# Visualizing sentence embedding clusters

Julia Kroll
CS 765
Fall 2021

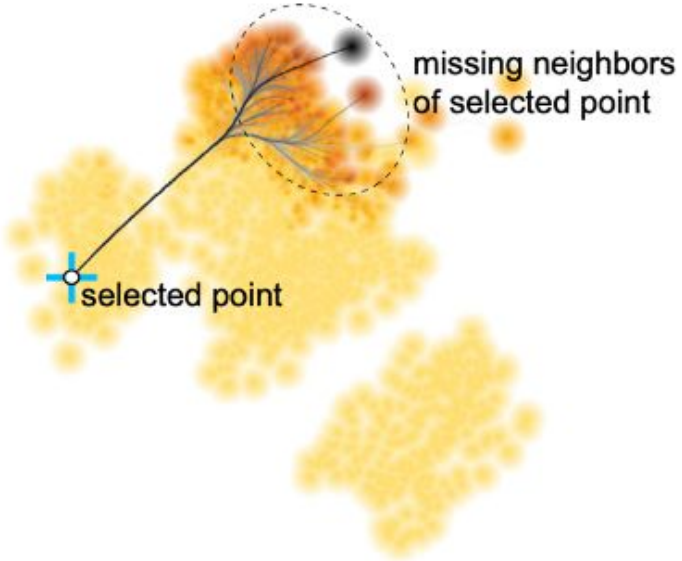
# Motivation

"Cluster sizes in a t-SNE plot mean nothing"

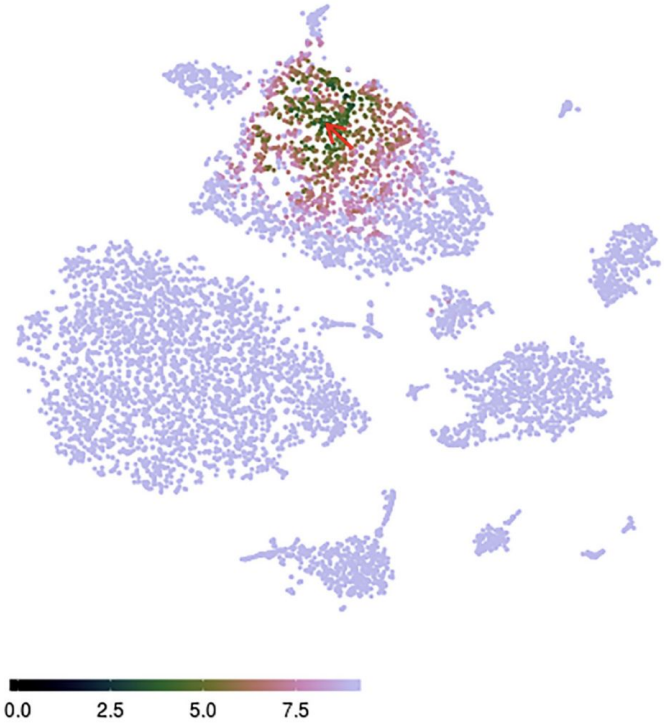
– *“How to Use t-SNE Effectively”*

LII
SUPER BOWL

# Related work



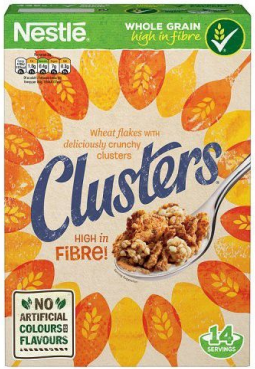

Martins, Coimbra, Minghim & Telea (2014)




Ovchinnikova & Anders (2020)

# Goals

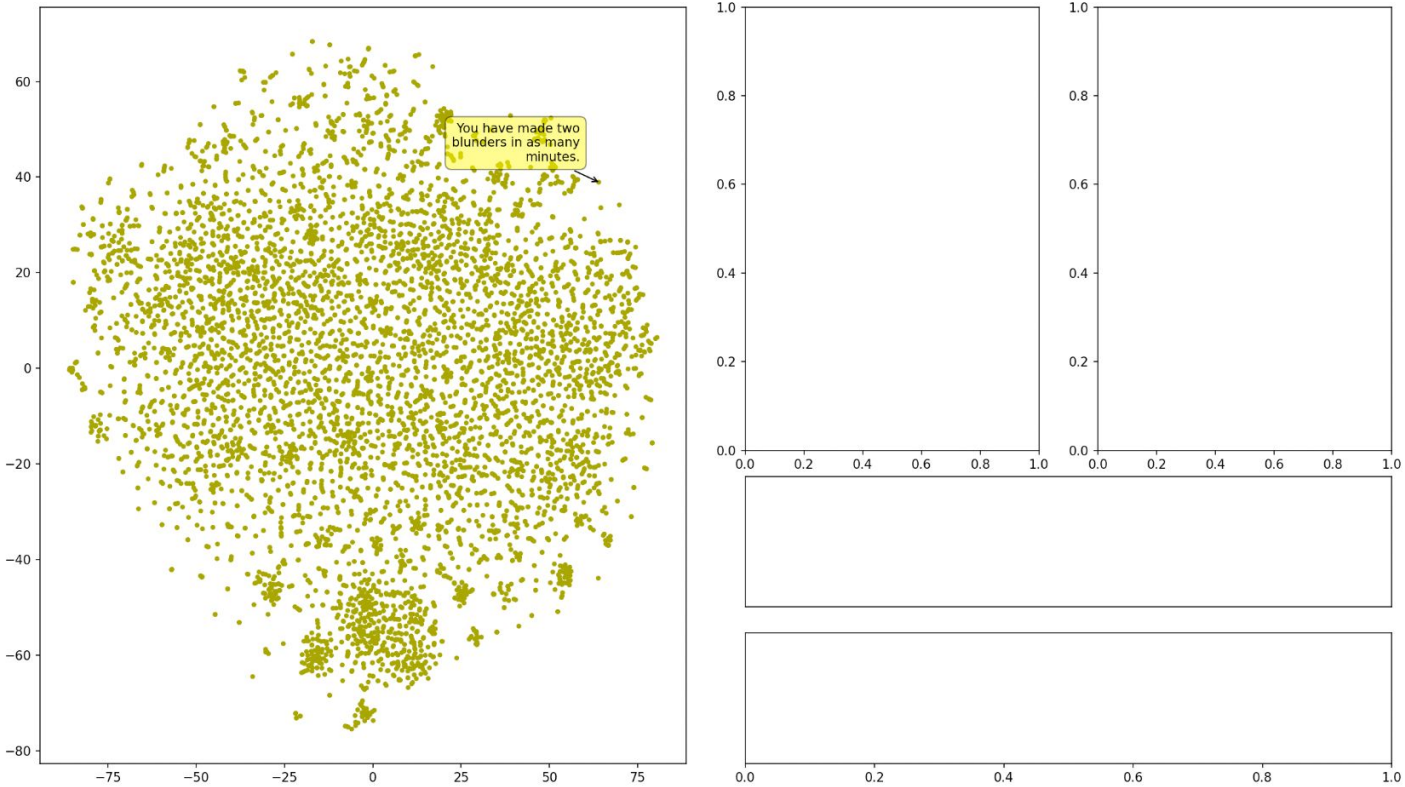
- Compare cluster sizes and distances
  - Density?
  - Distribution?
  - Outliers?
- Interactively select two clusters
- Sentence embeddings (384 dimensions) and t-SNE (2D)



# Demo

Figure 1

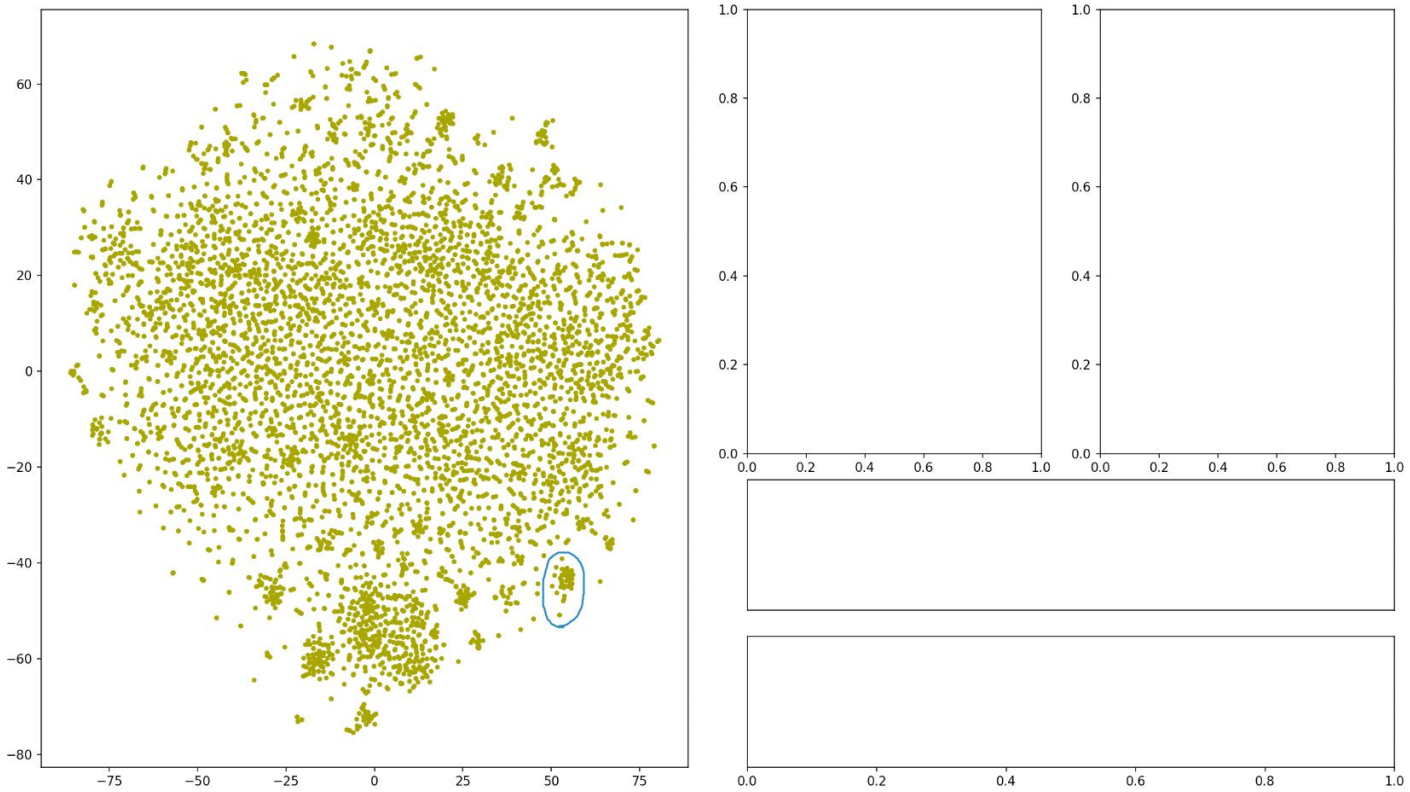
Select A cluster...




x=64.2 y=39.2

Figure 1

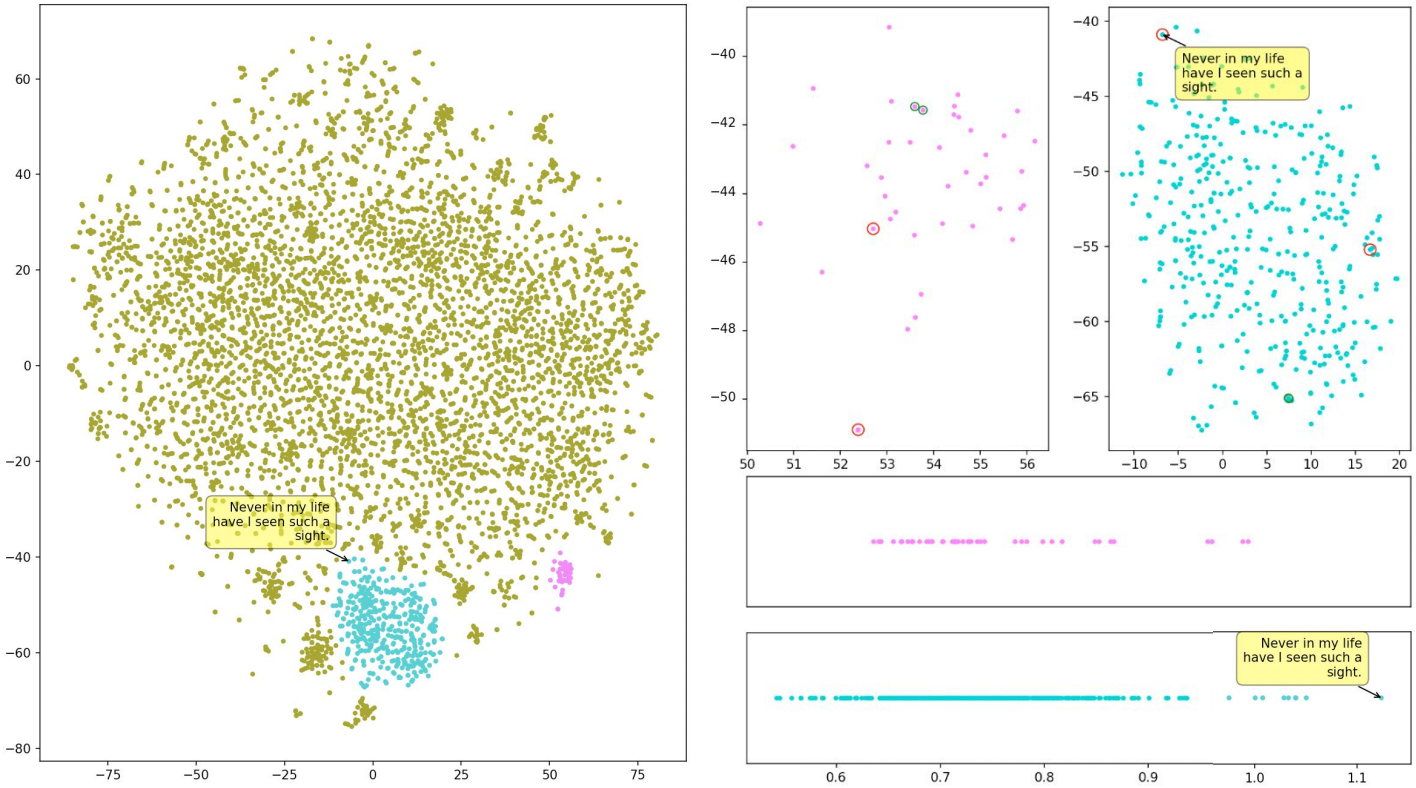
Select A cluster...




x=53.4 y=−53.4

Figure 1

Finished cluster selection (X to exit)




x=−6.72 y=−41.06

# Q&A