Data Journalism

Data Exploration Tools

Angelica Lo Duca
angelica.loduca@iit.cnr.it

Classification of tools

- Coding yes or no
- Functions: data cleansing, data visualization, data analysis, machine learning, visual analytics

Tools for non developers

Tool	Free	Online	Desktop	Maps	Graphs
ChartBlocks	the basic version - YES the advanced version - NO	X	-	-	X
Chartio	https://chartio.com/	X	-	X	X
Data Wrapper	YES	X	X	X	X
Dive MIT	YES	X		X	X
Domo	NO, but there is a trial version	X	-	X	X

%

Tool	Free	Online	Desktop	Maps	Graphs
Google Fusion Tables	YES	X	-	X	X
Google Spreadsheet	YES	X	-	-	X
IBM Watson Analytics	NO, but there is a trial version	X	-	X	X
Infogram	the basic version - YES the advanced version - NO	X	-	X	X
Microsoft Excel	Online - YES Desktop - NO	X	X	-	X
Microsoft Power BI	YES	X	X	X	X

%

Tool	Free	Online	Desktop	Maps	Graphs
Qlikview	NO, but there is a trial version		X	X	X
Raw	YES	X	X	-	X
Salesforce Einstein Analytics Platform	NO, but there is a trial version	X	-	-	X
Sisense	NO	X	-	-	X
Tableau	NO, but there is a trial version and a public version	-	X	X	X
Zoho Reports	NO, but there is a trial version	X	-	-	X

Tools for developers

- Highcharts
- Bokeh
- d3.js
- Chart.js
- Google Charts
- Leaflet
- dygraphs
- R software
- Plotly

Let’s see

- ChartBlocks
- Data Wrapper
- Google Fusion Tables
- Google Spreadsheet

Data Examples

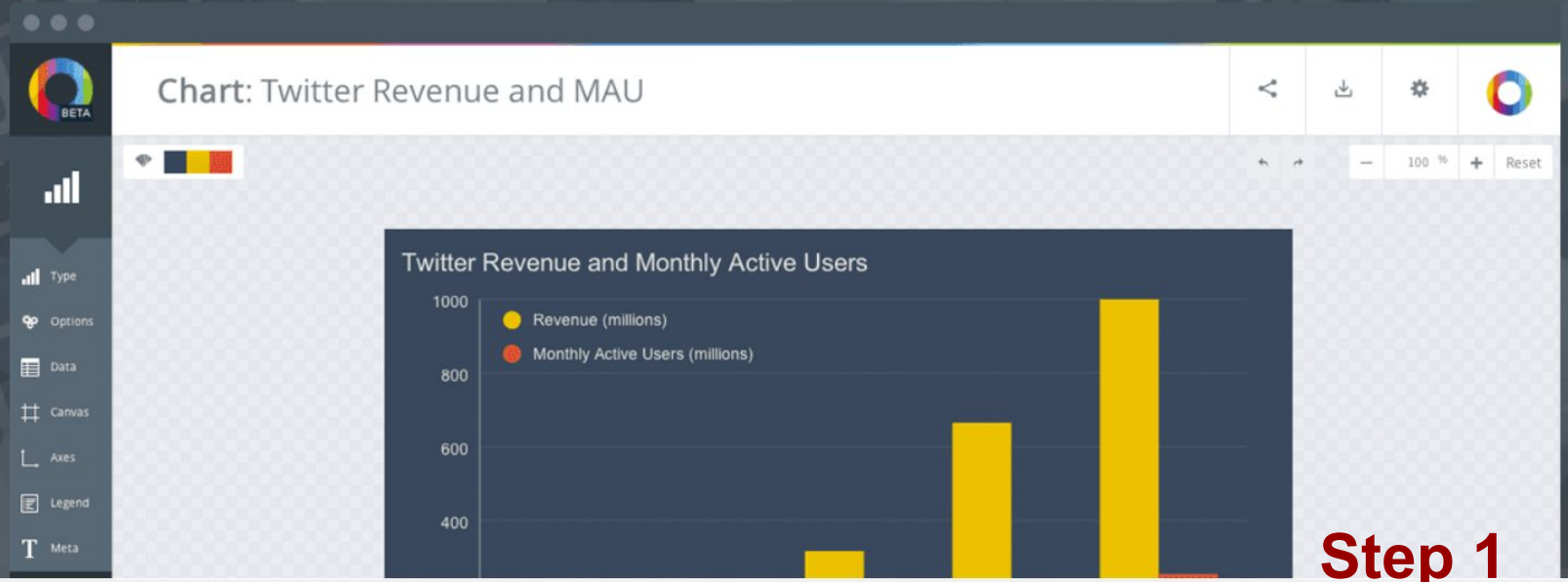
- **Bar Chart**- Played sports by men and women
  - This is a complex dataset, where sports are grouped by gender and year
- **Scatter Plot** - GDP vs Female FIFA Ranking
- **Map** - Female Teams Geography (A)


chartblocks

The online chart building tool

The world's easiest chart builder app. Design and share a chart in minutes.

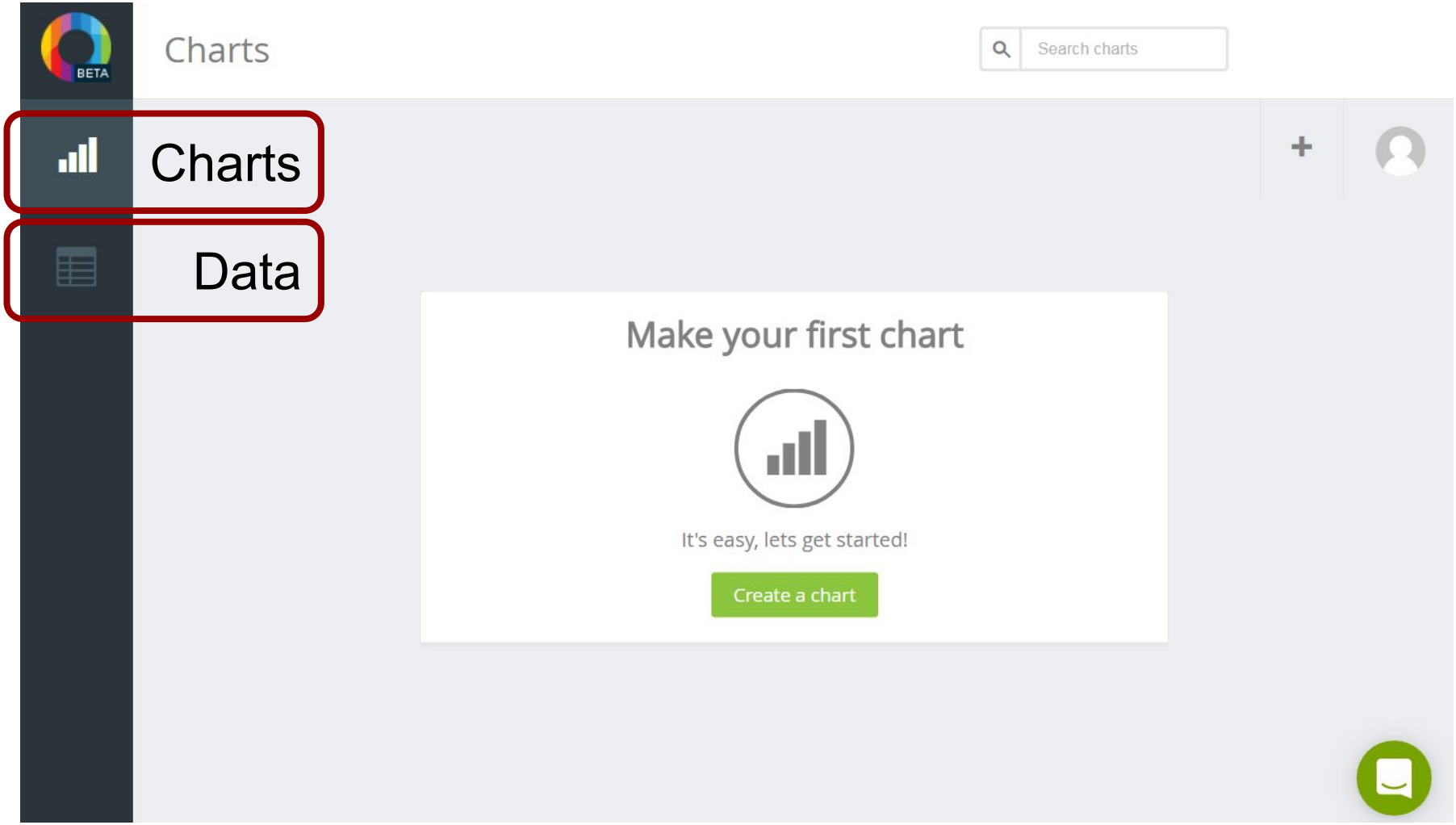
Get started | Features

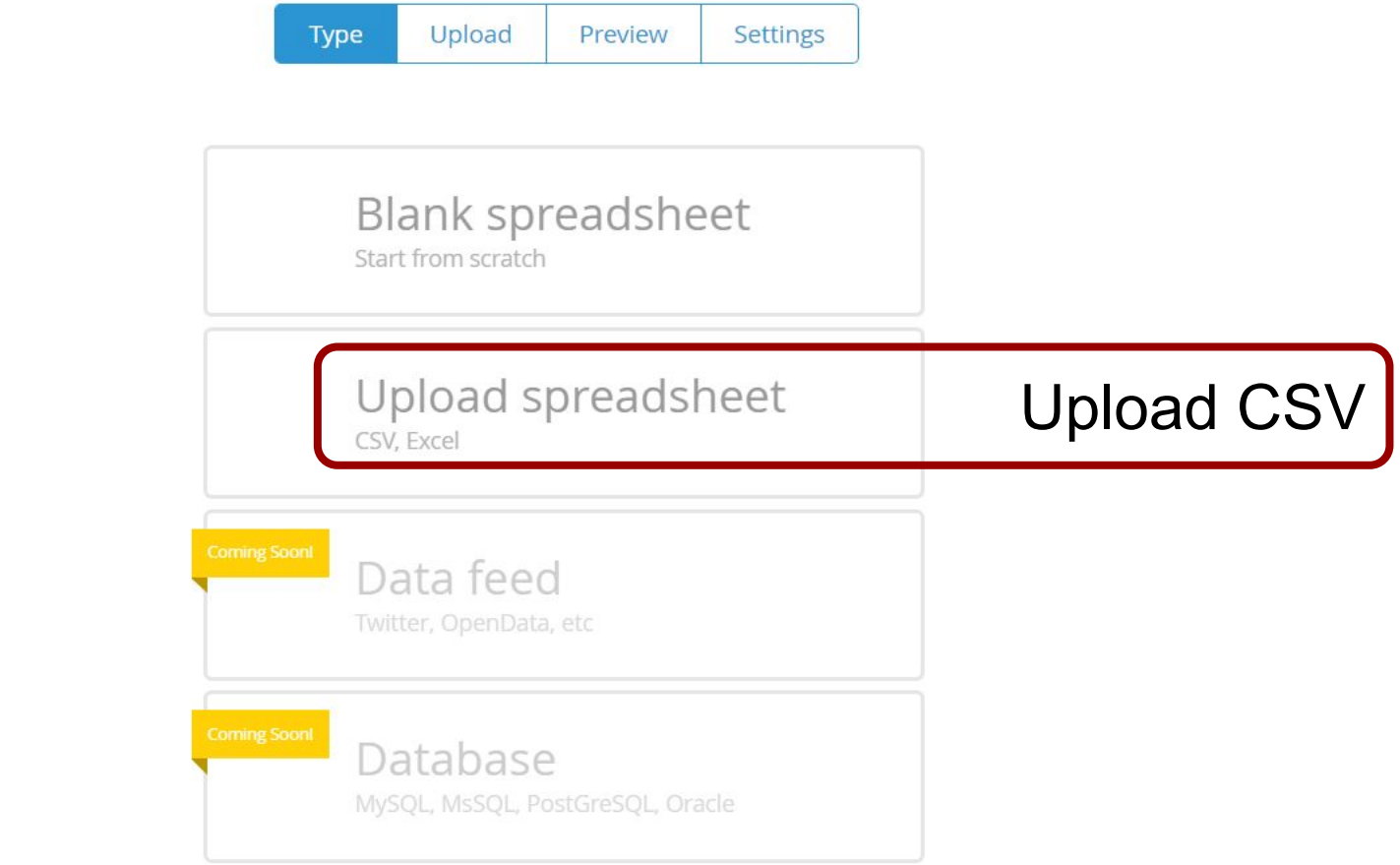


Step 1

Import your data, design your chart and then start sharing it. All within minutes.

Create a free account


Charts
Search charts
Charts
Data
Make your first chart
It's easy, lets get started!
Create a chart


Type
Upload
Preview
Settings
Blank spreadsheet
Start from scratch
Upload spreadsheet
CSV, Excel
Upload CSV
Coming Soon!
Data feed
Twitter, OpenData, etc
Coming Soon!
Database
MySQL, MsSQL, PostGreSQL, Oracle

Datasets: Create

← Previous | Type | Upload | Preview | Settings | Next →

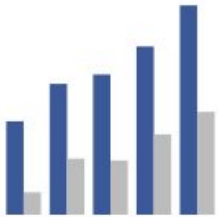
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	GDP	Ranking	Country														
2	561.7787463	106	Afghanistan														
3	4124.98239	71	Albania														
4	36988.62203	115	Andorra														
5	12440.32098	36	Argentina														
6	49755.31548	4	Australia														
7	44757.6349	21	Austria														
8	3878.709257	65	Azerbaijan														
9	41271.48215	22	Belgium														
10	22579.09342	72	Bahrain														

refresh

Sheet Worksheet ▾

Choose a type of chart

You will have the chance to customize your chart later



Column

Bar

Stacked

Line

Area

Stacked

Scatter

Pie

Donut

Election

Scatter Plot

Chart: Create

← Previous | Dataset | Type | Series | Template | Finish | Next →

Select one or more series

Create one or more series of data from your dataset. We've tried to pick the right settings for you.

My data is in rows ☑ Use first row as series name

	A Y	B X	C	D	E	F	G
1	GDP	Ranking	Country				
2	561.7787463	106	Afghanistan				
3	4124.98239	71	Albania				
4	36988.62203	115	Andorra				
5	12440.32098	36	Argentina				
6	49755.31548	4	Australia				
7	44757.6349	21	Austria				
8	3878.709257	65	Azerbaijan				
9	41271.48215	22	Belgium				
10	22579.09342	72	Bahrain				
11	4808.405425	59	Bosnia and He...				
12	4989.427763	52	Belarus				
13	3104.956089	79	Bolivia				
14	8649.948492	8	Brazil				
15	42183.2951	5	Canada				
16	8123.180873	16	China PR				

Series 1: GDP

X Column B

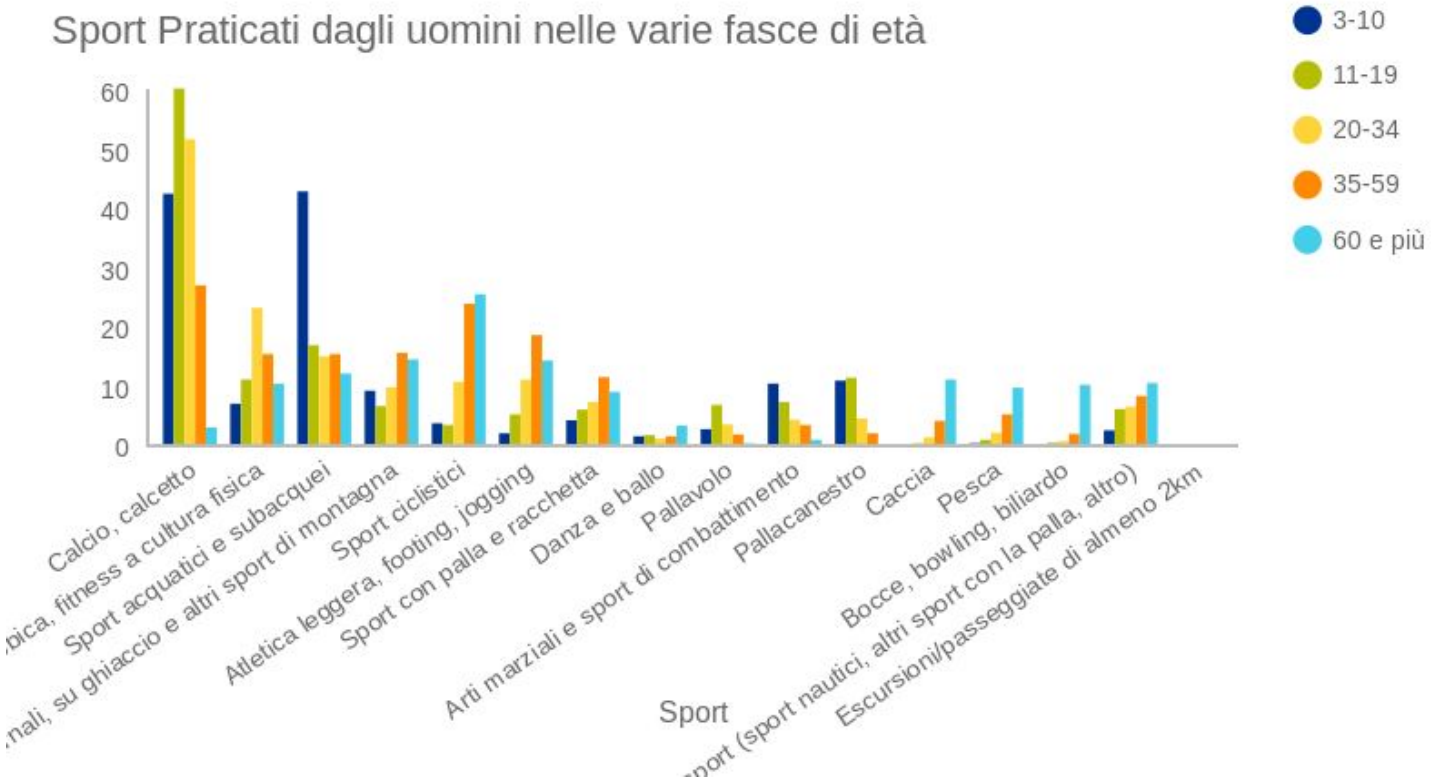
Y Column A

advanced

Select Columns

+ Add another ser

chartblocks



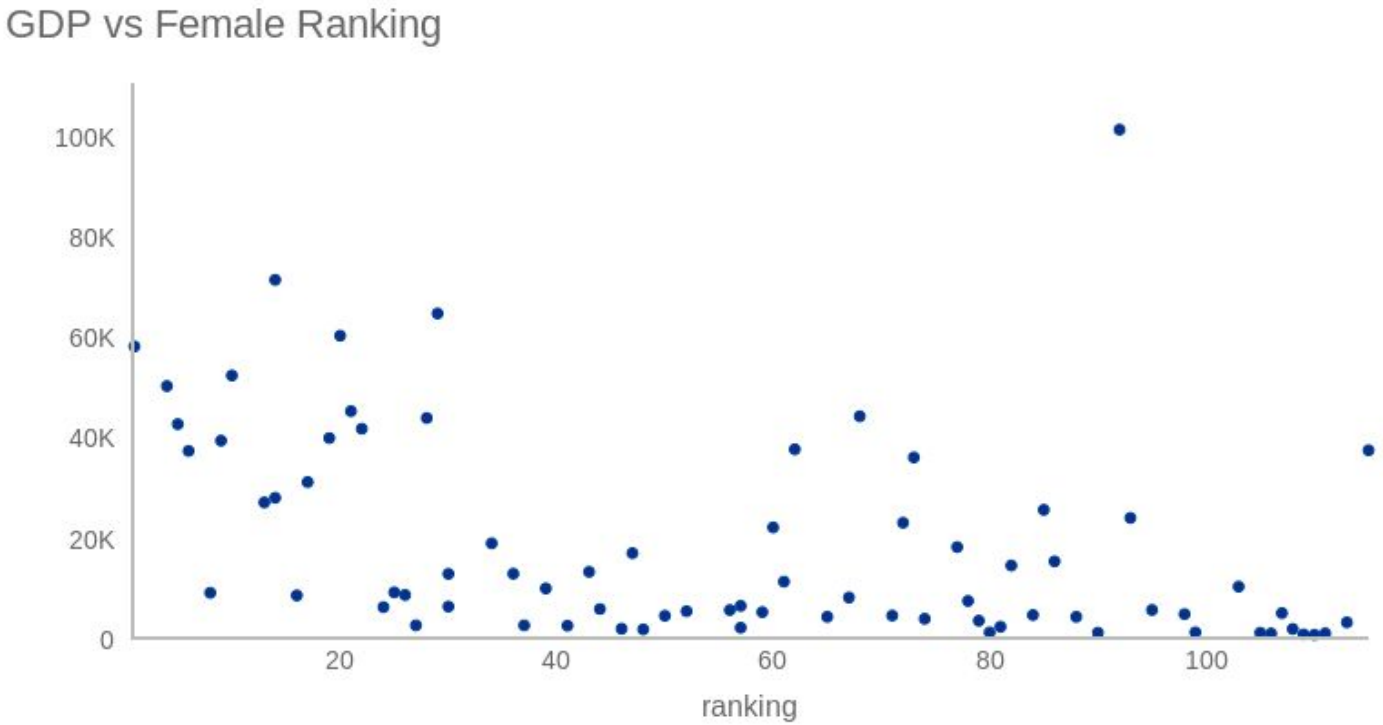
Histogram

Pros

-easy to draw
- both static and interactive charts

Cons

- reduced number of templates
- no complex filtering



Scatter Plot

Pros

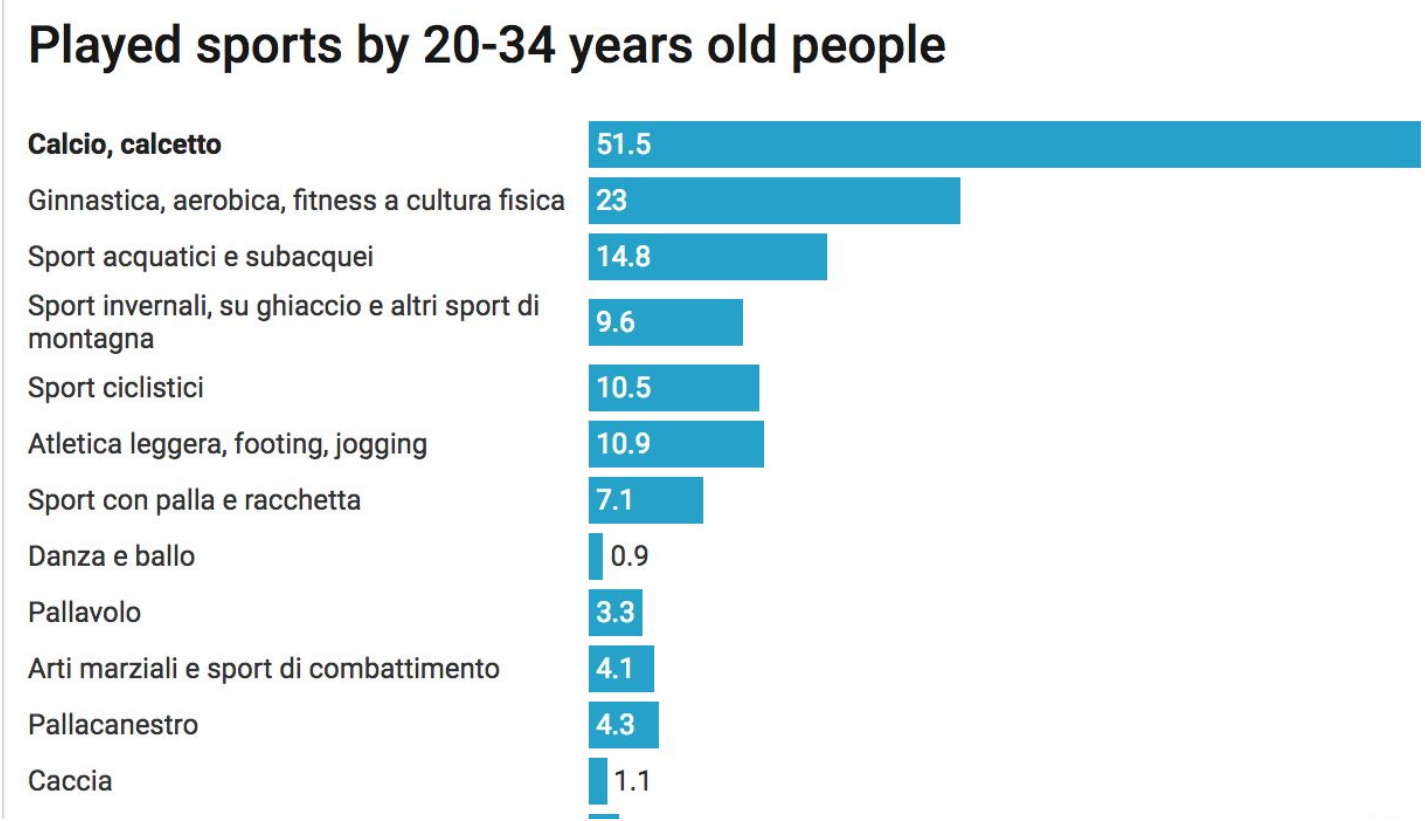
- easy to draw
- both static and interactive charts

Cons

- for each point it is not possible to add a specific label, i.e. the Country name
- no trending

Data Wrapper

Data Wrapper



Histogram

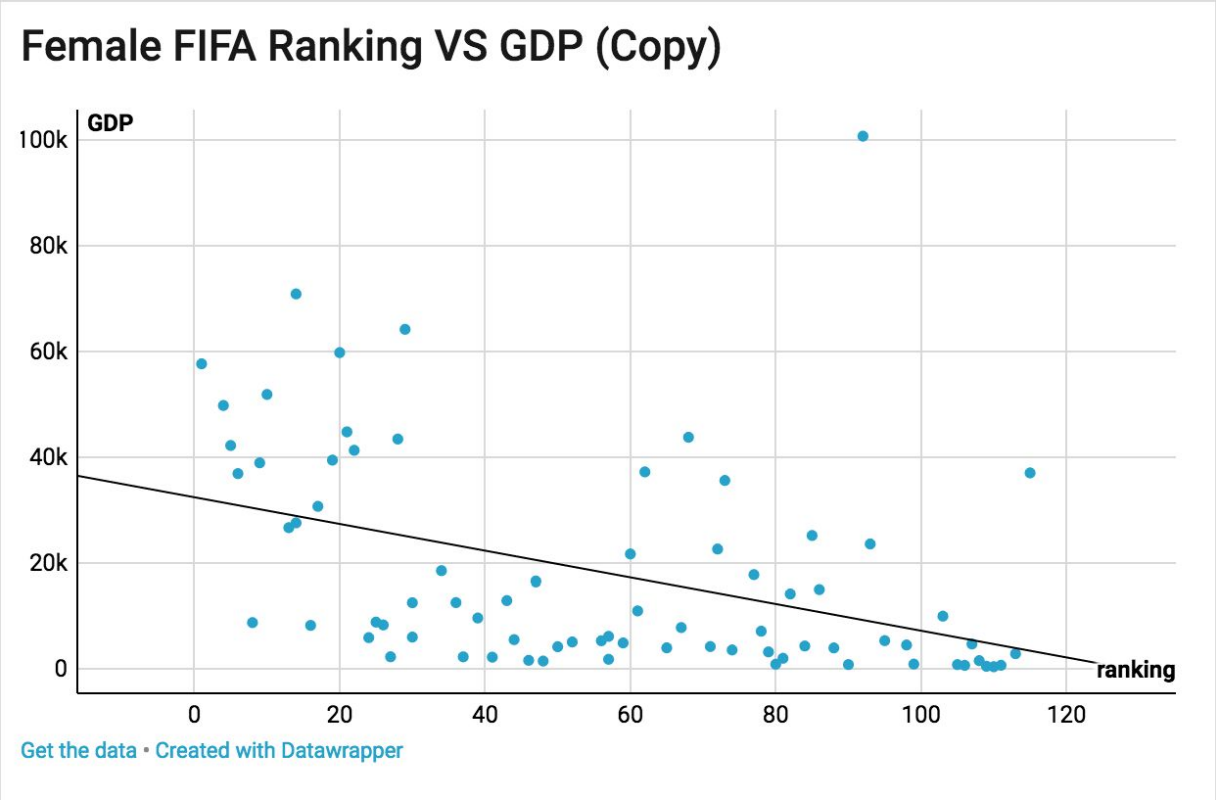
Pros

very fast to draw

Cons

- it is not possible to filter data, e.g. in the example men and women are listed together
- no multiple series

Data Wrapper



Scatter Plot

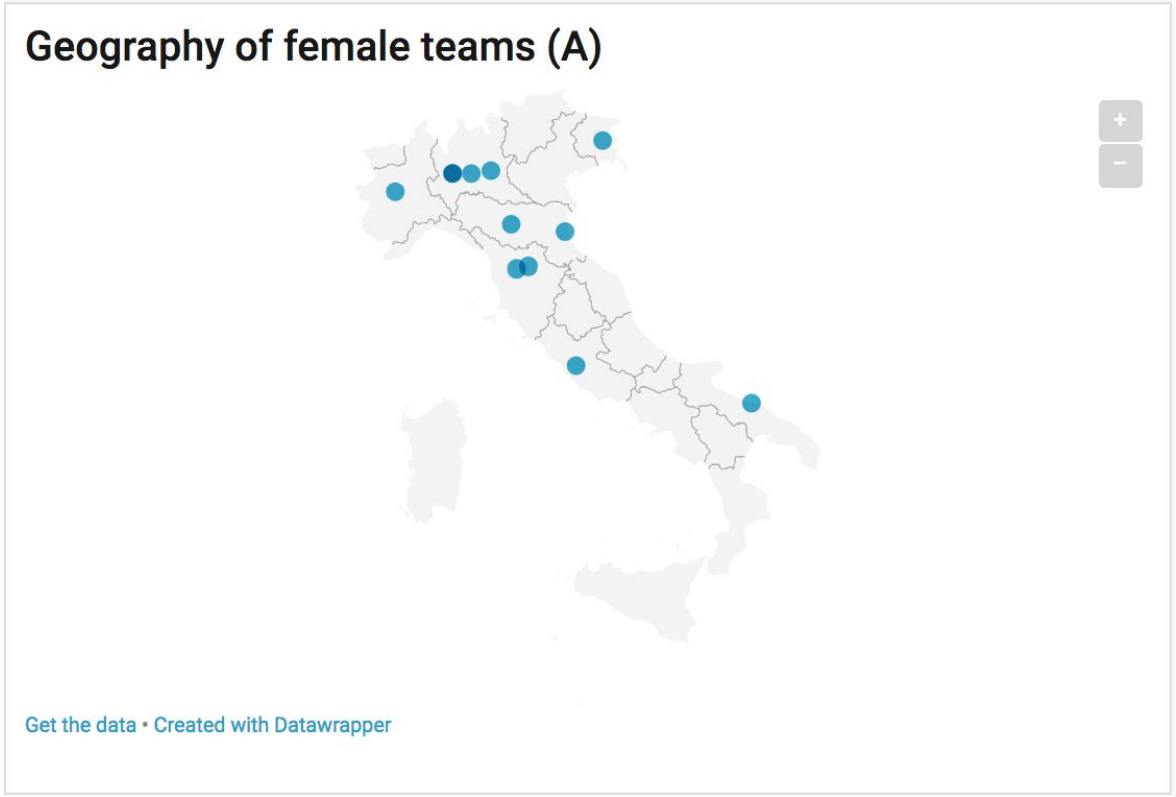
Pros

- very fast to draw
- trending line

Cons

- sometimes it does not work

Data Wrapper



Map

Pros

- very fast to draw
- interactive

Cons

- few features

Google Fusion Table

Una prima visualizzazione con Google Fusion Table

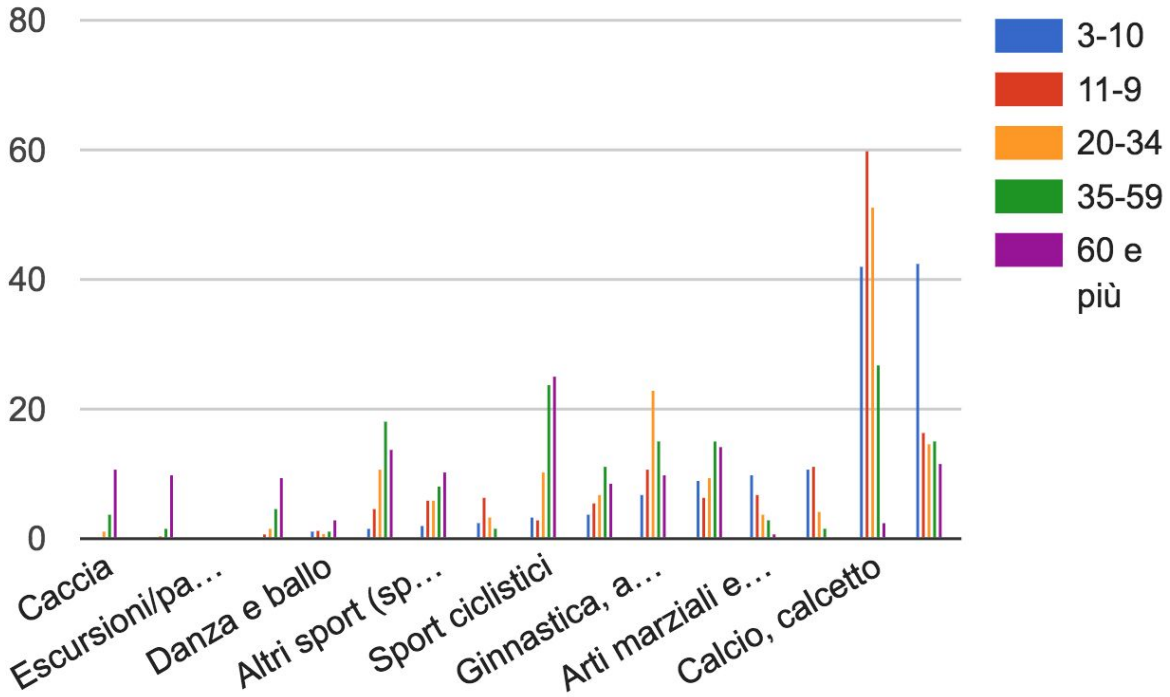
1) collego al sito https://fusiontables.google.com

2) carico il CSV

3) seleziono il tipo di grafico

4) seleziono i dati in base al risultato che voglio ottenere, ad esempio applico filtri ecc

Google Fusion Tables



Histogram

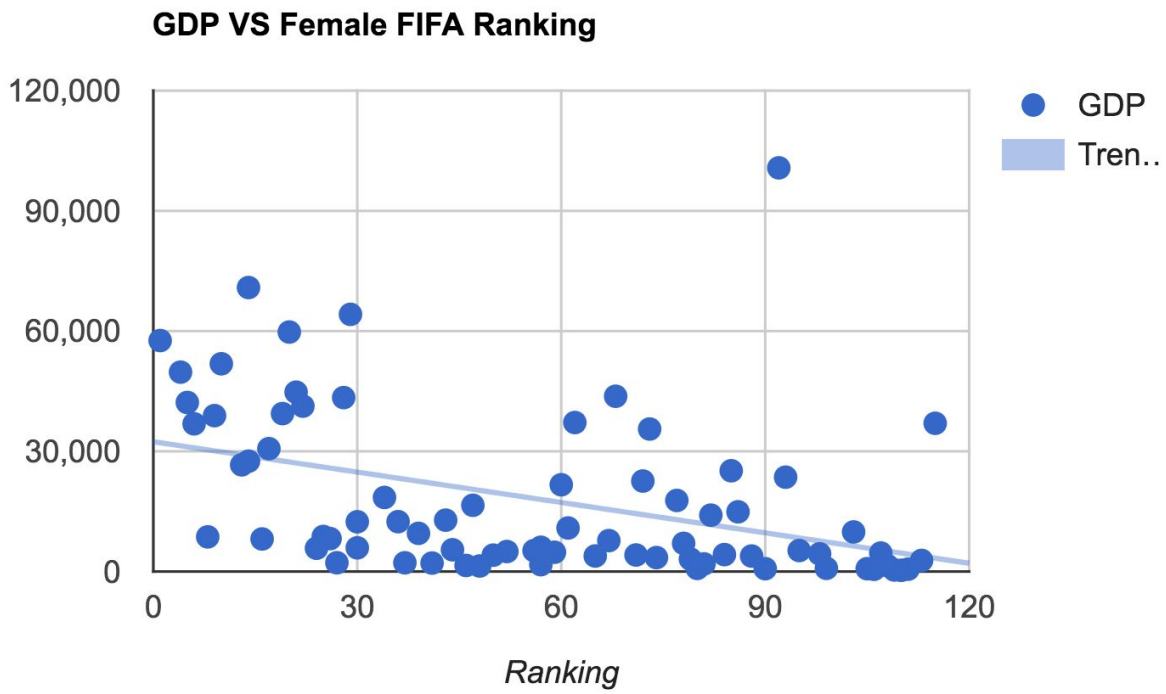
Pros

- data filtering
- very powerful
- interactive

Cons

- it is not simple to use

Google Fusion Tables



Scatter Plot

Pros

- data filtering
- very powerful
- trending line
- interactive

Cons

- it is not simple to use

Google Fusion Tables

Map

Pros

- data filtering
- very powerful
- integrated geocoding

Cons

- support only the territorial and satellite maps


tableau public

Esercizio

Esplorare il dataset uefa.csv (contenente i dati estratti dal PDF) e produrre qualche grafico interessante.

Budget Column

Number	Match
250,00	No
2,500	Yes
25,000	Yes
25,000,00	No

Regular Expression: **"€[0-9]*,(\b[0-9]{3}\b(,?))*$"**

Budget Column (2)

Number	Match
250,00	Yes
2,500	No
25,000	No
25,000,00	Yes

Regular Expression: **"€[0-9]*,(\b[0-9]{3}\b,)"**

```
FLOAT(IF CONTAINS([Budget for women’s football],"No") THEN
    "0"
ELSEIF ISNULL([Budget for women’s football]) THEN
    "0"
ELSEIF REGEXP_MATCH([Budget for women’s football],
"€[0-9]*,(\b[0-9]{3}\b(,?))*$") THEN
    REPLACE(REPLACE([Budget for women’s football], "€",""),",","")
ELSEIF REGEXP_MATCH([Budget for women’s football], "€[0-9]*,(\b[0-9]{3}\b,)")
THEN
    REPLACE(REPLACE(REPLACE(LEFT([Budget for women’s football],
    LEN([Budget for women’s football])-2),",","") + RIGHT([Budget for women’s
    football], 2), "€",""),",",".")
ELSE
    REPLACE(REPLACE([Budget for women’s football], "€",""),",",".")
END)
```

References

Tableau Tutorial (Italian)