

# Data Analysis Project

---

Nikolai Rutz, Luc Kury <[nikolai.rutz@stud.unibas.ch](mailto:nikolai.rutz@stud.unibas.ch), [luc.kury@unibas.ch](mailto:luc.kury@unibas.ch)>

Department of Mathematics and Computer Science, University of Basel

January 11, 2020

# Goals

---

## **Find connections between ...**

- Negative news headlines and influence on tourism, categorized by:
  - Political tensions
  - Terrorism (global)
  - Environmental disaster
  - Violence/Crime (local)
- Incoming tourism in a country and wealth of residential population

# The datasets

---

## **GDELT 1.0 Event Database**

The GDELT event database is an aggregation of news from worldwide media outlets.

- Source : GDELT
- Timeframe: 1979 - today (restricted to 2013 to 2018)
- Size : 18.0 GB (compressed)
- Format : CSV

## The datasets - continued

---

### **Nights spent at tourist accommodation establishments**

The data shows the number of foreign visitors who spent a night at tourist accommodation establishments in the corresponding European country.

- › Source : Eurostat
- › Timeframe: 1990 - 2018 (restricted to 2013 to 2018)
- › Size : 634.0 kB (compressed)
- › Format : XLS, CSV, TSV

## The datasets - continued

---

### **Mean and median income by broad group of citizenship**

This dataset contains average income values for European citizens, collected over the last ten years.

- › Source : Eurostat
- › Timeframe: 2013 - 2018
- › Size : 17.4 kB (compressed)
- › Format : XLS, CSV, TSV

# Analysis

---

## **GDELT 1.0 Event Database**

- Get sentiment of headline based on keywords and compare to given sentiment
- Categorize data by events (terrorism, disaster, etc.)

## **EuroStat, Nights spent at tourist accommodation establishments**

- Dataset can be imported as is

## **EuroStat, Mean and median income by broad group of citizenship**

- Dataset can be imported as is

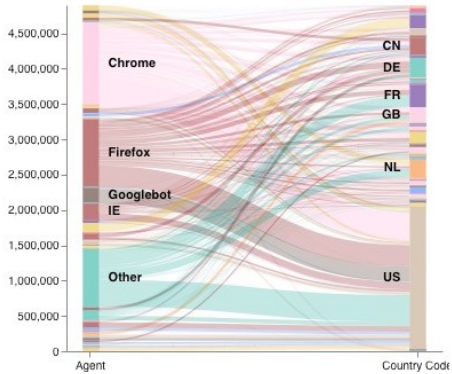
# Expectations

---

- Tourism suffers under negative headlines
- Different kind of headlines, produces different effects on tourism, e.g.:
  - Terrorism has more widespread effect
  - Environmental disaster have local effect
- Regions of higher wealth experience less tourism

# Visualization

---





Questions?