

# !Another DS Group

## Findings

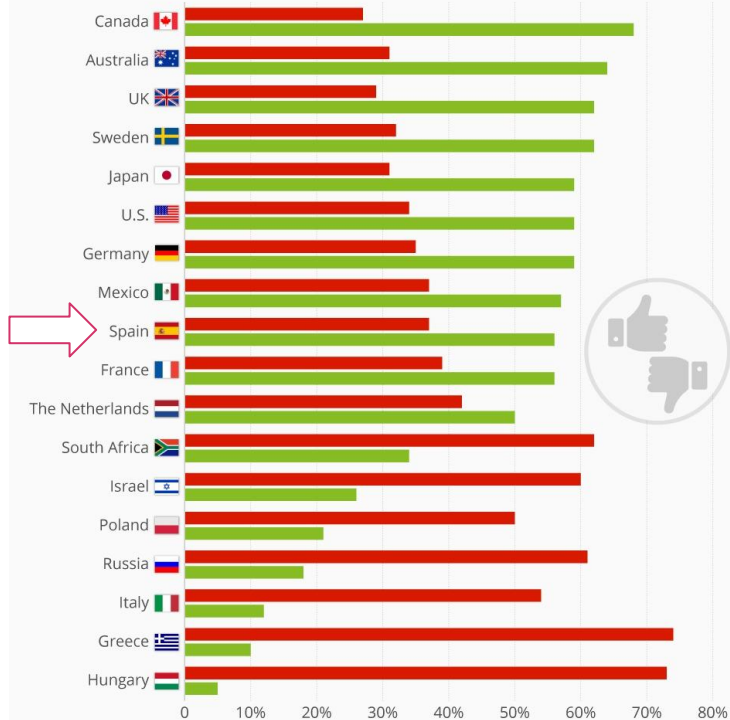


How do we  
understand **racism**?

## More Immigrant Countries See Diversity Positively

Share of citizens in top immigrant destinations welcoming and rejecting newcomers\*

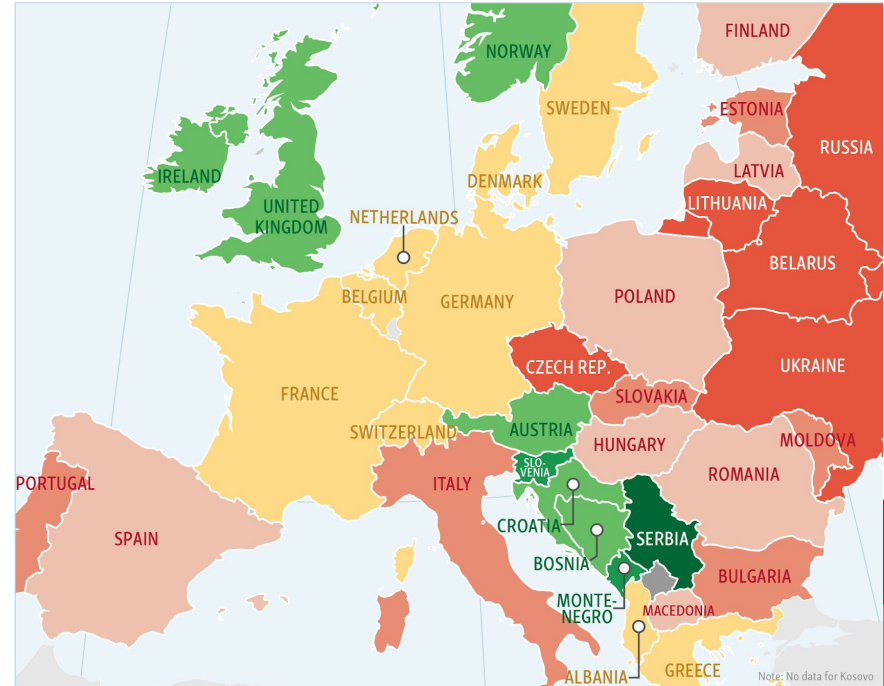
■ Immigrants are a burden on our country ■ Immigrants make our country stronger



Source: Elaborated by Statista, with Pew Research's data.

## Racism in the Western world

Dark skin is associated with...  
 More **positive** thoughts\*    .275    .350    .425    More **negative** thoughts  
 (Mean IAT score of white participants\*\*)

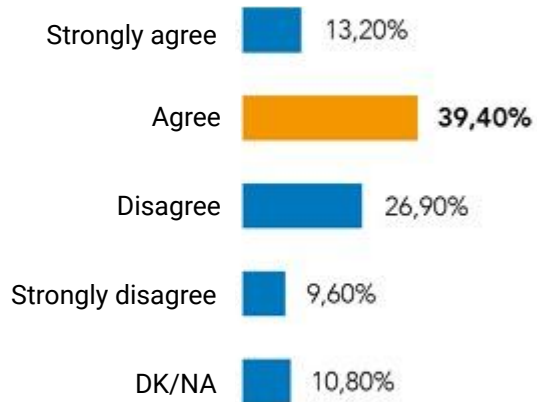


Source: Elaborated by RadioFreeEurope, with Project Implicit by Harvard University's data.

# Racism in the Hispanic world

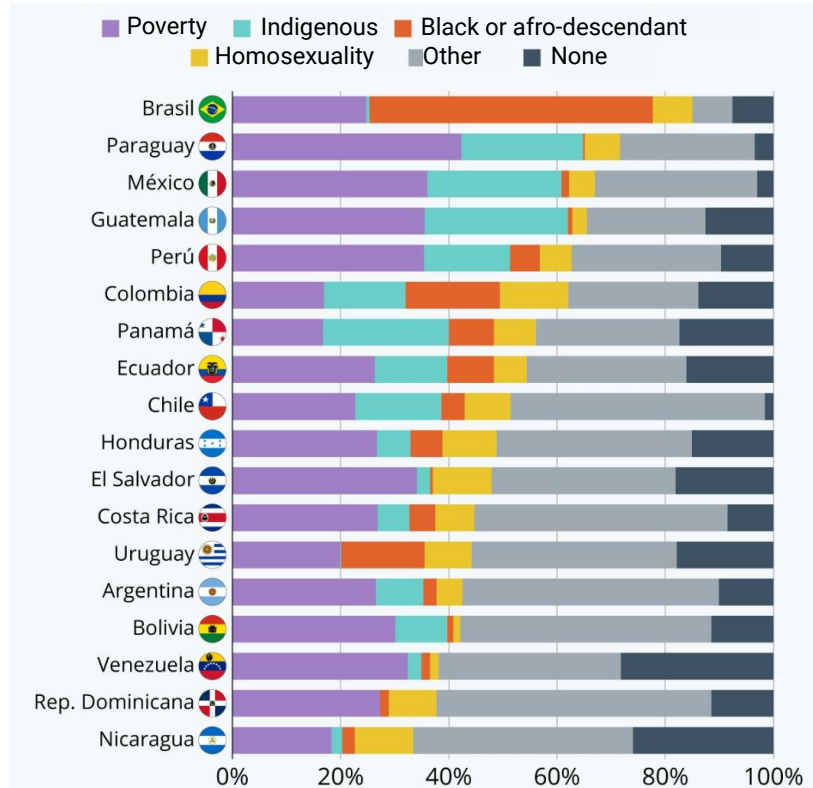
Percentage of agreement with the stereotype and fake news about the relationship between crime and migrants of the statement:

Spaniards **believe** that immigrants are the main cause of crime



Answers to question:

Which individuals or groups of people **do you think** are the **most discriminated** against in your country?



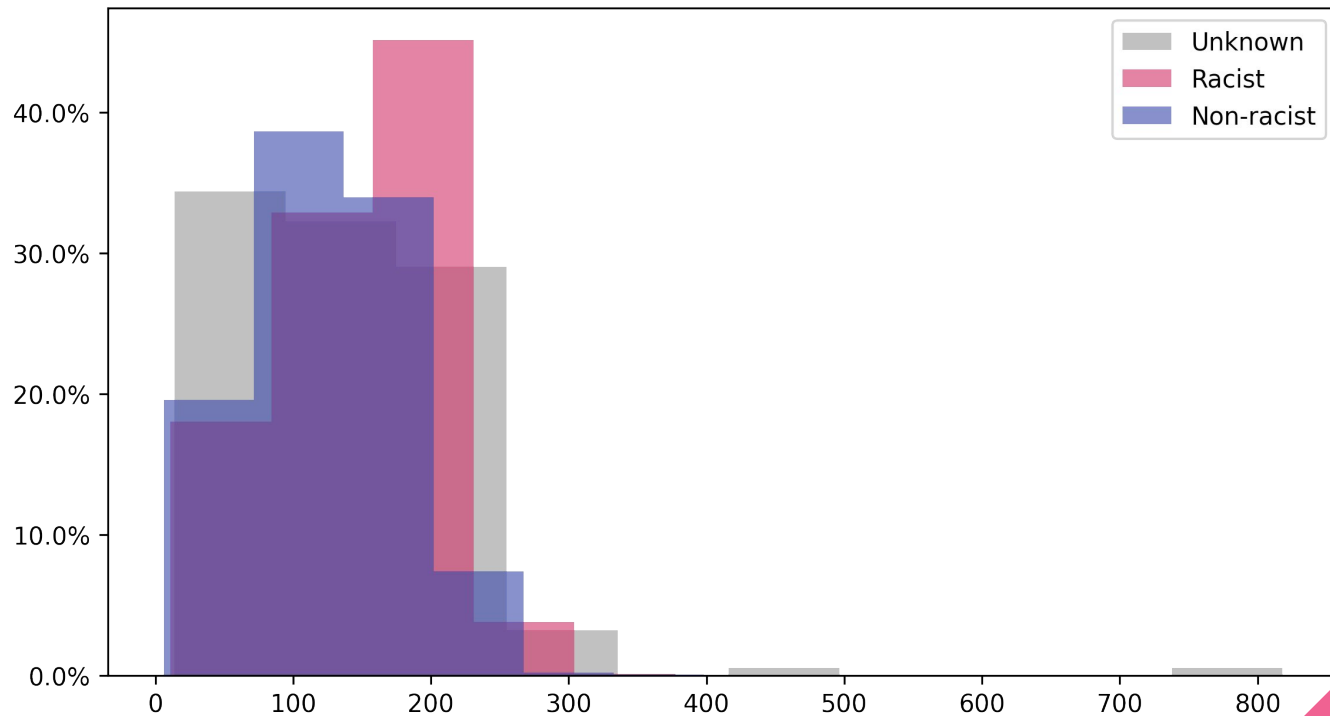
Source: Elaborated by Statista, with Latinobarómetro's data.

# Why this matters for Tweets?

- About racism:
  - Social **alarm** regarding relationship between immigration and crime is **triggered** by **media** (Bañón Hernández, 2003).
  - Racism and Xenophobia are **intertwined** (Mercedes-Fernández et al., 2017).
- About inherent bias in tweets:
  - Twitter API gives 1% of a **non-representative** universe of tweets.
  - Topics like economy + racism would not be properly reflected in tweet extractions.
  - Account **limited-windows** (Essential, Expanded or Academic access).

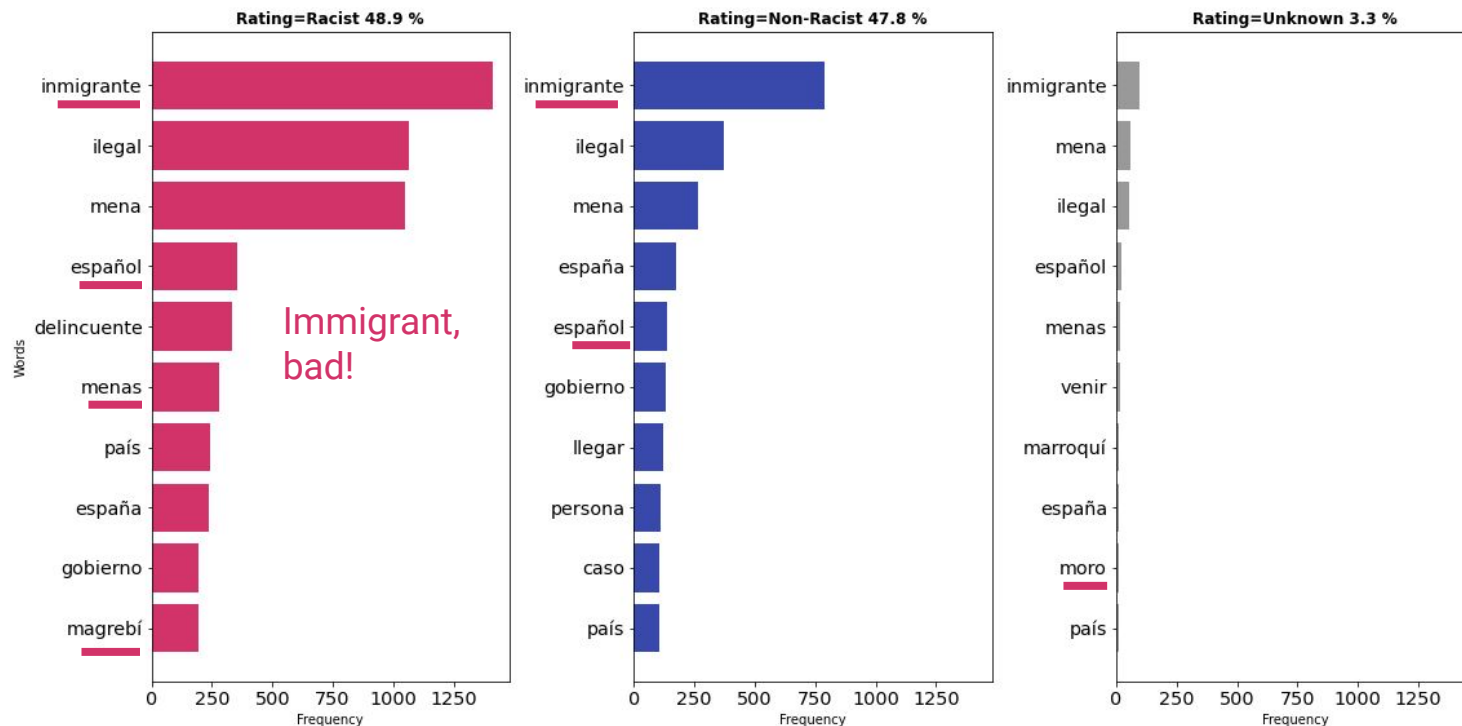


# Racist tweets are large!

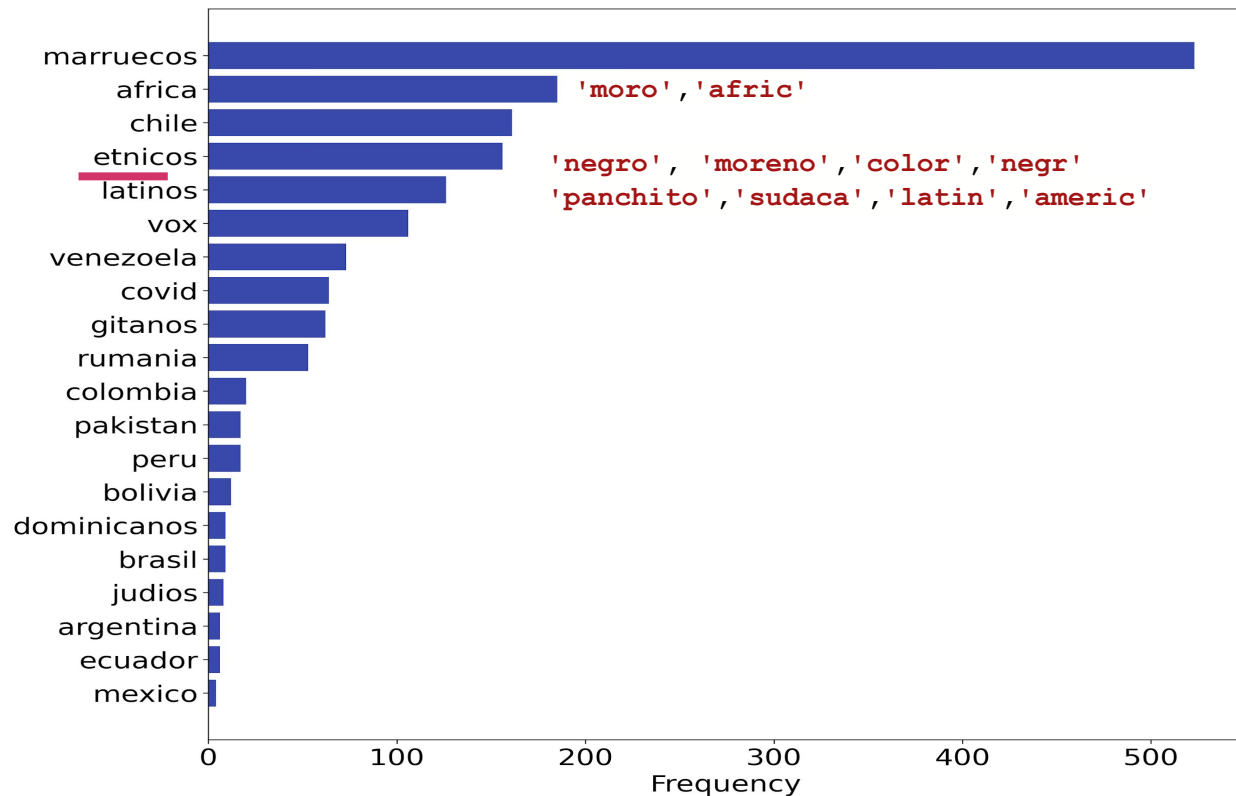


# Most frequent racist words in tweets - BCN Analytics

5672 original tweets



# Racism or Xenophobia?



There are more  
nationality-terms than  
ethnics

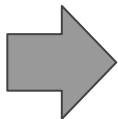
And sometimes are  
mixed!



# Are these tweets really racist, though?

16% of labels are controversial!

Original Unique Tweets Classification	Support Label
All racist	Definitely racist
Racist ~ non-racist	Maybe racist
Racist > non-racist	Maybe racist
Racist + unknown	Maybe racist
Racist < non-racist	Maybe not racist
Non-racist + unknown	Maybe not racist
All non-racist	Definitely not racist
All Unknown	Unknown

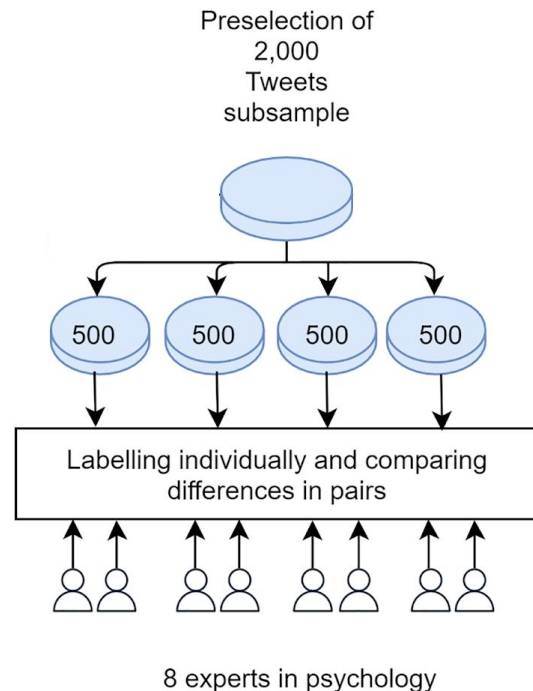


Final Label	Final consensus	Size	%
Racist	Definitely racist	2278	<b>40.1</b>
	Maybe racist	498	<b>8.8</b>
Non-racist	Maybe not racist	215	<b>3.8</b>
	Definitely not racist	2495	<b>47.0</b>
	Unknown	186	<b>3.3</b>

Be wary of  
F1-score >90%

# BETO's Research on Racism and Xenophobia

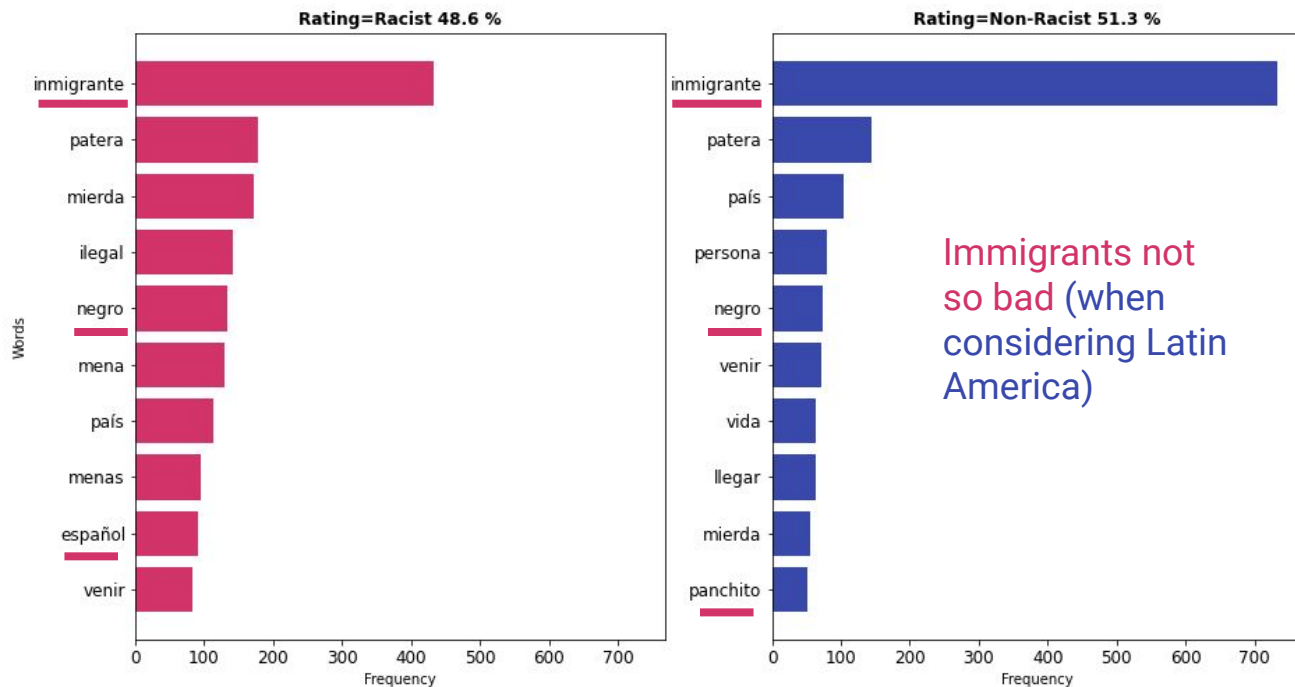
- Research done by Benítez-Andrades et al., 2022.
- Developed BETO: a transformer BERT-based model called BETO, using only Spanish text.
- **Iberoamerican** Spanish.
- **85.1%** of maximum **f1-score** obtained.



**Better data > Better model**

# Most frequent racist words - BETO's dataset

2000 original tweets



# What do our models say?

# Methodology I - (Started from Rule-based)

Simple rules

Considered racist if word  $x$  appear in tweets.

Where  $x$  could be any of:

- **Colectives:** negros, negras, chinos, chinas, mena, menas, catalanes, inmigrantes, rumanos, puertorriqueños, marroquíes, refugiados, ...
- **Words:** ilegal, moros, islamista, problemas, terror, vienen, chusma, delincuentes, criminales, violadores, paliza, agreden, expulsar, ...
- **Hashtags** = #menasfuera, #lañestáenpeligro, #vox, #ilegales, #okupas, ...

84.7% F1 in test

**72.9%** F1 in public <- baseline

# Methodology II - Models

- Ensemble pipeline: BERT + BETO + XGBoost to classify if a tweet is racist or not.
- NLTK and Spacy lemmatizers and stemmers for pre-processing.
- Used either word embeddings or tf-idf vectors.



# What do people say in racist tweets?

	Topic	Associated tokens
01	Immigrante = invasor que entra por Canarias	<ul style="list-style-type: none"><li>Immigrante, ilegal, llegar, canarias, entrar</li></ul>
02	Menas = carga económica para los españoles	<ul style="list-style-type: none"><li>Menas, dinero, mantener, españa, español</li></ul>
03	Latinoamericano = okupa y delincuente	<ul style="list-style-type: none"><li>Delincuente, país, chile, venezolano, okupa</li></ul>
04	Marroquí = violador y golpeador	<ul style="list-style-type: none"><li>Magrebí, violar, marroquí, chico, paliza</li></ul>

## Most used hashtags in racist tweets




Vox-related  
hashtags  
appear among  
racist tweets





# Conclusions

- A lot of **doubt** in defining what is racism even by professionals.
  - In **Spain** racism seems to be related with **immigration** while in **LatinoAmerica** racism is associated with **discrimination** by race.
  - The nationality that is most discriminated is **moroccan** and there is a general negative attitude towards **muslims**.
  - People of discriminated nationalities are recurrently associated by Spaniards to violence, illegal immigration, illegal house occupants and poverty.
  - Better **division** of train/val/test data.
- 

# References

- Benítez-Andrades et al. (2022), PeerJ Comput. Sci., DOI 10.7717/peerj-cs.906
- FERNÁNDEZ, M., VALBUENA, C., & CARO, R. (2016). Evolución del racismo, la xenofobia y otras formas de intolerancia en España. Informe 2016. OBERAXE. Ministerio de Trabajo e Inmigración, España.
- Red Acoge (2018), La autopercepción de discriminación. El impacto de la mirada del otro. Ministerio de Sanidad, Servicios Sociales e Igualdad, España.



# Thank you!

Q&A

---

!Another DS Group