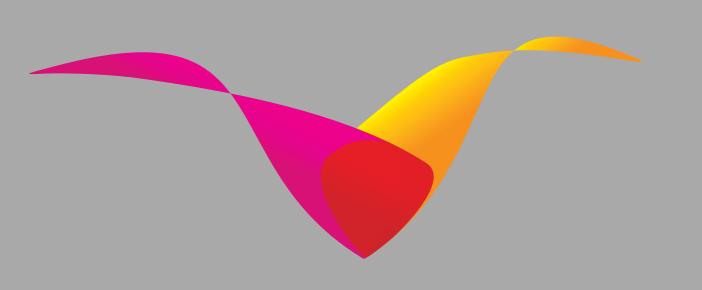
# Statistical Society of Canada Data Science and Analytics Section Workshop on Analysis of Unstructured Text

https://ssc-datascience.github.io
Dave Campbell @iamdavecampbell
Nathan Taback @NathanTaback

Part 1: 1:00 - 2:30 Eastern

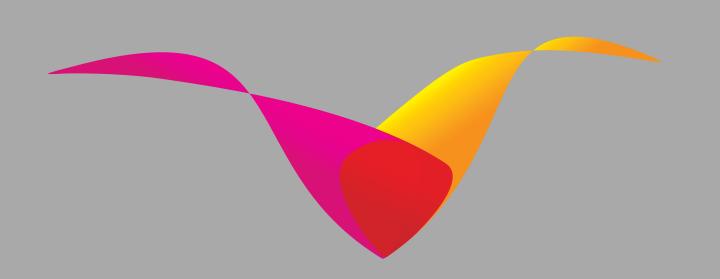
Part 2: 2:30 - 4:00 Eastern



#### Hi, we're new here!

 Goal: to connect researchers and practitioners, beginners and experts in Data Science and Analytics to exchange and develop ideas.

 Connecting industry, government, and academia at all levels while fostering discussion of advances, ideas, tools, and challenges.



#### This section looks fun, how can we join?

Get involved in one of many ways

- You're here, you're already involved. Thank you for joining our workshop!
- Join our section when you join the SSC (ssc.ca)
- The SSC is a charity which funds development of new talent, educational outreach, opening opportunities for individuals, and more, maybe you'd like to donate or convince your employer to do so:

https://www.canadahelps.org/en/charities/statistical-society-of-canadasociete-statistique-du-canada/



Workshop on Analysis of Unstructured Text

Part 1: 1:00-2:30 Eastern

R Libraries for Working with Text

Importing unstructured text into R

Regular expressions and text normalization

N-Grams

Word Vectors and Matrices

Part 2: 2:30-4:00 Eastern

Statistical and computational

models

Sentiment analysis

Topic Modelling

Word Embeddings

Open problems



#### Workshop on Analysis of Unstructured Text

Logistics;

We will open chat for questions, moderated by one of us

All code is available, try running code as we go through the workshop

https://github.com/ssc-datascience

https://ssc-datascience.github.io

