WEB SCIENCE PROJECT

MILAN DESIGN WEEK - FUORISALONE

Roberto Napoli – 861026

Analysis of the context

- Starting point: information on FuoriSalone (slides + explanation in class)
- FuoriSalone official website (2016)
- FuoriSalone official app (2017)

- Outcome: greater understanding of the datasets

Analysis of the datasets

- Analysis of each attribute of each dataset
- Definition of "key" attributes to link heterogeneous datasets:
  - Location: (latitude, longitude)/ID
  - User ID
  - Event ID
- Test of doubts on the datasets:
  - agenda.json, agenda_analytics.json

PROBLEM DEFINITION

- Determine the *top-k popular* events and get insights out of them
- Three perspectives:
  - Aggregated information but possible indicator of physical presence
    - Foursquare check-ins
    - Phone activities
  - Fine-grained user interest information but not direct physical presence
    - Agenda
    - Event actions

Tool usage

- Starting tool for data analysis:
  - Python + Python libraries (Pandas, Numpy, Matplotlib...)
- Doubt: will Python and its libraries be enough to execute all the analysis on large datasets and a 4GB laptop?
- Answer: yes, some minutes are needed to run all the code

Datasets

- fuorisalone_foursquare-history-2016-04.csv
- phone.csv
- agenda_analytics.json
- event.json
- event_actions_analytics.json
- location.json

Solution (1/3)

- Foursquare check-ins analysis methodology:
  - Selection of the check-ins during the days of the event
  - Comparison of each check-ins count record with the location of each event
    - Issue: precision of the geo-location number
      - From 4 digits of precision (~11.1m) to 3 digits (~111m)
  - Comparison with the dates of the events (day – time)
  - Ordering by decreasing number of users
 - Selection of the top-10 elements

Solution (2/3)

- User interest analysis methodology:
  - Aggregation of users by event
  - Selection of actions for event actions
  - Ordering by decreasing number of users
  - Selection of the top-10 elements

Solution (3/3)

- Phone activities analysis methodology:
  - Data cleansing on NaN values (*agerange, gender, activity* columns)
  - Selection of the phone activities during the days of the event
  - Comparison of each phone activity count record with the location of each event
    - Issue: precision of the geo-location number
      - From 4 digits of precision (~11.1m) to 3 digits (~111m)
  - Comparison with the dates of the events (day – time)
  - Ordering by decreasing number of users
  - Selection of the top-10 elements

Results (1/3)

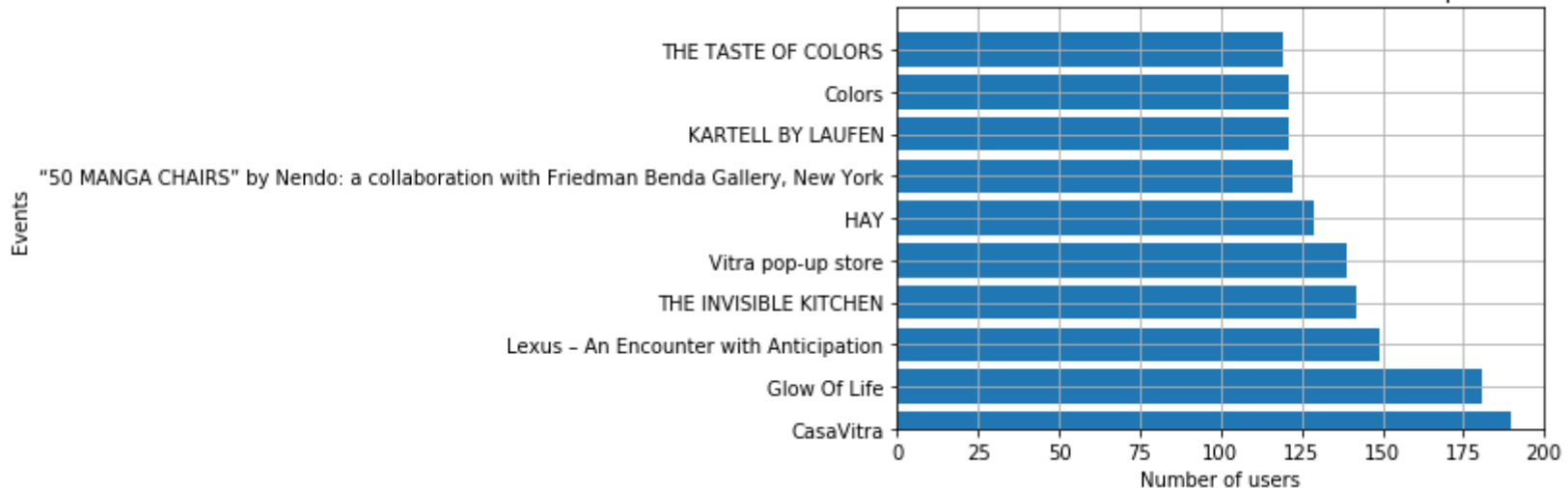- Foursquare check-ins:

Plot of the number of checkins
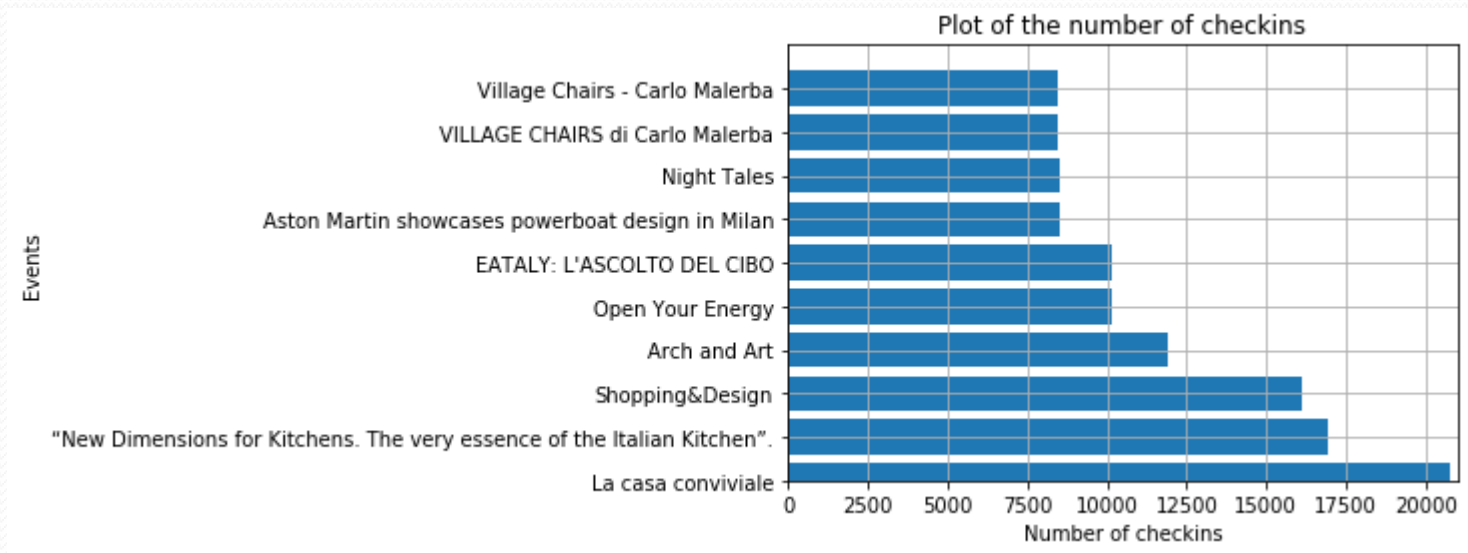
Results (2/3)

User Interest

- Agenda

Plot of the number of users who added an event to his personal agenda

Results (2/3)

User interest

- Map

Plot of the number of users who have looked on the map to find an event
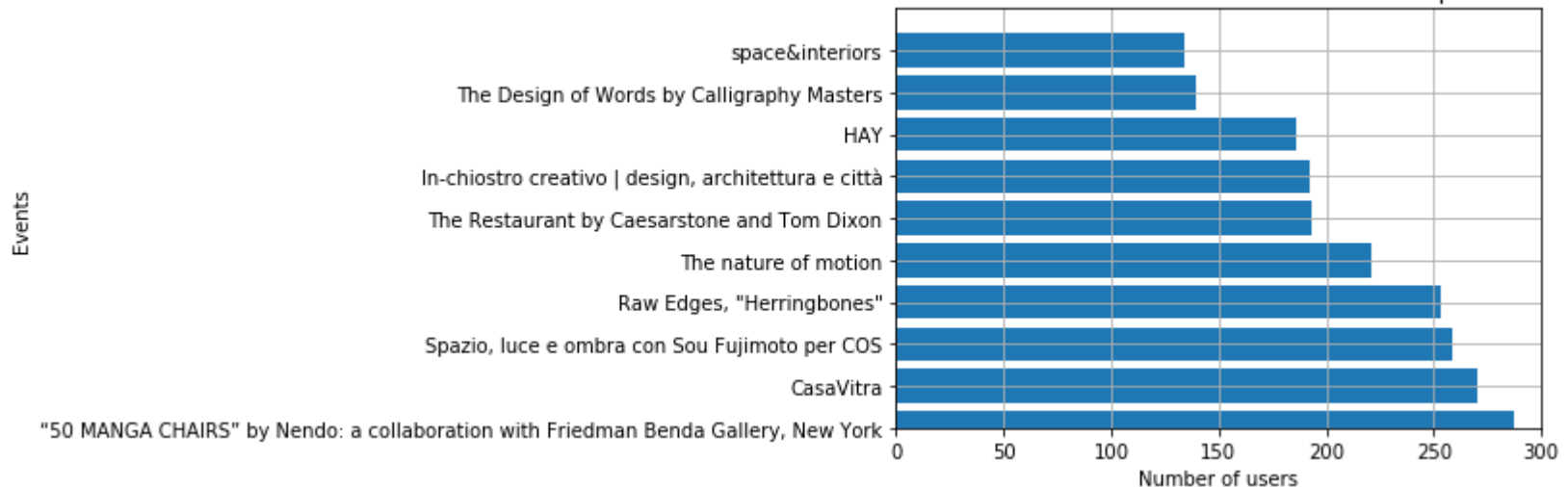
Results (2/3)

User interest

- Photos

Plot of the number of users who have looked at the photos of an event

Results (2/3)

User interest

- Share

Plot of the number of users who have shared an event on the social medias
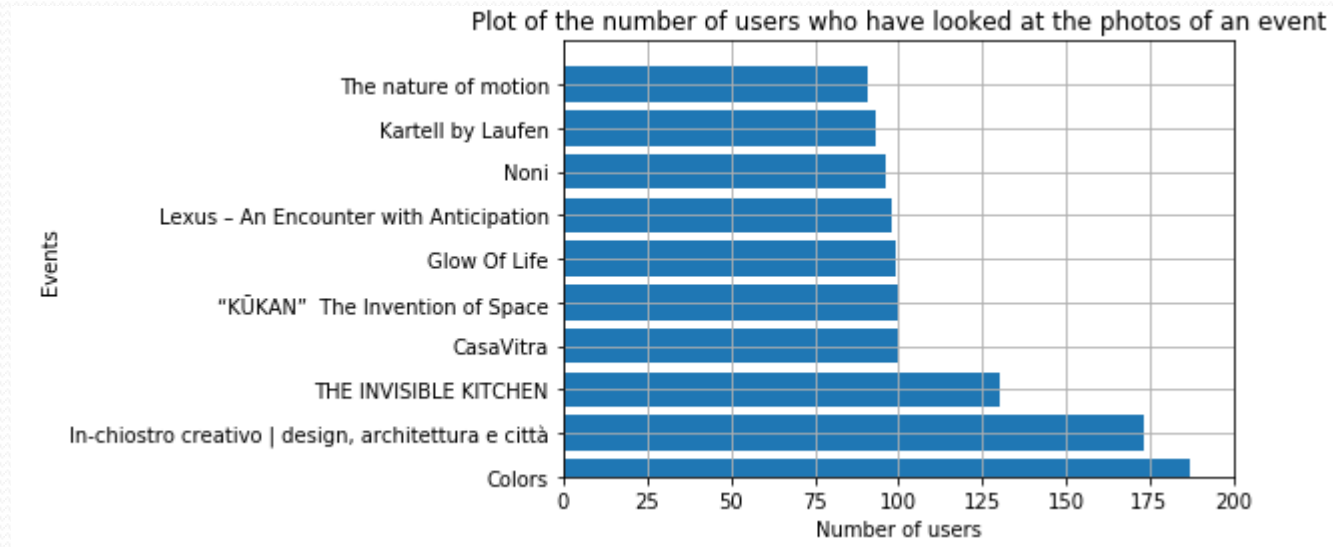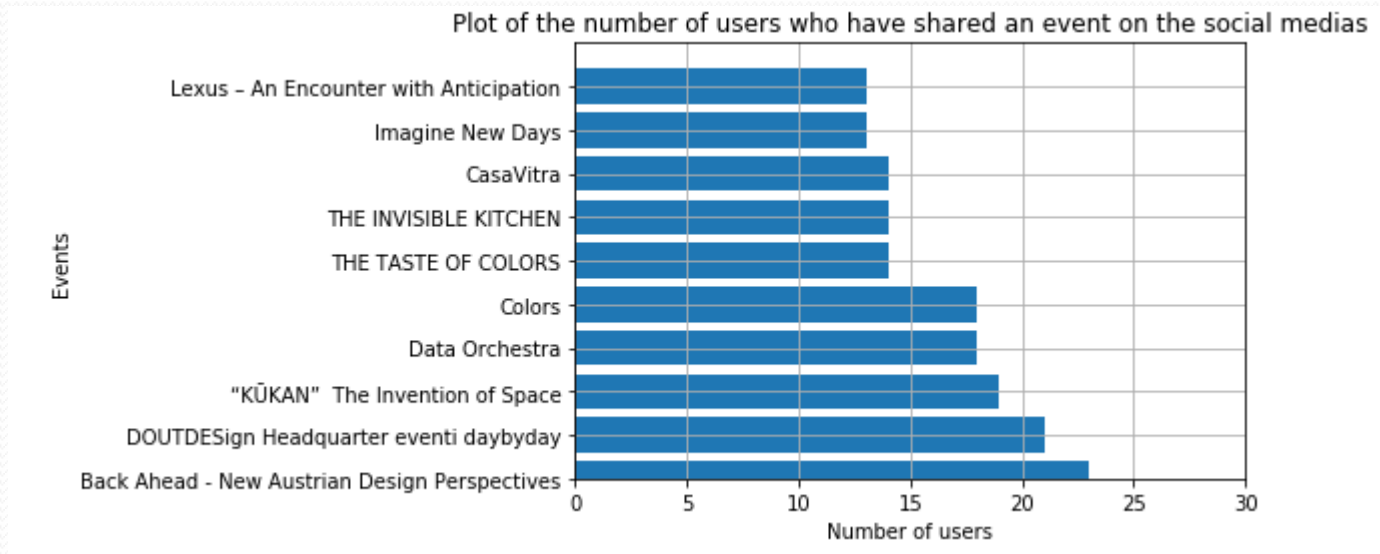
Results (2/3)
User interest

- Agenda and user interest

Plot of the number of users who have shown an interest to an event
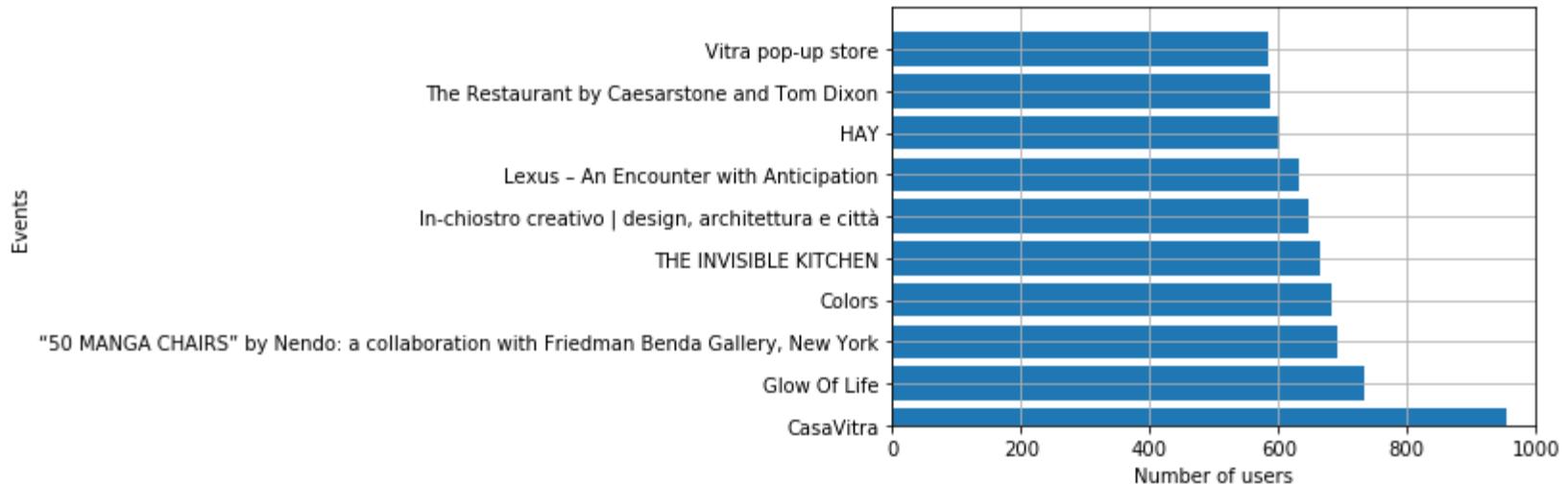
Results (3/3)

- Phone activities:

Plot of the number of phone activities
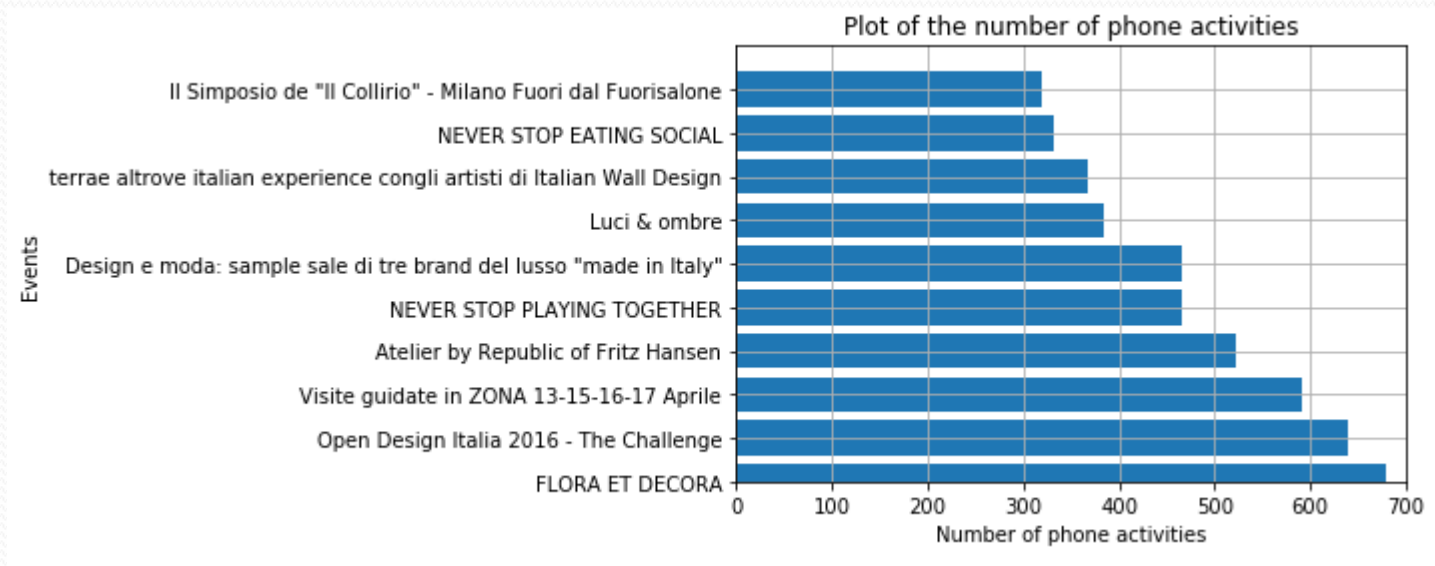
Results (3/3)

- Female phone activities:

Plot of the number of phone activities by females
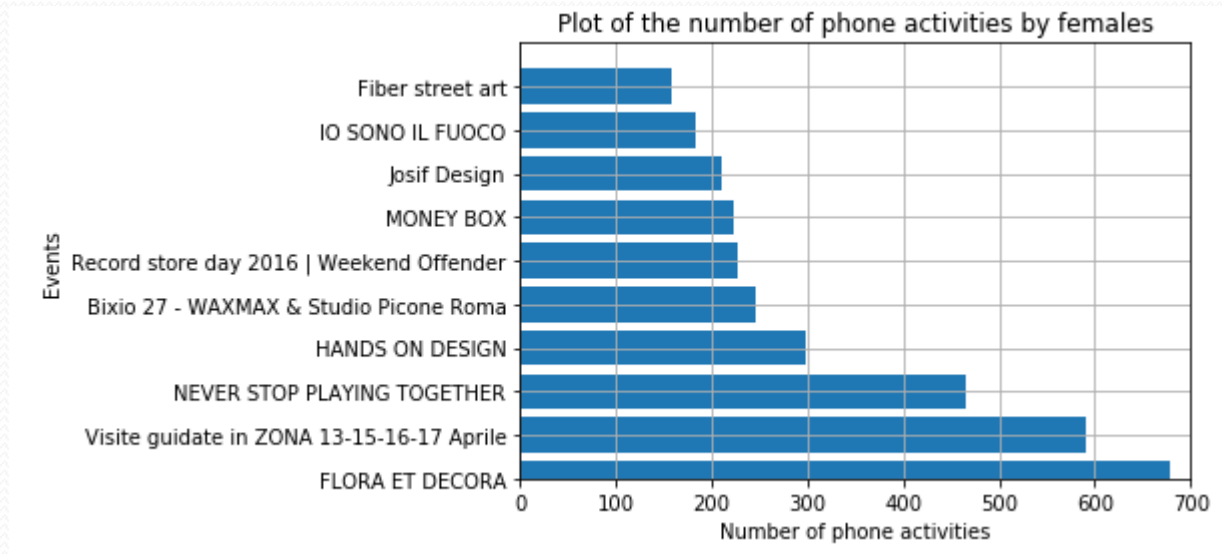
Results (3/3)

- Male phone activities:
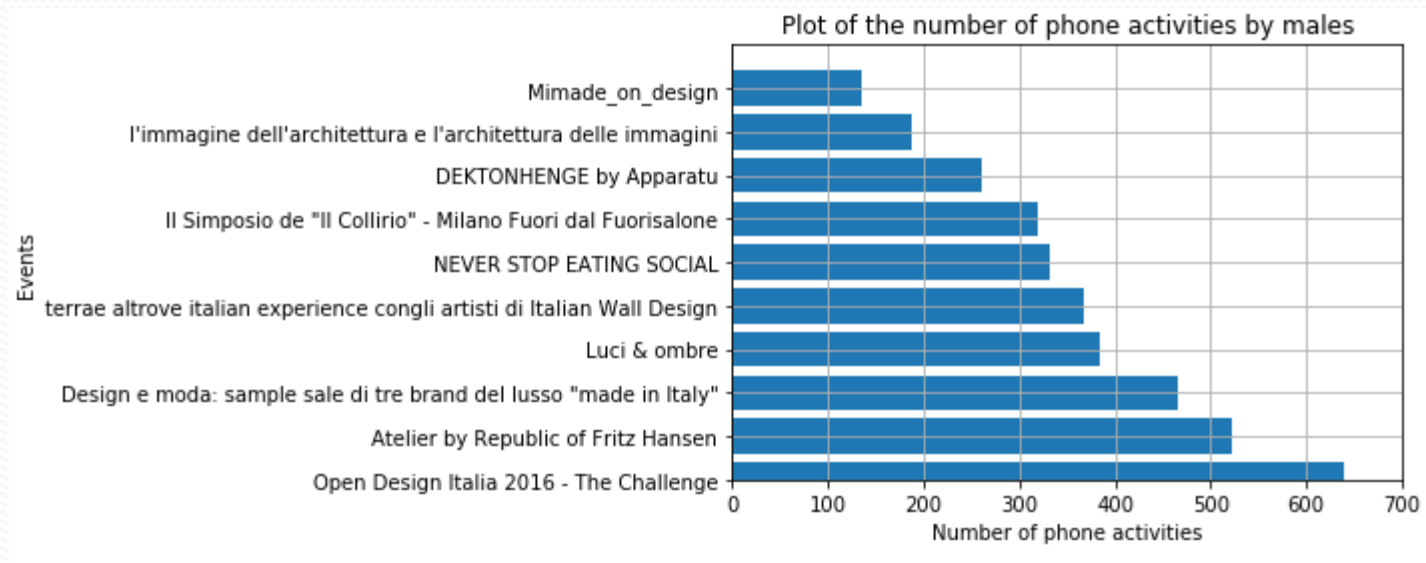
Results (3/3)

- Young phone activities:
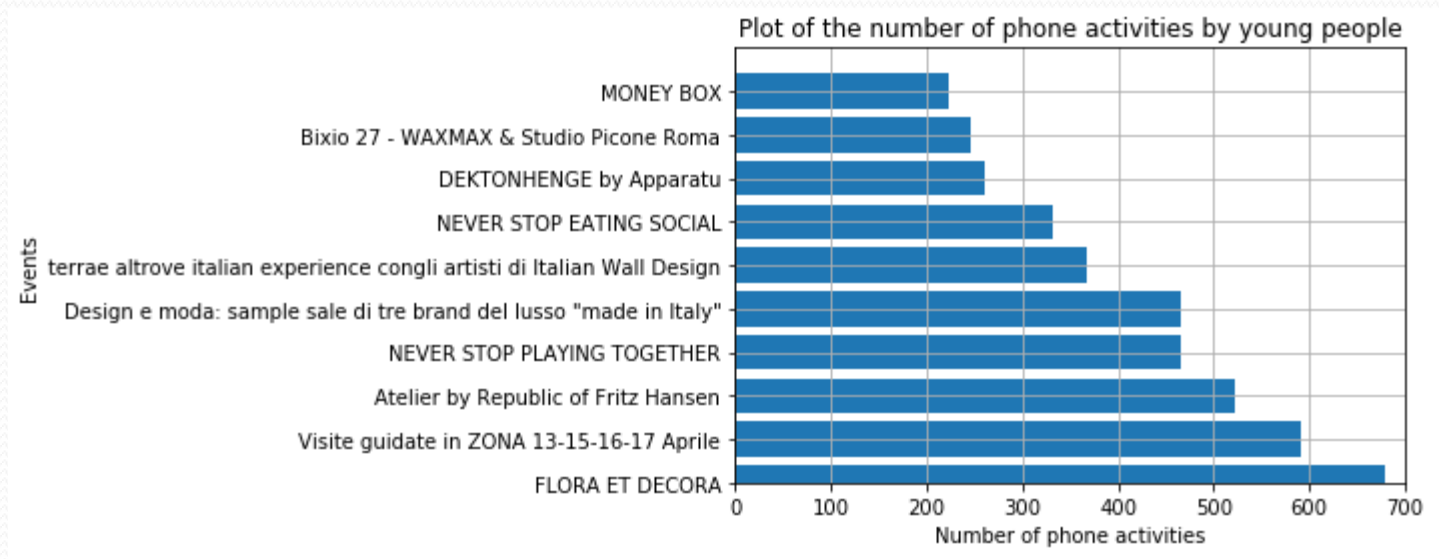
Results (3/3)

- Senior phone activities:

Plot of the number of phone activities by senior people
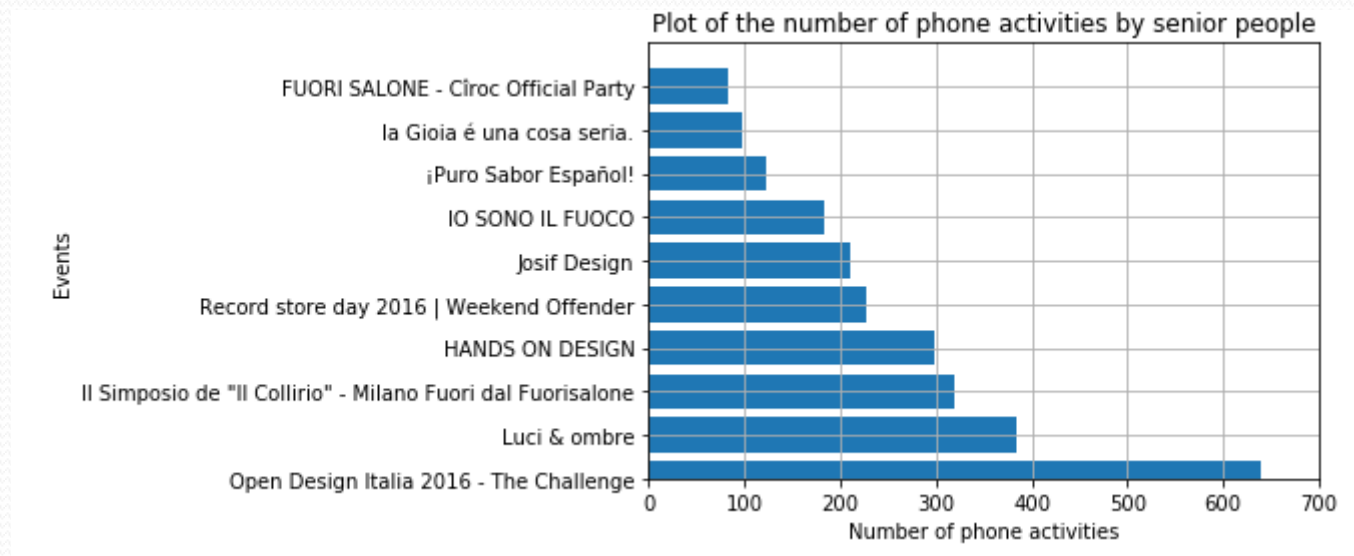
Results (4/9)

- Award-winning events
  - Shared events: 9/10
  - Events photos: 6/10
  - Agenda + user interest: 5/10
  - Agenda: 4/10
  - Events on map: 3/10
  - Check-ins: 0/10
  - Phone activities: 0/10

Results (5/9)

- Open-party events
  - Events photos: 4/10
  - Shared events: 3/10
  - Events on map: 3/10
  - Check-ins: 3/10
  - Agenda + user interest: 2/10
  - Agenda: 2/10
  - Phone activities: 2/10
    - Male: 3/10
    - Female: 2/10
    - Young: 2/10
    - Senior: 4/10

Results (6/9)

- Type of events

		type								
		esposizione	festa_cocktail	show	workshop	incontro	conferenza	presentazione	press_view	festa_riservata
dataset	check-ins	16,67%	16,67%	16,67%	16,67%	16,67%				16,67%
	agenda	75,00%	16,67%	8,33%						
	map	62,50%	18,75%		6,25%	6,25%	6,25%			
	photos	62,50%	25,00%	6,25%			6,25%			
	share	60,00%	20,00%	6,67%				6,67%	6,67%	
	user interest	71,43%	14,29%	7,14%	7,14%					
	phone activities	46,15%	15,38%		7,69%	7,69%		15,38%	7,69%	
	male phone	41,18%	17,65%			5,88%		17,65%	11,76%	5,88%
	female phone	58,33%	16,67%		8,33%				8,33%	8,33%
	young phone	43,75%	12,50%		6,25%			12,50%	18,75%	6,25%
	senior phone	45,45%	36,36%			9,09%				9,09%
	TOT	54,73%	18,92%	3,38%	4,05%	3,38%	1,35%	5,41%	5,41%	3,38%

Results (7/9)

- Brands names

	name	0
96	Vitra	7
47	LAGO	6
57	Miele	4
51	Lexus	4
15	Barovier&Toso	4
35	Gianlisa Handbags	3
105	evviva sangiovese di toscana	3
102	d'Officina	3
95	Valentina Brugnatelli Jewels	3
85	Studio Auriga	3

Results (7/9)

- Brands country

dataset	brand country	IT	JP	GB	DE	IL	CH	AT	NL	DK
	check-ins	91,67%		8,33%						
	agenda	60,00%	20,00%		20,00%					
	map	66,67%	11,11%	11,11%		11,11%				
	photos	66,67%	11,11%		11,11%		11,11%			
	share	44,44%	11,11%		11,11%			22,22%	11,11%	
	user interest	55,56%	11,11%	11,11%	11,11%	11,11%				
	phone	84,62%					7,69%			7,69%
	male phone	76,92%					15,38%			7,69%
	female phone	90,00%		10,00%						
	young phone	84,62%					7,69%			7,69%
	senior phone	88,89%		11,11%						
	TOT	75,68%	4,50%	4,50%	3,60%	1,80%	4,50%	1,80%	0,90%	2,70%

Results (8/9)

- Designers names

	name	surname	0
45	Paola	Navone	4
0	Alice	Tamburini	3
15	Elise	Brasca	3
61	Valentina	Brugnatelli	3
56	Studio	Roso	3
46	Paola	Silva Coronel	3
44	Oky Sato	Nendo	3
26	Jaime	Hayon	3
33	Marco	Garello	3
4	Architecture	Wednesday	3

Results (8/9)

- Designers country

dataset	designer country								
		IT	JP	GB	NL	DK	FI	FR	NO
	check-ins	91,67%		8,33%					
	agenda	50,00%	50,00%						
	map	25,00%	25,00%	50,00%					
	photos	25,00%	8,33%				66,67%		
	share	33,33%	33,33%		33,33%				
	user interest	50,00%	25,00%	25,00%					
	phone	50,00%		16,67%		16,67%		16,67%	
	male phone	66,67%		8,33%		8,33%		8,33%	8,33%
	female phone	100,00%							
	young phone	40,00%		20,00%		20,00%		20%	
	senior phone	100,00%							
	TOT	59,42%	7,25%	10,14%	1,45%	4,35%	11,59%	4,35%	1,45%

Results (8/9)

- Designers birthdate

		designer birthdate					
		1940-1950	1950-1960	1960-1970	1970-1980	1980-1990	1990-2000
dataset	check-ins	33,33%	33,33%	33,33%			
	agenda						
	map						
	photos				16,67%	83,33%	
	share						
	user interest						
	phone			50,00%	50,00%		
	male phone			25,00%	25,00%	25,00%	25,00%
	female phone						
	young phone			50,00%	50,00%		
	senior phone						
	TOT	7,14%	7,14%	21,43%	14,29%	42,86%	7,14%

Results (8/9)

- Designers sex

		designer sex	
		male	female
	check-ins	50,00%	50,00%
	agenda	50,00%	50,00%
	map	75,00%	25,00%
	photos	27,27%	72,73%
	share	50,00%	50,00%
dataset	user interest	50,00%	50,00%
	phone	28,57%	71,43%
	male phone	33,33%	66,67%
	female phone	100,00%	
	young phone	16,67%	83,33%
	senior phone	100,00%	
	TOT	42,65%	57,35%

Results (9/9)

- Categories (1/2)

		Architecture	Art	Exteriors	Conferences	Fashion	Food & beverage	Furniture	Games	Graphic Design
	check-ins	3,00	4,00		1,00	1,00	1,00	4,00		2,00
	agenda	1,00					2,00	2,00		2
	map	2,00	3,00		1,00		2,00	4,00		1
	photos	2,00	3,00		1,00		2,00	2,00		
	share	1,00	2,00		1,00		3,00	4,00		1
dataset	user interest	1,00	1,00	2,00	1,00		3,00	3,00	1,00	1,00
	phone	1,00	1,00	2	1,00	1,00	3,00	2,00	1,00	1,00
	male phone	2,00	3,00	3		1,00	3,00	4,00	1	2
	female phone	3,00	3,00	3	1,00	2,00		3,00	1	1,00
	young phone	2,00	2,00	2	1,00	2,00	2,00	3,00	1,00	1,00
	senior phone	1,00	3,00			1,00	2,00	2,00		2,00
	TOT	19,00	25,00	12,00	8,00	8,00	23,00	33,00	5,00	14,00

Results (9/9)

- Categories (2/2)

category										
Installations	Jewels	Kitchen	Lightning	Living	Other	Party	Product Design	Special materials	Technology	Bathroom
4,00		1,00	1,00	1,00	2,00		3,00	1,00	2,00	1
4,00		1,00	1,00	1,00		1,00	7,00	1,00	2,00	
5,00		1,00	1,00	1,00	1,00	2	7,00	1,00		1
4,00	1		3,00	2,00	1,00		6,00	2,00	3,00	
5,00		1,00	3,00	4,00	1,00	1,00	4,00	1,00	4,00	
6,00	1,00	1,00	2,00	1,00		2,00	7,00		2,00	
1,00	1,00		1,00	1,00	1,00	3,00	3,00	1,00		
4,00	1		1,00	2,00	1,00	3,00	5,00	2,00	1,00	
3,00	1	1,00	1,00	2,00		2	3,00	1,00	1,00	
3,00	1,00		1,00	1,00	1,00	1,00	3,00	1,00		
2,00	1,00		1,00	1,00		3,00	3,00	1,00	1,00	
41,00	7,00	6,00	16,00	17,00	8,00	18,00	51,00	12,00	16,00	2