Package: remdata (via r-universe)

May 28, 2026

Type Package

Title A Collection of Empirical and Simulated Relational Event Data Sequences

Version 0.2.0

Description Empirical and simulated data for relational event analyses. Each dataset consists of a relational event sequence and optional actor attributes. Individual datasets are redistributed under their original licenses as documented in inst/DATA_LICENSES.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

RoxygenNote 7.3.3

Depends R (>= 4.0.0)

LazyDataCompression xz

NeedsCompilation no

Author Joris Mulder [aut, cre], Roger Leenders [aut], Diana Karimova [aut], Marlyne Meijerink-Bosman [aut], Giuseppe Arena [aut], Rumana Lakdawala [aut], Mahdi Shafiee Kamalabad [ctb], Fabio Generoso Vieira [ctb]

Maintainer Joris Mulder <j.mulder3@tilburguniversity.edu>

Repository https://cran.r-universe.dev

Date/Publication 2026-05-28 17:07:12 UTC

RemoteUrl https://github.com/cran/remdata

RemoteRef HEAD

RemoteSha cb6b567a745346a205d315228e6e4461442cc6cb

Contents

ants . . . 2
baboons_obs . . . 4
baboons_sens . . . 5
brazilian_states_trade . . . 6
dnc . . . 7
highschool2011 . . . 8
highschool2012 . . . 9
highschool2013 . . . 10
history . . . 12
hospitalward . . . 13
households . . . 14
hypertext . . . 15
ICCSS17 . . . 16
infectious . . . 18
info . . . 19
Malawi . . . 19
manufacturing . . . 20
primaryschool . . . 21
randomREH . . . 22
randomREHsmall . . . 23
RFIDvalidity . . . 24
SFHHconference . . . 25
team4 . . . 26
wikitalkbr . . . 27
workplace2013 . . . 28
workplace2015 . . . 30
WS16 . . . 31

Index **33**

`ants` *Time-stamped interactions between ants*

Description

This dataset contains eight sequences with the time-stamped interactions between ants of four colonies during two filming session of approximately 1800-second intervals per colony. Interactions are defined here as antenna-body contact between individual ants.

Usage

```
data(ants)
```

Format

Named list with nine elements. Eight elements corresponds to a film session of a colony (e.g., colony11 = colony 1, film session 1) and contains a list with the elements 'edgelist' and 'attributes'. The 9th element, named 'colonies', contains attribute information about the colonies.

Details

Number of actors and events per colony and filming session Colony 1, session 1: 90 ants, 1911 events Colony 1, session 2: 72 ants, 1820 events Colony 2, session 1: 73 ants, 975 events Colony 2, session 2: 70 ants, 1917 events Colony 3, session 1: 12 ants, 78 events Colony 3, session 2: 6 ants, 104 events Colony 6, session 1: 35 ants, 652 events Colony 6, session 2: 34 ants, 367 events

edgelist

Dataframe with the time-stamped interactions between the ants within the respective colony during the film session.

time Numeric value with the time of the interaction in seconds since the beginning of the film session.

actor1 String with the name of the ant initiating the interaction by extending antenna to contact another ant's body or antenna.

actor2 String with the name of the ant receiving the interaction by being touched by another ant's antenna.

posx When recorded, numeric value with the position in pixels of the actor1 ant's head centroid.

posy When recorded, numeric value with the position in pixels of the actor1 ant's head centroid.

attributes

Dataframe with all the ants in the filming. Queen ants have a Q in the name.

name String with the name of the ant.

initialx Position in pixels of the ant's head centroid.

initialy Position in pixels of the ant's head centroid.

queen Logical whether or not the ant is a queen ant.

colony Integer id of the colony (1,2,3,6).

session Integer id of the filming session (1,2).

colonies

Dataframe with information about the 4 colonies.

id Integer id of the colony (1,2,3,6).

session Integer id of the filming session (1,2).

area Area of the nest in pixels^2.

length Mean body length of ants in pixels.

minx Coordinates in pixels of interaction bounding box.

miny Coordinates in pixels of interaction bounding box.

maxx Coordinates in pixels of interaction bounding box.

maxy Coordinates in pixels of interaction bounding box.

Source

doi:10.1371/journal.pone.0020298.s008

References

Blonder B, Dornhaus A (2011) Time-Ordered Networks Reveal Limitations to Information Flow in Ant Colonies. PLOS ONE 6(5): e20298. doi:10.1371/journal.pone.0020298

`baboons_obs` *Baboon's interactions.*

Description

This dataset contains the sequence of 5377 interactions between 20 Guinea baboons living in an enclosure of a Primate Center in France. The 'edgelist' element contains all directed interactions between pairs of baboons registered by an observer between June 13th 2019 and July 10th 2019. Moreover, the type of behavior and its duration is recorded. A distinction is made between 'POINT' events (without duration) and 'STATE' events (with duration > 0). Note that the edgelist contains event without a recipient and events directed towards the "self".

Usage

```
data(baboons_obs)
```

Format

Dataframe with 5377 rows and 7 columns.

time POSIXct value with the time stamps of the onset of the events.

actor1 Factor with the names of the sender of the events.

actor2 Factor with the names of the recipient of the events.

behavior Factor with the behaviors of the baboons.

category Factor with the classifications of the behaviors.

duration Numeric value indicating the duration of the events in seconds. The duration of events that lasted longer than 300 seconds (five minutes) were not recorded (i.e., they were recorded as lasting 300 seconds but may have lasted longer).

point Factor that indicates whether the event is a POINT event (YES) or a STATE event (NO).

Source

https://sociopatterns.org/datasets/baboons-interactions/

References

Geladi et al. (2020) Proc. R. Soc. A. 476:20190737 doi:10.1098/rspa.2019.0737

See Also

baboons_sens

`baboons_sens` *Baboon's interactions.*

Description

This dataset contains the sequence of 31861 interactions between 13 Guinea baboons living in an enclosure of a Primate Center in France, between June 13th 2019 and July 10th 2019. Wearable proximity sensors recorded with a 20-second time interval every contact between baboons that was active in the last 20 second window. A contact between two baboons is defined as the set of successive time-windows of 20 seconds during which the baboons are detected in contact. In the 'edgelist' element, these records are collected in a relational event sequence with for every contact between a pair of baboons the time of onset and its duration. Time is expressed in Unix time. The events in the sequence are undirected. Multiple events can occur at the same time point and the individuals can be in contact with multiple others at the same time.

Usage

```
data(baboons_sens)
```

Format

Dataframe with 31861 rows and 4 columns.

time POSIXct value indicating the timestamp for the onset of the event.

actor1 String with the name of the sender of the event.

actor2 String with the name of the recipient of the event.

duration Numeric value indicating the duration of the event in seconds.

Source

https://sociopatterns.org/datasets/baboons-interactions/

References

Geladi et al. (2020) Proc. R. Soc. A. 476:20190737 doi:10.1098/rspa.2019.0737,

See Also

baboons_obs

`brazilian_states_trade`

Short description of the data – Network of trades between Brazilian states and countries in the European Union. Relational event sequence with 28035 events among 55 actors. The events happened from January first 2015 until March first 2021. Long description – Trades happened between Brazilian states and countries in the European Union, thus it is a directed network. Here, only Brazilian states can be senders and only EU countries can be receivers, so this network contains 27 sender (number of Brazilia states) and 28 receivers (number of countries in the EU, which contained the UK until 2020).

Description

Short description of the data – Network of trades between Brazilian states and countries in the European Union. Relational event sequence with 28035 events among 55 actors. The events happened from January first 2015 until March first 2021. Long description – Trades happened between Brazilian states and countries in the European Union, thus it is a directed network. Here, only Brazilian states can be senders and only EU countries can be receivers, so this network contains 27 sender (number of Brazilia states) and 28 receivers (number of countries in the EU, which contained the UK until 2020).

Usage

```
brazilian_states_trade
```

Format

This data is a list that includes

1. **edgelist** A data frame with 28035 rows and 3 columns

   time This variable represents the date in which the trade occurred.

   actor1 Brazilian state sending goods (sender).

   actor2 European country receiving goods (receiver).

2. **actors** A list containg two named vectors: sender and receiver. The sender vector contains the names of the 27 actors that can act as a sender. The receiver contains a list with the 28 actors that act as receivers.

3. **attributes** A data frames with 28035 rows and 4 columns.

   state It contains the names of the states.

   time It contains the date of the trade.

   country It contains the name of the EU country.

   value It contains the value of the trade in US dollar.

Source

https://comexstat.mdic.gov.br/

dnc *Democratic National Committee (DNC) emails Network*

Description

This dataset describes the temporal directed network of emails in the 2016 Democratic National Committee email leak. Any directed edge in the sequence represents an email sent by one person (the sender) to another one (the recipient). Considering that an email can have any number of recipients, for any email with multiple receivers the single event is splitted into as many dyadic events as the number of recipients in the email. As a consequence of this, the resulting network contains about the double of the emails in the original data.

Usage

```
data(dnc)
```

Format

'dnc' is a List of two objects.

1. **edgelist**: a List containing a `data.frame` describing a network of 34421 events among 1866 actors (users in the website)

   time timestamp of the email (interval timing).

   actor1 ID (integer value) of the actor that is the sender of the email.

   actor2 ID (integer value) of the actor that receives the email.

2. **actors**: a vector of actors ID's.

Source

http://konect.cc/networks/dnc-temporalGraph/

References

Jérôme Kunegis. KONECT – The Koblenz Network Collection. In Proc. Int. Conf. on World Wide Web Companion, pages 1343–1350, 2013. https://dl.acm.org/doi/10.1145/2487788.2488173

`highschool2011` *Contacts between high school students.*

Description

This dataset contains the sequence of 10685 face-to-face contacts between high school students and their teachers in Marseilles, France during 4 days in December 2011. The data collection involved 118 students of three different classes and 8 teachers. Wearable proximity sensors recorded with a 20-second time interval every face-to-face contact between individuals that was active in the last 20 second window. A contact between two individuals is defined as the set of successive time-windows of 20 seconds during which the individuals are detected in contact. In the 'edgelist' element, these records are collected in a relational event sequence with for every contact between a pair of individuals the time of onset and its duration. Time is measured in seconds. The events in the sequence are undirected. Multiple events can occur at the same time point and the individuals can be in contact with multiple other individuals at the same time. Information on the gender and class/role of the individuals is available in the 'attributes' element.

Usage

```
data(highschool2011)
```

Format

List with elements 'edgelist' and 'attributes'.

edgelist

Dataframe with 10685 rows and 4 columns that contains the relational event sequence with face-to-face contacts between the students and teachers.

time Numeric value that indicates the onset of the contact in seconds.

actor1 Integer value with the ID of the first actor in the contact.

actor2 Integer value with the ID of the second actor in the contact.

duration Numeric value that indicates the duration of the contact in seconds.

attributes

Dataframe with 126 rows and 4 columns describing the gender and class/role of the students and teachers.

id Integer value with the ID of the individual.

time Change time of the attributes. Since the attributes are constant over the course of the study period, this value is equal to 0 for all entries.

class String with the class/role of the individual ('PC' stands for classes that focus on physics and chemistry and 'PSI' stands for classes that focus on engineering studies).

gender String with the gender of the individual ('M' = male, 'F' = female).

Source

https://sociopatterns.org/datasets/high-school-dynamic-contact-networks/

References

Fournet & Barrat (2014) PloS one 9(9) e107878 doi:10.1371/journal.pone.0107878

See Also

highschool2012

`highschool2012` *Contacts between high school students.*

Description

This dataset contains the sequence of 19774 face-to-face contacts between high school students and their teachers in Marseilles, France during 7 days in November 2012 (from a Monday to the Tuesday of the following week). The data collection involved 180 students of five different classes. Wearable proximity sensors recorded with a 20-second time interval every face-to-face contact between students that was active in the last 20 second window. A contact between two individuals is defined as the set of successive time-windows of 20 seconds during which the individuals are detected in contact.In the 'edgelist' element, these records are collected in a relational event sequence with for every contact between a pair of students the time of onset and its duration. Time is measured in seconds and expressed in UNIX ctime. The events in the sequence are undirected. Multiple events can occur at the same time point and the students can be in contact with multiple other students at the same time. Information on the gender and class of the students is available in the 'attributes' element.

Usage

```
data(highschool2012)
```

Format

List with elements 'edgelist' and 'attributes'.

edgelist

Dataframe with 19774 rows and 4 columns that contains the relational event sequence with face-to-face contacts between the students.

time Numeric value indicating the timestamp for the onset of the contact expressed in UNIX ctime.

actor1 Integer value with the ID of the first actor in the contact.

actor2 Integer value with the ID of the second actor in the contact.

duration Numeric value indicating the duration of the contact in seconds.

attributes

Dataframe with 180 rows and 4 columns describing the gender and class of the students.

id Integer value with the ID of the student.

time Change time of the attributes. Since the attributes are constant over the course of the study period, this value is equal to 0 for all entries.

class String with the class of the student ('MP' stands for classes that focus on mathematics and physics, 'PC' stands for classes that focus on physics and chemistry and 'PSI' stands for classes that focus on engineering studies).

gender String with the gender of the student ('M' = male, 'F' = female).

Source

https://sociopatterns.org/datasets/high-school-dynamic-contact-networks/

References

Fournet & Barrat (2014) PloS one 9(9) e107878 doi:10.1371/journal.pone.0107878

See Also

highschool2011

---

highschool2013 *Contacts between high school students.*

Description

This dataset contains the sequence of 67,613 face-to-face contacts between high school students and their teachers in Marseilles, France from December 2nd until December 6th, 2013. The data collection involved 329 students of nine different classes. Wearable proximity sensors recorded with a 20-second time interval every face-to-face contact between students that was active in the last 20 second window. A contact between two students is defined as the set of successive time-windows of 20 seconds during which the students are detected in contact. In the 'edgelist' element, these records are collected in a relational event sequence with for every contact between a pair of students the time of onset and its duration. Time is measured in seconds. The events in the sequence are undirected. Multiple events can occur at the same time point and the students can be in contact with multiple other students at the same time. On the 4th day of the study, students were asked to report the list of other students they had contact with during the day in the high school and give the approximate aggregated duration of these contacts. In total, 119 students reported on their contacts. The resulting directed reported contact network is available in the 'contacts' element. Moreover, in total 133 students reported the name of their friends in the high school. The resulting directed friendship network is available in the 'friendship' element. Information on whether or not the students are linked on Facebook is available in the 'facebook' element. Information on the gender and class of the students is available in the 'attributes' element.

Usage

```
data(highschool2013)
```

Format

List with elements 'edgelist', 'attributes', 'contacts', 'friendship' and 'facebook'.

edgelist

Dataframe with 67613 rows and 4 columns that contains the relational event sequence with face-to-face contacts between the students.

time Numeric value that indicates the timestamp for the onset of the contact expressed in UNIX ctime.

actor1 Integer value with the ID of the first actor in the contact.

actor2 Integer value with the ID of the second actor in the contact.

duration Numeric value that indicates the duration of the contact in seconds.

attributes

Dataframe with 329 rows and 4 columns describing the gender and class of the students.

id Integer value with the ID of the individual.

time Change time of the attributes. Since the attributes are constant over the course of the study period, this value is equal to 0 for all entries.

class String with the class of the individual.

gender String with the gender of the individual ('M' = male, 'F' = female). For 7 students, the gender is missing.

contacts

Matrix with the directed reported contacts network; 119 students reported on the duration of their contacts with other students in the high school on the 4th day of the study period (1 = at most 5 minutes, 2 = between 5 and 15 minutes, 3 = between 15 minutes and 1 hour, 4 = more than one hour). Row names and column names correspond to the respective actors.

friendship

Matrix with the directed reported friendships network; 133 reported on their friends (1 = did report a friendship, 0 = did not report a friendship). Row names and column names correspond to the respective actors.

facebook

Matrix with the undirected Facebook links (1 = are friends on Facebook, 0 = are not). Row names and column names correspond to the respective actors.

Source

https://sociopatterns.org/datasets/high-school-contact-and-friendship-networks/

References

Mastrandrea et al. (2015) PloS one 10(9) e0136497 doi:10.1371/journal.pone.0136497

See Also

highschool2011, highschool2012

`history` *Simulated relational event history*

Description

A dataset containing a small example of a relational event history. Data is simulated.

Usage

```
data(history)
```

Format

A dataframe with 115 rows and 5 variables:

time time of the event since onset of observation (e.g., in minutes)

actor1 the first actor involved in the event

actor2 the second actor involved in the event

setting the setting for the event

weight the intensity of the event (e.g., based on the duration)

Source

Simulated relational event history for actors in a social network.

See Also

`info` for exogenous information on the actors in the social network.

`hospitalward` *Contacts between patients and health-care workers.*

Description

This dataset contains the sequence of 14037 face-to-face contacts between patients, patients and health-care workers and between health-care workers in a hospital ward in Lyon, France, from Monday, December 6, 2010 at 1:00 pm to Friday, December 10, 2010 at 2:00 pm. The data collection involved 29 patients and 46 health-care workers. Wearable proximity sensors recorded with a 20-second time interval every face-to-face contact between individuals that was active in the last 20 second window. A contact between two individuals is defined as the set of successive time-windows of 20 seconds during which the individuals are detected in contact. In the 'edgelist' element, these records are collected in a relational event sequence with for every contact between a pair of individuals the time of onset and its duration. Time of the events is expressed in seconds since onset of the study period. The events in the sequence are undirected. Multiple events can occur at the same time point and the individuals can be in contact with multiple other individuals at the same time. Information on the statuses of the individuals is available in the 'attributes' element.

Usage

```
data(hospitalward)
```

Format

List with elements 'edgelist' and 'attributes'.

edgelist

Dataframe with 14037 rows and 4 columns that contains the relational event sequence with face-to-face contacts between the individuals.

time Numeric value indicating the timestamp for the onset of the contact expressed in seconds since onset of the study period (Monday, December 6, 2010 at 1:00 pm).

actor1 Integer value with the ID of the first actor in the contact.

actor2 Integer value with the ID of the second actor in the contact.

duration Numeric value indicating the duration of the contact in seconds.

attributes

Dataframe with 75 rows and 3 columns describing the status of the individuals.

id Integer value with the ID of the individual.

time Change time of the attributes. Since the attributes are constant over the course of the study period, this value is equal to 0 for all entries.

status String with the status of the individual ('NUR' = paramedical staff, i.e., nurses and nurses' aides; 'PAT' = patient, 'MED' = medical doctor; 'ADM' = administrative staff).

Source

https://sociopatterns.org/datasets/hospital-ward-dynamic-contact-network/

References

Vanhems et al. (2013) PloS one 8(9) e73970 doi:10.1371/journal.pone.0073970

`households` *Interactions between members of 5 Kenyan households*

Description

This dataset contains five sequences with the interactions between the members of five households of rural Kenya. The households are named 'E', 'F', 'L', 'H' and 'B'. Each of the five households was followed for three days between April 24 and May 12, 2012. During these three days, wearable proximity sensors recorded every face-to-face interaction among the members of the households. Members of households L, F and E wore the sensors during partly overlapping time windows. In the 'edgelist' elements of 'households', the proximity records are collected in a relational event sequence with for every contact between a pair of individuals the time of onset in hours since the start of the study period for the respective household and its duration in seconds. Only contact events recorded between 6am and 8pm are reported. The number of recorded interactions is 13473 among 17 members for household E, 1397 among 8 members for household F, 1365 among 6 members for household L, 10376 among 29 members for household H, and 5814 among 15 members for household B. The events in the sequence are undirected. Multiple events can occur at the same time point and the individuals can be in contact with multiple other individuals at the same time. In the 'attributes' elements of 'households', information on the age and the sex of the individuals is provided.

Usage

```
data(households)
```

Format

Named list of five elements, each element corresponds to a household and contains a list with the elements 'edgelist' and 'attributes':

edgelist

Dataframe with the face-to-face contacts between the members of the respective household.

time Numeric value that indicates the onset of the contact in hours since onset of the study period. Hours are indicated by integers corresponding to the end of the time interval (e.g, hour 0-1 with 1).

actor1 Integer value with the ID of the first actor in the contact.

actor2 Integer value with the ID of the second actor in the contact.

duration Numeric value that indicates the duration of the contact in seconds.

household String indicating the household of which the actors in the edgelist are members ('E', 'F', 'L', 'H' or 'B'). Note that this value is the same for each edgelist.

attributes

Dataframe describing the age and sex of the members of the respective household.

id Integer value with the ID of the individual.

time Change time of the attributes. Since the attributes are constant over the course of the study period, this value is equal to 0 for all entries.

age Age category of the individual (0 = 0-5 years, 1 = 6-14 years, 2 = 15-19 years, 3 = 20-49 years, 4 = 50 years and older).

sex String with the sex of the individual ('M' = male, 'F' = female).

Source

https://sociopatterns.org/datasets/kenyan-households-contact-network/

References

Kiti et al. (2016) EPJ Data Science 5(21) doi:10.1140/epjds/s1368801600842

---

hypertext *Interactions at the ACM hypertext 2009 conference*

Description

This dataset contains the sequence of 9865 face-to-face contacts between attendees of the ACM Hypertext 2009 conference hosted by the Institute for Scientific Interchange Foundation in Turin, Italy, from June 29th to July 1st 2009. The data collection involved 113 conference attendees. Wearable proximity sensors recorded with a 20-second time interval every face-to-face contact between individuals that was active in the last 20 second window. A contact between two individuals is defined as the set of successive time-windows of 20 seconds during which the individuals are detected in contact. In the 'hypertext' edgelist, these records are collected in a relational event sequence with for every contact between a pair of individuals the time of onset and its duration. Time of the events is expressed in seconds since onset of the study period. The events in the sequence are undirected. Multiple events can occur at the same time point and the individuals can be in contact with multiple other individuals at the same time.

Usage

```
data(hypertext)
```

Format

Dataframe with 9865 rows and 4 columns that contains the relational event sequence with face-to-face contacts between the individuals.

time Numeric value indicating the timestamp for the onset of the contact expressed in seconds since onset of the study period (8 am on June 29th 2009).

actor1 Integer value with the ID of the first actor in the contact.

actor2 Integer value with the ID of the second actor in the contact.

duration Numeric value indicating the duration of the contact in seconds.

Source

https://sociopatterns.org/datasets/hypertext-2009-dynamic-contact-network/

References

Isella et al. (2011) Journal of Theoretical Biology 271, 166 doi:10.1016/j.jtbi.2010.11.033

`ICCSS17` *International Conference on Computer Social Science 2017 (ICCSS17)*

Description

This dataset contains the sequence of 70910 face-to-face contacts between visitors of the International Conference on Computer Social Science 2017 (ICCSS17) in Cologne, July 10 to 13, 2017. The data collection involved 262 of the 339 conference attendees. Wearable proximity sensors recorded every face-to-face contact between individuals (physical proximity within approximately 1.5m) with a 20-second time resolution. A contact between two individuals is defined as the set of successive time-windows of 20-second during which the individuals are detected in close proximity. In the 'edgelist' element, these records are collected in a relational event sequence with for every contact between a pair of individuals the time of onset and its duration. Time is measured in seconds. The events in the sequence are undirected. Multiple events can occur at the same time point and the individuals can be in contact with multiple other individuals at the same time. Socio-demographic information about the participants (seniority, disciplinary background, country of residence, mother tongue, age, gender, role at conference) was collected through a survey and available in the 'attributes' element.

Usage

```
data(ICCSS17)
```

Format

List with elements 'edgelist' and 'attributes'

edgelist

Dataframe with 70910 rows and 4 columns that contains the relational event sequence with face-to-face contacts between the participants.

time POSIXct value that indicates the onset of the contact.

actor1 Integer value with the ID of the first actor in the contact.

actor2 Integer value with the ID of the second actor in the contact.

duration Numeric value that indicates the duration of the contact in seconds.

attributes

Dataframe with 262 rows and 9 columns describing the socio-demographic information about the participants.

id Integer value with the ID of the individual.

time Change time of the attributes. Since the attributes are constant over the course of the study period, this value is equal to 0 for all entries.

age Age range for the individual. One of 'younger than 30 years', '30 to 39 years', '40 to 49 years', or '50 years and older'.

gender Gender of the individual, 'Female' or 'Male'.

country Code for the country of residence of the individual.

language Code for the mother tongue of the individual.

seniority Academic seniority of the individual at the time of the conference. One of 'Bachelor/Master student', 'PhD student', 'Postdoctoral researcher', 'Junior/Assistant professor', 'Associate/Full professor', 'Other'

background Academic background of the individual. One of 'Social/ Political/Behavioral Sciences', 'Computer/Information Science', 'Math/ Physics/Biology', 'Media/Communication/Linguistics', or 'Others'.

role Role in the conference of the individual. One of 'Speaker', 'Poster presenter', 'Participation only', 'Staff'

previous Did the individual participate to a previous edition? (Yes or No).

Source

https://psycharchives.org/en/item/26710d25-c387-4506-b5f5-60d1bd3009a5

References

Genois, M., Zens, M., Oliveira, M., Lechner, C. M., Schaible, J., & Strohmaier, M. (2023). Combining sensors and surveys to study social interactions: A case of four science conferences. *Personality Science*, 4, Article e9957. doi:10.5964/ps.9957

`infectious` *Contacts between visitors of the artscience exhibition 'INFECTIOUS: STAY AWAY' on May 21st, 2009.*

Description

This dataset contains the sequence of 1332 face-to-face contacts between 96 visitors of the artscience exhibition 'INFECTIOUS: STAY AWAY' on May 21st, 2009. Wearable proximity sensors recorded with a 20-second time interval every face-to-face contact between visitors that was active in the last 20 second window. A contact between two individuals is defined as the set of successive time-windows of 20 seconds during which the individuals are detected in contact. In the 'infectious' edgelist, these records are collected in a relational event sequence with for every contact between a pair of individuals the time of onset and its duration. Time of the events is expressed in UNIX ctime. The events in the sequence are undirected. Multiple events can occur at the same time point and the individuals can be in contact with multiple other individuals at the same time.

Usage

```
data(infectious)
```

Format

Dataframe with 1332 rows and 4 columns that contains the relational event sequence with face-to-face contacts between the visitors.

time Numeric value indicating the timestamp for the onset of the contact expressed in UNIX ctime.

actor1 Integer value with the ID of the first actor in the contact.

actor2 Integer value with the ID of the second actor in the contact.

duration Numeric value indicating the duration of the contact in seconds.

Source

https://sociopatterns.org/datasets/infectious-exhibition-dynamic-contact-networks/

References

Isella et al. (2011) Journal of Theoretical Biology 271, 166 doi:10.1016/j.jtbi.2010.11.033

info *Simulated exogenous information on actors in a social network.*

Description

A dataset containing exogenous information on the actors in the social network of a relational event history. Data is simulated.

Usage

```
data(info)
```

Format

A dataframe with 10 rows and 5 variables:

id numeric id of the actor

time numeric value, describes when the value of the covariate changes, if it changes

age dichotomized age of the actor (e.g., 0 = below 25, 1 = 25 or older)

sex dichotomized sex of the actor (e.g., 0 = male, 1 = female)

extraversion standardized extraversion score of the actor

agreeableness standardized agreeableness score of the actor

Source

Simulated exogenous information on actors in a social network.

See Also

`history` for the relational event history.

Malawi *Contact patterns in a village in rural Malawi*

Description

This dataset contains the sequence of 35117 face-to-face contacts between 86 individuals living in a village in rural Malawi. The data collection was conducted between 16h December 2019 and 10th January 2020. Wearable proximity sensors recorded with a 20-second time interval every face-to-face contact between individuals that was active in the last 20 second window. A contact between two individuals is defined as the set of successive time-windows of 20 seconds during which the individuals are detected in contact. In the Malawi data object, these records are collected in a relational event sequence, with for every contact between a pair of individualsthe time of onset and its duration. The events in the sequence are undirected. Multiple events can occur at the same time point and the individuals can be in contactwith multiple other individuals at the same time.

Usage

```
data(Malawi)
```

Format

dataframe with 35117 rows and 5 columns.

time Numeric value: time in seconds since onset of the observation period.

actor1 Numeric value: ID of the first actor in the event.

actor2 Numeric value: ID of the second actor in the event.

duration Numeric value: duration of the event in seconds.

day Integer: day of the month.

Source

https://sociopatterns.org/datasets/contact-patterns-in-a-village-in-rural-malawi/

References

Ozella, L., Paolotti, D., Lichand, G. et al. Using wearable proximity sensors to characterize social contact patterns in a village of rural Malawi. EPJ Data Sci. 10, 46 (2021). doi:10.1140/epjds/s1368802100302w

---

manufacturing *Manufacturing emails Network*

Description

Internal email communication network between employees of a mid-sized manufacturing company in the period from 2010-01-01 to 2010-09-30. Any directed edge in the sequence represents an email sent by one person (the sender) to another one (the recipient). Considering that an email can have any number of recipients, for any email with multiple receivers the single event is splitted into as many dyadic events as the number of recipients in the email. Additional information about who in the company reports to whom is also available.

Usage

```
data(manufacturing)
```

Format

'manufacturing' is a List of three objects.

1. **edgelist**: a List containing a `data.frame` describing a network of 82876 events among 167 actors.

   time timestamp of the email (interval timing).

actor1 ID (integer value) of the actor that is the sender of the email.

actor2 ID (integer value) of the actor that receives the email.

2. **actors**: a vector of actors ID's.
3. **attributes**: a List of one object 'reportsto', that is a `data.frame` of two columns describing the which actor ('ID') reports to whom ('ReportsToID'). Actor ID = 86 is CEO (the only loop in this `data.frame`).

Source

doi:10.7910/DVN/6Z3CGX

References

Nurek, Mateusz, and Radosław Michalski. "Combining Machine Learning and Social Network Analysis to Reveal the Organizational Structures." Applied Sciences 10, no. 5 (2020): 1699. doi:10.3390/app10051699

`primaryschool` *Contacts in a primary school.*

Description

This dataset contains the sequence of 77521 face-to-face contacts between primary school students and their teachers in Lyon, France on Thursday October 1st and Friday October 2nd 2009. The data collection involved 232 students of five different grades of each two classes (i.e., 10 classes in total) and 10 teachers. Wearable proximity sensors recorded with a 20-second time interval every face-to-face contact between individuals that was active in the last 20 second window. A contact between two individuals is defined as the set of successive time-windows of 20 seconds during which the individuals are detected in contact. In the 'edgelist' element, these records are collected in a relational event sequence with for every contact between a pair of individuals the time of onset and its duration. Time is measured in seconds. The events in the sequence are undirected. Multiple events can occur at the same time point and the students can be in contact with multiple other students at the same time. Information on the gender and class/role of the students and teachers is available in the 'attributes' element.

Usage

```
data(primaryschool)
```

Format

List with elements 'edgelist' and 'attributes'.

edgelist

Dataframe with 77521 rows and 4 columns that contains the relational event sequence with face-to-face contacts between the students.

time Numeric value indicating the timestamp for the onset of the contact in seconds since onset of the study period.

actor1 Integer value with the ID of the first actor in the contact.

actor2 Integer value with the ID of the second actor in the contact.

duration Numeric value indicating the duration of the contact in seconds.

attributes

Dataframe with 242 rows and 4 columns describing the gender and class/role of the students and teachers.

id Integer value with the ID of the individual.

time Change time of the attributes. Since the attributes are constant over the course of the study period, this value is equal to 0 for all entries.

class String with the grade and class/role of the students and teachers.

gender String with the gender of the student ('M' = male, 'F' = female). Note that for 15 individuals, the gender is unknown.

Source

https://sociopatterns.org/datasets/primary-school-temporal-network-data/

References

Gemmetto et al. (2014) BMC Infectious Diseases 14, 695 doi:10.1186/s1287901406959, Stehle et al. (2011) PloS one 6(8) e23176 doi:10.1371/journal.pone.0023176

---

randomREH *Random Relational Event History*

Description

A randomly generated sequence of relational events with 20 actors and 9915 events. Each event type is associated to one of the three following sentiments: *conflict*, *competition* and *cooperation*.

Usage

```
randomREH
```

Format

`data(randomREH)` will load a list containing following objects:

`edgelist` a `data.frame` that contains the random sequence of events. Columns of the edgelist are:

- `time` the timestamp indicating the time at which each event occurred;
- `actor1` the name of the actor that generated the relational event;
- `actor2` the name of the actor that received the relational event;
- `type` the type of the relational event.

`actors` names of actors interacting in the dynamic network.

`types` names of event types observed in the network and describing the sentiment of the interaction (*conflict*, *competition* and *cooperation*).

`origin` starting time point (`t_0`) prior to the first observed event (`t_1`), the class of this object must be the same as the one of the time column in the edgelist.

`omit_dyad` a list where each element describes an alteration of the riskset which takes place at specific time points and for certain actors and/or types.

---

randomREHsmall *Random Relational Event History (small)*

Description

A subset from the randomly generated sequence of relational events 'randomREH', with 5 actors and 586 events (without event types).

Usage

```
randomREHsmall
```

Format

`data(randomREHsmall)` will load a list containing following objects:

`edgelist` a `data.frame` that contains the random sequence of events. Columns of the edgelist are:

- `time` the timestamp indicating the time at which each event occurred;
- `actor1` the name of the actor that generated the relational event;
- `actor2` the name of the actor that received the relational event;

`actors` names of actors interacting in the dynamic network.

`origin` starting time point (`t_0`) prior to the first observed event (`t_1`), the class of this object must be the same as the one of the time column in the edgelist.

`omit_dyad` a list where each element describes an alteration of the riskset which takes place at specific time points and for certain actors and/or types.

`RFIDvalidity` *Face-to-face contacts between participants of an experiment*

Description

This dataset contains two sequences with two types of measures of the same social interactions between participants of an experiment, advertised as an after-work get-together event. Participants were 11 individuals, staff and students of a Swiss university. During the event, participants engaged in social interactions with other participants. RFID badges and a video recorded the face-to-face interactions between the participants. RFID data was by the authors transformed into a time-stamped edgelist, available in the 'RFID' element. Two individiuals coded the interactions in the video, these are available in the 'video' element.

Usage

```
data(RFIDvalidity)
```

Format

List with elements 'RFID' and 'video'

RFID

Dataframe with 1,168 RFID-recorded face-to-face interactions.

time POSIXct value that indicates the onset of the contact.

actor1 Integer value with the ID of the first actor in the contact.

actor2 Integer value with the ID of the second actor in the contact.

duration Numeric value that indicates the duration of the contact in seconds.

video

Dataframe with 396 human-coded video-taped face-to-face interactions.

time POSIXct value that indicates the onset of the contact.

actor1 Integer value with the ID of the first actor in the contact.

actor2 Integer value with the ID of the second actor in the contact.

duration Numeric value that indicates the duration of the contact in seconds.

Source

https://osf.io/rrhxe/

References

Timon Elmer, Krishna Chaitanya, Prateek Purwar, Cristoph Stadtfeld (2019) The validity of RFID badges measuring face-to-face interactions. Behavior Research Methods 51:2120-2138. doi:10.3758/s134280181180y

SFHHconference *SFHH conference interactions.*

Description

This dataset contains the sequence of 26040 face-to-face contacts between 403 participants of the 2009 SFHH conference in Nice, France (June 4-5, 2009) . Wearable proximity sensors recorded with a 20-second time interval every face-to-face contact between individuals that was active in the last 20 second window. Following Genois & Barrat (2018), a contact between two individuals is defined as the set of successive time-windows of 20 seconds during which the individuals are detected in contact. In the 'SFHHconference' object, these records are collected in a relational event sequence with for every contact between a pair of individuals the time of onset and its duration. Time is expressed in seconds since the onset of the study period. The events in the sequence are undirected. Multiple events can occur at the same time point and the individuals can be in contact with multiple other individuals at the same time.

Usage

```
data(SFHHconference)
```

Format

Dataframe with 26040 rows and 4 columns.

time Numeric value that indicates the onset of the contact in seconds since onset of the study period.

actor1 Integer value with the ID of the first actor in the contact.

actor2 Integer value with the ID of the second actor in the contact.

duration Numeric value that indicates the duration of the contact in seconds.

Source

https://sociopatterns.org/datasets/sfhh-conference-data-set/

References

Cattuto et al. (2010) PloS ONE 5(7) e11596 doi:10.1371/journal.pone.0011596, Genois & Barrat (2018) EPJ Data Science 7:11 doi:10.1140/epjds/s1368801801401, Stehle et al. (2011) BMC Medicine 9(87) doi:10.1186/17417015987

team4 *Interactions between members of a research team*

Description

This dataset contains the sequence of 11,607 interactions between nine members of a research team. The dataset is part of the GEDII project, a project with the aim to develop new tools and methods for doing research on the impact of gender diversity in research & development teams. Eight teams have been participating in the GEDII case studies, the current data set contains data of 'team 4'. This research team works in the field of biomedical engineering in a public research center. Sociometric badges recorded every interaction between the team members for five consecutive days. The dataset contains both face-to-face interactions detected with infrared sensors and proximity data detected with bluetooth sensors. To create the relational event history in the 'edgelist' element, an event between a pair of team members is defined as the set of seconds during which the members are detected in contact, while they are not in the preceding nor in the next 25 second time window. The events in the sequence are undirected. Multiple events can occur at the same time point and the team members can be in contact with multiple other team members at the same time. Information on the gender, age, educational level, tenure and role of the individual within the team is available in the 'attributes' element. Moreover, the advice seeking network and friendship network for the team members is available in the 'advice' and 'social' elements.

Usage

```
data(team4)
```

Format

List with elements 'edgelist', 'attributes', 'advice' and 'social'

edgelist

Dataframe with the 11,607 face-to-face contacts between the members of the research team.

time POSIXct value that indicates the onset of the contact.

actor1 Integer value with the ID of the first actor in the contact.

actor2 Integer value with the ID of the second actor in the contact.

duration Numeric value that indicates the duration of the contact in seconds.

sensor String with the sensor that detected the contact (infrared or bluetooth)

signal_strength Numeric value with the average strength of the bluetooth signal during the contact. Higher values indicate stronger signal, which is shown to be related to, but not equal to, physical distance between the individuals wearing the sensors (see Chaffin et al., 2017 in Organizational Research Methods). Contacts detected with infrared sensors have signal strength "NA".

attributes

Dataframe describing a number of attributes of the 9 team members.

id Integer value with the ID of the individual.

time Change time of the attributes. Since the attributes are constant over the course of the study period, this value is equal to 0 for all entries.

role Role within the team.

gender String with the gender of the individual.

hqual Educational level.

tenure Months since joining the team.

age Age in years.

advice

Matrix with advice seeking network. The team members rated each other in a round-robin design by answering the question "I consult this person for work related advice" (1 = never, 2 = rarely, 3 = sometimes, 4 = very often, 5 = always). Row names and column names correspond to the respective actors.

social

Matrix with friendship network. The team members rated each other in a round-robin design by answering the question "I spend time socially with this person outside the lab/office" (1 = never, 2 = some times a year, 3 = some times a month, 4 = some times a week, 5 = daily). Row names and column names correspond to the respective actors.

Source

https://zenodo.org/records/3446071

References

Jörg Müller, Elisabeth Günther, Anne Humbert (2018). GEDII Wearable Sensors Dataset of 8 Research Teams. doi:10.5281/zenodo.1434706

---

wikitalkbr *Breton Wikipedia Talk Network*

Description

This data contain the communication network of the Breton Wikipedia. In this temporal network, a directed edges between two users mean that one user wrote a message on the talk page of another user (self-loops are excluded).

Usage

```
data(wikitalkbr)
```

Format

'wikitalkbr' is a List of two objects.

1. **edgelist**: a List containing a `data.frame` describing a network of 10391 events among 1049 actors (users in the website)

   time timestamp of the interaction (interval timing).

   actor1 ID (integer value) of the actor that writes an answer/comment.

   actor2 ID (integer value) of the actor that receives the answer/comment.

2. **actors**: a vector of actors ID's.

Source

http://konect.cc/networks/wiki_talk_br/

References

Jérôme Kunegis. KONECT – The Koblenz Network Collection. In Proc. Int. Conf. on World Wide Web Companion, pages 1343–1350, 2013. https://dl.acm.org/doi/10.1145/2487788.2488173

`workplace2013` *Contacts in a workplace.*

Description

This dataset contains the sequence of 4592 face-to-face contacts between employees measured in an office building in France, from June 24 to July 3, 2013. The data collection involved 92 employees of five different departments. Wearable proximity sensors recorded with a 20-second time interval every face-to-face contact between individuals that was active in the last 20 second window. A contact between two individuals is defined as the set of successive time-windows of 20 seconds during which the individuals are detected in contact. In the 'edgelist' element, these records are collected in a relational event sequence with for every contact between a pair of individuals the time of onset and its duration. Time is expressed in seconds since the onset of the study period, taken as 0:00 on June 24, 2013. The events in the sequence are undirected. Multiple events can occur at the same time point and the individuals can be in contact with multiple other individuals at the same time. Information on the department of the individuals is available in the 'attributes' element.

Usage

```
data(workplace2013)
```

Format

List with elements 'edgelist' and 'attributes'.

edgelist

Dataframe with 4592 rows and 4 columns that contains the relational event sequence with face-to-face contacts in the workplace.

time Numeric value that indicates the onset of the contact in seconds since 0:00 June 24, 2013.

actor1 Integer value with the ID of the first actor in the contact.

actor2 Integer value with the ID of the second actor in the contact.

duration Numeric value that indicates the duration of the contact in seconds.

attributes

Dataframe with 92 rows and 3 columns describing the department of the employees.

id Integer value with the ID of the individual.

time Change time of the attributes. Since the attributes are constant over the course of the study period, this value is equal to 0 for all entries.

department String with the department of the employee ('DISQ' = Department of Scientific Direction, 'DMCT' = Department of Chronic Diseases and Traumatisms, 'DSE' = Department of Health and Environment, 'SRH' = Human Resources, and 'SFLE' = Logistics).

Source

https://sociopatterns.org/datasets/contacts-in-a-workplace/

References

Genois et al. (2015) Network Science 3(3) pp. 326-347 doi:10.1017/nws.2015.10, Genois & Barrat (2018) EPJ Data Science 7:11 doi:10.1140/epjds/s1368801801401

See Also

workplace2015

`workplace2015` *Contacts in a workplace.*

Description

This dataset contains the sequence of 33751 face-to-face contacts between employees measured in an office building in France, during two weeks in 2015. The data collection involved 232 employees of 12 different departments. Wearable proximity sensors recorded with a 20-second time interval every face-to-face contact between individuals that was active in the last 20 second window. Following Genois & Barrat (2018), a contact between two individuals is defined as the set of successive time-windows of 20 seconds during which the individuals are detected in contact. In the 'edgelist' element, these records are collected in a relational event sequence with for every contact between a pair of individuals the time of onset and its duration. Time is expressed in seconds since the onset of the study period. The events in the sequence are undirected. Multiple events can occur at the same time point and the individuals can be in contact with multiple other individuals at the same time. Information on the department of the individuals is available in the 'attributes' element.

Usage

```
data(workplace2015)
```

Format

List with elements 'edgelist' and 'attributes'.

edgelist

Dataframe with 33751 rows and 4 columns that contains the relational event sequence with face-to-face contacts in the workplace.

time Numeric value that indicates the onset of the contact in seconds since onset of the study period.

actor1 Integer value with the ID of the first actor in the contact.

actor2 Integer value with the ID of the second actor in the contact.

duration Numeric value that indicates the duration of the contact in seconds.

attributes

Dataframe with 232 rows and 3 columns describing the department of the employees.

id Integer value with the ID of the individual.

time Change time of the attributes. Since the attributes are constant over the course of the study period, this value is equal to 0 for all entries.

department String with the department of the employee.

Source

https://sociopatterns.org/datasets/test/

References

Genois & Barrat (2018) EPJ Data Science 7:11 doi:10.1140/epjds/s1368801801401

See Also

workplace2013

WS16 *GESIS Computational Social Science Winter Symposium 2016 (WS16)*

Description

This dataset contains the sequence of 53164 face-to-face contacts between visitors of the GESIS Computational Social Science Winter Symposium 2016 (WS16) in Cologne on November 30 and December 1, 2016. The data collection involved 138 of the 149 conference attendees. Wearable proximity sensors recorded every face-to-face contact between individuals (physical proximity within approximately 1.5m) with a 20-second time resolution. A contact between two individuals is defined as the set of successive time-windows of 20-second during which the individuals are detected in close proximity. In the 'edgelist' element, these records are collected in a relational event sequence with for every contact between a pair of individuals the time of onset and its duration. Time is measured in seconds. The events in the sequence are undirected. Multiple events can occur at the same time point and the individuals can be in contact with multiple other individuals at the same time. Socio-demographic information about the participants (seniority, disciplinary background, country of residence, mother tongue, age, gender, role at conference) was collected through a survey and available in the 'attributes' element.

Usage

```
data(WS16)
```

Format

List with elements 'edgelist' and 'attributes'

edgelist

Dataframe with 53164 rows and 4 columns that contains the relational event sequence with face-to-face contacts between the participants.

time POSIXct value that indicates the onset of the contact.

actor1 Integer value with the ID of the first actor in the contact.

actor2 Integer value with the ID of the second actor in the contact.

duration Numeric value that indicates the duration of the contact in seconds.

attributes

Dataframe with 138 rows and 9 columns describing the socio-demographic information about the participants.

id Integer value with the ID of the individual.

time Change time of the attributes. Since the attributes are constant over the course of the study period, this value is equal to 0 for all entries.

age Age range for the individual. One of 'younger than 30 years', '30 to 39 years', or '40 years and older'.

gender Gender of the individual, 'Female' or 'Male'.

country Code for the country of residence of the individual.

language Code for the mother tongue of the individual.

seniority Academic seniority of the individual at the time of the conference. One of 'Bachelor student', 'Master student', 'PhD student', 'Postdoctoral researcher', 'Assistant/associate professor', 'Full professor', 'Other'

background Academic background of the individual. One of 'Social Science, Humanities, Political Science', 'Computer Science, Information Science', 'Physics, Mathematics, Statistics', 'Media, Communication, Other'.

role Role in the conference of the individual. One of 'Speaker', 'Poster presenter', 'Participation only', 'Staff'

previous Did the individual participate to a previous edition? (Yes or No).

Source

https://psycharchives.org/en/item/26710d25-c387-4506-b5f5-60d1bd3009a5

References

Genois, M., Zens, M., Oliveira, M., Lechner, C. M., Schaible, J., & Strohmaier, M. (2023). Combining sensors and surveys to study social interactions: A case of four science conferences. *Personality Science*, 4, Article e9957. doi:10.5964/ps.9957

Index

∗ **datasets**
ants, 2
baboons_obs, 4
baboons_sens, 5
brazilian_states_trade, 6
dnc, 7
highschool2011, 8
highschool2012, 9
highschool2013, 10
hospitalward, 13
households, 14
hypertext, 15
ICCSS17, 16
infectious, 18
Malawi, 19
manufacturing, 20
primaryschool, 21
randomREH, 22
randomREHsmall, 23
RFIDvalidity, 24
SFHHconference, 25
team4, 26
wikitalkbr, 27
workplace2013, 28
workplace2015, 30
WS16, 31
∗ **dataset**
history, 12
info, 19

ants, 2

baboons_obs, 4, *5*
baboons_sens, 5, *5*
brazilian_states_trade, 6

dnc, 7

highschool2011, 8, *10*, *12*
highschool2012, 9, *9*, *12*
highschool2013, 10
history, 12, *19*
hospitalward, 13
households, 14
hypertext, 15

ICCSS17, 16
infectious, 18
info, *12*, 19

Malawi, 19
manufacturing, 20

primaryschool, 21

randomREH, 22
randomREHsmall, 23
RFIDvalidity, 24

SFHHconference, 25

team4, 26

wikitalkbr, 27
workplace2013, 28, *31*
workplace2015, *29*, 30
WS16, 31