Package 'onadata'

January 24, 2022

Title Data Sets for Keith McNulty's Handbook of Graphs and Networks in People Analytics

Version 0.1

Description Data sets for network analysis related to People Analytics.
Contains various data sets from the book 'Handbook of Graphs and Networks in People Analytics' by Keith McNulty (2021).

License MIT + file LICENSE

Encoding UTF-8

LazyData true

URL https://ona-book.org

RoxygenNote 7.1.1

Depends R (>= 2.10)

NeedsCompilation no

Author Keith McNulty [aut, cre] (<https://orcid.org/0000-0002-2332-1654>)

Maintainer Keith McNulty <keith.mcnulty@gmail.com>

Repository CRAN

Date/Publication 2022-01-24 20:22:42 UTC

R topics documented:

`caviar_end` *Caviar end data*

Description

Network edgelist data as at the end of the Operation Caviar investigation into drug trafficking in Canada

Usage

```
caviar_end
```

Format

A dataframe with 72 rows and 3 variables:

from An individual under surveillance

to An individual under surveillance

weight The number of intercepted communications between the individuals

Source

Carlo Morselli

Examples

```
caviar_end
```

`caviar_middle` *Caviar middle data*

Description

Network edgelist data as at the middle of the Operation Caviar investigation into drug trafficking in Canada

Usage

```
caviar_middle
```

Format

A dataframe with 50 rows and 3 variables:

from An individual under surveillance

to An individual under surveillance

weight The number of intercepted communications between the individuals

Source

Carlo Morselli

Examples

```
caviar_middle
```

`caviar_start` *Caviar start data*

Description

Network edgelist data as at the start of the Operation Caviar investigation into drug trafficking in Canada

Usage

```
caviar_start
```

Format

A dataframe with 26 rows and 3 variables:

from An individual under surveillance

to An individual under surveillance

weight The number of intercepted communications between the individuals

Source

Carlo Morselli

Examples

```
caviar_start
```

---

chinook_customers *Chinook customer data*

Description

Extract of data on customers of a music sales company

Usage

```
chinook_customers
```

Format

A dataframe with 59 rows and 4 variables:

CustomerId Customer ID number

FirstName Customer First Name

LastName Customer Last Name

SupportRepId ID of Sales Rep assigned to customer

Source

Chinook Database on Github

Examples

```
chinook_customers
```

chinook_employees *Chinook employee data*

Description

Extract of data on employees of a music sales company

Usage

```
chinook_employees
```

Format

A dataframe with 8 rows and 4 variables:

EmployeeId Employee ID number

FirstName Employee First Name

LastName Employee Last Name

ReportsTo ID of Employee who they report to

Source

Chinook Database on Github

Examples

```
chinook_employees
```

chinook_invoices *Chinook invoice data*

Description

Extract of data on customer invoices from a music sales company

Usage

```
chinook_invoices
```

Format

A dataframe with 412 rows and 2 variables:

InvoiceId Invoice ID number

CustomerId CustomerID number

Source

Chinook Database on Github

Examples

```
chinook_invoices
```

`chinook_items` *Chinook item sales data*

Description

Extract of data on items sold by a music sales company

Usage

```
chinook_items
```

Format

A dataframe with 2240 rows and 2 variables:

InvoiceId ID number of invoice containing the item

TrackId ID number of the item

Source

Chinook Database on Github

Examples

```
chinook_items
```

`dolphins` *Bottlenose dolphin social network*

Description

Edgelist of network of frequent interaction between bottlenose dolphins in Doubtful Sound, New Zealand

Usage

```
dolphins
```

Format

A dataframe with 159 rows and 2 variables:

from Dolphin ID

to Dolphin ID

Source

Lusseau et al, 2003

Examples

```
dolphins
```

---

email_edgelist *Email edgelist*

Description

Edgelist of network of email communications at a large European research institution

Usage

```
email_edgelist
```

Format

A dataframe with 24929 rows and 2 variables:

from ID of sender

to ID of receiver

Source

Leskovec et al, 2007

Examples

```
email_edgelist
```

email_vertices *Email vertices*

Description

Vertex data of network of email communications at a large European research institution

Usage

```
email_vertices
```

Format

A dataframe with 1005 rows and 2 variables:

id Vertex ID of individual

dept Department of individual

Source

Leskovec et al, 2007

Examples

```
email_vertices
```

eu_referendum *EU referendum data*

Description

Data on voting in the UK EU membership referendum in 2016

Usage

```
eu_referendum
```

Format

A dataframe with 382 rows and 4 variables:

Region UK Region

Area_Code UK Area Code

Remain Number of votes to remain in the EU

Leave Number of votes to leave the EU

Source

UK Electoral Commission

Examples

```
eu_referendum
```

---

friends_tv_edgelist *Friends TV edgelist*

Description

Edgelist of network of characters of US TV Show Friends based on appearing in the same scene

Usage

```
friends_tv_edgelist
```

Format

A dataframe with 2976 rows and 3 variables:

from Friends character

to Friends character

weight Number of scenes with both characters

Source

Keith McNulty

Examples

```
friends_tv_edgelist
```

g14_edgelist *G14 edgelist*

Description

Edgelist of small network of 14 vertices

Usage

```
g14_edgelist
```

Format

A dataframe with 18 rows and 3 variables:

from Vertex ID

to Vertex ID

weight Edge weight

Examples

```
g14_edgelist
```

karate *Zachary's Karate Club edgelist*

Description

Edgelist of network of social interactions between members of a karate club

Usage

```
karate
```

Format

A dataframe with 78 rows and 2 variables:

from Member ID

to Member ID

Source

Zachary, 1977

Examples

```
karate
```

koenigsberg *Bridges of Koenigsberg edgelist*

Description

Edgelist of network of places connected by bridges in the city of Koenigsberg

Usage

```
koenigsberg
```

Format

A dataframe with 7 rows and 2 variables:

from Place name

to Place name

Source

Euler, 1736

Examples

```
koenigsberg
```

lesmis *Les Miserables character network*

Description

Edgelist of network of characters in Victor Hugo's Les Miserables based on appearance in the same chapter

Usage

```
lesmis
```

Format

A dataframe with 254 rows and 3 variables:

from Character name

to Character name

weight Number of chapters both characters appear in

Source

Knuth, 1993

Examples

```
lesmis
```

`londontube_edgelist` *London Tube network edgelist*

Description

Edgelist of network of London Tube/Underground stations

Usage

```
londontube_edgelist
```

Format

A dataframe with 406 rows and 4 variables:

from Station ID

to Station ID

line Name of line connecting stations

linecolor Official color of line connecting stations

Examples

```
londontube_edgelist
```

`londontube_vertices` *London Tube network vertices*

Description

Vertices of network of London Tube/Underground stations

Usage

```
londontube_vertices
```

Format

A dataframe with 302 rows and 4 variables:

id Station ID

name Station name

latitude Station latitude

longitude Station longitude

Examples

```
londontube_vertices
```

`madmen_edges` *Mad Men network edgelist*

Description

Edgelist of network of romantic relationships between characters of the TV show Mad Men

Usage

```
madmen_edges
```

Format

A dataframe with 39 rows and 3 variables:

Name1 Character name

Name2 Character name

Married Whether the relationship was part of a marriage

Examples

```
madmen_edges
```

madmen_vertices *Mad Men network vertices*

Description

Vertices of network of romantic relationships between characters of the TV show Mad Men

Usage

```
madmen_vertices
```

Format

A dataframe with 45 rows and 3 variables:

label Character name

Gender Character gender

Main Whether the character is a main character

Examples

```
madmen_vertices
```

netscience *Network Science collaboration network*

Description

Edgelist of network of academic collaboration between network scientists

Usage

```
netscience
```

Format

A dataframe with 2742 rows and 3 variables:

from Scientist name

to Scientist name

weight Measure of strength of collaboration

Source

Newman, 2006

Examples

```
netscience
```

ontariopol_edgelist *Ontario politician Twitter interaction network edglist*

Description

Edgelist of Twitter interaction network of Ontario province politicians

Usage

```
ontariopol_edgelist
```

Format

A dataframe with 6095 rows and 3 variables:

from Politician ID

to Politician ID

weight Number of Twitter interactions

Source

Christopher Belanger

Examples

```
ontariopol_edgelist
```

ontariopol_vertices *Ontario politician Twitter interaction network vertices*

Description

Vertices of Twitter interaction network of Ontario province politicians

Usage

```
ontariopol_vertices
```

Format

A dataframe with 108 rows and 4 variables:

id Politician ID

screen_name Politician Twitter screen name

name Politician name

party Party affiliation

Source

Christopher Belanger

Examples

```
ontariopol_vertices
```

`park_reviews` *Yelp park reviews*

Description

Data on Yelp reviews of dog parks in Phoenix, AZ

Usage

```
park_reviews
```

Format

A dataframe with 231 rows and 4 variables:

park_id Park ID

user_id User ID

park_name Park name

stars Number of stars awarded by user

Examples

```
park_reviews
```

`pizza` *Random Acts of Pizza*

Description

Data on altruistic acts by Reddit users fulfiling random requests for pizza

Usage

```
pizza
```

Format

A dataframe with 400 rows and 5 variables:

requester ID of the requester

responder ID of the individual who responded by ordering pizza for the requester

request_id ID of the request

requester_votes Number of Reddit votes made by the requester

requester_subreddits Number of subreddits which the requester is a member of

Source

Althoff et al, 2014

Examples

```
pizza
```

---

s50_edges *Teenage Friends and Lifestyle Study network edgelist*

Description

Edgelist of friend network of teenage girls in Scotland

Usage

```
s50_edges
```

Format

A dataframe with 122 rows and 2 variables:

from Person ID

to Person ID

Source

Pearson & Michell, 2009

Examples

```
s50_edges
```

s50_vertices *Teenage Friends and Lifestyle Study network vertices*

Description

Vertices of friend network of teenage girls in Scotland

Usage

```
s50_vertices
```

Format

A dataframe with 50 rows and 5 variables:

id Person ID

smoke Frequency of smoking from 1 (Never) to 3 (Regularly)

alcohol Frequency of drinking alcohol from 1 (Never) to 5 (More than once a week)

drugs Frequency of cannabis use from 1 (Never) to 4 (Regularly)

sport Frequency of sporting activity from 1 (Not regularly) to 2 (Regularly)

Source

Pearson & Michell, 2009

Examples

```
s50_vertices
```

schoolfriends_edgelist *Schoolfriends network edgelist*

Description

Edgelist of network of schoolfriends in a French high school

Usage

```
schoolfriends_edgelist
```

Format

A dataframe with 2105 rows and 3 variables:

from Person ID

to Person ID

type Whether the friendship is a known Facebook connection or if it was reported by `from` person

Source

Mastandrea et al, 2015

Examples

```
schoolfriends_edgelist
```

`schoolfriends_vertices`

Schoolfriends network vertices

Description

Vertices of network of schoolfriends in a French high school

Usage

```
schoolfriends_vertices
```

Format

A dataframe with 329 rows and 3 variables:

id Person ID

class School class of person

gender Gender of person

Source

Mastandrea et al, 2015

Examples

```
schoolfriends_vertices
```

wikivote *Wikipedia administrator voting network*

Description

Edgelist of network of votes for Wikipedia administrators

Usage

```
wikivote
```

Format

A dataframe with 103688 rows and 2 variables:

from ID of voter

to ID of vote recipient

Examples

```
wikivote
```

workfrance_edgelist *Workplace network edgelist*

Description

Edgelist of network of interactions between people in a French office building based on location sensor technology

Usage

```
workfrance_edgelist
```

Format

A dataframe with 932 rows and 3 variables:

from Person ID

to Person ID

mins Number of minutes spent co-located

Source

Génois et al, 2015

Examples

```
workfrance_edgelist
```

`workfrance_vertices` *Workplace network vertices*

Description

Vertices of network of interactions between people in a French office building based on location sensor technology

Usage

```
workfrance_vertices
```

Format

A dataframe with 211 rows and 2 variables:

id Person ID

dept Department of person

Source

Génois et al, 2015

Examples

```
workfrance_vertices
```

Index