Identifying content topics of opportunity for **PlateaPR**

presented by **Tania Hernandez Colon**

Goal

- Better understand our audience.
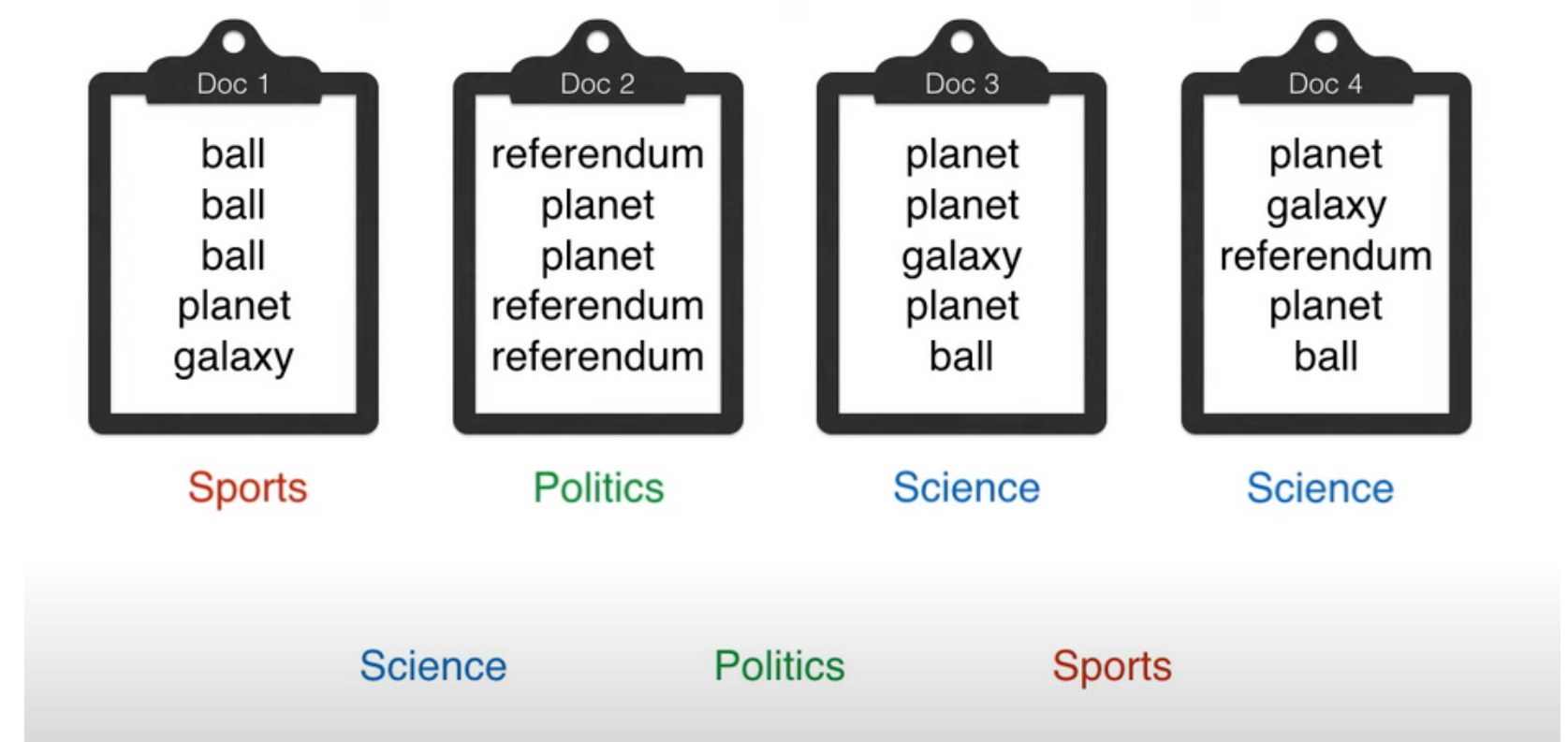- Identify topics resonates the most with the audience.

Approach

- Obtain the blog articles and sessions associated with those
- Classify the posts with topics using Latent Dirichlet Allocation (LDA)
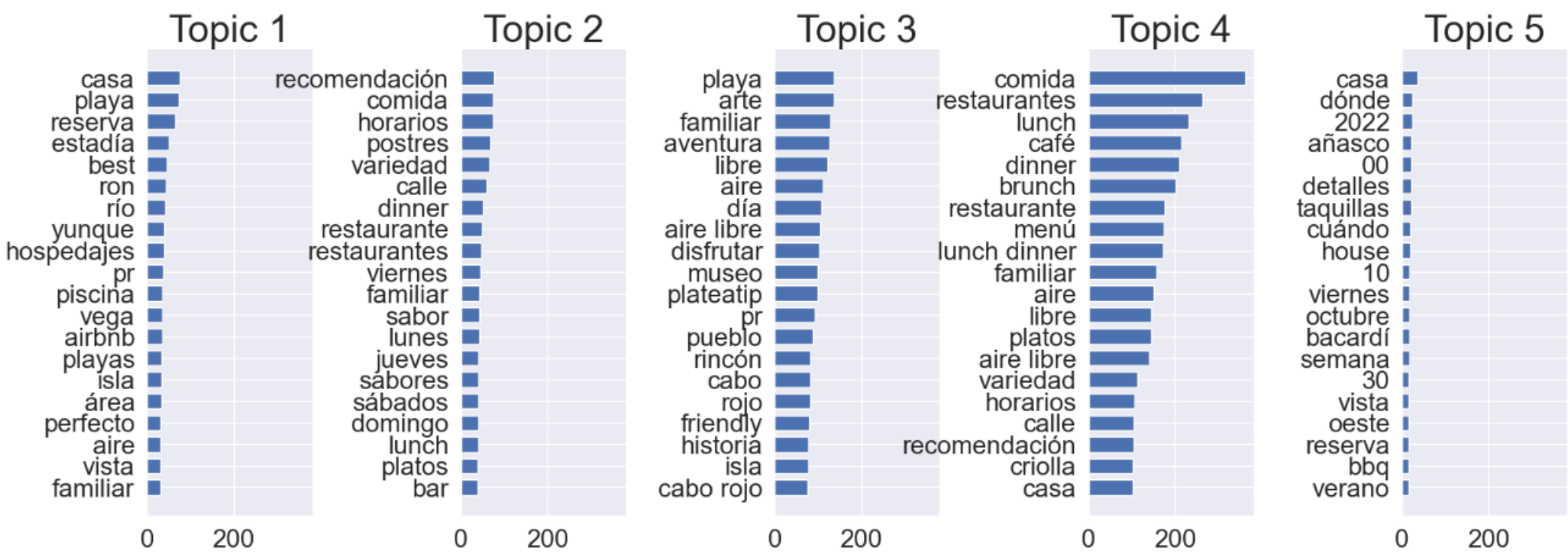- Identify which topics bring more sessions

LDA

At a high level the LDA attempts to model two probability distributions given data that is "words present in documents".

1. The probability of a topic over words
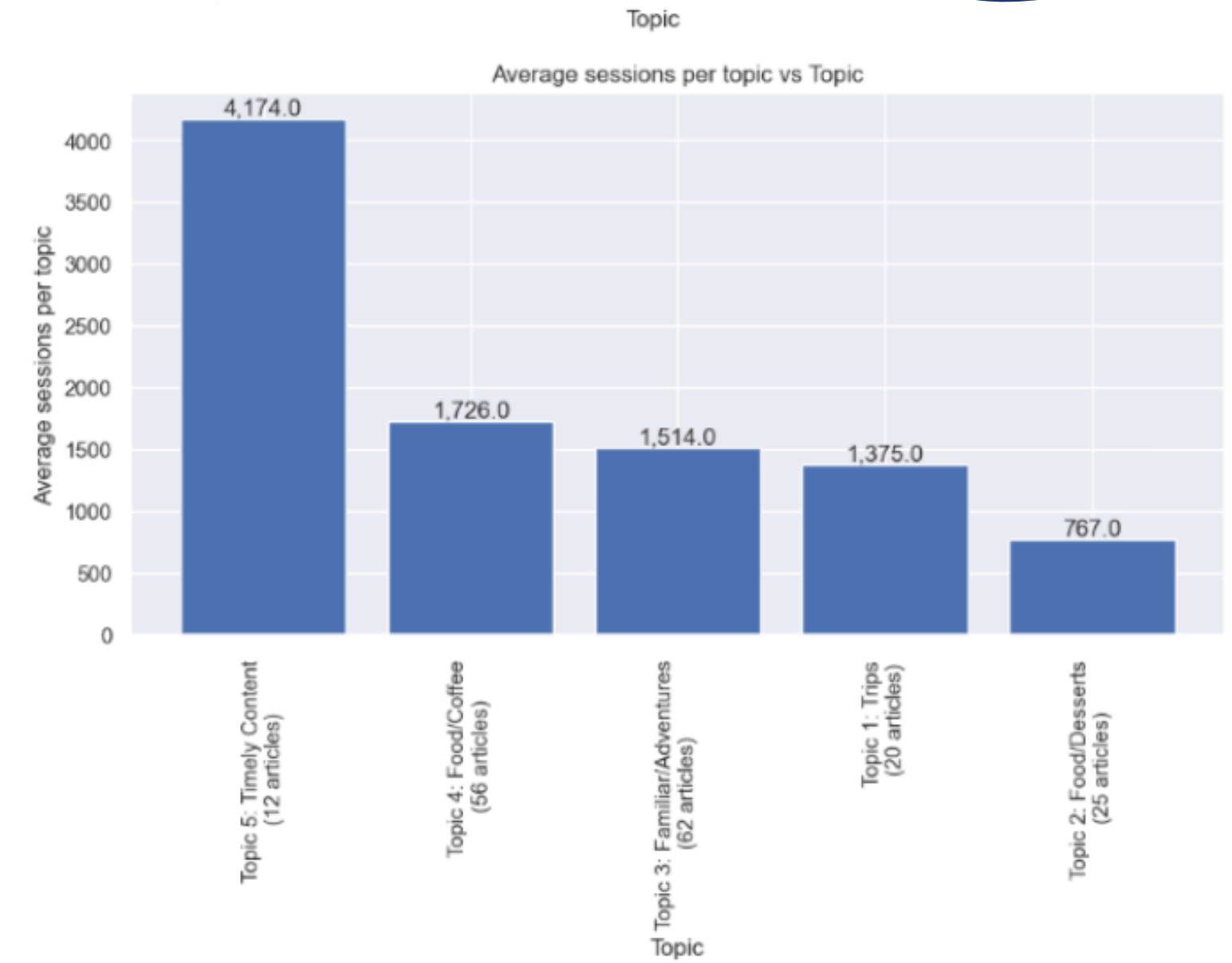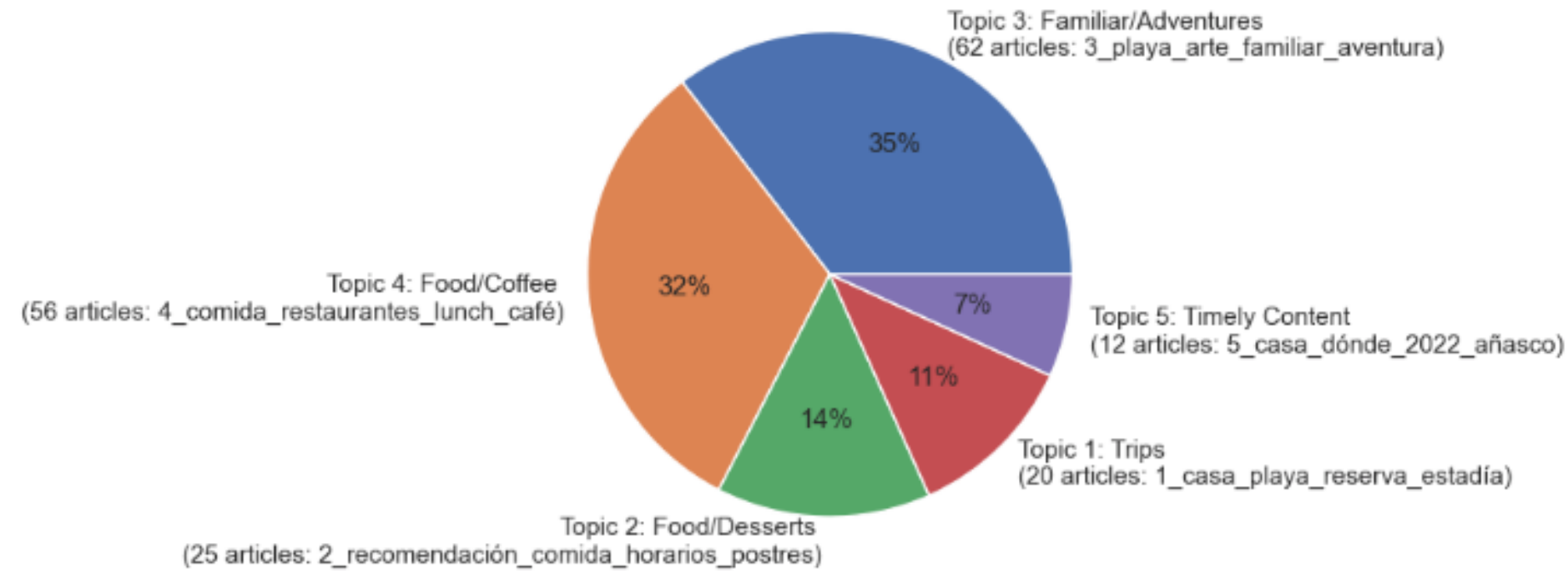2. The probability of a document over topics

Results

Topic 1	Featuring house, beach, reserve, stay. A broader topic could be "Trips".
Topic 2	Featuring recommendation, food, hours, desserts. A broader topic could be "Food/Desserts"
Topic 3	Featuring beach, art, family, adventures, outdoors. A broader topic could be "Family/ Adventures"
Topic 4	Featuring food, restaurants, lunch, coffee dinner "Food/Coffee"
Topic 5	Featuring the words house, where, 2022. Since this chart had words that did not make a lot of sense together, we proceeded to take a look the topics to see if there is a trend. A trend was identified of being content related to events and holidays and therefore we called it "Timely content"

Results

Questions?