**Data Journalism Training**

**Bolivia assignments**

<https://github.com/peterverweij/EJC2_data>

**DAY 1**

**1. introduction**

- who is who? Any skills already?

- goals of the training: intro in ddj, theory and skills

- method: hands-on subgroups

**2. what is ddj?** Intro movie, what is your conclusion

what is ddj in ppt, steps and examples

**3. let’s do all in one; to check**

Pop Bolivia from wikipedia: <https://en.wikipedia.org/wiki/Bolivia>

Is that a good source?

a. scrape the table with chrome:

- chrome browser , extensions. table capture

have chrome browser and search for table capture extension, in chrome store, add to chrome

- largest city and towns, highlight and scrape, copy, paste in excel; sheets 2

- rework to sheet 1 and do the assignments in the sheet

- find total pop boli and put that in a cel, what is your conclusion from perc?

- import spreadsheet (blad 1) in tableau and make a bar graph and a map ( set city names as geo, and edit map location.

- which type of graph?

- save as, (register first) at the bottom choose download export as pic and embed, or choose share for embed

b. scraping with outwit

open outwit and paste the wiki link

go to data and click tables

select row 62-71, export as xls

open in excel en clean: header column, delete col, del text and set to number

explain difference text and number

do the same calculation

c. Question:

What do you think: do you need an intro to excel basics: simple formula, total average, filter, order, graph, copy paste special?

About csv and xls how to import?

Or Basic excel training use coke.xls, if needed show ribbon and most import jobs; if needed show how to imput text and figures in cells and command line. Example: coke.xls

or If not needed:

continue with creating subgroups and choose topic for each group

4. Let’s check your searching and finding data for DDJ

- find population of the Netherlands;

\* how do you search, engine and keywords, what does more words imply- wider or smaller

\* check credibility by url

\* how to search on a page for a word or number ctr-F

\* advanced searches using file type and or site

bolivia gdp filetype:xls

bolivia site:bbc.co.uk

\* using exact search:

" the second largest city of Bolivia is"

" the S&P rating of Bolivia is" look at #2

**Which are your standard data sources? Store and exchange sources in the newsroom?**

Standard sources:

international

<https://databank.worldbank.org/data/source/world-development-indicators>

try load data for bolivia, GDP growth , unemployment , government debt

<https://tradingeconomics.com/mozambique/rating>

explore trading economics

<https://data.imf.org/?sk=388DFA60-1D26-4ADE-B505-A05A558D9A42&sId=1479329334655>

IMF financial data

<https://datacatalog.worldbank.org/dataset/global-financial-inclusion-global-findex-database>

Financial inclusion

<https://www.who.int/countries/bol/en/>health data

<https://www.cia.gov/library/publications/the-world-factbook/geos/bl.html>

basic bol data

National:

Population data: <http://worldpopulationreview.com/countries/bolivia-population/>

**Bolivia**

wiki pedia:

scrape tables from webpage: <https://en.wikipedia.org/wiki/2019_Bolivian_general_election>

voter turnout database : <https://www.idea.int/data-tools>

choose Bolivia and choose turnout; scrape tables

calc key figures and make graph: why is a line not a good viz?

Members of the parliament als os for the senat

<http://www.diputados.bo/legisladores>

scrape table and create pivot table

make scraper for details of each member

National bureau of statistics: <https://www.ine.gob.bo/>

census database

**5. Working with pdf**

Open Barclays pdf; Mozambique p.74, table p.77

scrape key market data using tabula

export file is .csv

import in excel

other solution services on de web: pdf to excel

<https://www.pdftoexcel.com/>

**6. Work in small groups on the project; report tomorrow 9.00**

**Day 2**

**0. Report subgroups, topics discussion**

**1. Repeat day** 1: practice scraping a table, clean in xls, and make viz in tableau

use: <https://en.wikipedia.org/wiki/Uruguay>

population per provinces for 2017, and density

You can use that for example for mapping election data; victims of natural disasters etc

**2. Scraping financial data using trading economics**

- what are the sovereign ratings for south american countries

- scrape web page, select countries, export as excel and clean

- import in tableau and make graph and map

- use pivot tables to do calculations in excel and tableau

- is Bolivia rating high or low, what is high or low (average)

**3. viz with Flourish**

<https://flourish.studio/>

make fast viz with flourish

my examples: <https://app.flourish.studio/projects>

Try also your yesterday data in Flourish

**4. Work in projects**

**5. some statistic**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Nominal  just categories | Ordinal  categories with order | ratio/interval  measurement scale |
| Var 1 | modal ; median  range  bar graph. Pie graph |  | Average  standard dev  histogram  line graph |
| Var 2 | Relationship  perc diff  chi square  double bar graph |  | Relationship  correlation  regression  trend line  scatter diagram |

Explore statistics using GDP and life expectation for Bolivia; data wolrldbank <https://data.worldbank.org/country/bolivia>

download data from Worldbank, clean up and calc and viz

or use life\_gdp.xlsx

**6. More excel/advanced:**

example: mayors.xls

split column

merge column

playing with pivot tables, association and correlation

DISCUSS DIFFERENT GRAPH TYPES IN RELATIONS WITH DATA

<https://flourish.studio/2018/09/28/choosing-the-right-visualisation/>

**Moving chart**

download worldbank health spending and life expectancy for south American countries, from 2000

I use life\_gdp xls in Flourish and tableau

**7. work on projects**

**Alternatives**

**1.building an advanced scraper** in outwit for data from various web pages

using trading economics, south america countries, key financial data; gdp per cap, inflation, interest, debt to gdp

- about html tags

- building scraper:tradingeconomics.com\_bol\_scraper: <https://tradingeconomics.com/bolivia/indicators> ; change country name!

- running scrapper list of countries using auto explore: <https://tradingeconomics.com/countries>, list and select countries; only 100 allowed, build url list south America countries; 1 download a list in html: outwithub list...html; change the countries country list.html; select countries and run scraper

- export data to xls and clean

make viz based on xls in tableau

**1.1 Big Mac index for South American data**

**2. Mapping election data**

**3. Analyzing data members of parliament**

**4. finalize work in projects**

**5. presentation of projects using digital projection: ppt, viz and xls**

**6. evaluation: what do you take home?**