# Information Visualization

# CHECKPOINT I: Visualization Proposal

G21

**1. Domain**

ERASMUS is a European Union program which supports students to study at a school abroad for a certain time. Students can pick their destination from a list of partner schools offered by their school.

This project aims to visualise the flow of international students within the ERASMUS programme. More precisely it explores which destination students of different countries and degrees choose as their target destination.

Our visualisation can help future ERASMUS students to decide which country is the best for them by showing which countries are popular among students of similar degrees as well as which students decide to visit such countries. It also can be useful for schools, countries and European citizens to visualize how ERASMUS programme is utilized. For example, schools can see popular destination and can try to conclude new agreements with schools from this destination.

**2. Dataset**

**From Student\_Mobility\_2013-14.xlsx**

https://data.europa.eu/euodp/en/data/dataset/erasmus-mobility-statistics-2013-14

Our project is based on the dataset containing information about students involved in the ERASMUS university exchange programme. Each row of the dataset describes a student on exchange. Dataset provided by the EU Open Data Portal is available for years 2009 – 2014 (accessible on the URL listed above). It consists of 272 497 records.

The dataset contains information about the sending and receiving institution, the area of studies, sending and receiving country code, information about the degree, as well as participant's gender, language and identification whether the participant requires any special needs. City of the sending institution is not available however should be obtainable from another dataset.

The dataset contains missing values as well as encoding problems which need to be sorted out in the data preparation phase. We will also use other data sources to find out the financial situation in countries, distances etc. These data can be obtained from <https://ec.europa.eu/eurostat/data/database>.

**EUC\_for\_academic\_year\_2013\_2014.xls**

This dataset contains information about all universities participating in ERASMUS program for the year 2013-2014.

URL: <https://data.europa.eu/euodp/en/data/dataset/erasmus-mobility-statistics-2013-14/resource/9ad29b2b-d63f-4660-ba21-772925d57362>

**allCountries.csv**

Contains information about the cities - most importantly **coordinates** as well as the country where the city belongs and also the **localized names** of the cities: e.g. localized names for the Prague are PRG, Praag, Prag, Praga, Pragae, Prago, Prague, Praha and so on.

URL: <https://download.geonames.org/export/dump/allCountries.zip>

**countryInfo.txt**

Contains mapping between ISO (CZ) / ISO3 (CZE) shortcuts and basic information about counties as is the country name, area, population, currency etc.

URL: <https://download.geonames.org/export/dump/countryInfo.txt>

* + - 1. **3. Example Questions**

1. The work will try to answer basic questions about the reasons why students pick their final ERASMUS destination. It will try to discover if there exist any patterns in selected areas (for example gender).
2. Which countries are popular for being an Erasmus destination? Is there any difference between bachelor and master’s degrees?
3. Does gender effects the choice?
4. What is the flow of students doing Erasmus between countries?
5. Does distance from the home country to the Erasmus country matter?
6. How does difference in cost of living in home and target country effect the selection?
7. Do students prefer to go to bigger or smaller cities? Is there any correlation between the size of their home university city and the receiving university city?

**4. Data Sample**

**from Student\_Mobility\_2013-14.xlsx**Action,CallYear,ProjectNumber,MobiilityID,SendingCountry,ReceivingCountry,MobilityType,SpecialNeeds,SubjectAreaCode,SubjectAreaName,CombinedMobilityYesNo,StartDate,EndDate,DurationInMonths,DurationInDays,SubsistenceTravel,LevelOfStudy,ParticipantID,ParticipantGender,ParticipantType,Language,SendingPartnerErasmusID,SendingPartnerName,HostingPartnerErasmusID,HostingPartnerName,HostingPartnerCountry,HostingPartnerCity

ERA02,2013,2013-1-FR1-ERA02-50121,576529,FR,IE,Mob-SMP,0.0,441,Physics,,02-JUN-2014 00.00.00,02-SEP-2014 00.00.00,3,2,720.0,Second Cycle,576529,F,Students,EN,F STRASBO48,UNIVERSITE DE STRASBOURG,,Dublin City University,IE,Dublin

ERA04,2013,2013-1-PT1-ERA04-16705,16705-MOB-00060,PT,CH,Mob-SMP,0.0,520,Engineering and engineering trades,,01-OCT-2013 00.00.00,31-MAR-2014 00.00.00,6,0,3096.0,Second Cycle,257433481,M,Students,FR,P LISBOA02,Universidade de Lisboa,,Paul Scherrer Institut,CH,Villigen PSI

**From cost-of-living-2016.csv**City,Country,Cost.of.Living.Index,Rent.Index,Cost.of.Living.Plus.Rent.Index,Groceries.Index,Restaurant.Price.Index,Local.Purchasing.Power.Index,Milk(regular)(1 liter),Monthly.Pass,Apartment(1.bedroom).in.City.Centre,"Internet(10 Mbps, Unlimited Data, Cable/ADSL)",Cappuccino(regular),Water(0.33 liter bottle),Eggs(12),Water(1.5 liter bottle),Domestic Beer (0.5 liter bottle),One-way Ticket (Local Transport),"Basic (Electricity, Heating, Water, Garbage) for 85m2 Apartment","Cinema, International Release, 1 Seat",Apples (1kg)

Northampton,United Kingdom,67.16,26.85,48.18,59.6,74.13,110.82,1.2,65.53,744.44,22.08,3.65,1.55,2.1,1.08,1.44,2.95,168.42,12.19,2.62

Allentown,PA,86.59,32.41,61.08,91.75,74.59,82.12,1.1,50.0,883.8,52.84,4.16,1.28,2.39,2.0,2.0,2.0,259.81,11.0,8.34

**From allCountries.csv**geonameid,name,asciiname,alternatenames,latitude,longitude,feature class,feature code,country code,cc2,admin1 code,admin2 code,admin3 code,admin4 code,population,elevation,dem,timezone,modification date

3286426,Crvica,Crvica,Crvica,44.01608,19.58014,P,PPLL,BA,,02,,,,0,,418,Europe/Sarajevo,2019-01-10

2931198,Elendhof,Elendhof,,49.07394,12.50164,S,FRM,DE,,2.0,93.0,9372.0,9372125.0,0,,611,Europe/Berlin,2013-02-19

**From CountryInfo.txt**ISO,ISO3,ISO-Numeric,fips,Country,Capital,Area(in sq km),Population,Continent,tld,CurrencyCode,CurrencyName,Phone,Postal Code Format,Postal Code,Regex,Languages,geonameid,neighbours, EquivalentFipsCode

NL,NLD,528,NL,Netherlands,Amsterdam,41526.0,16645000,EU,.nl,EUR,Euro,31,#### @@,^(\d{4}[A-Z]{2})$,"nl-NL,fy-NL",2750405,"DE,BE",

NR,NRU,520,NR,Nauru,Yaren,21.0,10065,OC,.nr,AUD,Dollar,674,,,"na,en-NR",2110425,,