

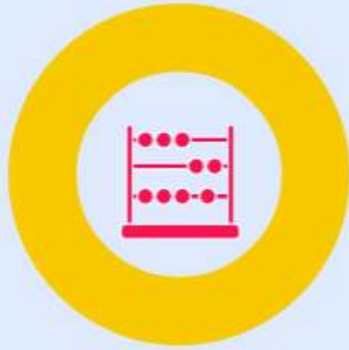
Proyecto: Conteo y clasificación por audio.

Universidad Nacional Autónoma de México
Licenciatura en Ciencia de datos

- Martiñón Luna Jonathan José
- Tapia López José de Jesús



Objetivos



Sensar

A través de la identificación de audio, poder estimar la cantidad de personas que se encuentren en ciertas áreas



Agrupar

Separar la cantidad de personas en: Hombres y Mujeres



Simplificar

Dejar de lado el uso de sensores o contadores manuales y dejar el trabajo a un simple micrófono.

Utilidad



Conocimiento sobre horarios con mayor afluencia



Reconocimiento sobre la población que se alcanza



Información sobre el límite de afluencia (Hablando de la pandemia)

Ventajas y desventajas



Ética

Tener un micrófono que a pesar de tener como objetivo sensar, al mismo tiempo está grabando conversaciones



Facilidad de uso

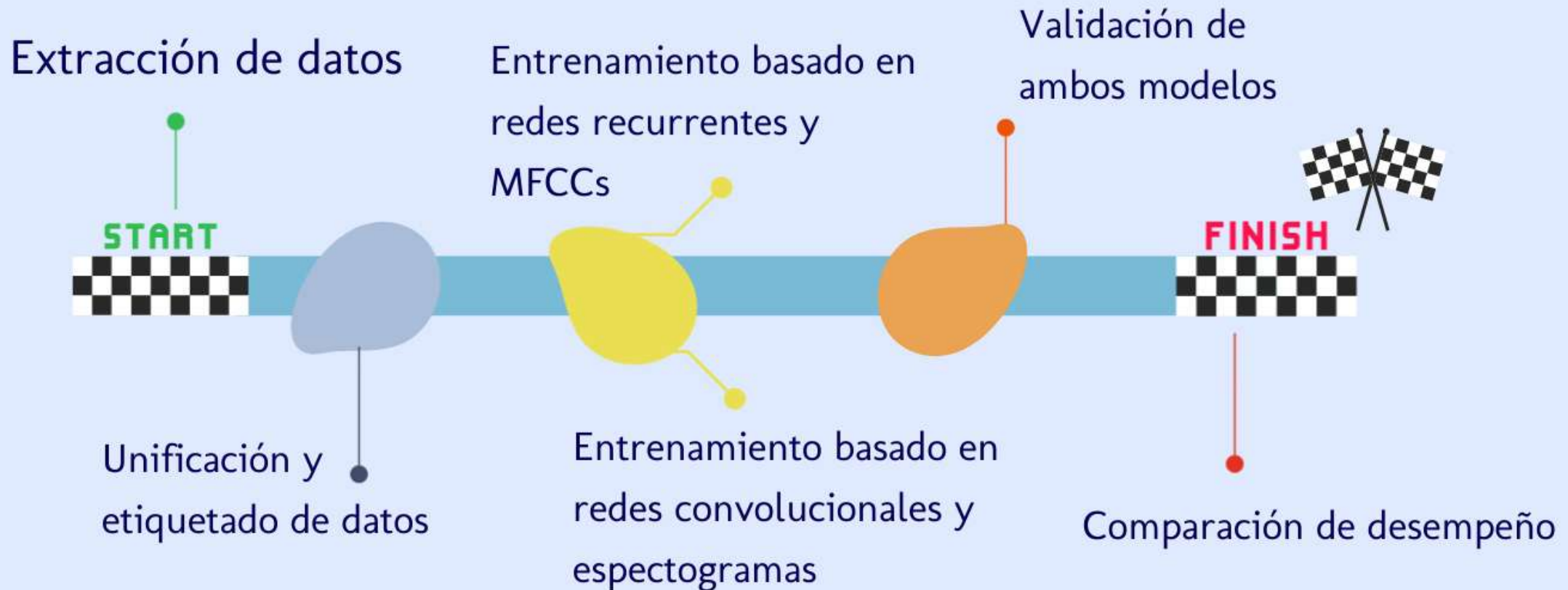
Resulta menos complicado trabajar con un micrófono y un modelo que instalar sensores o contar manualmente.

Información

Proporciona información de sectores (Hombres y mujeres) que habitúan un área



Forma de Trabajo



1

Se descargó el conjunto de audios individuales de:

<https://www.openslr.org/12>

OpenSLR

[Home](#) [Resources](#)

LibriSpeech ASR corpus

Identifier: SLR12

Summary: Large-scale (1000 hours) corpus of read English speech

Category: Speech

License: CC BY 4.0

Downloads (use a mirror closer to you):

[dev-clean.tar.gz](#) [337M] (development set, "clean" speech) Mirrors: [\[China\]](#)

[dev-other.tar.gz](#) [314M] (development set, "other", more challenging, speech) Mirrors: [\[China\]](#)

[test-clean.tar.gz](#) [346M] (test set, "clean" speech) Mirrors: [\[China\]](#)

[test-other.tar.gz](#) [328M] (test set, "other" speech) Mirrors: [\[China\]](#)

[train-clean-100.tar.gz](#) [6.3G] (training set of 100 hours "clean" speech) Mirrors: [\[China\]](#)

[train-clean-360.tar.gz](#) [23G] (training set of 360 hours "clean" speech) Mirrors: [\[China\]](#)

Datos

- Duración: Variada
- Idioma: Inglés

Audios

	waveform	Sample rate	Utterance	Speaker_id	Chapter_id	Utterancia id
Tipo	Tensor	Int	String	Int	Int	Int
Ejemplo	tensor([[0.0003 ,..., 0.0016]])	16000	"He Hoped..."	1089	134686	0

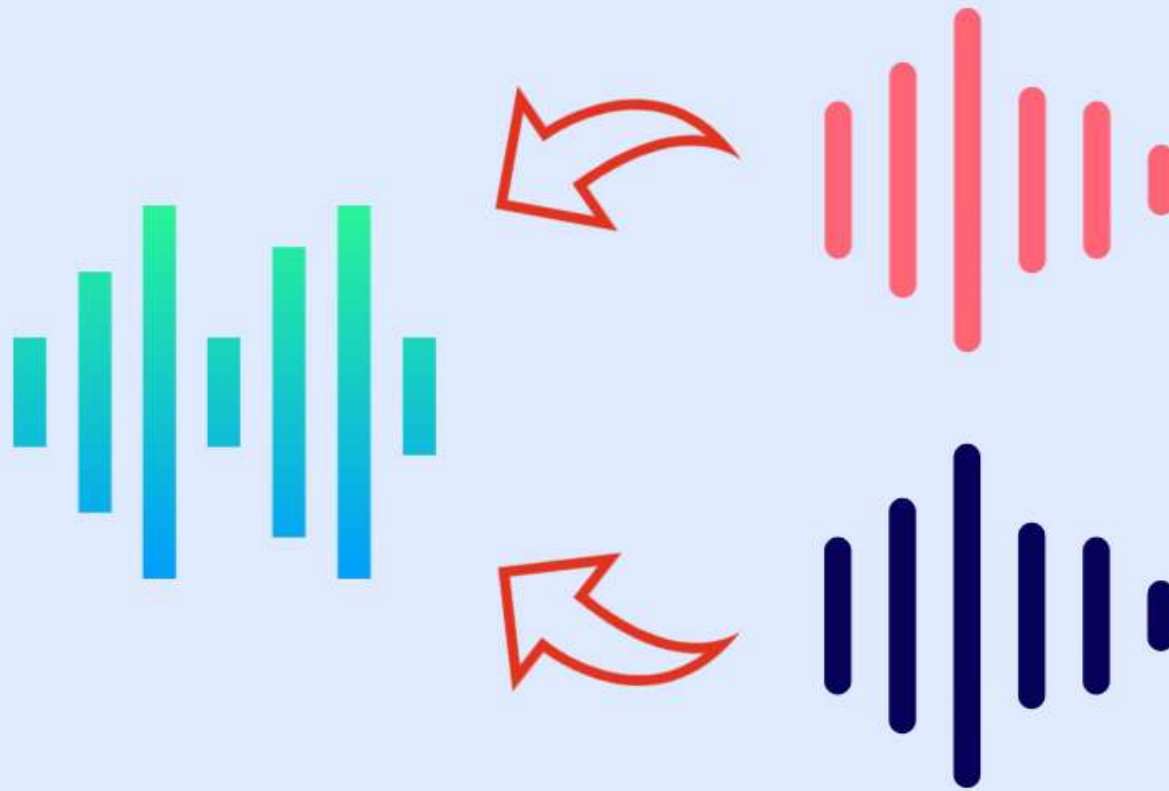
Datos

Metadatos

	ID	Sexo	Conjunto	Duración	Nombre
Tipo	Int	Bool (F,M)	Texto	Float	Text
Ejemplo	16	F	train-clean-360	25.03	Kristin LeMoine

2

Mezcla y etiquetado de audio Hombre - Mujer





Separación en:

- Entrenamiento
- Validación
- Prueba

	Entrenamiento	Validación	Prueba
Peso	3.5 GB	364 MB	364 MB
Registros	100,000	10,000	10,000

Datos



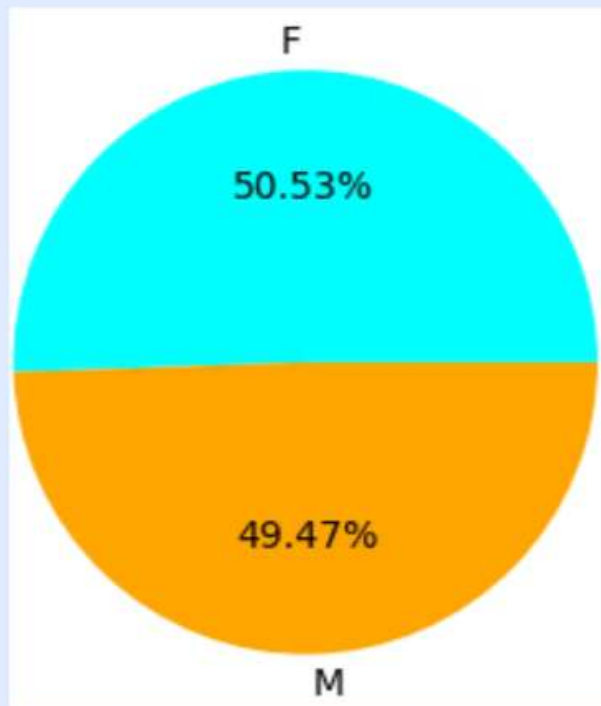


Ejemplo

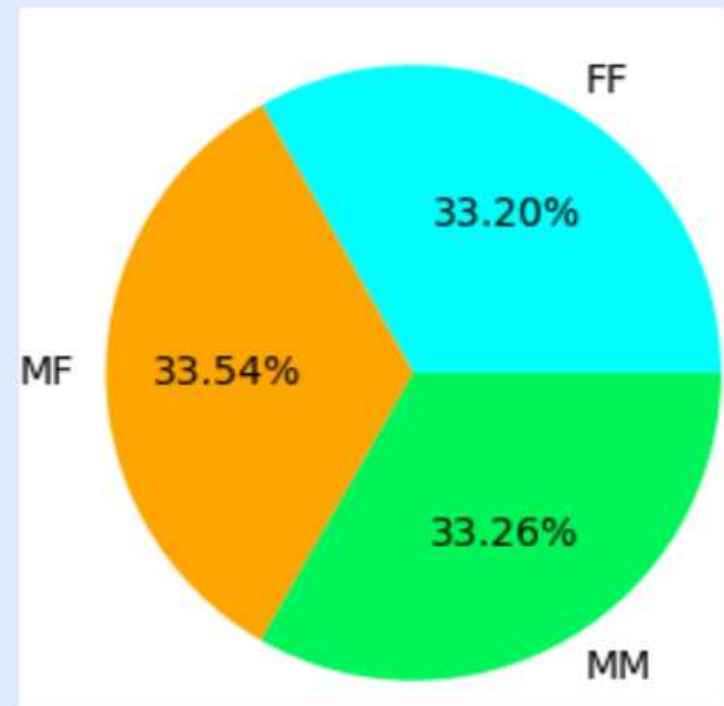
Datos

	Waveform	Speakers	Speakers sex	F	M
Registro	Tensor[[0.000 1,...,0.003]]	3	MMF	1	2

Distribución en Train Hombres - Mujeres

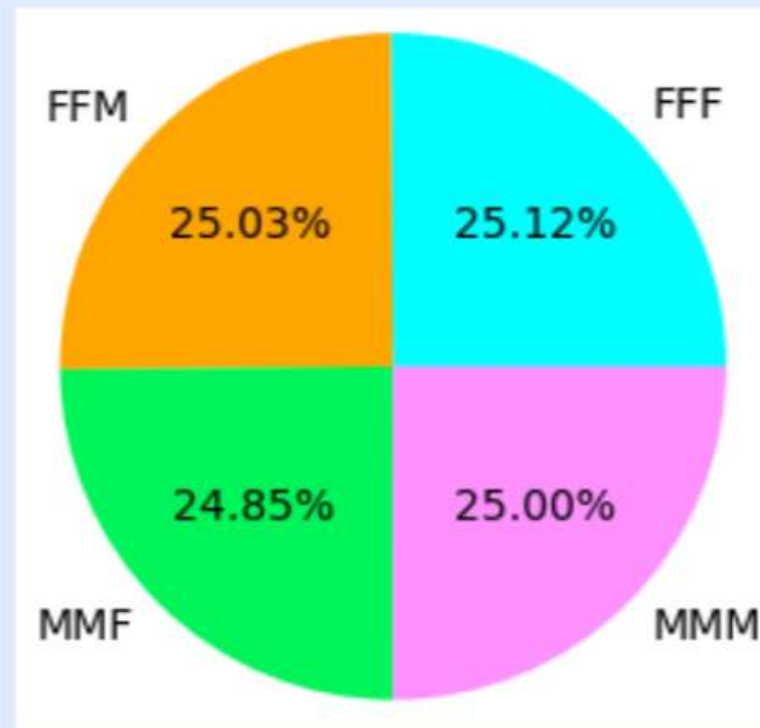


Datos



Distribución en Train Hombres - Mujeres

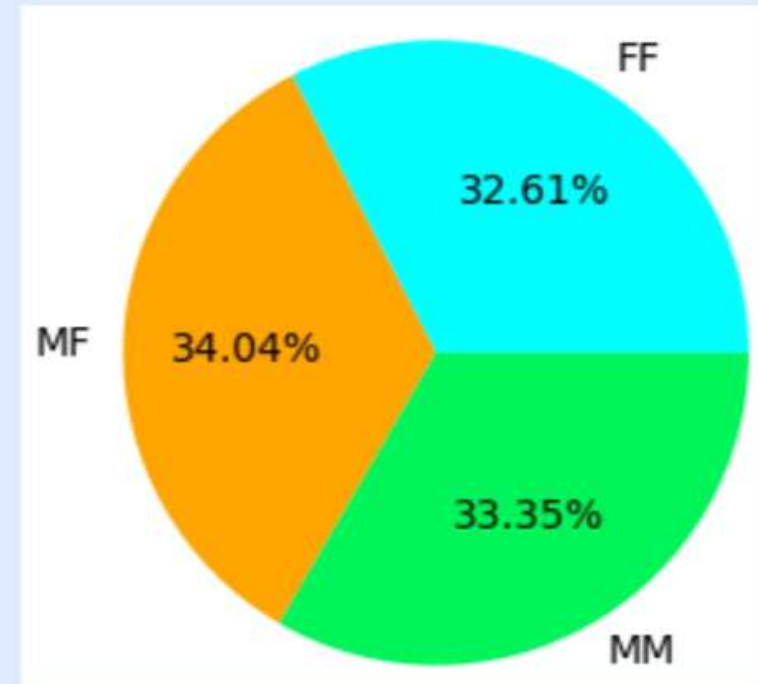
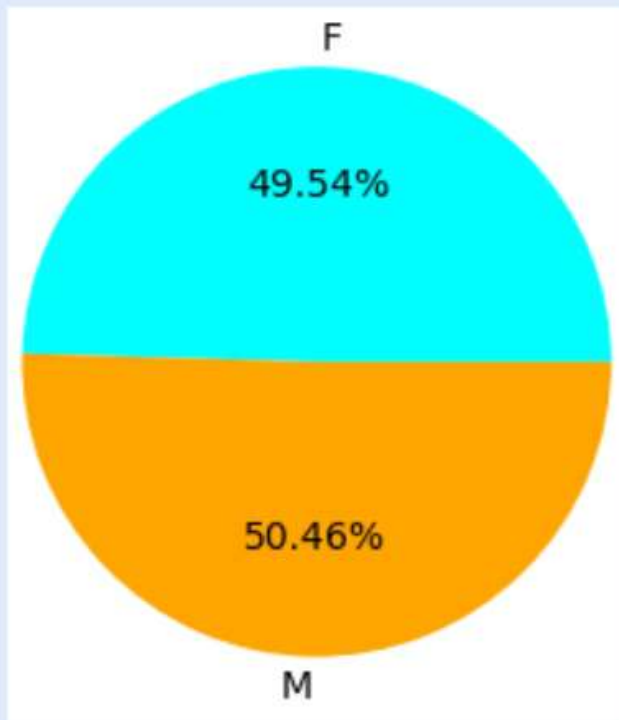
Datos





Distribución en Validación Hombres - Mujeres

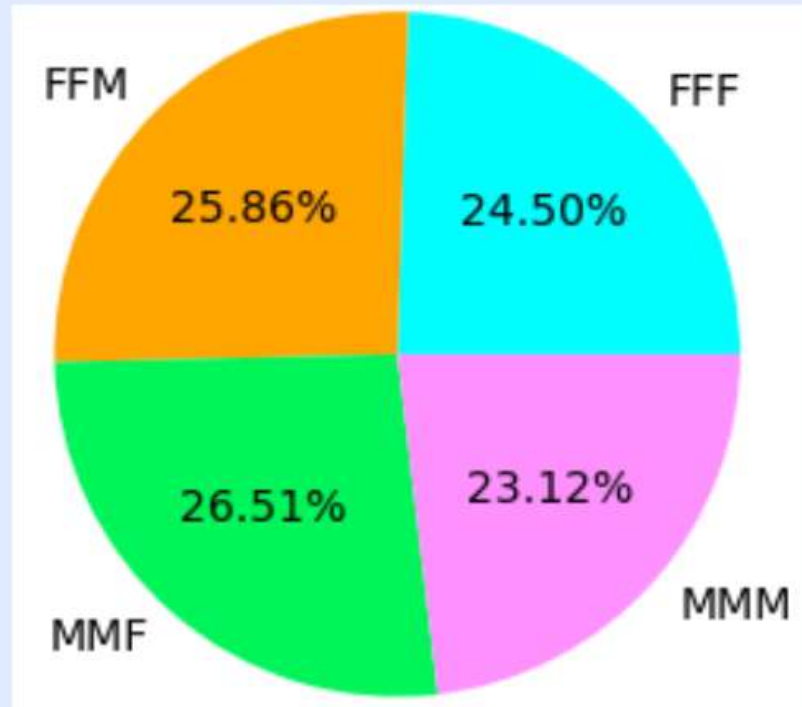
Datos





Distribución en Validación Hombres - Mujeres

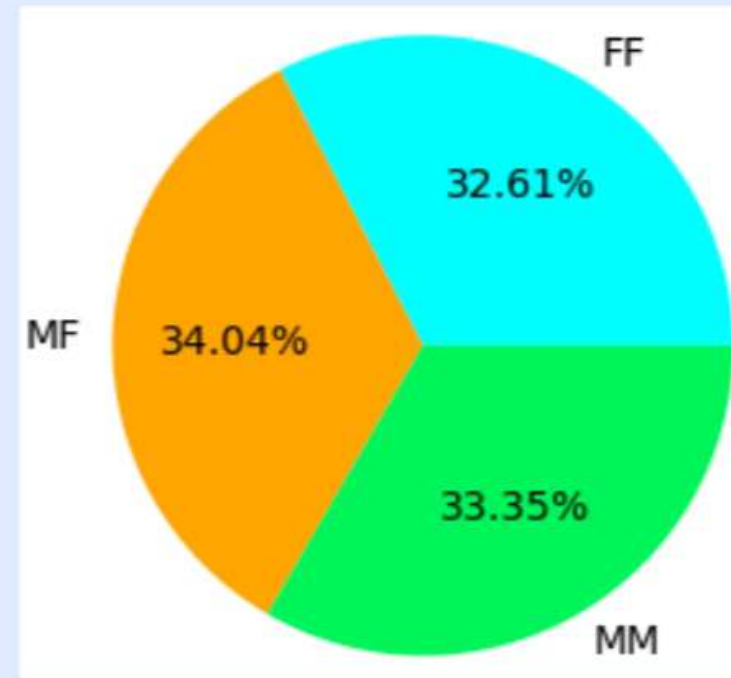
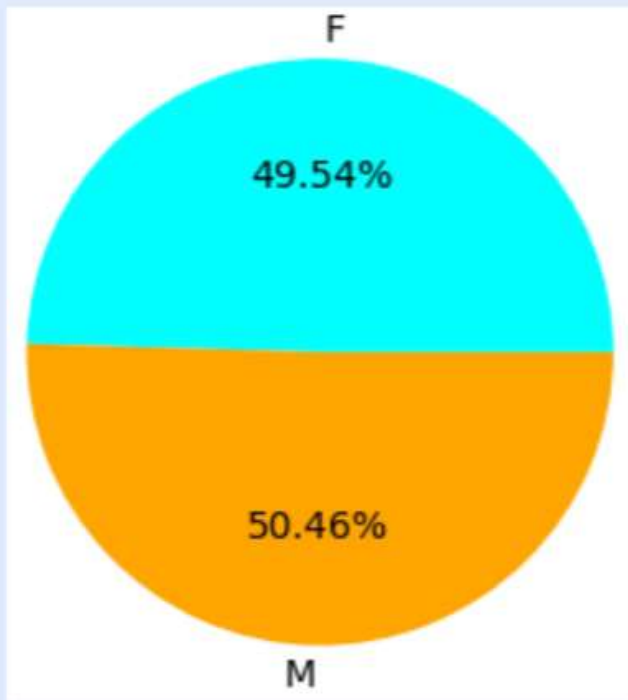
Datos





Distribución en Prueba Hombres - Mujeres

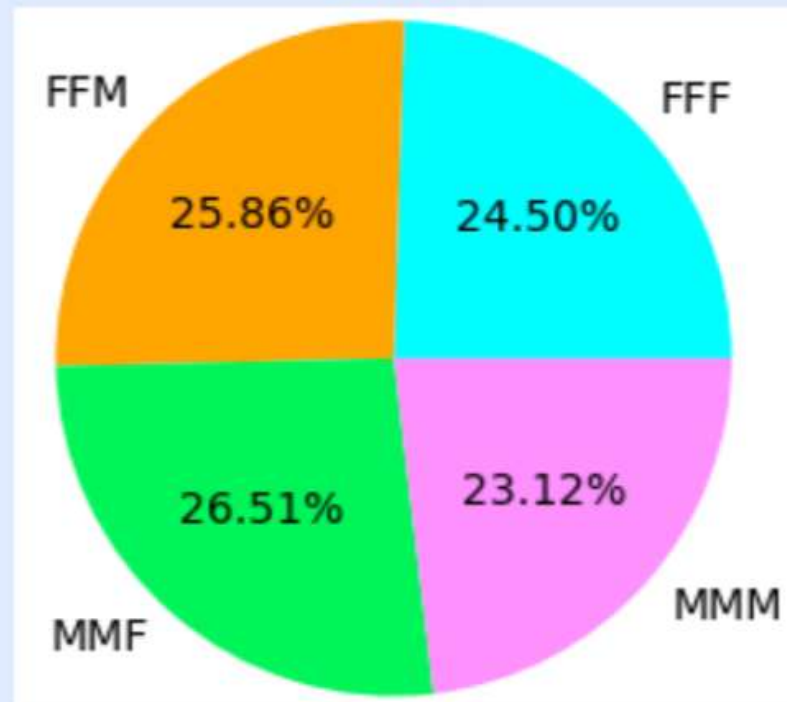
Datos





Distribución en Prueba Hombres - Mujeres

Datos



Seguimos trabajando...

