KNOWLEDGE AND THE WEB 2016-2017
Homework 3 - group 8

Alexander Tang
Tom De Groote
Joran Van de Woestijne
Lien Michiels

1   Research Question

One of the major issues that lead to the vote for Brexit, is the rising amount of immigrants entering the UK each year. The **Vote Leave campaign** says: "A vote for leave will be a vote to cut immigration".[1] We wish to investigate whether this argument is well-founded or not. We approach this problem by analyzing the following:

- Compare yearly amount of immigrants in the UK with other countries in the EU.

- Compare yearly crime rate in the UK with other countries. Research if there's a correlation between amount of immigrants and the crime rate (in particular categories).

2   Description Vocabularies

3   Expected Issues

1. **Missing data:** When there's not enough data, we cannot perform a (meaningful) statistical research. For that reason, we first explored some possible options. Analyzing the available data first allows us to gain insight in choosing a good research question.

2. **Inconsistent data:** Data from different sources may differ. We've already found some inconsistency between data from *Eurostat*[2] and the *UK government*[3], regarding the amount of asylum seekers each year. Our current plan is to find more databases (ie. *DBpedia*) to compare with. This may provide some insight on which database is 'more likely' to be correct. If the data is relatively the same, then we could use the mean to provide a compromise.

3. **Bad data:** Some outliers in the data that are either highly unlikely or impossible may be truncated from the dataset. Data that's badly formatted may in some cases be restored to it's obviously intended value, or if it cannot be restored then it will be removed.

---

[1] https://www.theguardian.com/politics/2016/jun/27/eu-referendum-reality-check-leave-campaign-promises
[2] http://ec.europa.eu/eurostat
[3] https://data.gov.uk/