

Predicting women's preferences for caesarian delivery from online forum discussions

Michele Tizzoni

@mtizzoni

ISI Foundation



INSTITUTE
FOR SCIENTIFIC INTERCHANGE
FOUNDATION



AMLD

January 27, 2020



C-Section: a public health concern

- Optimising the use of caesarean section (CS) is of global concern.
- **Underuse** leads to maternal and perinatal mortality and morbidity.
- **Overuse** of CS has not shown benefits and can create harm.
- Worldwide, **the frequency of CS continues to increase**, and interventions to reduce unnecessary CSs have shown little success.
- Identifying the underlying factors for the continuing increase in CS use could improve the efficacy of interventions.

*source: Betrán, Ana Pilar, et al. "Interventions to reduce unnecessary caesarean sections in healthy women and babies." *The Lancet* 392.10155 (2018): 1358-1368.*

- Of the 6.2 million unnecessary caesareans done each year, 50% are in Brazil and China.

- Brazil also has the 10th highest number of preterm births in the world.

C-Section in Brazil





UNICEF Campaign

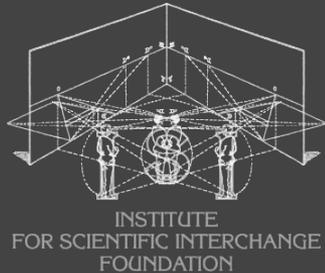
- April 19 - June 20, 2017
- a series of informative videos
- a dedicated website with educational materials and Frequently Asked Questions,
- Facebook posts, promoted through the UNICEF Brazil Facebook page
- printed material and advertising boards.

<https://www.unicef.org/brazil/relatorios/quem-espera-espera>

Team

unicef  | for every child

Natalia Adler
Paulo Henrique Alves



Alberto Ceria
André Panisson
Daniela Paolotti
Kyriaki Kalimeri

The Dataset



- BabyCenter is a popular website in Brazil that provides information on **conception, pregnancy and maternity** in general
- The app version of the platform has been downloaded by more than **500,000 users**
- The dataset consisted of comments related to threads which contain the words **“C-Section”** or **“childbirth”**

Thread Name



TENTANTES GRAVIDEZ BEBÊ CRIANÇA SUA VIDA ENXOVAL, QUARTO & FESTAS NOMES COMUNIDADE



Bebês do mesmo mês

[Ver todos](#)

Categorias

[Índice dos grupos](#)

[Tentantes](#)

[Gravidez](#)

[Bebês](#)

[Fotos e vídeos](#)

[Nomes de bebê](#)

[Pais](#)

[Gêmeos ou mais](#)

[Alimentação](#)

[Compras e enxoval](#)

[Maiorzinhos](#)

[Necessidades especiais](#)

[Amigos do BabyCenter](#)

[dizQdiz de mãe](#)

Cesárea



Lehhsimplicio

19/07/17

Thread Creator

Meninas vocês que vão fazer cesárea, quanto custou ?

♥ 0



Compartilhar



Thread Content

(5) comentários

[Comentar este tópico](#)



LulukaRS

19/07/17

Comment Creator

No meu caso vou fazer pelo plano! Não vai custar nada,... acho q tem um tópico mais antigo sobre preço das cesarianas,... de repente da uma procurada nos tópicos anteriores,... talvez ajude!

Responder

♥ 1

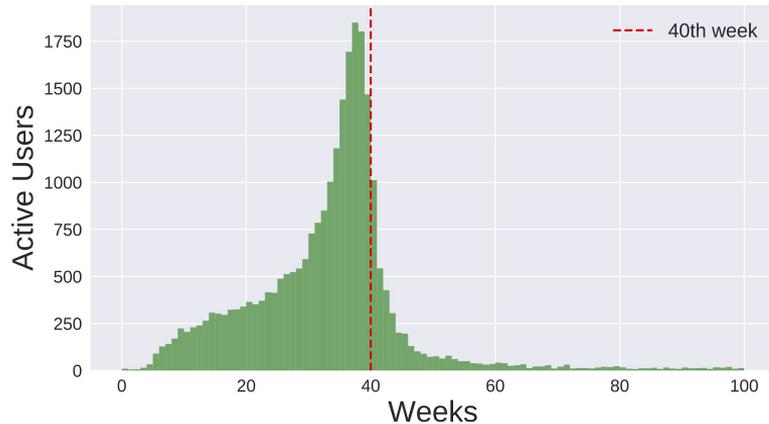
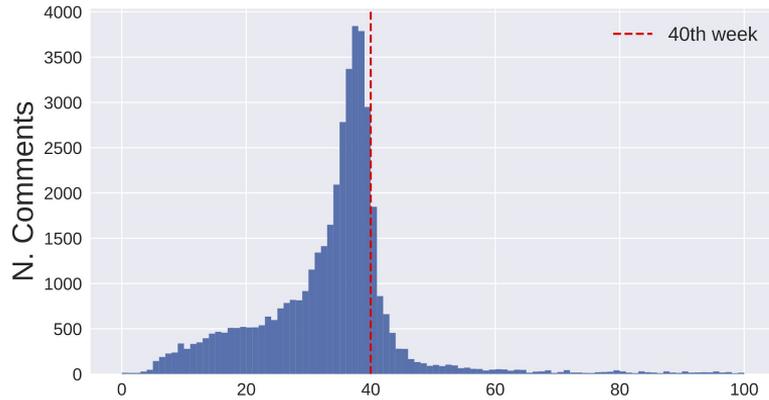


Comment Content

Objectives

- Analysis of the overall volume and the temporal evolution of term frequencies in posts published along the gestation period.
- Differences in the evolution of term frequencies between C-Section and natural childbirth users.
- Measure differences in term frequencies to identify the moment in which the delivery choice is taken.
- Predict choice of delivery based on term frequencies.

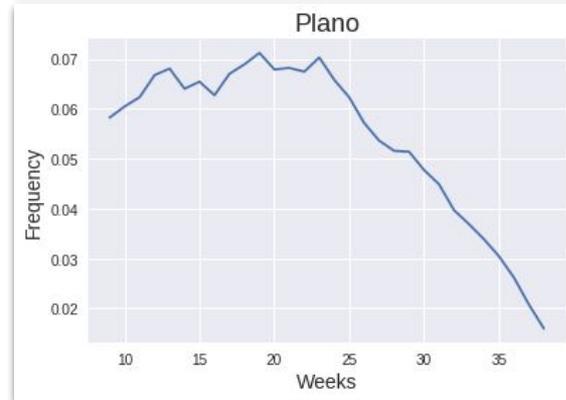
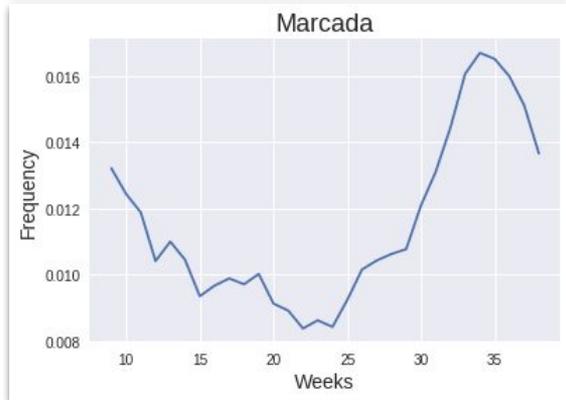
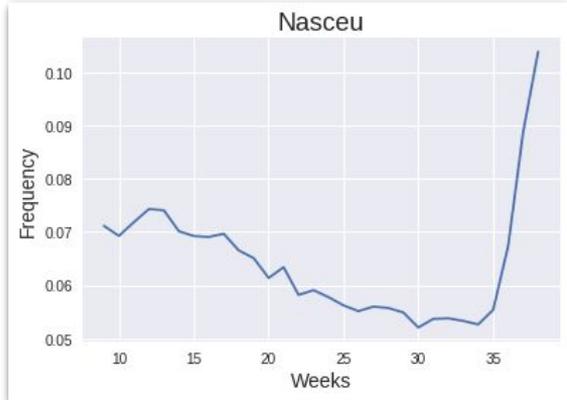
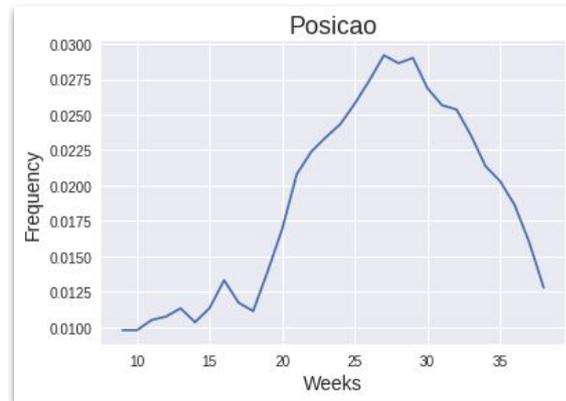
Activity along the gestation period



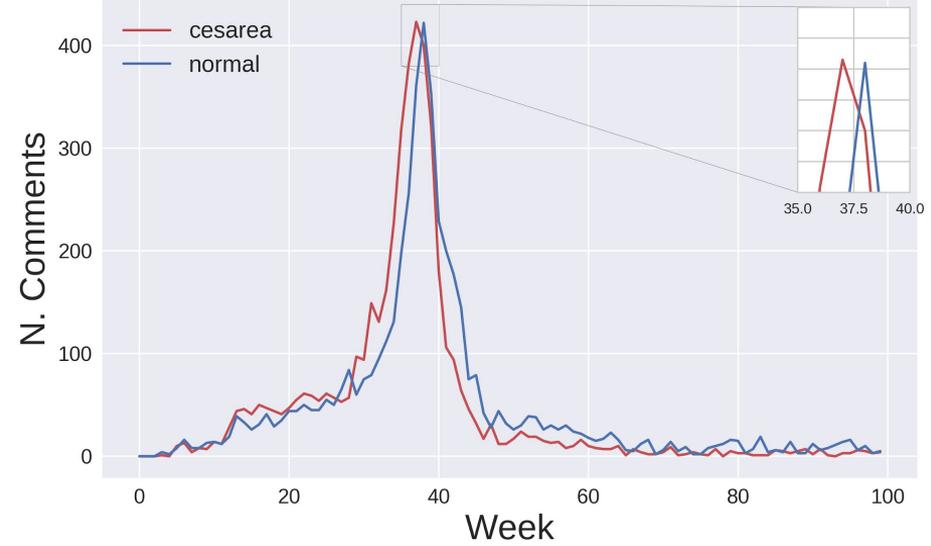
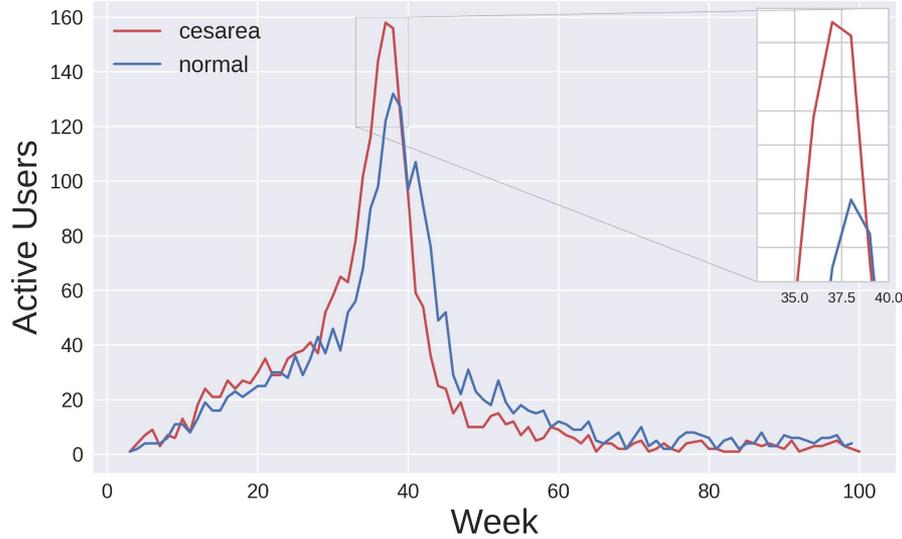
We detect the state of pregnancy for 8,410 users through regular expressions.

Users' activity is **not uniform** along the gestation period but increases significantly as the due date approaches.

Terms frequency



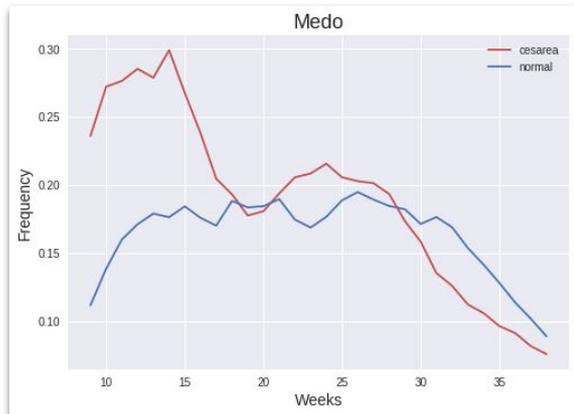
Disaggregated Analysis



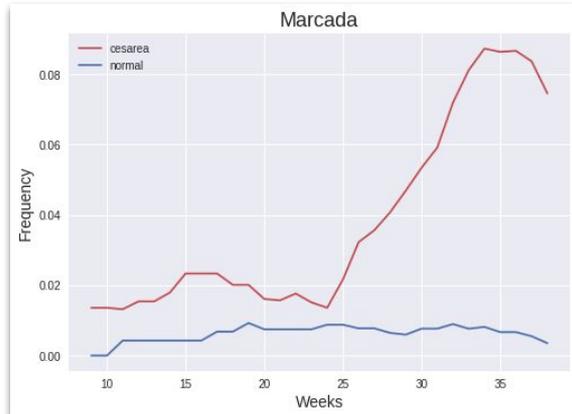
We identify through regular expressions the delivery choice for 619 users:

- **55%** C-Sections
- **45%** natural delivery

Disaggregated Terms Frequency



Fear



Booked



Wanted

normal

Classification

We build a Naive Bayes text classifier to predict a user delivery choice. The training set is defined by a window $[Start_week, End_week]$.

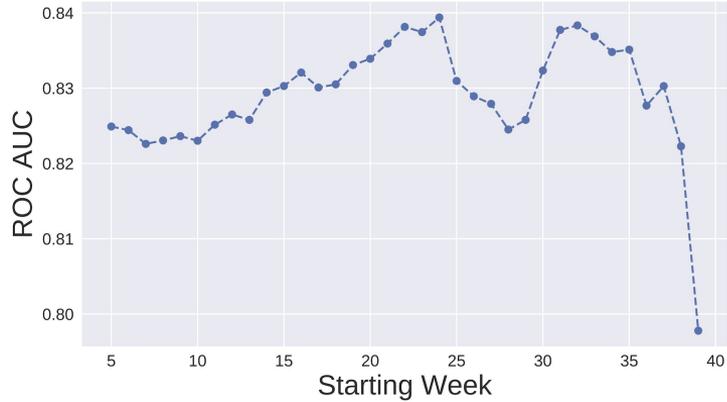
Experiment 1

The training set has a fixed $End_week=40$ and the performance (ROC AUC) of the classifier is evaluated by changing the $Start_week > 5$.

Experiment 2

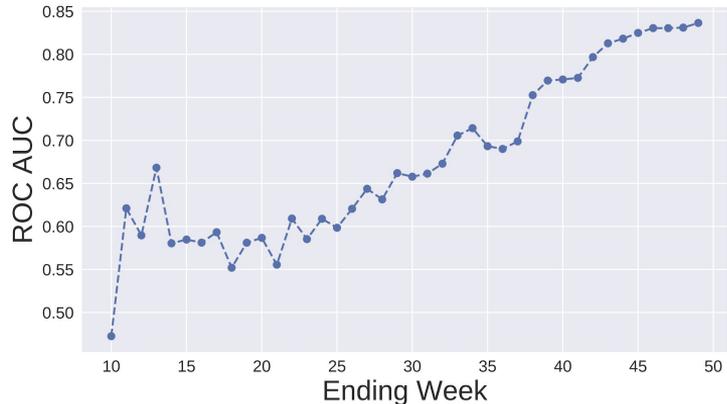
The training set has a fixed $Start_week=5$ and the performance (ROC AUC) of the classifier is evaluated by changing the $End_week > 5$.

Classification



Experiment 1

The training set has a fixed *End_week*=40 and the performance of the classifier is evaluated by changing the *Start_week*>5.



Experiment 2

The training set has a fixed *Start_week*=5 and the performance of the classifier is evaluated by changing the *End_week*>5.

Conclusions

- Internet forums are highly informative to understand attitudes and behaviors related to maternity and parenting.
- Standard NLP techniques are useful to identify topics pertaining to delivery choices.
- A user's delivery preference can be identified from textual content.
- Results could be used to promote awareness campaigns.

Thank you!



Questions?

Michele Tizzoni
@mtizzoni

ISI Foundation

