

## Part 1: Business Question

*Formulate a research or business question regarding the ongoing COVID-19 pandemic, and solve your question using Twitter Data.*

*Goal: explain the idea to C-level executives.*

- Trending topics/hashtags: Do some sort of Sentiment Analysis (positive, neutral, negative) on hashtags, try to find positive messages among all the news on the COVID-19 virus, try to identify trending topics, analyze the evolution of certain topics through time since the start of the corona crisis, ....
  - Identify spammers: Try to identify fake accounts, for example, accounts that are used for propaganda/fake news/....
- How to identify corona deniers (WHAT?)
  - Identify corona deniers (WHO?)
  - Find hotspots of deniers on map (WHERE?)
  - Visualize over time (WHEN?)

## Part 2: Product

*Turn your solution into a small product.*

*Goal: deliver a reusable app for fellow data engineers.*

- Filter on hashtags (#coronaFake, #coronascam, #coronahoax, #spamdemic, ...)
- Classify people based on their last tweets, using NLP
- Visualize clusters of deniers on a map using the locations of tweets, using dash.py
- Visualize over time, using dash.py:
  - the number of hashtags of deniers
  - vs the number of acceptors
  - vs the number of tweets regarding hot/trending topics before Covid-19.