



UNIVERSIDAD
DE SALAMANCA

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

Graphs

Graphs

Remarks

Final

Studying coincidences with network analysis and other statistical tools

M. Escobar_(modesto@usal.es)

Universidad de Salamanca

2014 Nordic and Baltic Stata Users Group meeting

5th September



The aims of this presentation are:

- To show *coincidence analysis*, which is a statistical framework to study concurrence of events in large sets of scenarios combining network analysis with multivariate statistics.
- To present *coin*, an ado program that is able to perform this analysis.
- As an example, an analysis of people in the picture albums of three eminent characters in the early 20th century will be presented.
- I also applied this analysis to
 - Audience figures.
 - Content analysis of media and textbooks.
 - Multiresponse analysis in questionnaires.



Coincidence analysis

Definition

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

Graphs

Graphs

Remarks

Final

- Coincidence analysis is a set of techniques whose object is to detect which people, subjects, objects, attributes or events tend to appear at the same time in different delimited spaces.
- These delimited spaces are called n scenarios, and are considered as units of analysis (i).
- In each scenario a number of J events X_j may occur (1) or may not (0) occur.
- We call incidence matrix (\mathbf{X}) an $n \times J$ matrix composed by 0 and 1, according to the incidence or not of every event X_j .
- In order to make comparative analysis of coincidences, these scenarios may be classified in H sets



3 grades of coincidence

Mere and probable events

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

Graphs

Graphs

Remarks

Final

- Two events (X_j and X_k) are defined as 1) **merely** coincident if they occur in the same scenario at least once:

$$[\exists_i (x_{ij} = 1 \wedge x_{ik} = 1)] \vee f_{jk} \geq 1$$

- Additionally, two events (X_j and X_k) are defined as 2) **probably** coincident if they occur more frequently than if they are independent:

$$f_{jk} > \frac{f_{jj} f_{kk}}{n}$$



3 grades of coincidence (cont.)

Statistically probable events

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

Graphs

Graphs

Remarks

Final

- And two coincidences are 3) **statistically probable** if the joint frequency of their events meets one of the following inequalities:

$$P(r_{jk} \leq 0) < c$$

$$P(\theta_{jk} \leq 1) < c$$

$$P(p(X_j) - p(X_j|X_k) \leq 0) < c$$

- where r_{jk} is the Haberman residual, θ_{jk} is the odd ratio, and the third equation represents a one tailed Fisher exact test. Furthermore, c is the selected level of significance, normally 0.05)



Adjacencies

Definition for statistically probable events

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

Graphs

Graphs

Remarks

Final

- Two events j and k can be considered adjacent according to the following rule:

$$A[j, k] = 1 \Leftrightarrow [P(r_{jk} \leq 0) < c] \wedge j \neq k$$

- Therefore, a $J \times J$ matrix \mathbf{A} may be elaborated with 0 valued diagonal elements and 1 in the case where r_{jk} is significantly below the level c . Other elements should also be 0.
- From \mathbf{A} the $J \times J$ distance matrix \mathbf{D} , with geodesics (shortest paths between nodes), can be obtained.



Adjacencies (cont.)

Definition for mere and probable coincidences

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

Graphs

Graphs

Remarks

Final

- By extension, other adjacency matrices can be elaborated following

- Mere coincidence criterion

$$A[j, k] = 1 \Leftrightarrow f_{jk} \geq 1$$

- Or probable coincidence criterion

$$A[j, k] = 1 \Leftrightarrow [\mathbb{P}(r_{jk} \leq 0) < 0.5] \wedge j \neq k$$



Bar plots

Definition

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

Graphs

Graphs

Remarks

Final

- A incidences plot is the representation of the frequencies of the events with proportional to their size bars.
- A coincidences plot is a composed graph of incidences and coincidences. Every event has its own coincidences plot.



Bar plots of incidences/coincidences

Different patterns of coincidences

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

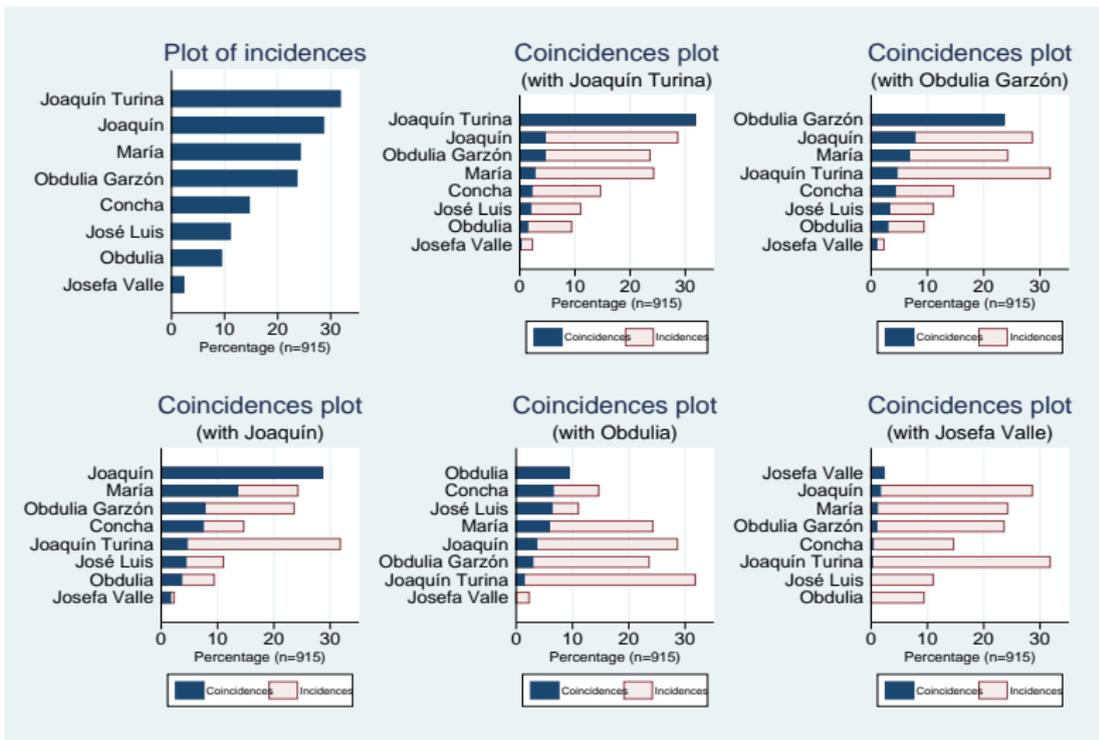
Masó

Graphs

Graphs

Remarks

Final





Graph Definition

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

Graphs

Graphs

Remarks

Final

- “A graph \mathcal{G} consist of two sets of information: a set of Nodes (events), $\mathcal{N} = \{n_1, n_2, \dots, n_g\}$, and a set of lines (coincidences), $\mathcal{L} = \{l_1, l_2, \dots, l_L\}$ between pair of nodes ”. (Wasserman and Faust 1994).
- A non trivial problem is where to draw each node, i.e, the spatial distribution of the nodes.



Spatial distribution of nodes

Five alternatives

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

Graphs

Graphs

Remarks

Final

- Network (Moreno 1934) and coincidence (Escobar 2009) analyses, based on Haberman residuals of \mathbf{F} (circular coordinates).
- Other mappings of the adjacency matrix, based on Haberman residuals, can be used: multidimensional scaling (MDS) and cluster analysis.
- Or via correspondence analysis (Benzecri 1973), using matrix \mathbf{X} as input and obtaining only column coordinates (incidents).
- Alternatively, we can obtain coordinates of events with principal component analysis (Pearson, 1901) using tetrachoric correlations (Everitt 1910).
- Or a biplot (Gabriel 1971) can be drawn with events as variables and suppressing scenarios (rows)



Multi-dimensional scaling graph of coincidences

Mere (· · ·), probable(- -) and statistically probable (—) coincidences

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

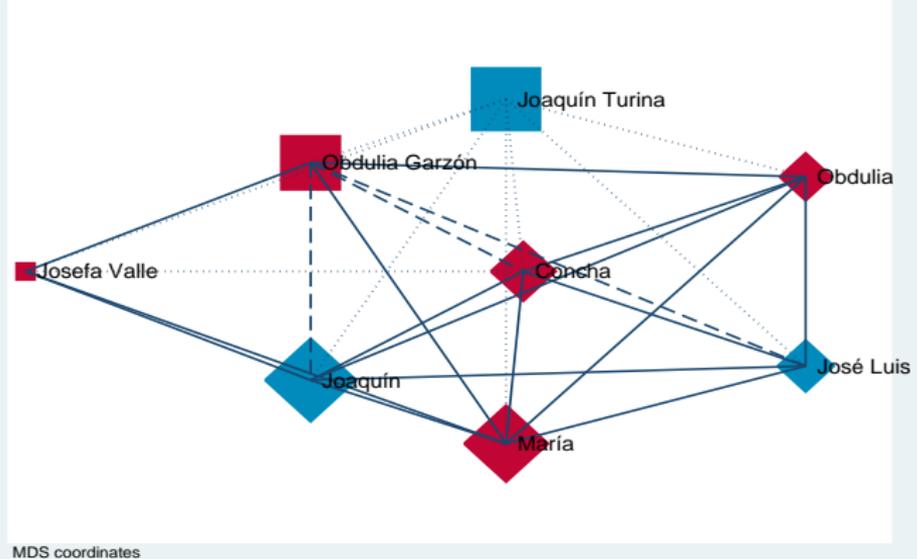
Graphs

Graphs

Remarks

Final

Representation of family members in Turina albums





Graph comparisons

MDS, Biplot, CA and PCA

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

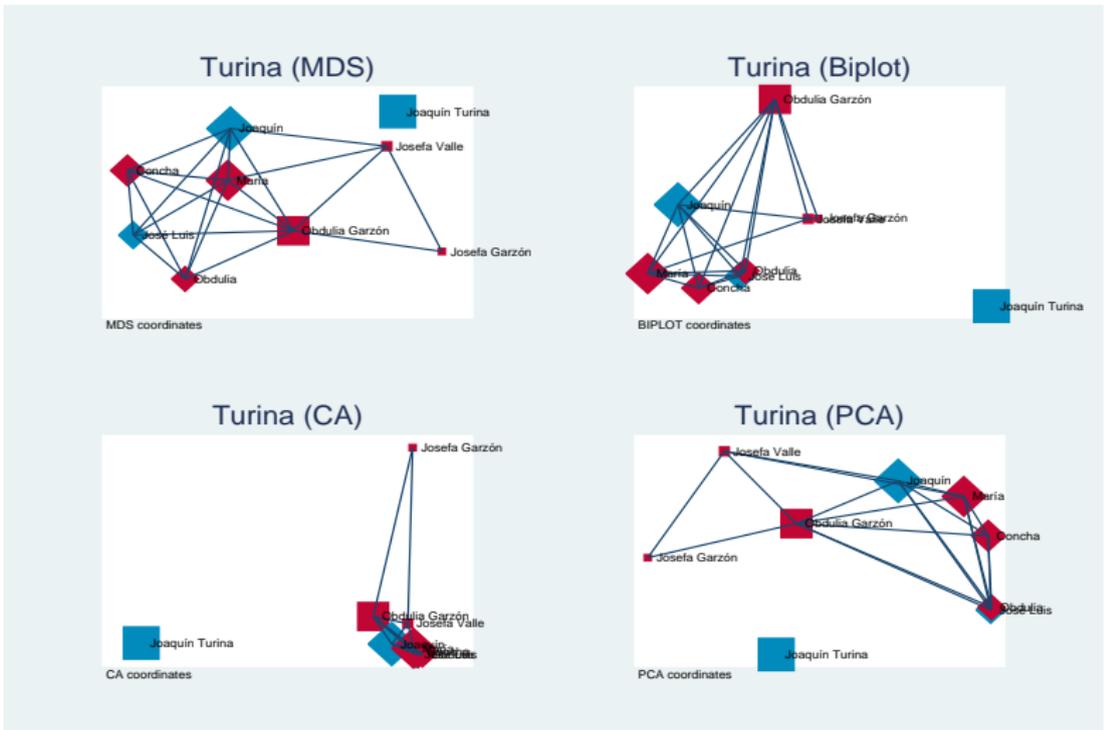
Masó

Graphs

Graphs

Remarks

Final





coin

What is it?

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

Graphs

Graphs

Remarks

Final

- coin is an ado program in its development phase, which is capable of performing coincidence analysis
- Its input is a dataset with scenarios as rows and events as columns.
- Its outputs are:
 - Different matrices (frequencies, percentages, residuals (3), distances, adjacencies and edges)
 - Several bar graphs, network graphs (circle, mds, pca, ca, biplot) and dendrograms (single, average, waverage, complete, wards, median, centroid)
 - Measures of centrality (degree, closeness, betweenness, information) (eigenvector and power)
 - Options to export to excel and .csv files
- Its syntax is simple, but flexible. Many options (output, bonferroni, p value, minimum, special event, graph control and options, ...)



Social network program

Stata program

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

Graphs

Graphs

Remarks

Final

- Stata has commands for mds, pca, biplot, ca, cluster, ...
- Although there are no tools for SNA in Stata, some advanced users have begun to write some routines. I wish to highlight the following works from which I have obtained insights:
 - Corten (2010) wrote a routine to visualize social networks [netplot]
 - Mihura (2012) created routines (SGL) to calculate networks centrality measures, including two Stata commands [netsis and netsummarize]
 - Recently, White (2013) presented a suite of Stata programs for network meta-analysis which includes the network graphs of Anna Chaimani in the UK users group meeting. And Grund (2013, 2014) announced a presentation on plotting and analyzing social networks in the Nordic and Baltic Stata Users Group.



```
coin varlist [ if ] [ in ] [ weight ] [ , options ]
```

Options can be classified into the following groups:

- **Outputs:** f, g, v, h, e, r, s, n, ph, o, po, pf, t, a, d, l, c, all, x, xy.
- **Controls:** head(*varlist*), variable(*varname*), ascending, descending, minimum (#), support(#), pvalue(#), levels(# # #), bonferroni, lminimum(#), iterations(#).
- **Plots**
 - Bar: bar, cbar(*varname*)
 - Graph: plot(circle|mds|ca|pca|biplot)
 - Dendrograms: dendrogram(single|complete|average|wards)



coin example (I)

Matrix of coincidences in the photograph albums of the Turina family

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

Graphs

Graphs

Remarks

Final

```
. coin Turina-Valle, f
```

```
915 scenarios. 18 probable coincidences amongst 8 events. Density: 0.64
```

```
8 events(n>=5): Turina Garzon Joaquin Maria Concha JoseLuis Obdulia Valle
```

	Frequencies	Tur-a	Gar-n	Joa-n	Maria	Con-a	Jos-s	Obd-a	Valle
Joaquín Turina		291							
Obdulia Garzón		42	216						
Joaquín		42	71	262					
María		25	62	124	222				
Concha		20	39	68	100	134			
José Luis		18	30	40	60	64	101		
Obdulia		13	27	33	54	60	58	86	
Josefa Valle		2	9	15	10	3	0	0	21



coin example (II)

Matrix of adjacencies in the photograph albums of the Turina family

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

Graphs

Graphs

Remarks

Final

```
. coin Turina-Valle, a
```

```
915 scenarios. 18 probable coincidences amongst 8 events. Density: 0.64
```

```
8 events(n>=5): Turina Garzon Joaquin Maria Concha JoseLuis Obdulia Valle
```

Adjacency matrix	Tur-a	Gar-n	Joa-n	Maria	Con-a	Jos-s	Obd-a	Valle
Joaquín Turina	0.0							
Obdulia Garzón	0.0	0.0						
Joaquín	0.0	1.0	0.0					
María	0.0	1.0	1.0	0.0				
Concha	0.0	1.0	1.0	1.0	0.0			
José Luis	0.0	1.0	1.0	1.0	1.0	0.0		
Obdulia	0.0	1.0	1.0	1.0	1.0	1.0	0.0	
Josefa Valle	0.0	1.0	1.0	1.0	0.0	0.0	0.0	0.0



coin example (III)

Centrality measures in the photograph albums of the Turina family

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

Graphs

Graphs

Remarks

Final

```
. coin Turina-Valle, c
```

```
915 scenarios. 18 probable coincidences amongst 8 events. Density: 0.64
```

```
8 events(n>=5): Turina Garzon Joaquin Maria Concha JoseLuis Obdulia Valle
```

Centrality measures	Degree	Close	Between	Inform
Joaquín Turina	0.00	.	0.00	.
Obdulia Garzón	0.86	1.00	0.07	0.16
Joaquín	0.86	1.00	0.07	0.16
María	0.86	1.00	0.07	0.16
Concha	0.71	0.86	0.00	0.14
José Luis	0.71	0.86	0.00	0.14
Obdulia	0.71	0.86	0.00	0.14
Josefa Valle	0.43	0.67	0.00	0.11



Study of picture collections

Approach

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

Graphs

Graphs

Remarks

Final

- The aim is to analyze the set of people in three photograph collections.
- The first step is to quantify the number of pictures of every person.
- However, it is not only important how many times they appear, but also with whom.
- These ideas are based on the interactionist theory of the self outlined by G. H. Mead.
- The pictures will be considered as scenarios.
- People are going to be considered as incidences (variables). Do they appear or don't they?



Data Sources

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

Graphs

Graphs

Remarks

Final

- The **Unamuno**'s archive contains around 1,117 pictures. A substantive part of them, 941, belonged to the familiar album.
 - This collection is from the “Casa-Museo Unamuno de la Universidad de Salamanca”.
- The **Turina**'s archive consists of 1,438 photographs from the family album, plus over 1,800 other photos stored in folders, and a collection of postcards acquired by Turina himself. The file set adds up to a total of 5,271 images.
 - The photos and Turina's archive come from the Spanish Library of Contemporary Music and Theatre (Juan March Foundation in Madrid).
- The **Masó**'s archive contains 237 family pictures. Its main character is Joan Masó (main photographer too).
 - These photos come from the “Fundació Rafael Masó's Archive ”(Girona).



Miguel de Unamuno

Biography

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

Graphs

Graphs

Remarks

Final

- Miguel de Unamuno was born in Bilbao in September 1864.
- In 1880 he moved to Madrid to study Philosophy and Languages, and got married to Concha Lizárraga.
- He obtained a doctorate in 1883, and in 1891 he obtained the post of professor of Greek in the “Universidad de Salamanca”.
- In the early 1900, he was appointed as the University Rector. Fourteen years later, he was removed by ministerial decree.
- In 1924 he was banished by General Primo de Rivera to Fuerteventura.
- He came back to the Universidad de Salamanca in 1930.
- During the II Republic (1931-1936), he was member of the Spanish Parliament, and was again appointed as Rector.



Unamuno's Pictures

Unamuno (1864-1936)

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

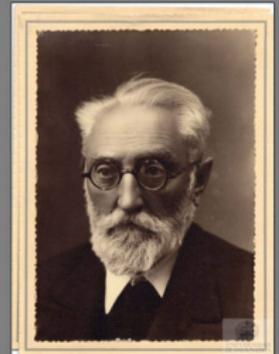
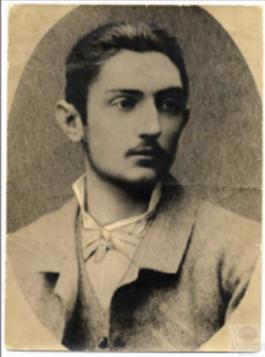
Masó

Graphs

Graphs

Remarks

Final





Nuclear family

Unamuno-Lizárraga family

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

Graphs

Graphs

Remarks

Final





Public pictures

Unamuno

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

Graphs

Graphs

Remarks

Final





People in Unamuno's Albums

Family and colleagues (egonet)

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

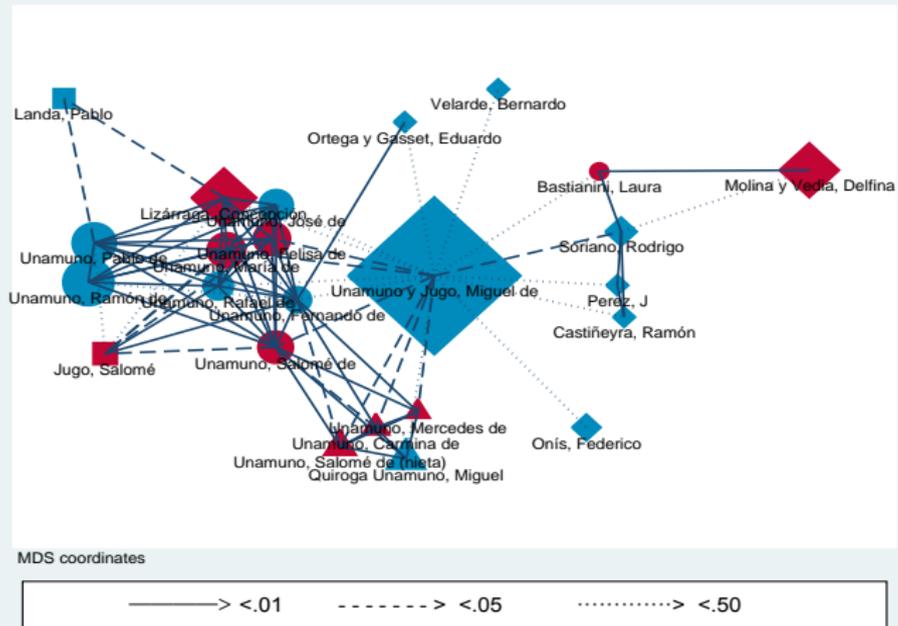
Graphs

Graphs

Remarks

Final

Unamuno albums





Joaquín Turina

Biography

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

Graphs

Graphs

Remarks

Final

- Joaquín Turina Pérez was born in Seville in December 1882,
- He studied music in Madrid and Paris, where he met artists such as Isaac Albéniz and Manuel de Falla. He returned to Madrid at the beginning of World War I.
- He was responsible for the management of the theater Eslava in Madrid and from 1919 he served as the conductor of the “Teatro Real” .
- In 1931 he became Professor of Composition at Conservatory of Madrid and in 1935 was appointed as a member of the “Real Academia de Bellas Artes de San Fernando” .
- He died in 1949 leaving behind musicals like *Fantastic Dances* and *Fancy Clock*. He also published academic works like *A Treatise on Musical Composition* (1946).



Pictures of Turina

Turina (1882-1949)

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

Graphs

Graphs

Remarks

Final





Family photos

Turina-Garzón family

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

Graphs

Graphs

Remarks

Final





Public pictures

Turina

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

Graphs

Graphs

Remarks

Final





People in Turina's Album

Family and colleagues

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

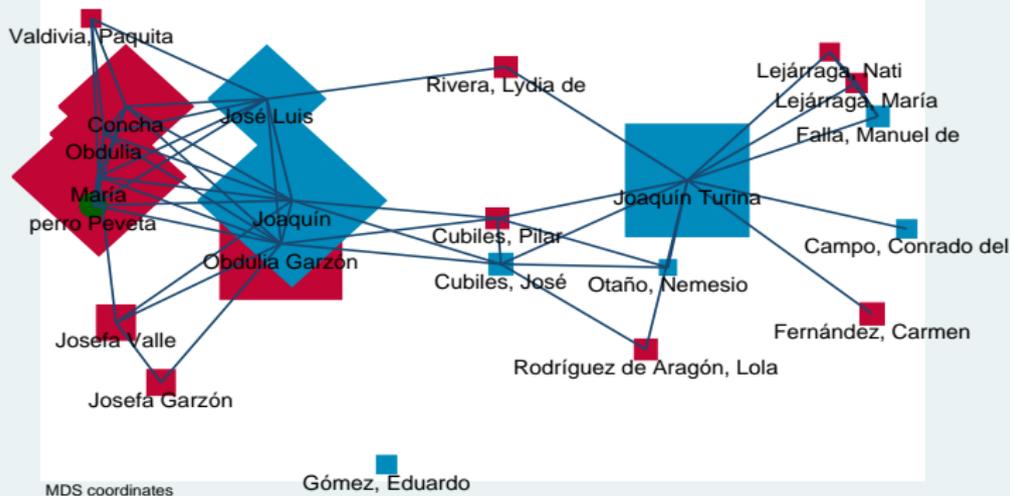
Graphs

Graphs

Remarks

Final

Turina albums





Rafael Masó

Biography

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

Graphs

Graphs

Remarks

Final

- Rafael Masó was born in Girona in August 1880. He was the second of eleven siblings.
- He was a architecture student in Barcelona where he moved to in 1900, was an admirer of Gaudí, and joined the Noucentisme, an alternative movement to Modernisme.
- Besides his architectural works, he was a Catalan nationalist, urban planner and promoter of art and literature
- His most outstanding works include the “Teixidor Flour Mill” (1910), the “Masó House” (1911), and the “Athenea cultural centre” (1912), all in Girona.
- He died in Girona in 1935, when he was 54.



Rafael Masó

(1880-1935)

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

Graphs

Graphs

Remarks

Final





Rafael Masó

Masó-Valentí and Masó-Bru families

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

Graphs

Graphs

Remarks

Final





Public pictures

Masó

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

Graphs

Graphs

Remarks

Final





People in Masó's Collection

Egnet of Rafael Masó

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

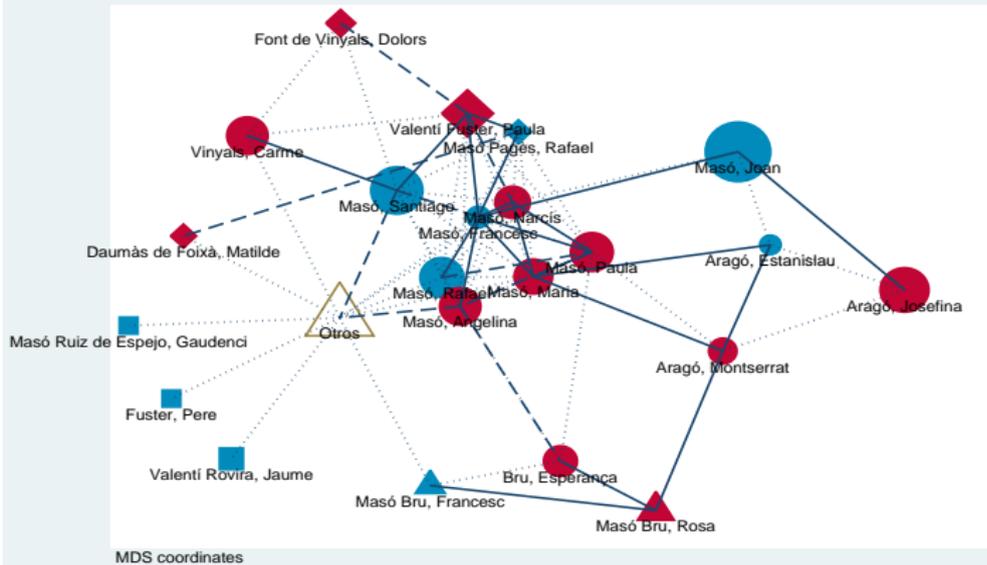
Graphs

Graphs

Remarks

Final

Maso albums





People in Unamuno's Albums

Family and colleagues (egonet)

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

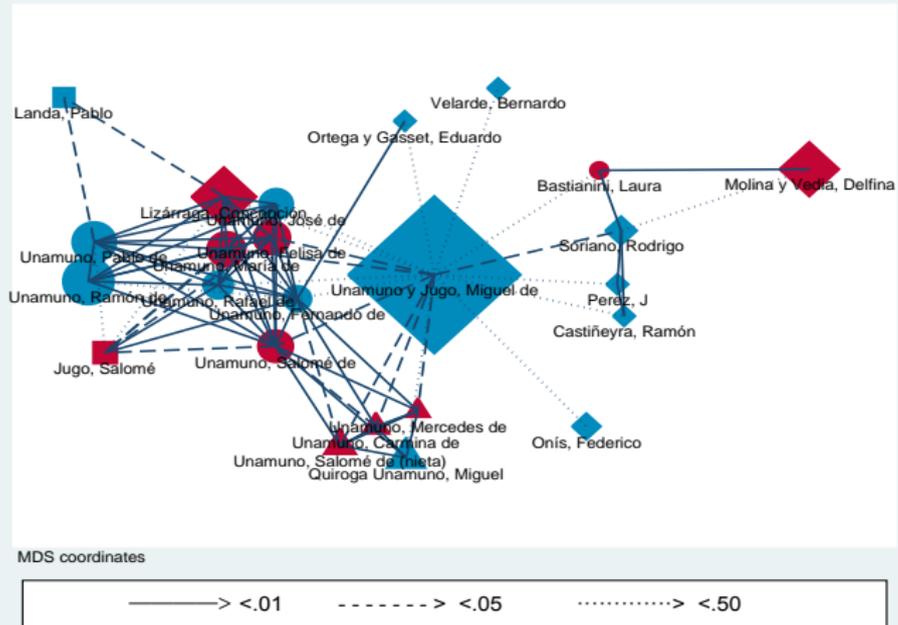
Graphs

Graphs

Remarks

Final

Unamuno albums





People in Turina's Album

Family and colleagues

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

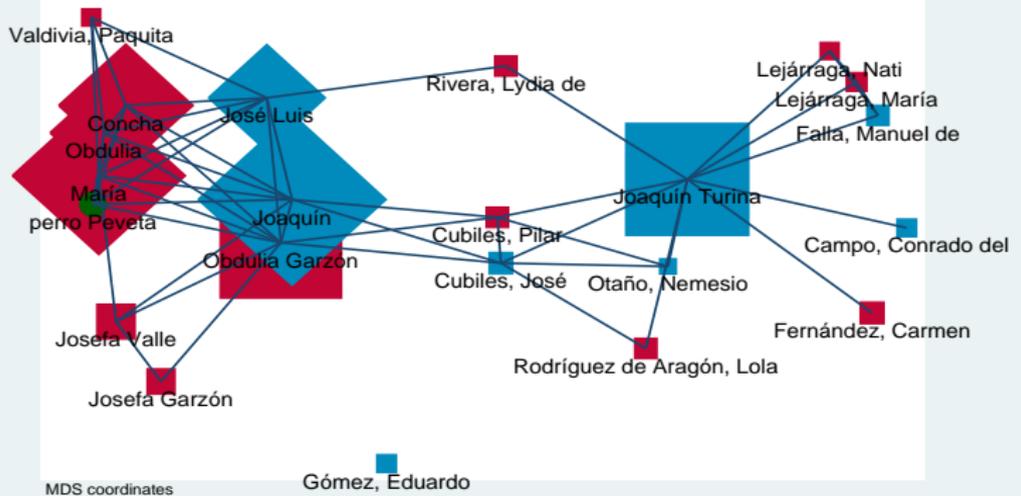
Graphs

Graphs

Remarks

Final

Turina albums





People in Masó's Collection

Egnet of Rafael Masó

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

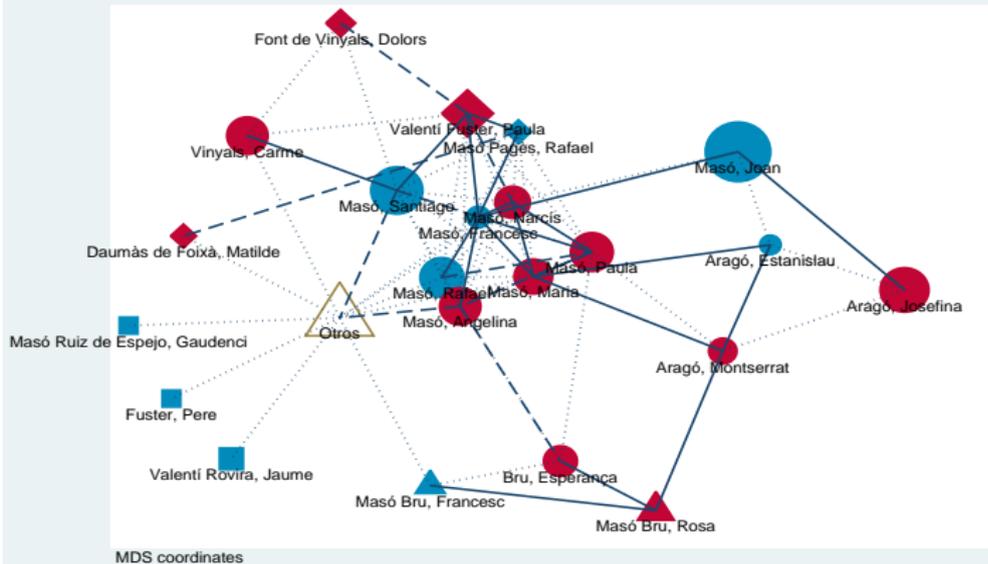
Graphs

Graphs

Remarks

Final

Maso albums





Maso's family

Graph with Gephi (after Stata export)

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

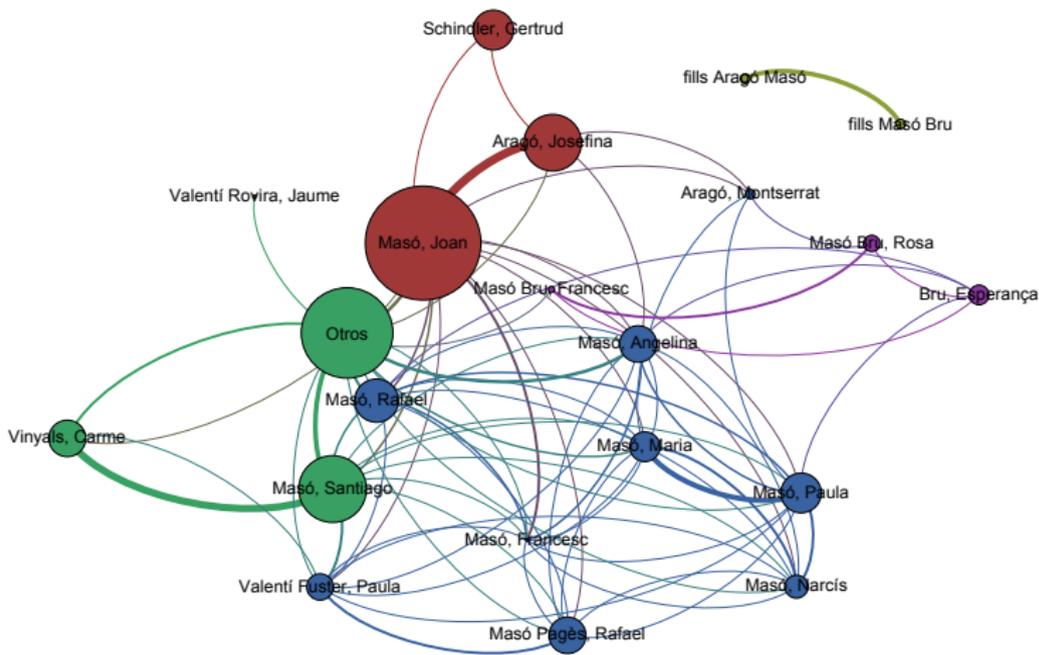
Masó

Graphs

Graphs

Remarks

Final





Remarks

About coincidence analysis

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

Graphs

Graphs

Remarks

Final

- I've proposed a manner of analyzing coincidences mixing different statistical tools.
- I think that the novelty of coincidence analysis is combining several techniques in order to represent reality graphically.
- This may also be useful in comparing different kinds of analysis with dichotomous variables.
- I think that the above approach could be extensively used with the aid of the coin and other forthcoming programs.



Availability of coin

Frame Subtitle

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

Graphs

Graphs

Remarks

Final

- If you are users of a version superior to the 11.2 of Stata, you can have a free copy of coin by typing:
 - `net install coin, from(http://sociocav.usal.es/stata/) replace`
- It is still a beta version, but it works reasonably well.
- Comments and criticisms will be welcome!!



UNIVERSIDAD
DE SALAMANCA

Last slide

Thanks

Coincidences

M.E

Introduction

Definitions

Coincidence

Types

Adjacency

Plots

Bars

Graphs

coin

Data

Unamuno

Turina

Masó

Graphs

Graphs

Remarks

Final

Tak for din opmærksomhed!
modesto@usal.es