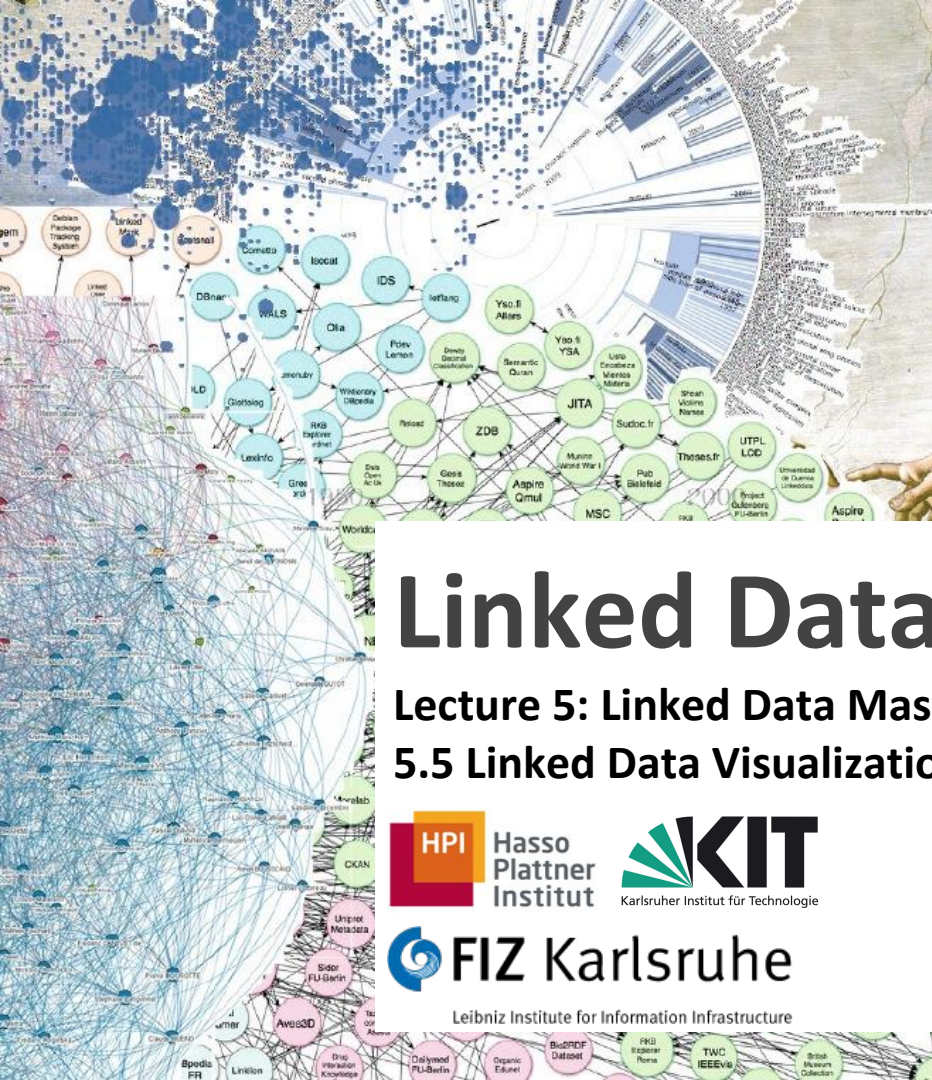




Linked Data Engineering

Lecture 5: Linked Data Mashups and Applications

5.5 Linked Data Visualization

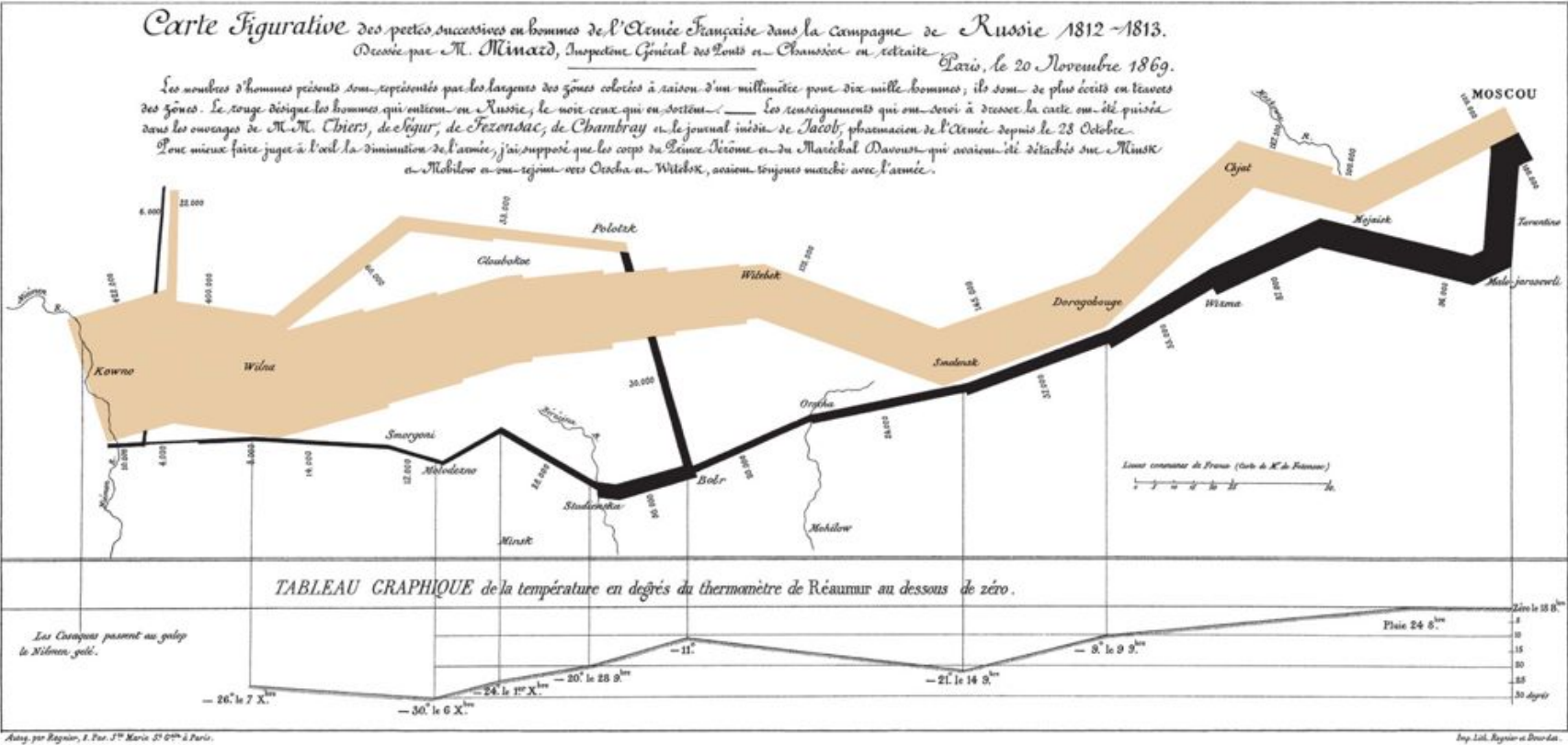


Leibniz Institute for Information Infrastructure

Prof. Dr. Harald Sack
FIZ Karlsruhe

Leibniz Institute for Information Infrastructure
Karlsruhe Institute of Technology

Autumn 2016



Charles Joseph Minard (1781-1870) - Napoleon's Russian Campaign

A Picture tells more than a 1000 words...

- Pictures have been used to convey information long before the development of writing
- A single picture can be processed (“understood”) much faster than a text page
- Human perception is processing in **parallel**, text analysis is limited by the **sequential** process of reading

Information Visualization

- **Information Visualization** is the study of (interactive) visual representations of abstract data to reinforce human cognition
- **Information graphics** or **infographics** are graphic visual representations of information, data or knowledge intended to present information quickly and clearly
 - a static form of information visualization
 - aims to emphasize specific findings gained from the visualized data
 - **Mandatory precondition:** Data Analysis

Hands On Data Visualization Example

- **Dataset:** DBpedia



SPARQL

- **Task:**

Draw a **map chart** which indicates the **number of authors per country**

Geo Chart



e.g. Google Charts

Task: Draw a Map Chart...

- **Dataset: DBpedia**

- Examine a representative example, e.g.
http://dbpedia.org/resource/Jules_Verne

About: Jules Verne

An Entity of Type `person`, from Named Graph : `http://dbpedia.org`, within Data Space : `dbpedia.org`

Jules Gabriel Verne (/vʒːrn/; French: [ʒyl vɛʁn]; 8 February 1828 – 24 March 1905) was a French novelist, poet, and playwright best known for his adventure novels and his profound influence on the literary genre of science fiction. Verne was born to bourgeois parents in the seaport of Nantes, where he was trained to follow in his father's footsteps as a lawyer, but quit the profession early in life to write for magazines and the stage.

Property	Value
<code>dbo:abstract</code>	- Jules Gabriel Verne (/vʒːrn/; French: [ʒyl vɛʁn]; 8 February 1828 – 24 March 1905) was a French novelist, poet, and playwright best known for his adventure novels and his profound influence on the literary genre of science fiction. Verne was born to bourgeois parents in the seaport of Nantes, where he was trained to follow in his father's footsteps as a lawyer, but quit the profession early in life to write for magazines and the stage. His collaboration with the publisher Pierre-Jules Hetzel led to the creation of the <i>Voyages extraordinaires</i>, a widely popular series of scrupulously researched adventure novels including <i>Journey to the Center of the Earth</i> (1864), <i>Twenty Thousand Leagues Under the Sea</i> (1870), and <i>Around the World in Eighty Days</i> (1873). Verne is generally considered a major literary author in France and most of Europe, where he has had a wide influence on the literary avant-garde and on surrealism. His reputation is markedly different in Anglophone regions, where he has often been labeled a writer of genre fiction or children's books, largely because of the highly abridged and altered translations in which his novels are often reprinted. Verne has been the second most-translated author in the world since 1979, ranking between Agatha Christie and William Shakespeare. He has sometimes been called the "Father of Science Fiction", a title that has also been given to H. G. Wells and Hugo Gernsback. ^(en) - Jules-Gabriel Verne (* 8. Februar 1828 in Nantes; † 24. März 1905 in Amiens) war ein französischer Schriftsteller. Bekannt wurde er vor allem durch seine Romane <i>Die Reise zum Mittelpunkt der Erde</i> (1864), <i>20.000 Meilen unter dem Meer</i> (1869–1870) sowie <i>Reise um die Erde in 80 Tagen</i> (1873). Neben Hugo Gernsback, Kurd Laßwitz und H. G. Wells gilt Jules Verne als einer der Begründer der Science-Fiction-Literatur. ^(de)
<code>dbo:activeYearsEndYear</code>	1905-01-01 ^(xsd:date)

Task: Draw a Map Chart...

- **SPARQL Query:**
 - **Country and Number of authors per country**
 - `?author rdf:type dbo:Writer .`
 - `?author dbo:birthPlace ?birthplace .`
 - `?birthplace dbo:country ?country .`
 - `?country rdfs:label ?country_name`
 - `FILTER (lang(?country_name)="en")`
 - `GROUP BY ?country_name`
 - `COUNT(?author)`
 - `xsd:string(?country_name)`

Task: Draw a Map Chart...

● SPARQL Query:

Virtuoso SPARQL Query Editor

[About](#) | [Namespace Prefixes](#) | [Inference rules](#) | [iSPARQL](#)

Default Data Set Name (Graph IRI)

Query Text

```
SELECT DISTINCT xsd:string(?country_name) COUNT(?author) as ?count WHERE {
?author rdf:type dbo:Writer ;
        dbo:birthPlace ?birthplace .
?birthplace dbo:country ?country .
?country rdfs:label ?country_name FILTER (lang(?country_name)="en") .
} GROUP BY ?country_name
ORDER BY DESC(?count)
```

(Security restrictions of this server do not allow you to retrieve remote RDF data, see [details](#).)

Results Format:

Execution timeout:

milliseconds (values less than 1000 are ignored)

Options:

☒ Strict checking of void variables ☐ Log debug info at the end of output (has no effect on some queries and output formats)

(The result can only be sent back to browser, not saved on the server, see [details](#))

Copyright © 2016 [OpenLink Software](#)
Virtuoso version 07.20.3217 on Linux (i686-generic-linux-glibc212-64), Single Server Edition

[DBpedia SPARQL query](#)

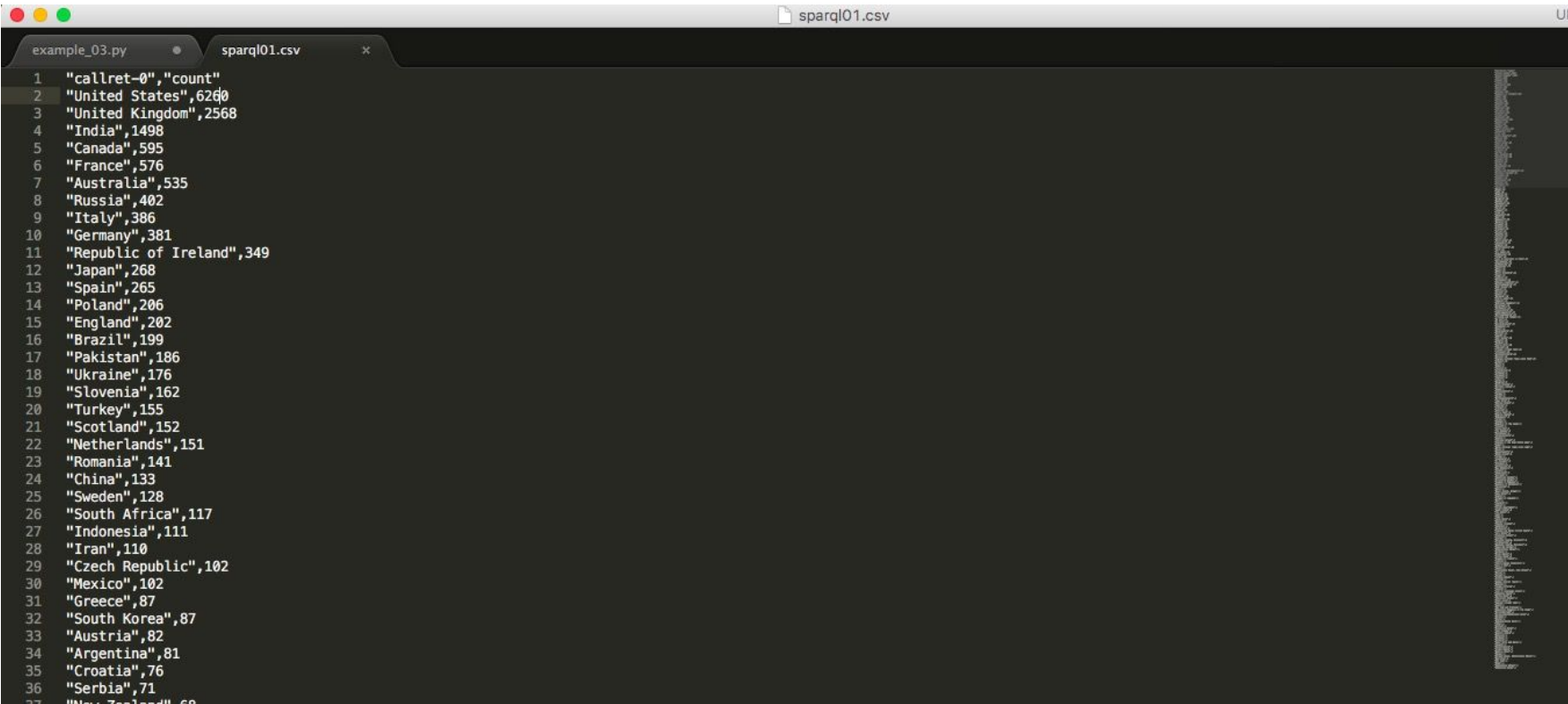
Task: Draw a Map Chart...

- **SPARQL Result:**
 - Save as CSV

callret-0	count
United States	6260
United Kingdom	2568
India	1498
Canada	595
France	576
Australia	535
Russia	402
Italy	386
Germany	381
Republic of Ireland	349
Japan	268
Spain	265
Poland	206
England	202
Brazil	199
Pakistan	186
Ukraine	176
Slovenia	162
Turkey	155
Scotland	152
Netherlands	151
Romania	141
China	133
Sweden	128
South Africa	117
Indonesia	111
Iran	110

Task: Draw a Map Chart...

- Pre-process CSV to fit in Google Geo Chart:



```

sparql01.csv
example_03.py sparql01.csv x
1 "callret-0","count"
2 "United States",6260
3 "United Kingdom",2568
4 "India",1498
5 "Canada",595
6 "France",576
7 "Australia",535
8 "Russia",402
9 "Italy",386
10 "Germany",381
11 "Republic of Ireland",349
12 "Japan",268
13 "Spain",265
14 "Poland",206
15 "England",202
16 "Brazil",199
17 "Pakistan",186
18 "Ukraine",176
19 "Slovenia",162
20 "Turkey",155
21 "Scotland",152
22 "Netherlands",151
23 "Romania",141
24 "China",133
25 "Sweden",128
26 "South Africa",117
27 "Indonesia",111
28 "Iran",110
29 "Czech Republic",102
30 "Mexico",102
31 "Greece",87
32 "South Korea",87
33 "Austria",82
34 "Argentina",81
35 "Croatia",76
36 "Serbia",71
37 "New Zealand",68
  
```

Task: Draw a Map Chart...

- **Pre-process CSV to fit in Google Geo Chart:**

Google Charts

HOME

GUIDES

REFERENCE

SUPPORT

Calendar Charts

Candlestick Charts

Column Charts

Combo Charts

Diff Charts

Donut Charts

Gantt Charts

Gauge Charts

GeoCharts

Histograms

Intervals

Line Charts

Maps

Org Charts

Pie Charts

Sankey Diagrams

Scatter Charts

Stepped Area Charts

Table Charts

Timelines

Tree Map Charts

Trendlines

Waterfall Charts

Word Trees

Miscellaneous Examples

```

<html>
<head>
  <script type="text/javascript" src="https://www.gstatic.com/charts/loader.js"></script>
  <script type="text/javascript">
    google.charts.load('upcoming', {'packages': ['geochart']});
    google.charts.setOnLoadCallback(drawRegionsMap);

    function drawRegionsMap() {

      var data = google.visualization.arrayToDataTable([
        ['Country', 'Popularity'],
        ['Germany', 200],
        ['United States', 300],
        ['Brazil', 400],
        ['Canada', 500],
        ['France', 600],
        ['RU', 700]
      ]);

      var options = {};

      var chart = new google.visualization.GeoChart(document.getElementById('regions_div'));

      chart.draw(data, options);
    }
  </script>
</head>
<body>
  <div id="regions_div" style="width: 900px; height: 500px;"></div>
</body>

```

<https://developers.google.com/chart/interactive/docs/gallery/geochart>

Task: Draw a Map Chart...

- **Pre-process CSV to fit in Google Geo Chart:**

- **CSV:**

- "United States",6260

- "United Kingdom",2568

- ...

- **Required Format:**

- ['United States', 6260],

- ['United Kingdom', 2568],

- ...

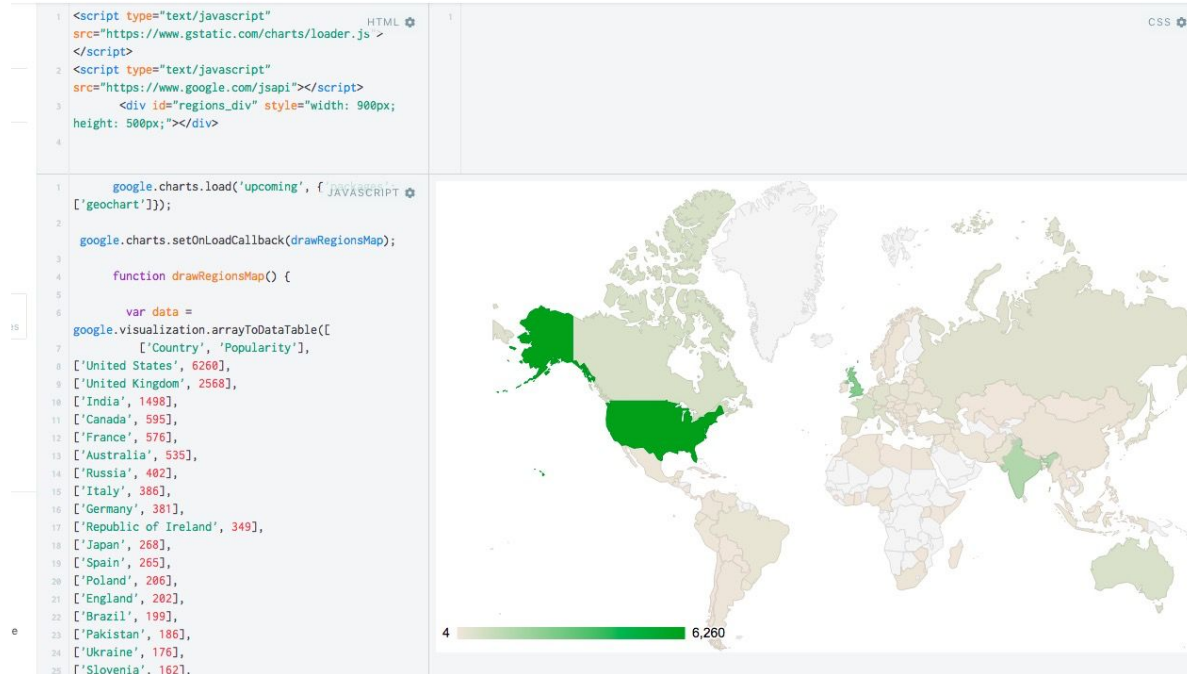
- E.g. use **sed** and **awk** UNIX-command line tool

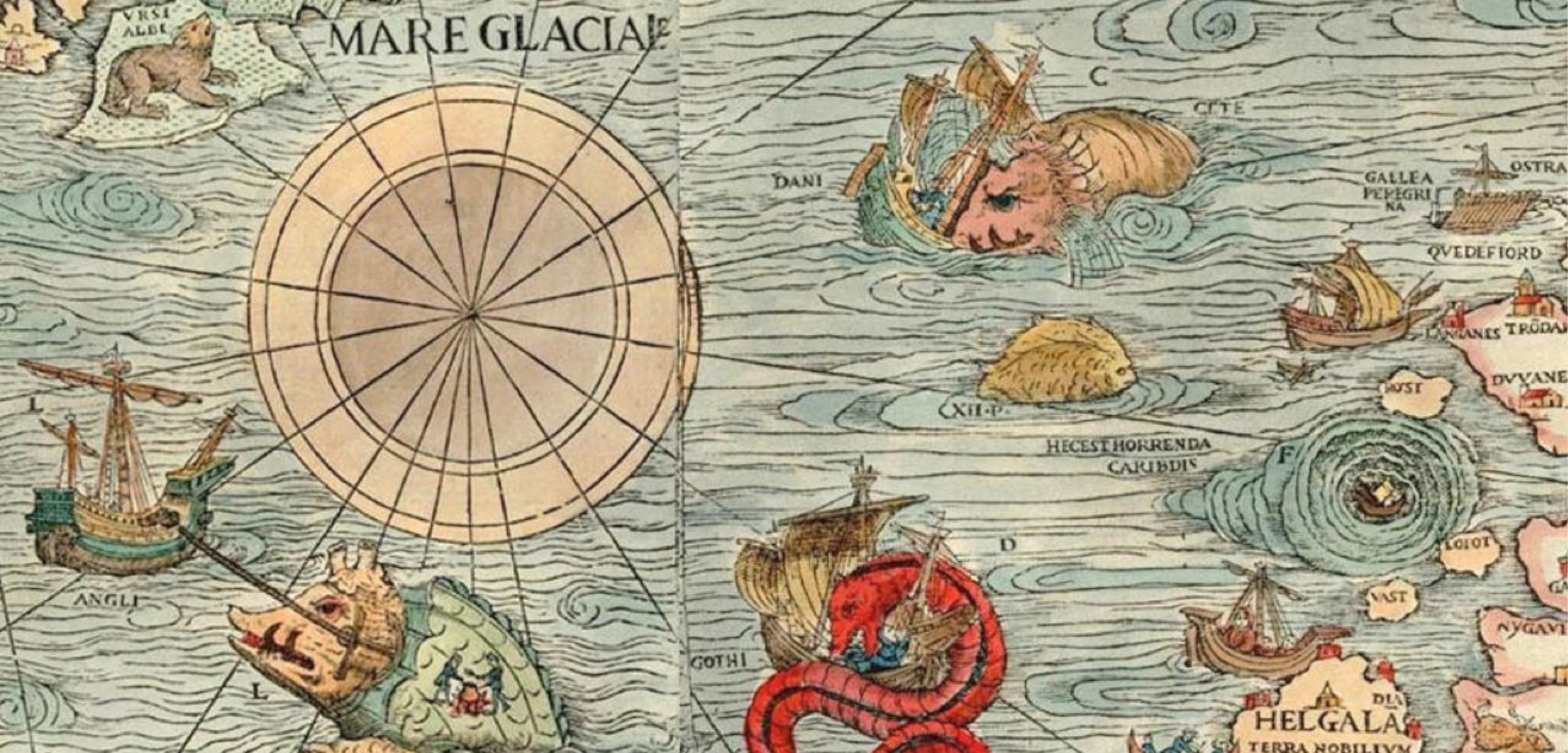
- ```
sed "s/\"/'/g" sparql01.csv | awk 'BEGIN {FS=","} ; { printf "[%s,%s],\n", $1, $2}' > new_sparql01
```

# Task: Draw a Map Chart...

- **Draw Geo Chart:**

- E.g. with **jsfiddle**, <https://jsfiddle.net/api/post/library/pure/>





**Next: 06 - Linked Data Analytics**

Lecture 5 - Linked Data Mashups & Applications - OpenHPI - Course Linked Data Engineering