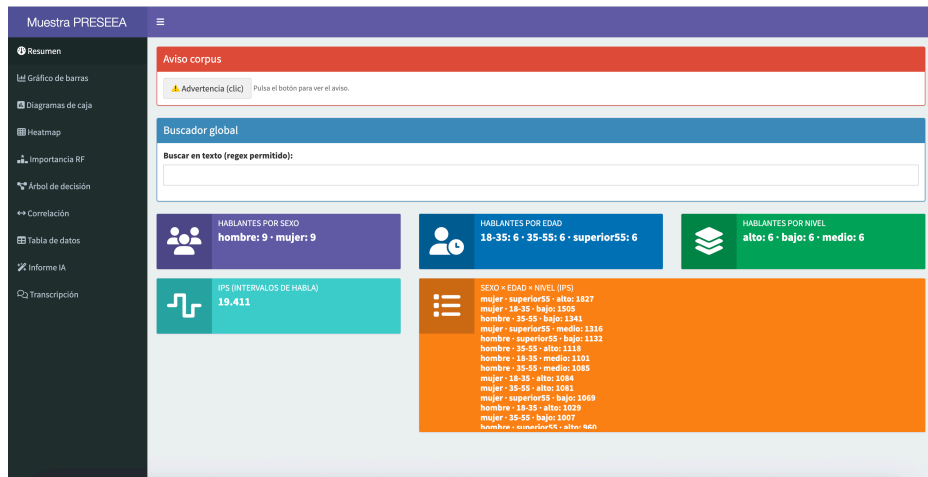


Figure 1: Imagen1

Diapositiva 2

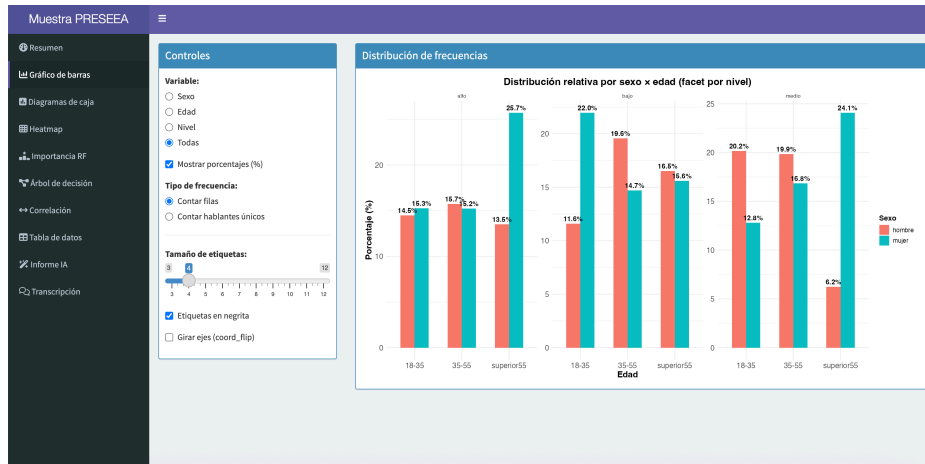


Figure 2: Imagen2

Diapositiva 3

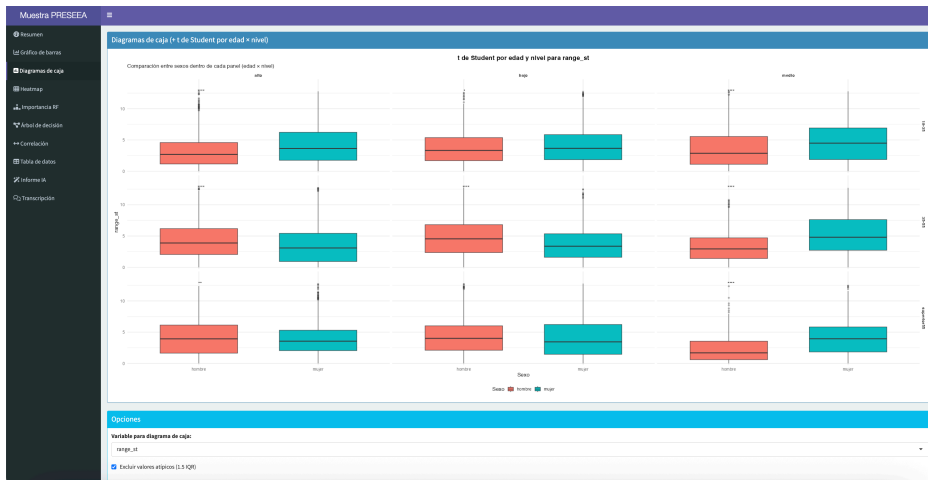


Figure 3: Imagen3

Diapositiva 4

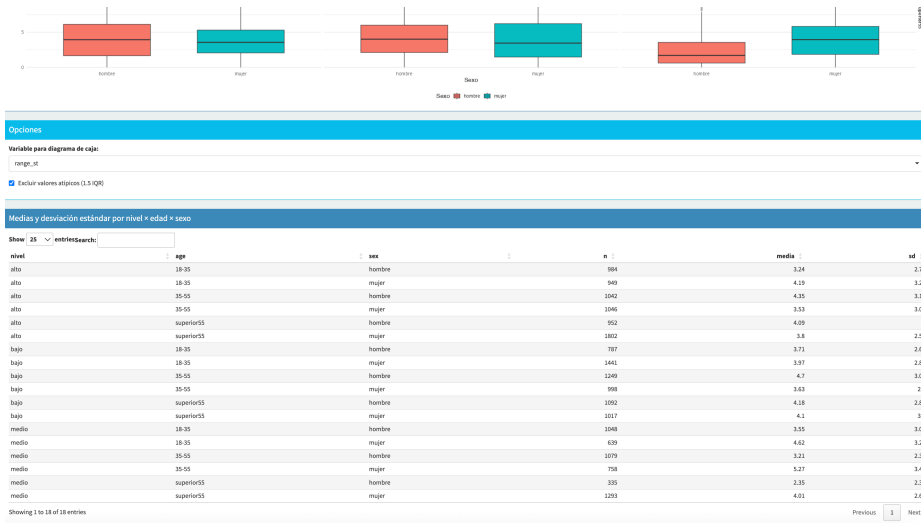


Figure 4: Imagen4

Diapositiva 5

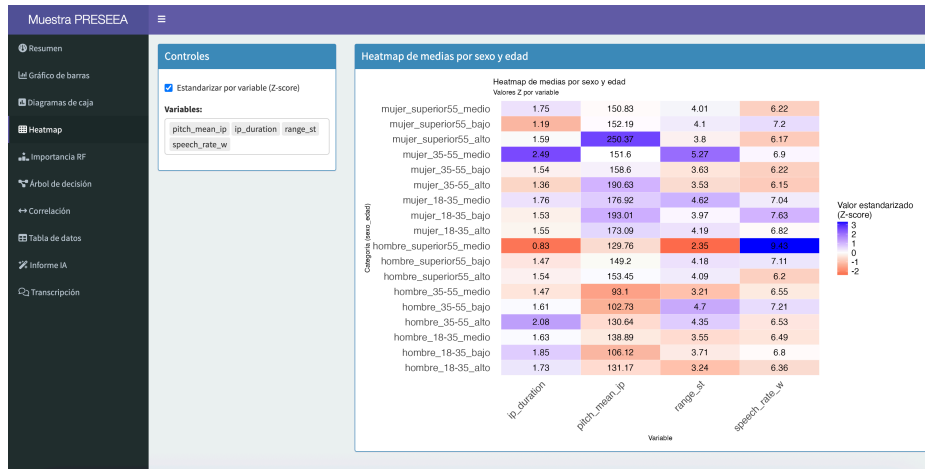


Figure 5: Imagen5

Diapositiva 6

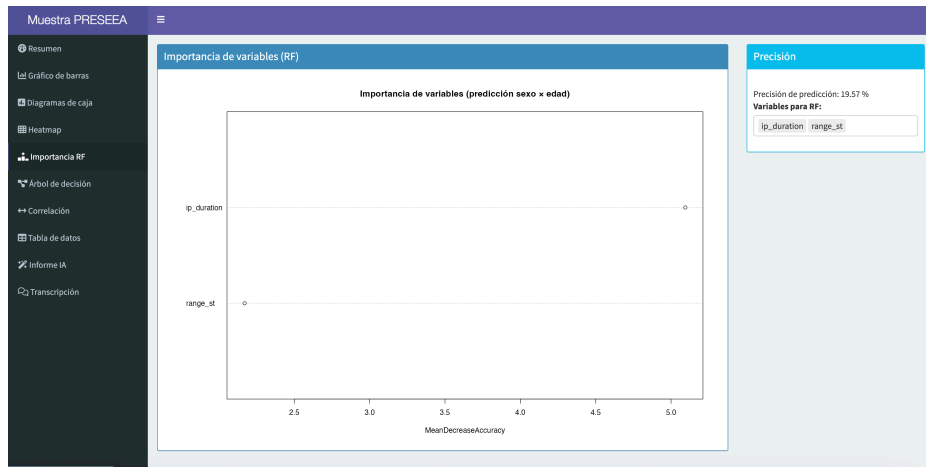


Figure 6: Imagen6

Diapositiva 7

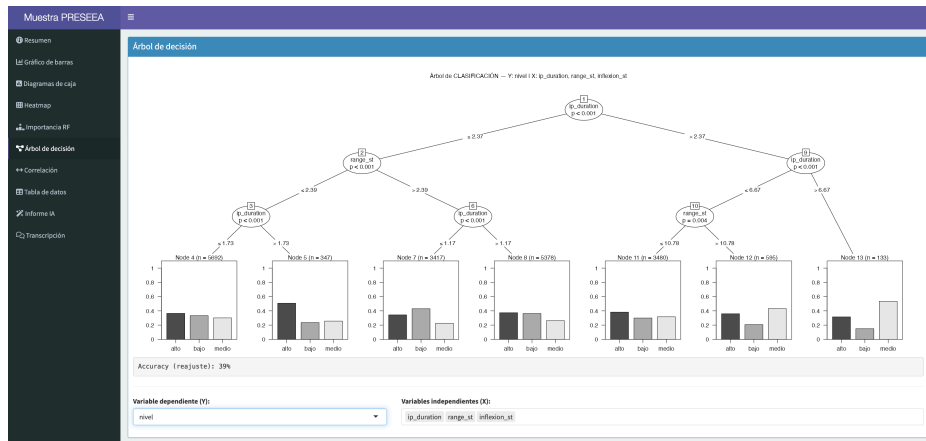


Figure 7: Imagen7

Diapositiva 8

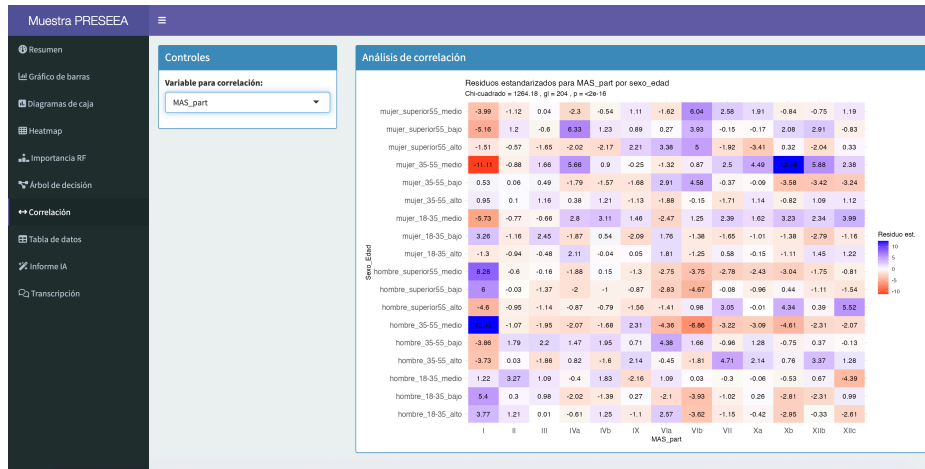


Figure 8: Imagen8

Diapositiva 9

Muestra PRESEEA

Resumen

Gráfico de barras

Diagramas de caja

Heatmap

Importancia RF

Árbol de decisión

Correlación

Tabla de datos

Informe IA

Transcripción

Filtros superiores

Sexo: All

Edad: All

MAS_part: All

TOBI: All

Nivel: All

Filename: All

Nº filas a mostrar: 100

Ordenar por: pitch_mean_ip

☐ Orden descendente

Search:

	filename	ip_id	sex	age	ip_tmin	ip_tmax	spk	ip_text	utterance	MAS_part	TOBI	pitch_mean_ip	ip_duration	range_sl
	All				All	All		All	All			All	All	All
23	VAL_H2M_2020_00119		29 hombre	35-55	82.69	84.73	spk01	pasaba allí los veranos pero ahora ya no	Yo de hecho de pequeño pasaba allí los veranos, pero ahora ya no nos vamos para prácticamente nada.	Xllc	L*+L%	62.86	2.04	
88	VAL_M2M_2020_0016		721 mujer	35-55	2697.49	2698.63	spk01	entonces ya nunca vuelve	entonces ya nunca vuelve nada a ser lo	Xb	L+H*+HL%	68.49	1.14	

Figure 9: Imagen9

Diapositiva 10

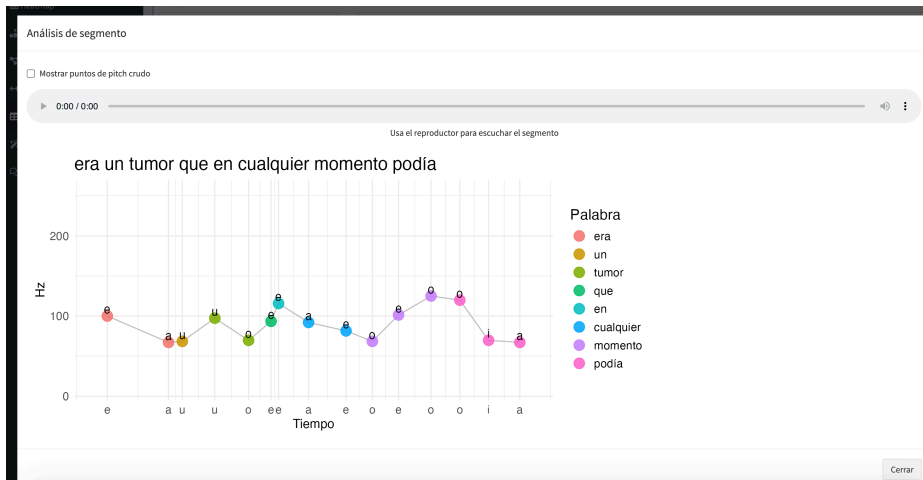


Figure 10: Imagen10

Diapositiva 11

Muestra PRESEEA

Resumen

Gráfico de barras

Diagramas de caja

Heatmap

Importancia RF

Árbol de decisión

Correlación

Tabla de datos

Informe IA

Transcripción

Opciones

Archivo (filename):
VAL_H1A_2020_003
☐ Mostrar 'utterances' (Whisper 3)

Transcripción

Show 50 entriesSearch:

source	filename	spk	start	end	duracion	n_ip	text
ip	VAL_H1A_2020_003	spk00	0.61	1.32	0.71	1	hola qué tal estás
ip	VAL_H1A_2020_003	spk01	2.39	3.12	0.73	1	muy bien muchas gracias
ip	VAL_H1A_2020_003	spk00	4.23	6.34	2.11	2	bueno / le hablo de tú de usted
ip	VAL_H1A_2020_003	spk01	7.8	10.4	2.6	3	me da igual // de tú mismo / más confiado
ip	VAL_H1A_2020_003	spk00	10.98	21.08	10.1	6	vale es que la verdad es que es /// un problema esto de la / forma de tratar /// pero tú por ejemplo / cómo // te diriges a los jóvenes
ip	VAL_H1A_2020_003	spk01	23.11	23.38	0.27	1	de tú
ip	VAL_H1A_2020_003	spk00	24.24	24.8	0.56	1	y a los mayores
ip	VAL_H1A_2020_003	spk01	25.24	27.24	2	2	de usted // siempre que guarda un respeto
ip	VAL_H1A_2020_003	spk00	28.19	34.28	6.09	2	y por ejemplo si tuvieras que preguntar por una calle aunque sea joven // cómo le hablarías
ip	VAL_H1A_2020_003	spk01	35.7	41.65	5.95	4	siempre suelo hablar de tú / a los jóvenes // entonces yo creo que en esa circunstancia // le hablaría de tú

Figure 11: Imagen11

Diapositiva 12

15	hombre 35-55 medio	I (757)	VII (37)	Via (22)
16	mujer 35-55 medio	I (281)	Xb (104)	VII (83)
17	hombre superior55 medio	I (259)	VII (9)	XIic (9)
18	mujer superior55 medio	I (591)	Vib (123)	VII (88)

Showing 1 to 18 of 18 entries

Previous 1 Next

Generación de informe con IA

OpenAI API Key

sk-...

Generar informe con IA

Exportar informe

Título del informe:

Informe PRESEEA Valencia

Autor/a:

Equipo PRESEEA Valencia

Descargar HTML

Descargar PDF

El informe incluye nota: elaborado gracias al proyecto ECOS/C-N, e indica si se usó búsqueda o todo el corpus.

Figure 12: Imagen12

Diapositiva 13

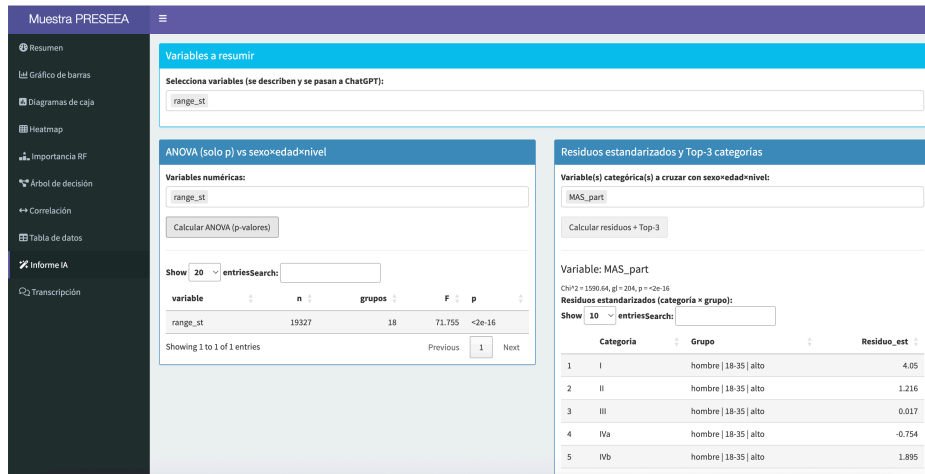


Figure 13: Imagen13