

# Data Analysis Project - Final Report

---

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 12, 2020

# Recap: Project Idea

---

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

- Negative news headlines and influence on tourism, categorized by:
  - Political tensions
  - Terrorism
  - Environmental disaster
  - Violence/Crime
- 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.

## **EuroStat Tourism data set**

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

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

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

# Data integration

---

## **Technologies used:**

- PostgreSQL
- Docker
- Python, Pandas
- Tableau

## **Integration step:**

Cleaning, Normalization, Unification, Transformation

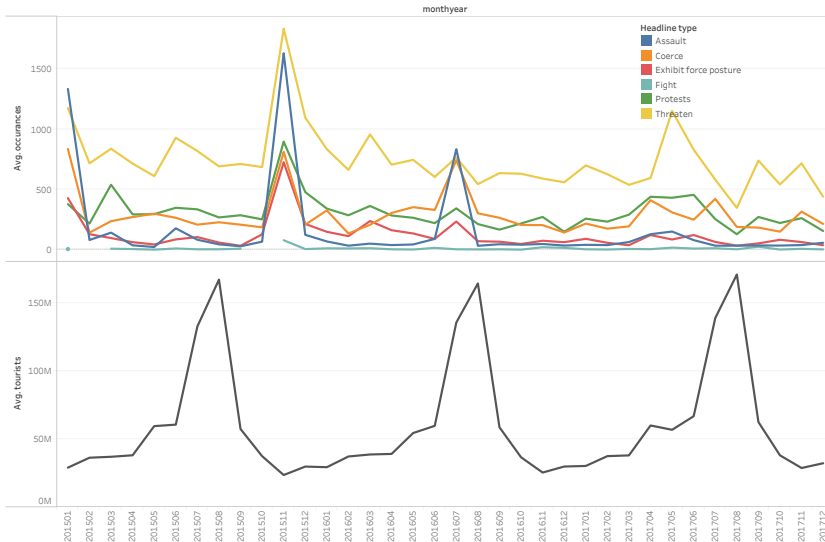
# Analysis & Visualization

---

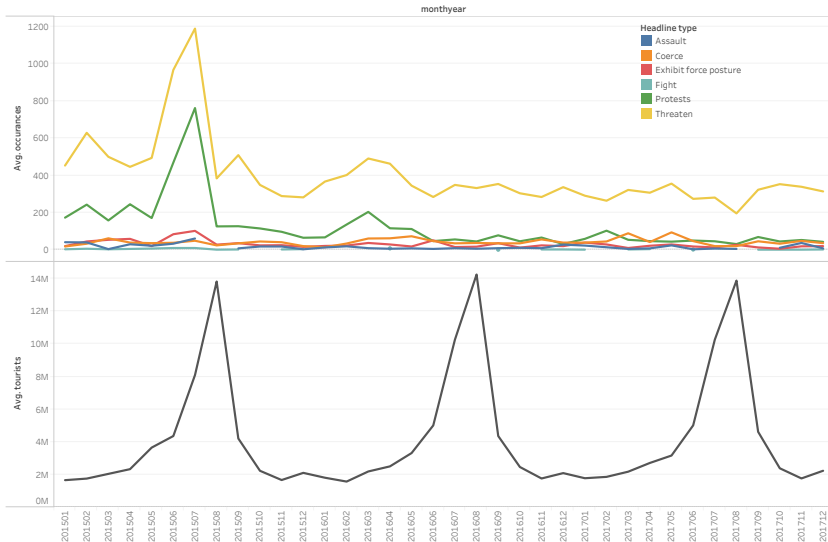
## **Changes in tourists for a specific country due to negative headlines**

- › eventbasecode attribute categorizes event (i.e. '13' => *THREATEN*)
- › adm1code attribute indicates the location (country) of the event
- › monthyear attribute represents the time the event occurred
- › cid attribute shows the country for tourism
- › time attribute indicates the month for a tourism value

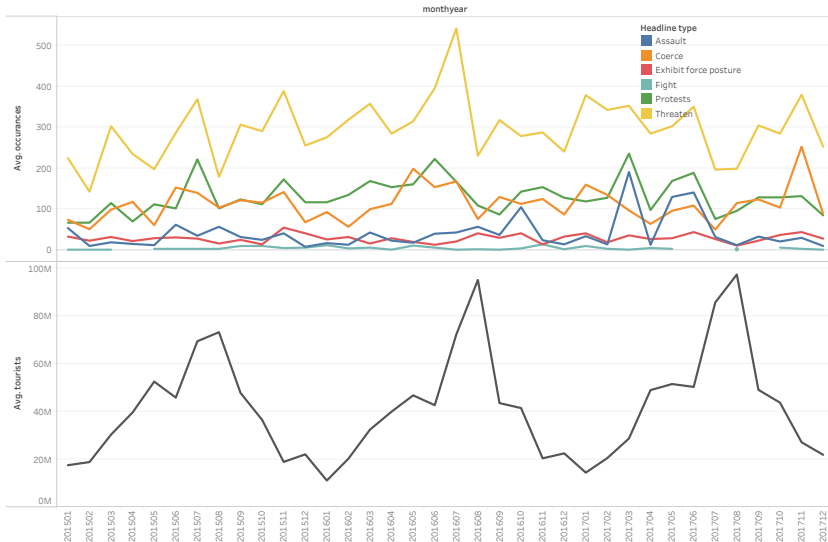
## Negative Headlines and Tourist count France



## Negative Headlines and Tourist count Greece



## Negative Headlines and Tourist count United Kingdom





# Analysis & Visualization

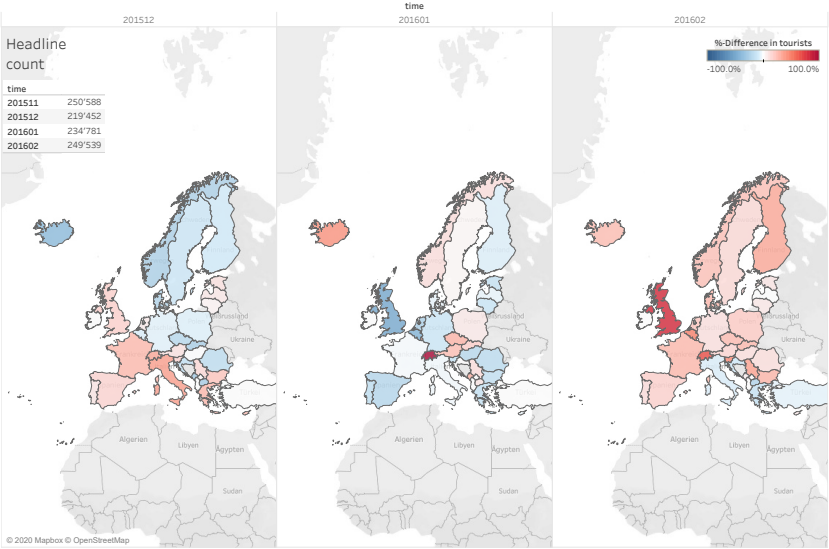
---

## Changes in tourism European Union

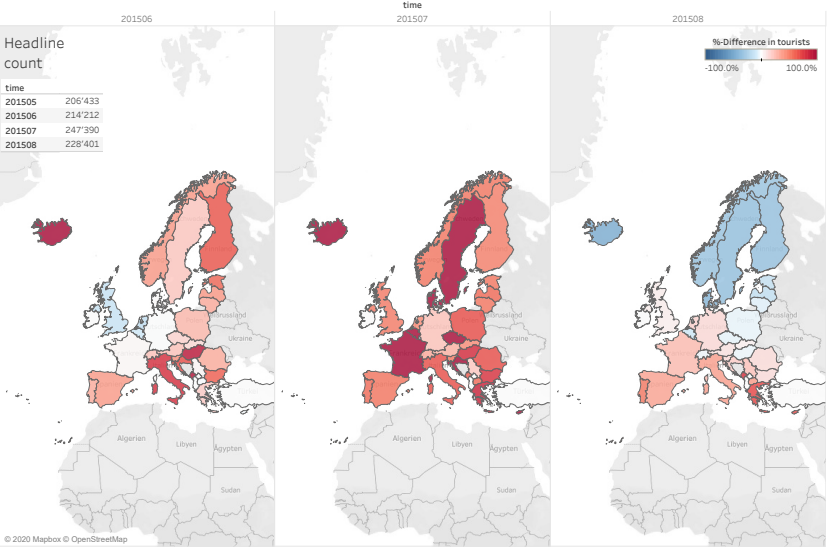
Overview of key events time frame

Start Date	End Date	Event
05/2015	08/2015	Bailout Crisis, Greece
11/2015	02/2016	Assassination Paris, France
06/2016	09/2016	Brexit, UK
09/2017	12/2017	Referendum Kurds, Turkey

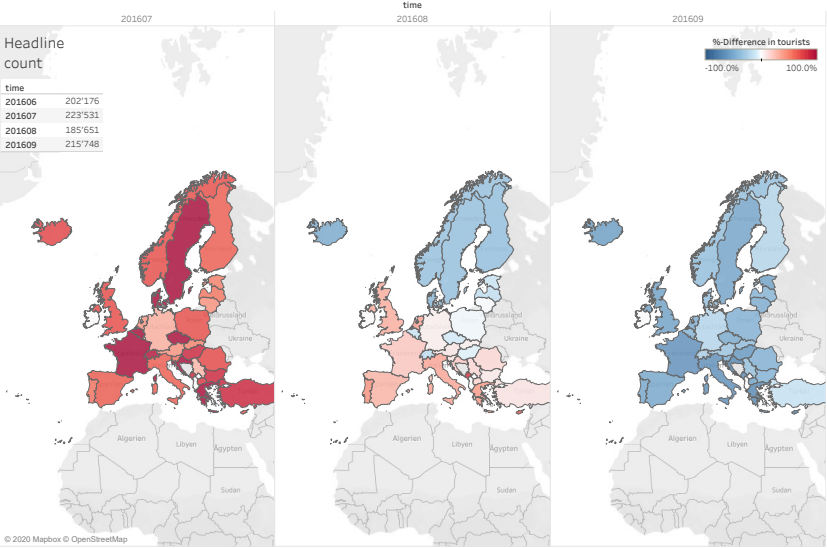
Changes in tourists count between 2015/11 and 2016/02



Changes in tourists count between 2015/05 and 2015/08



Changes in tourists count between 2016/06 and 2016/09



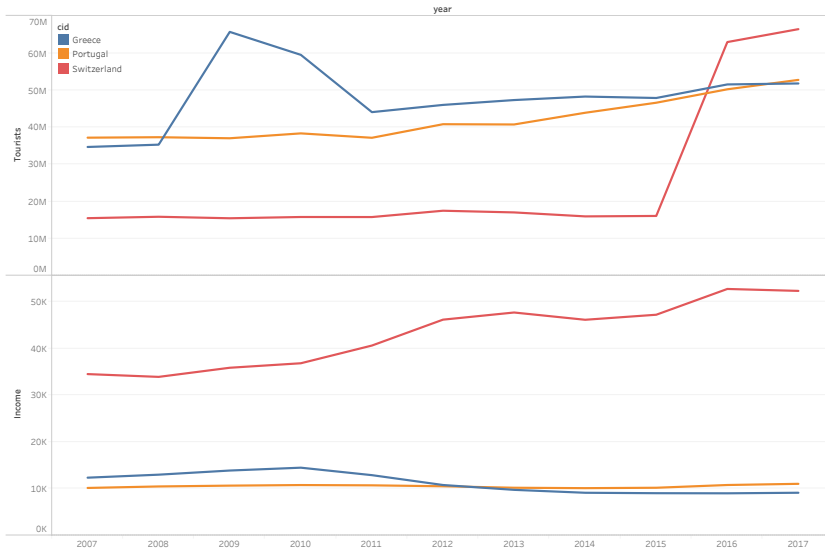
# Analysis & Visualization

---

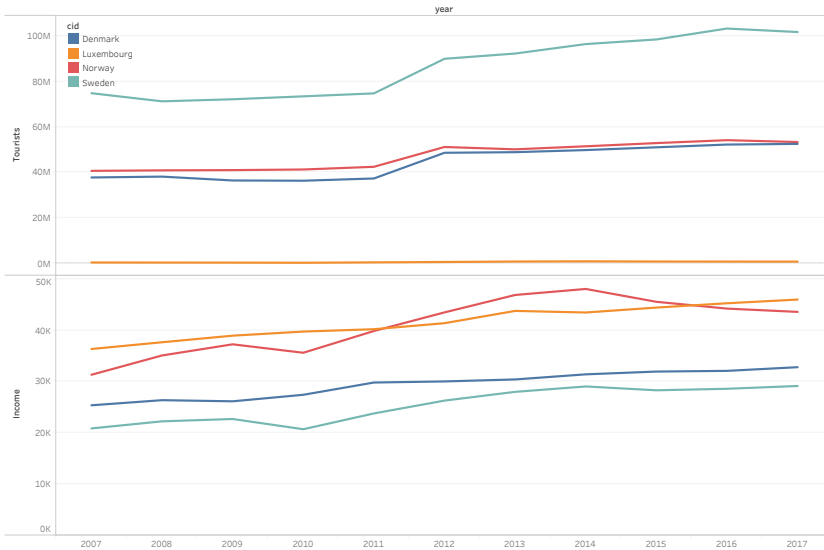
## **Influence of wealth on tourism in the European Union**

- › `cid` attribute shows the country for tourism and income
- › `value` attribute shows the income for a country
- › `time` attribute indicates the month for a tourism and the year for an income value

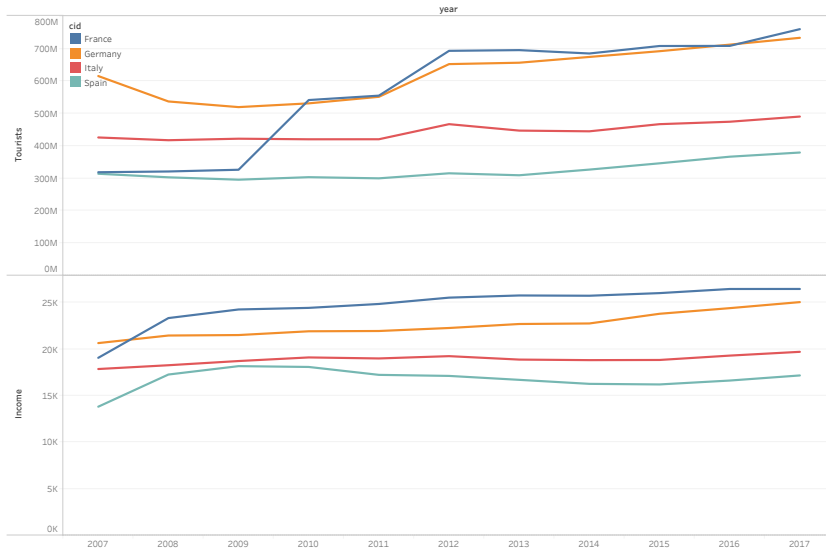
Tourist and Income per Year for Greece, Portugal, Switzerland



Tourist and Income per Year for Denmark, Luxembourg, Norway and Sweden



Tourist and Income per Year for France, Germany, Italy and Spain





# Conclusion

---

- › Correlation between negative headlines and tourism count is weak.
- › Seasons dictate tourism instead of headlines.
- › Growing income has direct impact on tourism (positive).
- › Weather data set would yield stronger correlation with tourism



Questions?