Package ‘Ecdat’

February 19, 2015

Version 0.2-7
Date 2013-04-25
Title Data sets for econometrics
Author Yves Croissant <yves.croissant@let.ish-lyon.cnrs.fr>
Maintainer Spencer Graves <spencer.graves@effectivedefense.org>
Depends R (>= 2.10), Ecfun
Suggests car, systemfit, sem, lme4, sandwich, gdata, RCurl, XML, tis
Description Data sets for econometrics
LazyData true
License GPL (>= 2)
URL http://www.r-project.org
Repository CRAN
Repository/R-Forge/Project ecdat
Repository/R-Forge/Revision 239
Repository/R-Forge/DateTimeStamp 2014-09-03 17:38:51
Date/Publication 2014-09-04 06:55:00
NeedsCompilation no

R topics documented:

  Accident ................................................................. 4
  Airline ........................................................................ 5
  Airq ........................................................................... 6
  bankingCrises ............................................................. 7
  Benefits ........................................................................ 8
  Bids ........................................................................... 9
  BudgetFood ................................................................. 10
  BudgetItaly ................................................................. 11
  BudgetUK ................................................................. 12
  Bwages ...................................................................... 13
**R topics documented:**

<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capm</td>
<td>14</td>
</tr>
<tr>
<td>Car</td>
<td>15</td>
</tr>
<tr>
<td>Caschool</td>
<td>16</td>
</tr>
<tr>
<td>Catsup</td>
<td>17</td>
</tr>
<tr>
<td>Cigar</td>
<td>18</td>
</tr>
<tr>
<td>Cigarette</td>
<td>19</td>
</tr>
<tr>
<td>Clothing</td>
<td>20</td>
</tr>
<tr>
<td>Computers</td>
<td>21</td>
</tr>
<tr>
<td>Consumption</td>
<td>22</td>
</tr>
<tr>
<td>CPSch3</td>
<td>23</td>
</tr>
<tr>
<td>Cracker</td>
<td>24</td>
</tr>
<tr>
<td>CRANpackages</td>
<td>25</td>
</tr>
<tr>
<td>Crime</td>
<td>26</td>
</tr>
<tr>
<td>CRSPday</td>
<td>27</td>
</tr>
<tr>
<td>CRSPmon</td>
<td>28</td>
</tr>
<tr>
<td>Diamond</td>
<td>29</td>
</tr>
<tr>
<td>DM</td>
<td>30</td>
</tr>
<tr>
<td>Doctor</td>
<td>31</td>
</tr>
<tr>
<td>DoctorAUS</td>
<td>32</td>
</tr>
<tr>
<td>DoctorContacts</td>
<td>33</td>
</tr>
<tr>
<td>Earnings</td>
<td>34</td>
</tr>
<tr>
<td>Electricity</td>
<td>35</td>
</tr>
<tr>
<td>Fair</td>
<td>36</td>
</tr>
<tr>
<td>Fatality</td>
<td>37</td>
</tr>
<tr>
<td>FinancialCrisisFiles</td>
<td>38</td>
</tr>
<tr>
<td>Fishing</td>
<td>39</td>
</tr>
<tr>
<td>Forward</td>
<td>40</td>
</tr>
<tr>
<td>FriendFoe</td>
<td>41</td>
</tr>
<tr>
<td>Garch</td>
<td>42</td>
</tr>
<tr>
<td>Gasoline</td>
<td>43</td>
</tr>
<tr>
<td>Griliches</td>
<td>44</td>
</tr>
<tr>
<td>Grunfeld</td>
<td>45</td>
</tr>
<tr>
<td>HC</td>
<td>46</td>
</tr>
<tr>
<td>Hdma</td>
<td>47</td>
</tr>
<tr>
<td>Heating</td>
<td>48</td>
</tr>
<tr>
<td>Hedonic</td>
<td>49</td>
</tr>
<tr>
<td>HI</td>
<td>50</td>
</tr>
<tr>
<td>Housing</td>
<td>51</td>
</tr>
<tr>
<td>Hstarts</td>
<td>52</td>
</tr>
<tr>
<td>Icecream</td>
<td>53</td>
</tr>
<tr>
<td>incomeInequality</td>
<td>54</td>
</tr>
<tr>
<td>IncomeUK</td>
<td>59</td>
</tr>
<tr>
<td>Index.Econometrics</td>
<td>60</td>
</tr>
<tr>
<td>Index.Economics</td>
<td>62</td>
</tr>
<tr>
<td>Index.Observations</td>
<td>65</td>
</tr>
<tr>
<td>Index.Source</td>
<td>68</td>
</tr>
<tr>
<td>Index.Time.Series</td>
<td>72</td>
</tr>
<tr>
<td>Irates</td>
<td>73</td>
</tr>
<tr>
<td>Topic</td>
<td>Page</td>
</tr>
<tr>
<td>------------------</td>
<td>------</td>
</tr>
<tr>
<td>Journals</td>
<td>74</td>
</tr>
<tr>
<td>Kakadu</td>
<td>75</td>
</tr>
<tr>
<td>Ketchup</td>
<td>77</td>
</tr>
<tr>
<td>Klein</td>
<td>78</td>
</tr>
<tr>
<td>LaborSupply</td>
<td>79</td>
</tr>
<tr>
<td>Labour</td>
<td>80</td>
</tr>
<tr>
<td>Longley</td>
<td>80</td>
</tr>
<tr>
<td>LT</td>
<td>81</td>
</tr>
<tr>
<td>Macrodat</td>
<td>82</td>
</tr>
<tr>
<td>Males</td>
<td>83</td>
</tr>
<tr>
<td>ManufCost</td>
<td>84</td>
</tr>
<tr>
<td>Mathlevel</td>
<td>85</td>
</tr>
<tr>
<td>MCAS</td>
<td>86</td>
</tr>
<tr>
<td>MedExp</td>
<td>87</td>
</tr>
<tr>
<td>Metal</td>
<td>88</td>
</tr>
<tr>
<td>Mishkin</td>
<td>89</td>
</tr>
<tr>
<td>Mode</td>
<td>90</td>
</tr>
<tr>
<td>ModeChoice</td>
<td>91</td>
</tr>
<tr>
<td>Mofa</td>
<td>92</td>
</tr>
<tr>
<td>Money</td>
<td>93</td>
</tr>
<tr>
<td>MoneyUS</td>
<td>94</td>
</tr>
<tr>
<td>Mpyr</td>
<td>95</td>
</tr>
<tr>
<td>Mroz</td>
<td>96</td>
</tr>
<tr>
<td>MunExp</td>
<td>97</td>
</tr>
<tr>
<td>MW</td>
<td>98</td>
</tr>
<tr>
<td>NaturalPark</td>
<td>99</td>
</tr>
<tr>
<td>Nerlove</td>
<td>100</td>
</tr>
<tr>
<td>nonEnglishNames</td>
<td>101</td>
</tr>
<tr>
<td>OFP</td>
<td>101</td>
</tr>
<tr>
<td>Oil</td>
<td>103</td>
</tr>
<tr>
<td>Orange</td>
<td>104</td>
</tr>
<tr>
<td>Participation</td>
<td>105</td>
</tr>
<tr>
<td>PatentsHGH</td>
<td>106</td>
</tr>
<tr>
<td>PatentsRD</td>
<td>107</td>
</tr>
<tr>
<td>PE</td>
<td>108</td>
</tr>
<tr>
<td>politicalKnowledge</td>
<td>109</td>
</tr>
<tr>
<td>Pound</td>
<td>111</td>
</tr>
<tr>
<td>PPP</td>
<td>112</td>
</tr>
<tr>
<td>Pricing</td>
<td>113</td>
</tr>
<tr>
<td>Produc</td>
<td>114</td>
</tr>
<tr>
<td>PSID</td>
<td>115</td>
</tr>
<tr>
<td>RetSchool</td>
<td>116</td>
</tr>
<tr>
<td>Schooling</td>
<td>118</td>
</tr>
<tr>
<td>Solow</td>
<td>119</td>
</tr>
<tr>
<td>Somerville</td>
<td>120</td>
</tr>
<tr>
<td>SP500</td>
<td>121</td>
</tr>
<tr>
<td>Star</td>
<td>122</td>
</tr>
<tr>
<td>Strike</td>
<td>123</td>
</tr>
</tbody>
</table>
Accident

Description

a cross-section

number of observations : 40

Usage

data(Accident)

Format

A dataframe containing :

type  ship type, a factor with levels (A,B,C,D,E)
constr year constructed, a factor with levels (C6064,C6569,C7074,C7579)
operate year operated, a factor with levels (O6074,O7579)
months measure of service amount
acc accidents
Source

References

See Also
Index.Source, Index.Economics, Index.Econometrics, Index.Observations

<table>
<thead>
<tr>
<th>Airline</th>
<th>Cost for U.S. Airlines</th>
</tr>
</thead>
</table>

Description
a panel of 6 observations from 1970 to 1984

number of observations : 90
observation : production units
country : United States

Usage
data(Airline)

Format
A dataframe containing :

airline  airline
year     year
cost     total cost, in $1,000
output   output, in revenue passenger miles, index number
pf       fuel price
lf       load factor, the average capacity utilization of the fleet

References

See Also
### Description

A cross-section from 1972

*number of observations*: 30

*observation*: regional

*country*: United States

### Usage

data(Airq)

### Format

A dataframe containing:

- **airq**: indicator of air quality (the lower the better)
- **vala**: value added of companies (in thousands of dollars)
- **rain**: amount of rain (in inches)
- **coas**: is it a coastal area?
- **dens**: population density (per square mile)
- **medi**: average income per head (in US dollars)

### References


### See Also

[Index.Source], [Index.Economics], [Index.Econometrics], [Index.Observations]
Description

A data.frame identifying which of 70 countries had a banking crisis each year 1800:2010. The first column is year. The remaining columns carry the names of the countries; those columns are 1 for years with banking crises and 0 otherwise.

Usage

data(bankingCrises)

Format

A data.frame

Details

This file was created using the following command:

bankingCrises <- readFinancialCrisisFiles(FinancialCrisisFiles)

This is documented further in the help file for readFinancialCrisisFiles.

This is an update of a subset of the data used to create Figure 10.1. Capital Mobility and the Incidence of Banking Crises, All Countries, 1800-2008, Reinhart and Rogoff (2009, p. 156).

The general upward trend visible in a plot of these data may be attributed to at least two different factors: (1) The gradual increase in the proportion of human labor that is monetized. (2) An increase in the general ability of cronies of those in power to gamble with other people’s money in forming and bankrupting financial institutions. The marked feature of this plot is the virtual absence of banking crises during the period of the Bretton Woods agreement, 1944, to 1971 when US President Nixon in effect canceled it by taking the US off the silver standard.

Author(s)

Spencer Graves

Source

http://www.reinhartandrogoff.com

References


See Also

readFinancialCrisisFiles
Examples

data(bankingCrises)
numberOfCrises <- rowSums(bankingCrises[-1], na.rm=TRUE)
plot(bankingCrises$year, numberOfCrises, type='b')

# Write to a file for Wikimedia Commons
svg('bankingCrises.svg')
plot(bankingCrises$year, numberOfCrises, type='b', cex.axis=2,
     las=1, xlab='', ylab='', bty='n', cex=0.5)
abline(v=c(1945, 1971), lty='dashed', col='blue')
text(1958, 14, 'Bretton Woods', srt=90, cex=2, col='blue')
dev.off()

Description

a cross-section from 1972

number of observations: 4877
observation: individuals
country: United States

Usage

data(Benefits)

Format

A time serie containing:

stateur state unemployment rate (in %)
statemb state maximum benefit level
state state of residence code
age age in years
tenure years of tenure in job lost
joblost a factor with levels (slack\_work,position\_abolished,seasonal\_job\_ended,other)
nwhite non-white?
school12 more than 12 years of school?
sex a factor with levels (male,female)
bluecol blue collar worker?
smsa lives in smsa?
married married?


dkids has kids ?
dykids has young kids (0-5 yrs) ?
yrdispl year of job displacement (1982=1,..., 1991=10)
rr replacement rate
head is head of household ?
ui applied for (and received) UI benefits ?

Source


References


See Also


Bids

<table>
<thead>
<tr>
<th>Bids Received By U.S. Firms</th>
</tr>
</thead>
</table>

Description

a cross-section

number of observations : 126

observation : production units

country : United States

Usage

data(Bids)
Format

A dataframe containing:

docno  doc no.
weeks weeks
numbids count
takeover delta (1 if taken over)
bidprem bid Premium
insthold institutional holdings
size size measured in billions
legrest legal restructuring
rearest real restructuring
finrest financial restructuring
regulatn regulation
whtknight white knight

Source


References


See Also

Index.Source, Index.Economics, Index.Econometrics, Index.Observations

<table>
<thead>
<tr>
<th>BudgetFood</th>
<th>Budget Share of Food for Spanish Households</th>
</tr>
</thead>
</table>

Description

a cross-section from 1980

number of observations : 23972

observation : households
country : Spain
Usage
data(BudgetFood)

Format
A dataframe containing:

- **wfood**: percentage of total expenditure which the household has spent on food
- **totexp**: total expenditure of the household
- **age**: age of reference person in the household
- **size**: size of the household
- **town**: size of the town where the household is placed categorised into 5 groups: 1 for small towns, 5 for big ones
- **sex**: sex of reference person (man, woman)

Source

References

See Also
Index.Source, Index.Economics, Index.Econometrics, Index.Observations

<table>
<thead>
<tr>
<th>BudgetItaly</th>
<th>Budget Shares for Italian Households</th>
</tr>
</thead>
</table>

Description
a cross-section from 1973 to 1992

- **number of observations**: 1729
- **observation**: households
- **country**: Italy

Usage
data(BudgetItaly)
BudgetUK

Format

A dataframe containing:

- `wfood` food share
- `whouse` housing and fuels share
- `wmisc` miscellaneous share
- `pfood` food price
- `phouse` housing and fuels price
- `pmisc` miscellaneous price
- `totexp` total expenditure
- `year` year
- `income` income
- `size` household size
- `pct` cellule weight

Source


References


See Also

*Index.Source, Index.Economics, Index.Econometrics, Index.Observations*

---

**BudgetUK**    

*Budget Shares of British Households*

Description

a cross-section from 1980 to 1982

- number of observations: 1519
- observation: households
- country: United Kingdon

Usage

`data(BudgetUK)`
Bwages

Format

A dataframe containing:

- `wfood` budget share for food expenditure
- `wfuel` budget share for fuel expenditure
- `wcloth` budget share for clothing expenditure
- `walc` budget share for alcohol expenditure
- `wtrans` budget share for transport expenditure
- `wother` budget share for other good expenditure
- `totexp` total household expenditure (rounded to the nearest 10 UK pounds sterling)
- `income` total net household income (rounded to the nearest 10 UK pounds sterling)
- `age` age of household head
- `children` number of children

Source


References


See Also

Index.Source, Index.Economics, Index.Econometrics, Index.Observations

Bwages  Wages in Belgium

Description

a cross-section from 1994

- number of observations: 1472
- observation: individuals
- country: Belgium

Usage

data(Bwages)
Format
A dataframe containing:

- **wage**: gross hourly wage rate in euro
- **educ**: education level from 1 [low] to 5 [high]
- **exper**: years of experience
- **sex**: a factor with levels (males,female)

Source
European Community Household Panel.

References

See Also
Index.Source, Index.Economics, Index.Econometrics, Index.Observations

---

**Capm** Stock Market Data

Description
monthly observations from 1960–01 to 2002–12

*number of observations*: 516

Usage
data(Capm)

Format
A time serie containing:

- **rfood**: excess returns food industry
- **rdur**: excess returns durables industry
- **rcon**: excess returns construction industry
- **rmrf**: excess returns market portfolio
- **rf**: riskfree return

Source
most of the above data are from Kenneth French’s data library at [http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library](http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library).
Car

References

See Also

<table>
<thead>
<tr>
<th>Car</th>
<th>Stated Preferences for Car Choice</th>
</tr>
</thead>
</table>

Description
a cross-section

number of observations : 4654

observation : individuals
country : United States

Usage
data(Car)

Format
A dataframe containing :

choice choice of a vehicle among 6 propositions
college college education ?
hsg2 size of household greater than 2 ?
coml5 commute lower than 5 miles a day ?
typez body type, one of regcar (regular car), sportuv (sport utility vehicle), sportcar, stwagon (station wagon), truck, van, for each proposition z from 1 to 6
fuelz fuel for proposition z, one of gasoline, methanol, cng (compressed natural gas), electric.
pricez price of vehicle divided by the logarithm of income
rangez hundreds of miles vehicle can travel between refuelings/rechargings
accz acceleration, tens of seconds required to reach 30 mph from stop
speedz highest attainable speed in hundreds of mph
pollutionz tailpipe emissions as fraction of those for new gas vehicle
sizez 0 for a mini, 1 for a subcompact, 2 for a compact and 3 for a mid-size or large vehicle
spacez fraction of luggage space in comparable new gas vehicle
costz cost per mile of travel (tens of cents) : home recharging for electric vehicle, station refueling otherwise
stationz fraction of stations that can refuel/recharge vehicle
Source

References

See Also
*Index.Source, Index.Economics, Index.Econometrics, Index.Observations*

---

**Caschool**

*The California Test Score Data Set*

**Description**
a cross-section from 1998-1999

- *number of observations*: 420
- *observation*: schools
- *country*: United States

**Usage**
data(Caschool)

**Format**
A dataframe containing:

- **distcod**: district code
- **county**: county
- **district**: district
- **grspan**: grade span of district
- **enrltot**: total enrollment
- **teachers**: number of teachers
- **calwpct**: percent qualifying for CalWorks
- **mealpct**: percent qualifying for reduced-price lunch
- **computer**: number of computers
- **testscr**: average test score \((\text{read.scr}+\text{math.scr})/2\)
- **comptsu**: computer per student
- **expnstu**: expenditure per student
- **str**: student teacher ratio
**Catsup**

- *avginc* district average income
- *elpct* percent of English learners
- *readscr* average reading score
- *mathscr* average math score

**Source**


**References**


**See Also**

Index.Source, Index.Economics, Index.Econometrics, Index.Observations

---

**Catsup**

*Choice of Brand for Catsup*

**Description**

a cross-section

- *number of observations*: 2798
- *observation*: individuals
- *country*: United States

**Usage**

`data(Catsup)`

**Format**

A dataframe containing:

- *id*  individuals identifiers
- *choice*  one of heinz41, heinz32, heinz28, hunts32
- *disp.z*  is there a display for brand z?
- *feat.z*  is there a newspaper feature advertisement for brand z?
- *price.z*  price of brand z
Source

References

See Also
[Index.Source, Index.Economics, Index.Econometrics, Index.Observations]

---

<table>
<thead>
<tr>
<th>Cigar</th>
<th>Cigarette Consumption</th>
</tr>
</thead>
</table>

**Description**
a panel of 46 observations from 1963 to 1992

- **number of observations**: 1380
- **observation**: regional
- **country**: United States

**Usage**
data(Cigar)

**Format**
A dataframe containing:

- **state**: state abbreviation
- **year**: the year
- **price**: price per pack of cigarettes
- **pop**: population
- **pop16**: population above the age of 16
- **cpi**: consumer price index (1983=100)
- **ndi**: per capita disposable income
- **sales**: cigarette sales in packs per capita
- **pimin**: minimum price in adjoining states per pack of cigarettes
Source


References


See Also


---

**Cigarette**

*The Cigarette Consumption Panel Data Set*

Description

a panel of 48 observations from 1985 to 1995

*number of observations*: 528
*observation*: regional
*country*: United States

Usage

data(Cigarette)

Format

A dataframe containing:

*state*  state
*year*   year
*cpi*    consumer price index
*pop*    state population
*packpc* number of packs per capita
*income* state personal income (total, nominal)
*tax*    average state, federal, and average local excise taxes for fiscal year
*avgprs* average price during fiscal year, including sales taxes
*taxs*   average excise taxes for fiscal year, including sales taxes
Source
Professor Jonhatan Gruber, MIT.

References

See Also

| Clothing | Sales Data of Men's Fashion Stores |

Description
a cross-section from 1990
number of observations : 400
observation : production units
country : Netherland

Usage
data(Clothing)

Format
A dataframe containing :
tsales annual sales in Dutch guilders
sales sales per square meter
margin gross-profit-margin
nown number of owners (managers)
full number of full-timers
part number of part-timers
aux number of helpers (temporary workers)
hour total number of hours worked
hourspw number of hours worked per worker
inv1 investment in shop-premises
inv2 investment in automation.
size sales floorspace of the store (in m$^2$).
start year start of business
References


See Also

`Index.Source, index.Economics, Index.Econometrics, Index.Observations`

---

### Computers

*Prices of Personal Computers*

#### Description

- A cross-section from 1993 to 1995
- *number of observations*: 6259
- *observation*: goods
- *country*: United States

#### Usage

`data(Computers)`

#### Format

A dataframe containing:

- **price**: price in US dollars of 486 PCs
- **speed**: clock speed in MHz
- **hd**: size of hard drive in MB
- **ram**: size of Ram in MB
- **screen**: size of screen in inches
- **cd**: is a CD-ROM present?
- **multi**: is a multimedia kit (speakers, sound card) included?
- **premium**: is the manufacturer was a "premium" firm (IBM, COMPAQ)?
- **ads**: number of 486 price listings for each month
- **trend**: time trend indicating month starting from January of 1993 to November of 1995.

#### Source

Consumption

References

See Also
Index.Source, Index.Economics, Index.Econometrics, Index.Observations

---

Consumption: Quarterly Data on Consumption and Expenditure

Description
quarterly observations from 1947-1 to 1996-4

*number of observations*: 200

*observation*: country

*country*: Canada

Usage
data(Consumption)

Format
A time serie containing:

*yd* personal disposable income, 1986 dollars

*ce* personal consumption expenditure, 1986 dollars

References

See Also
Description

a cross-section from 1998

number of observations : 11130

observation : individuals

country : United States

Usage

data(CPSch3)

Format

A dataframe containing :

year  survey year
ahe  average hourly earnings
sex  a factor with levels (male,female)

Source


References


See Also

Index.Source, Index.Economics, Index.Econometrics, Index.Observations
### Cracker

**Choice of Brand for Crackers**

### Description

- a cross-section
- *number of observations*: 3292
- *observation*: individuals
- *country*: United States

### Usage

```r
data(Cracker)
```

### Format

A dataframe containing:

- **id**: individuals identifiers
- **choice**: one of sunshine, kleeblter, nabisco, private
- **disp.z**: is there a display for brand z?
- **feat.z**: is there a newspaper feature advertisement for brand z?
- **price.z**: price of brand z

### Source


### References


### See Also

*Index.Source, Index.Economics, Index.Econometrics, Index.Observations*
CRANpackages

**Description**

Data casually collected on the number of packages on the Comprehensive R Archive Network (CRAN) at different dates.

NOTE: This could change in the future. See Details below.

**Usage**

```r
data(CRANpackages)
```

**Format**

A data.frame containing:

- **Version** an ordered factor of the R version number primarily in use at the time. This was taken from archives of the major releases at https://svn.r-project.org/R/branches/R-1-3-patches/tests/internet.Rout.save, ... https://svn.r-project.org/R/branches/R-3-1-branch/tests/internet.Rout.save
- **Date** an object of class Date giving the date on which the count of the number of CRAN packages was determined.
- **Packages** an integer number of packages on the CRAN mirror checked on the indicated Date.
- **Source** A factor giving the source (person) who collected the data.

**Details**

This seems to provide the most widely available source for data on the growth of CRAN, manually recorded by John Fox and Spencer Graves. For a discussion of these and related data, see Fox (2009).

For more detail, see the CRAN packages data on Github maintained by Hadley Wickham. This contains the description file of every package uploaded to CRAN prior to the date of Hadley’s most recent update. The current maintainer of the Ecdat and Ecfun packages would consider contributions along the following lines:

1. It might be nice to have a more complete dataset or datasets showing CRAN growth. This might include code fitting multiple models and predicting future growth with error bounds computed using Bayesian Model Averaging. These model fits might make an interesting addition to the examples in this help file. With a little more effort, it might make an interesting note for R Journal. Functions written to fit those models might be added to the Ecfun package.
2. It might be nice to have a function in Ecfun to download the CRAN packages data from Github and convert it to a format suitable for updating this dataset.

The current maintainer for Ecdat and Ecfun (Spencer Graves) might be willing to accept code and documentation for this but is not ready to do it himself at the present time.
Source


Examples

```r
plot(Packages~Date, CRANpackages, log='y')
# almost exponential growth
```

<table>
<thead>
<tr>
<th>Crime</th>
<th>Crime in North Carolina</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Description

a panel of 90 observations from 1981 to 1987

*number of observations*: 630

*observation*: regional

*country*: United States

Usage

data(Crime)

Format

A dataframe containing:

- **county** county identifier
- **year** year from 1981 to 1987
- **crmrte** crimes committed per person
- **prbarr** 'probability' of arrest
- **prbconv** 'probability' of conviction
- **prbpris** 'probability' of prison sentence
- **avgsen** average sentence, days
- **polpc** police per capita
- **density** people per square mile
- **taxpc** tax revenue per capita
- **region** one of 'other', 'west' or 'central'
- **smsa** 'yes' or 'no' if in SMSA
- **pctmin** percentage minority in 1980
- **wcon** weekly wage in construction
wtuc  weekly wage in trns, util, commun
wtrd  weekly wage in whole sales and retail trade
wfir  weekly wage in finance, insurance and real estate
wser  weekly wage in service industry
wmfg  weekly wage in manufacturing
wfed  weekly wage of federal employees
wsta  weekly wage of state employees
wloc  weekly wage of local governments employees
mix   offence mix: face-to-face/other
pctymle percentage of young males

Source


References


See Also


CRSPday

Daily Returns from the CRSP Database

Description

daily observations from 1969-1-03 to 1998-12-31
number of observations : 2528
observation : production units
country : United States

Usage

data(CRSPday)
Format

A dataframe containing:

- `year` the year
- `month` the month
- `day` the day
- `ge` the return for General Electric, Permno 12060
- `ibm` the return for IBM, Permno 12490
- `mobil` the return for Mobil Corporation, Permno 15966
- `crsp` the return for the CRSP value-weighted index, including dividends

Source


References


See Also

- `Index.Source`
- `Index.Economics`
- `Index.Econometrics`
- `Index.Observations`
- `Index.Time.Series`

---

**CRSPmon**

*Monthly Returns from the CRSP Database*

Description

- monthly observations from 1969-1 to 1998-12
- `number of observations` : 360
- `observation` : production units
- `country` : United States

Usage

- `data(CRSPmon)`
Format

A time serie containing:

- **ge** the return for General Electric, Permno 12060
- **ibm** the return for IBM, Permno 12490
- **mobil** the return for Mobil Corporation, Permno 15966
- **crsp** the return for the CRSP value-weighted index, including dividends

Source


References


See Also

Source


References


See Also

*Index.Source, Index.Economics, Index.Econometrics, Index.Observations*

---

<table>
<thead>
<tr>
<th>DM</th>
<th>DM Dollar Exchange Rate</th>
</tr>
</thead>
</table>

Description

weekly observations from 1975 to 1989

*number of observations*: 778

*observation*: country

*country*: Germany

Usage

data(DM)

Format

A dataframe containing:

- **date**: the date of the observation (19850104 is January, 4, 1985)
- **s**: the ask price of the dollar in units of DM in the spot market on friday of the current week
- **f**: the ask price of the dollar in units of DM in the 30-day forward market on friday of the current week
- **s30**: the bid price of the dollar in units of DM in the spot market on the delivery date on a current forward contract

Source

Bekaert, G. and R. Hodrick (1993) “On biases in the measurement of foreign exchange risk premi-


References

**Doctor**

**See Also**


---

### Description

- a cross-section from 1986
- number of observations: 485
- observation: individuals
- country: United States

### Usage

```r
data(Doctor)
```

### Format

A dataframe containing:

- `doctor` the number of doctor visits
- `children` the number of children in the household
- `access` is a measure of access to health care
- `health` a measure of health status (larger positive numbers are associated with poorer health)

### Source


### References


**See Also**

- `Index.Source`, `Index.Economics`, `Index.Econometrics`, `Index.Observations`
Description

a cross-section from 1977–1978

number of observations : 5190

observation : individuals

country : Australia

Usage

data(DoctorAUS)

Format

A dataframe containing :

sex  sex
age  age
income  annual income in tens of thousands of dollars
insurance  insurance contract (medlevy : mediban levy, levyplus : private health insurance, freepoor : government insurance due to low income, freerepa : government insurance due to old age disability or veteran status
illness  number of illness in past 2 weeks
actdays  number of days of reduced activity in past 2 weeks due to illness or injury
hscore  general health score using Goldberg’s method (from 0 to 12)
chcond  chronic condition (np : no problem, la : limiting activity, nla : not limiting activity)
doctorco  number of consultations with a doctor or specialist in the past 2 weeks
nondocco  number of consultations with non-doctor health professionals (chemist, optician, physiotherapist, social worker, district community nurse, chiropodist or chiropractor) in the past 2 weeks
hospadmi  number of admissions to a hospital, psychiatric hospital, nursing or convalescent home in the past 12 months (up to 5 or more admissions which is coded as 5)
hospdays  number of nights in a hospital, etc. during most recent admission: taken, where appropriate, as the mid-point of the intervals 1, 2, 3, 4, 5, 6, 7, 8-14, 15-30, 31-60, 61-79 with 80 or more admissions coded as 80. If no admission in past 12 months then equals zero.
medecine  total number of prescribed and nonprescribed medications used in past 2 days
prescrib  total number of prescribed medications used in past 2 days
nonprese  total number of nonprescribed medications used in past 2 days
Source


References


See Also

`Index.Source, Index.Economics, Index.Econometrics, Index.Observations`

---

**DoctorContacts**

### Description

A cross-section from 1977–1978

*number of observations*: 20186

### Usage

data(DoctorContacts)

### Format

A time serie containing:

- **mdu**: number of outpatient visits to a medical doctor
- **lc**: log(coinsurance+1) where coinsurance rate is 0 to 100
- **idp**: individual deductible plan?
- **lpi**: log(annual participation incentive payment) or 0 if no payment
- **fmde**: log(max(medical deductible expenditure)) if IDP=1 and MDE>1 or 0 otherwise
- **physlim**: physical limitation?
- **ndisease**: number of chronic diseases
- **health**: self-rate health (excellent, good, fair, poor)
- **line**: log of annual family income (in \$)
- **lfam**: log of family size
- **educdec**: years of schooling of household head
- **age**: exact age
- **sex**: sex (male, female)
- **child**: age less than 18?
- **black**: is household head black?
Source


References


See Also


---

**Earnings**

**Earnings for Three Age Groups**

**Description**

a cross-section from 1988-1989

*number of observations*: 4266

*observation*: individuals

*country*: United States

**Usage**

data(Earnings)

**Format**

A dataframe containing:

*age* age groups, a factor with levels (g1,g2,g3)

*y* average annual earnings, in 1982 US dollars

**Source**


**References**


**Electricity**

**See Also**

*Index.Source, Index.Economics, Index.Econometrics, Index.Observations*

---

**Electricity**

*Cost Function for Electricity Producers*

**Description**

- a cross-section from 1970 to 1970
  - number of observations: 158
  - observation: production units
  - country: United States

**Usage**

`data(Electricity)`

**Format**

A dataframe containing:

- **cost** total cost
- **q** total output
- **pl** wage rate
- **sl** cost share for labor
- **pk** capital price index
- **sk** cost share for capital
- **pf** fuel price
- **sf** cost share for fuel

**Source**


**References**


**See Also**

*Index.Source, Index.Economics, Index.Econometrics, Index.Observations*
**Fair**

*Extramarital Affairs Data*

**Description**

A cross-section

*number of observations*: 601

*observation*: individuals

*country*: United States

**Usage**

`data(Fair)`

**Format**

A dataframe containing:

- **sex**: a factor with levels (male,female)
- **age**: age
- **ym**: number of years married
- **child**: children ? a factor
- **religious**: how religious, from 1 (anti) to 5 (very)
- **education**: education
- **occupation**: occupation, from 1 to 7, according to hollingshead classification (reverse numbering)
- **rate**: self rating of marriage, from 1 (very unhappy) to 5 (very happy)
- **nbaffairs**: number of affairs in past year

**Source**


**References**


**See Also**

[Index.Source], [Index.Economics], [Index.Econometrics], [Index.Observations]
Fatality

Drunk Driving Laws and Traffic Deaths

**Description**

a panel of 48 observations from 1982 to 1988

- number of observations: 336
- observation: regional
- country: United States

**Usage**

data(Fatality)

**Format**

A dataframe containing:

- state  state ID code
- year  year
- mrall  traffic fatality rate (deaths per 10000)
- beertax  tax on case of beer
- mlda  minimum legal drinking age
- jaild  mandatory jail sentence?
- comserd  mandatory community service?
- vmiles  average miles per driver
- unrate  unemployment rate
- perinc  per capita personal income

**Source**

Pr. Christopher J. Ruhm, Department of Economics, University of North Carolina.

**References**


**See Also**

Description

FinancialCrisisFiles is an object of class financialCrisisFiles created by the financialCrisisFiles function in Ecfun. It describes files containing data on financial crises downloadable from http://www.reinhartandrogoff.com/data/browse-by-topic/topics/7/.

Usage

data(FinancialCrisisFiles)

Details

Reinhart and Rogoff (http://www.reinhartandrogoff.com) provide numerous data sets analyzed in their book, "This Time Is Different: Eight Centuries of Financial Folly". Of interest here are data on financial crises of various types for 70 countries spanning the years 1800 - 2010, downloadable from http://www.reinhartandrogoff.com/data/browse-by-topic/topics/7/.

The function financialCrisisFiles in Ecfun produces a list of class financialCrisisFiles describing four different Excel files in very similar formats with one sheet per Country and a few extra descriptor sheets. The data object FinancialCrisisFiles is the default output of that function.

Value

FinancialCrisisFiles is a list with components carrying the names of files to be read. Each component is a list of optional arguments to pass to do.call(read.xls, ...) to read the sheet with name = name of that component.

This corresponds to the files downloaded from http://www.reinhartandrogoff.com/data/browse-by-topic/topics/7/ in January 2013 (except for the fourth, which was not available there because of an error with the web site but instead was obtained directly from Prof. Reinhart).

Author(s)

Spencer Graves

Source

http://www.reinhartandrogoff.com

References

See Also

read.xls

---

**Fishing**  
*Choice of Fishing Mode*

**Description**

A cross-section

*number of observations*: 1182

*observation*: individuals

*country*: United States

**Usage**

data(Fishing)

**Format**

A dataframe containing:

- **mode**: recreation mode choice, on of: beach, pier, boat and charter
- **price**: price for chosen alternative
- **catch**: catch rate for chosen alternative
- **pbeach**: price for beach mode
- **ppier**: price for pier mode
- **pboat**: price for private boat mode
- **pcharter**: price for charter boat mode
- **cbeach**: catch rate for beach mode
- **cpier**: catch rate for pier mode
- **cboat**: catch rate for private boat mode
- **ccharter**: catch rate for charter boat mode
- **income**: monthly income

**Source**


**References**

**Forward**

**Exchange Rates of US Dollar Against Other Currencies**

**Description**

monthly observations from 1979–01 to 2001–12

*number of observations*: 276

**Usage**

data(Forward)

**Format**

A time serie containing:

- **usdp** exchange rate USD/British Pound Sterling
- **usdeuro** exchange rate US D/Euro
- **eurobp** exchange rate Euro/Pound
- **usdp1** 1 month forward rate USD/Pound
- **usdeuro1** 1 month forward rate USD/Euro
- **eurobp1** 1 month forward rate Euro/Pound
- **usdp3** 3 month forward rate USD/Pound
- **usdeuro3** month forward rate USD/Euro
- **eurobp3** month forward rate Euro/Pound

**Source**

Datastream.

**References**


**See Also**

Description

a cross-section from 2002–03
number of observations : 227
observation : individuals
country : United States

Usage

data(FriendFoe)

Format

A dataframe containing :

sex  contestant’s sex
white is contestant white ?
age  contestant’s age in years
play contestant’s choice : a factor with levels "foe" and "friend". If both players play "friend", they share the trust box, if both play "foe", both players receive zero prize, if one of them play "foe" and the other one "friend", the "foe" player receive the entire trust bix and the "friend" player nothing
round  round in which constestant is eliminated, a factor with levels ("1","2","3")
season season show, a factor with levels ("1","2")
cash the amount of cash in the trust box
sex1 partner’s sex
white1 is partner white ?
age1 partner’s age in years
play1 partner’s choice : a factor with levels "foe" and "friend"
win money won by contestant
win1 money won by partner

Source


References

See Also

Index.Source, Index.Economics, Index.Econometrics, Index.Observations

Garch

Description

daily observations from 1980–01 to 1987–05–21

number of observations : 1867

observation : country
country : World

Usage

data(Garch)

Format

A dataframe containing:

date date of observation (yymmd)
day day of the week (a factor)
dm exchange rate Dollar/Deutsch Mark
ddm dm-dm(-1)
bp exchange rate of Dollar/British Pound
cd exchange rate of Dollar/Canadian Dollar
dy exchange rate of Dollar/Yen
sf exchange rate of Dollar/Swiss Franc

References


See Also

Index.Source, Index.Economics, Index.Econometrics, Index.Observations,
Index.Time.Series
<table>
<thead>
<tr>
<th>Gasoline</th>
<th>Gasoline Consumption</th>
</tr>
</thead>
</table>

**Description**

a panel of 18 observations from 1960 to 1978

- *number of observations*: 342
- *observation*: country
- *country*: OECD

**Usage**

data(Gasoline)

**Format**

A dataframe containing:

- **country**: a factor with 18 levels
- **year**: the year
- **lgaspcar**: logarithm of motor gasoline consumption per auto
- **lincomep**: logarithm of real per-capita income
- **lrpmg**: logarithm of real motor gasoline price
- **lcarpcap**: logarithm of the stock of cars per capita

**Source**


**References**


**See Also**

Griliches  Wage Data

Description

a cross-section from 1980

*number of observations*: 758

*observation*: individuals

*country*: United States

Usage

data(Griliches)

Format

A dataframe containing:

- **rns**: residency in the southern states (first observation)?
- **rns80**: same variable for 1980
- **mrt**: married (first observation)?
- **mrt80**: same variable for 1980
- **smsa**: residency in metropolitan areas (first observation)?
- **smsa80**: same variable for 1980
- **med**: mother's education in years
- **iq**: IQ score
- **kww**: score on the “knowledge of the world of work” test
- **year**: year of the observation
- **age**: age (first observation)
- **age80**: same variable for 1980
- **school**: completed years of schooling (first observation)
- **school80**: same variable for 1980
- **expr**: experience in years (first observation)
- **expr80**: same variable for 1980
- **tenure**: tenure in years (first observation)
- **tenure80**: same variable for 1980
- **lw**: log wage (first observation)
- **lw80**: same variable for 1980
Source


References


See Also

Index.Source, Index.Economics, Index.Econometrics, Index.Observations

Grunfeld Investment Data

Description

a panel of 10 observations from 1935 to 1954

number of observations : 200
observation : production units
country : United States

Usage

data(Grunfeld)

Format

A dataframe containing :

firm observation
year date
inv gross Investment
value value of the firm
capital stock of plant and equipment

Source


References

**HC**

*Heating and Cooling System Choice in Newly Built Houses in California*

**Description**

A cross-section

- number of observations: 250
- observation: households
- country: California

**Usage**

`data(HC)`

**Format**

A dataframe containing:

- **depvar** heating system, one of gcc (gas central heat with cooling), ecc (electric central resistance heat with cooling), erc (electric room resistance heat with cooling), hpc (electric heat pump which provides cooling also), gc (gas central heat without cooling), ec (electric central resistance heat without cooling), er (electric room resistance heat without cooling)
- **ich.z** installation cost of the heating portion of the system
- **icca** installation cost for cooling
- **och.z** operating cost for the heating portion of the system
- **occa** operating cost for cooling
- **income** annual income of the household

**References**

Kenneth Train’s home page: [http://elsa.berkeley.edu/~train/](http://elsa.berkeley.edu/~train/).

**See Also**

The Boston HDMA Data Set

Description

a cross-section from 1997-1998

number of observations: 2381

observation: individuals

country: United States

Usage

data(Hdma)

Format

A dataframe containing:

- dir  debt payments to total income ratio
- hir  housing expenses to income ratio
- lvr  ratio of size of loan to assessed value of property
- ccs  consumer credit score from 1 to 6 (a low value being a good score)
- mcs  mortgage credit score from 1 to 4 (a low value being a good score)
- pbr  public bad credit record?
- dmi  denied mortgage insurance?
- self  self employed?
- single is the applicant single?
- uria 1989 Massachusetts unemployment rate in the applicant’s industry
- condominium is unit condominium?
- black  is the applicant black?
- deny mortgage application denied?

Source

Federal Reserve Bank of Boston.


References


See Also

Index.Source, Index.Economics, Index.Econometrics, Index.Observations
Heating System Choice in California Houses

Description

a cross-section

*number of observations*: 900

*observation*: households

*country*: California

Usage

data(Heating)

Format

A dataframe containing:

- **idcase**: id
- **depvar**: heating system, one of gc (gas central), gr (gas room), ec (electric central), er (electric room), hp (heat pump)
- **ic.z**: installation cost for heating system z (defined for the 5 heating systems)
- **oc.z**: annual operating cost for heating system z (defined for the 5 heating systems)
- **pb.z**: ratio oc.z/ic.z
- **income**: annual income of the household
- **agehed**: age of the household head
- **rooms**: numbers of rooms in the house

References

Kenneth Train’s home page: [http://elsa.berkeley.edu/~train/](http://elsa.berkeley.edu/~train/).

See Also

*Index.Source, Index.Economics, Index.Econometrics, Index.Observations*
**Hedonic Prices of Census Tracts in Boston**

### Description

a cross-section

*number of observations*: 506

*observation*: regional

*country*: United States

### Usage

data(Hedonic)

### Format

A dataframe containing:

- **mv**: median value of owner-occupied homes
- **crim**: crime rate
- **zn**: proportion of 25,000 square feet residential lots
- **indus**: proportion of nonretail business acres
- **chas**: is the tract bounds the Charles River?
- **nox**: annual average nitrogen oxide concentration in parts per hundred million
- **rm**: average number of rooms
- **age**: proportion of owner units built prior to 1940
- **dis**: weighted distances to five employment centers in the Boston area
- **rad**: index of accessibility to radial highways
- **tax**: full value property tax rate ($/\$10,000)
- **ptratio**: pupil/teacher ratio
- **blacks**: proportion of blacks in the population
- **lstat**: proportion of population that is lower status
- **townid**: town identifier

### Source


References


See Also

*Index.Source, Index.Economics, Index.Econometrics, Index.Observations*

---

**HI**  
*Health Insurance and Hours Worked By Wives*

### Description

a cross-section from 1993  
*number of observations*: 22272  
*observation*: individuals  
*country*: United States

### Usage

data(HI)

### Format

A dataframe containing:

- `whrswk` hours worked per week by wife
- `hhi` wife covered by husband’s HI?
- `whi` wife has HI thru her job?
- `hhi2` husband has HI thru own job?
- `education` a factor with levels, "&lt;9years", "9-11years", "12years", "13-15years", "16years", ">16years"
- `race` one of white, black, other
- `hispanic` hispanic?
- `experience` years of potential work experience
- `kidslt6` number of kids under age of 6
- `kids618` number of kids 6–18 years old
- `husby` husband’s income in thousands of dollars
- `region` one of other, northcentral, south, west
- `wght` sampling weight
Source


References


See Also

`Index.Source, Index.Economics, Index.Econometrics, Index.Observations`

<table>
<thead>
<tr>
<th>Housing</th>
<th>Sales Prices of Houses in the City of Windsor</th>
</tr>
</thead>
</table>

Description

a cross-section from 1987  
*number of observations*: 546  
*observation*: goods  
*country*: Canada

Usage

data(Housing)

Format

A dataframe containing:

- **price**: sale price of a house  
- **lotsize**: the lot size of a property in square feet  
- **bedrooms**: number of bedrooms  
- **bathrms**: number of full bathrooms  
- **stories**: number of stories excluding basement  
- **driveway**: does the house has a driveway?  
- **recroom**: does the house has a recreational room?  
- **fullbase**: does the house has a full finished basement?  
- **gashw**: does the house uses gas for hot water heating?  
- **airco**: does the house has central air conditioning?  
- **garagepl**: number of garage places  
- **prefarea**: is the house located in the preferred neighbourhood of the city?
**Source**


**References**


**See Also**

`Index.Source, Index.Economics, Index.Econometrics, Index.Observations`

---

**Hstarts**

*Housing Starts*

**Description**

quarterly observations from 1960-1 to 2001-4

*number of observations*: 168

*observation*: country

*country*: Canada

**Usage**

`data(Hstarts)`

**Format**

A time serie containing:

- `hs` the log of urban housing starts in Canada, not seasonally adjusted, CANSIM series J6001, converted to quarterly

- `hssa` the log of urban housing starts in Canada, seasonally adjusted, CANSIM series J9001, converted to quarterly. Observations prior to 1966:1 are missing

**References**


**See Also**

**Icecream**

<table>
<thead>
<tr>
<th>Icecream</th>
<th>Ice Cream Consumption</th>
</tr>
</thead>
</table>

**Description**

four–weekly observations from 1951–03–18 to 1953–07–11

*number of observations*: 30

*observation*: country

*country*: United States

**Usage**

data(Icecream)

**Format**

A time serie containing:

**cons** consumption of ice cream per head (in pints);

**income** average family income per week (in US Dollars);

**price** price of ice cream (per pint);

**temp** average temperature (in Fahrenheit);

**Source**


**References**


**See Also**

Description

Data on quantiles of the distributions of family incomes in the United States. This combines three data sources:

1. US Census Table F-1 for the central quantiles
2. Piketty and Saez for the 95th and higher quantiles
3. Gross Domestic Product and implicit price deflators from MeasuringWorth.com

Usage
data(incomeInequality)

Format

A data.frame containing:

- **Year** numeric year 1947:2012
- **Number.thousands** number of families in the US
- **quintile1, quintile2, median, quintile3, quintile4, p95** quintile1, quintile2, quintile3, quintile4, and p95 are the indicated quantiles of the distribution of family income from US Census Table F-1. The media is computed as the geometric mean of quintile2 and quintile3. This is accurate to the extent that the lognormal distribution adequately approximates the central 20 percent of the income distribution, which it should for most practical purposes.

**P90, P95, P99, P99.5, P99.9, P99.99** The indicated quantiles of family income per Piketty and Saez


**P95IRSvsCensus** ratio of the estimates of the 95th percentile of distributions of family income from the Piketty and Saez analysis of data from the Internal Revenue Service (IRS) and from the US Census Bureau. The IRS has ranged between 72 and 98 percent of the Census Bureau figures for the 95th percentile of the distribution, with this ratio averaging around 75 percent since the late 1980s. However, this systematic bias is modest relative to the differences between the different quantiles of interest in this combined dataset.

**personsPerFamily** average number of persons per family using the number of families from US Census Table F-1 and the population from MeasuringWorth.com.

**realGDPperFamily** personsPerFamily * realGDPperCap

**mean.median** ratio of realGDPperFamily to the median. This is a measure of skewness and income inequality.
Details
For details on how this data frame was created, see "F1.PikettySaez.R" in system.file('scripts', package='fda'). This provides links for files to download and R commands to read those files and convert them into an updated version of incomeInequality. This is a reasonable thing to do if it is more than 2 years since max(incomeInequality$year). All data are in constant 2012 dollars.

Author(s)
Spencer Graves

Source


Examples
##
## Rato of IRS to census estimates for the 95th percentile
##
data(incomeInequality)
plot(P95IRSvsCensus~Year, incomeInequality, type='b')
# starts ~0.74, trends rapidly up to ~0.97,
# then drifts back to ~0.75
abline(h=0.75)
abline(v=1989)
# check
sum(is.na(incomeInequality$P95IRSvsCensus))
# The Census data runs to 2011; Pikety and Saez runs to 2010.
quantile(incomeInequality$P95IRSvsCensus, na.rm=TRUE)
# 0.72 ... 0.98
##
## Persons per Family
##
plot(personsPerFamily~Year, incomeInequality, type='b')
quantile(incomeInequality$personsPerFamily)
# ranges from 3.72 to 4.01 with median 3.84
# -- almost 4
##
## GDP per family
##
plot(realGDPperFamily~Year, incomeInequality, type='b', log='y')

## Plot the mean then the first quintile, then the median,
## 99th, 99.9th and 99.99th percentiles
##
## plotCols <- c(21, 3, 5, 11, 13:14)
kcols <- length(plotCols)
plotColors <- c(1:6, 8:13)[1:kcols] # omit 7=yellow
plotLty <- 1:kcols

matplot(incomeInequality$Year, incomeInequality[plotCols]/1000,
        log='y', type='l', col=plotColors, lty=plotLty)

### Growth broadly shared 1947 - 1970, then began diverging
### The divergence has been most pronounced among the top 1%
### and especially the top 0.01%

## Growth rate by quantile 1947-1970 and 1970 - present
##
## keyYears <- c(1947, 1970, 2010)
## (iYears <- which(is.element(incomeInequality$Year, keyYears)))

(dyYears <- diff(keyYears))
k <- length(keyYears)
(lblYrs <- paste(keyYears[-k], keyYears[-1], sep='-', collapse=''))

(growth <- sapply(incomeInequality[iYears,], function(x, labels=lblYrs){
  dx <- exp(diff(log(x)))
  names(dx) <- labels
  dx
})(iYears))

# as percent
(gr <- round(100*(growth-1), 1))

# The average annual income (realGDPperFamily) doubled between
# 1970 and 2010 (increased by 101 percent), while the median household
# income increased only 23 percent.

## Income lost by each quantile 1970-2010
## relative to the broadly shared growth 1947-1970
##
## (lostGrowth <- (growth[, 'realGDPperFamily']-growth[, plotCols]))
## # 1947-1970: The median gained 20% relative to the mean,
## # while the top 1% lost ground
## # 1970-2010: The median lost 79%, the 99th percentile lost 29%,
## # while the top 0.1% gained

(lostIncome <- (lostGrowth[2,] *
               incomeInequality[iYears[2], plotCols]))
The median family lost $39,000 per year in income relative to what they would have with the same economic growth broadly shared as during 1947–1970. That's slightly over $36,500 per year = $100 per day.

\[
gryr \leftarrow \text{growth}^{(1/\text{Years})}
\]
\[
gryr. \leftarrow \text{round}(100 \times (gryr-1), 1)
\]

## Regression line: linear spline

\[
\text{varyg} \leftarrow c(3:14, 21)
\]
\[
\text{Varyg} \leftarrow \text{names(incomeInequality)[varyg]}
\]
\[
\text{str(F01ps} \leftarrow \text{reshape(incomeInequality}[c(1, varyg)], idvar='Year',
  \text{id=F1.PikettySeaz$Year,}
  \text{time=Varyg, timevar='ptile',}
  \text{varying=list(Varyg), direction='long'})
\]
\[
\text{names(F01ps}[2:3] \leftarrow c('variable', 'value')
\]
\[
\text{F01ps$variable} \leftarrow \text{factor(F01ps$variable)
}\]

# linear spline basis function with knot at 1970
\[
\text{F01ps$1970p} \leftarrow \text{pmax}(0, \text{F01ps$Year-1970})
\]
\[
\text{table(nas} \leftarrow \text{is.na(F01ps$value))
\]

# 6 NAs, one each of the Piketty-Saez variables in 2011
\[
\text{F01i} \leftarrow \text{F01ps}[!nas, ]
\]

# formula:
# \( \log(\text{value}/1000) \sim b \times \text{Year} + \) (for each variable:
#   different intercept + (different slope after 1970))

\[
\text{Fit} \leftarrow \text{lm(\log(\text{value}/1000)} \sim \text{Year+variable+1970p, F01i})
\]
\[
\text{anova(Fit)
}\]

# all highly significant
# The residuals may show problems with the model,
# but we will ignore those for now.

# Model predictions
\[
\text{str(Pred} \leftarrow \text{predict(Fit))
\]

## Combined plot

# Plot to a file? Wikimedia Commons prefers svg format.
\[
\text{svg('incomeInequality8.svg')}
\]

# If you want software to convert svg to another format such as png,
# consider GIMP (www.gimp.org).

# Base plot

# Leave extra space on the right to label with growth since 1970
\[
\text{op} \leftarrow \text{par(mar=c(5, 4, 4, 5)*0.1)}
\]
```r
matplot(incomeInequality$Year, incomeInequality[plotCols]/1000, 
  log="y", type='l', col=plotColors, lty=plotLty, 
  xlab='', ylab='', las=1, axes=FALSE, lwd=2)
axis(1, at=seq(1950, 2010, 10), 
  labels=c(1950, NA, 1970, NA, 1990, NA, 2010), cex.axis=1.5)
yat <- c(10, 50, 100, 500, 1000, 5000, 10000)
axis(2, yat, labels=c("$10K", "$50K", "$100K", "$500K", 
  "$1M", "$5M", "$10M"), las=1, cex.axis=1.2)

# Label the lines
pctls <- paste(c(20, 40, 50, 60, 80, 90, 95, 99, 99.5, 99.9, 99.99), 
  
  '%', sep='')
lineLb10 <- c('Year', 'families K', 'pctls', 
  'realGDP.M', 'GDP deflator', 'pop-K', 'realGDPperfFamily', 
  '95 pct(IRS / Census)', 'size of household', 
  'average family income', 'mean/median')

(lineLbs <- lineLb10[plotCols])
setl75 <- (incomeInequality$Year==1975)
laby <- incomeInequality[setl75, plotCols]/1000

text(1973.5, c(1.2, 1.2, 1.3, 1.5, 1.9)*laby[-1], lineLbs[-1], cex=1.2)
text(1973.5, 1.2*laby[1], lineLbs[1], cex=1.2, srt=10)

##
## Add lines + points for the knots in 1970
##
## End <- numeric(kcols)
F01names <- names(incomeInequality)
for(i in seq(length=kcols)){
  seli <- as.character(F01$variable) == F01names[plotCols[i]]
  # with(F01[seli, ], lines(Year, exp(Pred[seli]), col=plotColors[i]))
  yri <- F01$Year[seli]
  predi <- exp(Pred[seli])
  lines(yri, predi, col=plotColors[i])
  End[i] <- predi[length(predi)]
  setl70i <- (yri==1970)
  points(yri[setl70i], predi[setl70i], col=plotColors[i])
}

##
## label growth rates
##
table(setl70. <- (incomeInequality$Year>1969))
(lastYrs <- incomeInequality[setl70., 'Year'])
(lastYr. <- max(lastYrs)+4)
#text(lastYr., End, gR., xpd=NA)
text(lastYr., End, paste(gr[2, plotCols], '%', sep=''), xpd=NA)
text(lastYr.+7, End, paste(grYr.[2, plotCols], '%', sep=''), xpd=NA)

##
## Label the presidents
```
Load data and plot:

```r
data(IncomeUK)
(m99.95 <- with(IncomeInequality, sqrt(P99.9*P99.99)/1000))
```

Plot:

```r
(text(1949, 5000, 'Truman')
text(1956.8, 5000, 'Eisenhower', srt=90)
text(1963, 5000, 'Kennedy', srt=90)
text(1966.8, 5000, 'Johnson', srt=90)
text(1971, 5*m99.95[24], 'Nixon', srt=90)
text(1975, 5*m99.95[28], 'Ford', srt=90)
text(1978.5, 5*m99.95[32], 'Carter', srt=90)
text(1985.1, m99.95[38], 'Reagan')
text(1991, 0.94*m99.95[44], 'GHW Bush', srt=90)
text(1997, m99.95[50], 'Clinton')
text(2005.1, 1.1*m99.95[58], 'GW Bush', srt=90)
text(2010, 1.2*m99.95[62], 'Obama', srt=90)
```

Set margins and dev.off():

```r
par(op) # reset margins
dev.off() # for plot to a file
```

**IncomeUK**

*Seasonally Unadjusted Quarterly Data on Disposable Income and Expenditure*

**Description**

Quarterly observations from 1971–1 to 1985–2

- **number of observations**: 58
- **observation**: country
- **country**: United Kingdom

**Usage**

`data(IncomeUK)`

**Format**

A time serie containing:

- **income**: total disposable income (million Pounds, current prices)
- **consumption**: consumer expenditure (million Pounds, current prices)

**References**

See Also


\underline{Index.Econometrics \quad Econometric fields}

\begin{description}
\item[b] binomial model
  \begin{itemize}
  \item \texttt{Benefits}: Unemployment of Blue Collar Workers
  \item \texttt{Hdma}: The Boston HDMA Data Set
  \item \texttt{Mroz}: Labor Supply Data
  \item \texttt{Participation}: Labor Force Participation
  \item \texttt{Train}: Stated Preferences for Train Traveling
  \end{itemize}
\item[c] censored and truncated model
  \begin{itemize}
  \item \texttt{Fair}: Extramarital Affairs Data
  \item \texttt{HI}: Health Insurance and Hours Worked By Wives
  \item \texttt{Mofa}: International Expansion of U.S. Mofa’s (majority–owned Foreign Affiliates in
  Fire (finance, Insurance and Real Estate)
  \item \texttt{Tobacco}: Households Tobacco Budget Share
  \item \texttt{Workinghours}: Wife Working Hours
  \end{itemize}
\item[count d] count data
  \begin{itemize}
  \item \texttt{Accident}: Ship Accidents
  \item \texttt{Bids}: Bids Received By U.S. Firms
  \item \texttt{Doctor}: Number of Doctor Visits
  \item \texttt{DoctorAUS}: Doctor Visits in Australia
  \item \texttt{DoctorContacts}: Contacts With Medical Doctor
  \item \texttt{OFP}: Visits to Physician Office
  \item \texttt{PatentsHGH}: Dynamic Relation Between Patents and R\&D
  \item \texttt{PatentsRD}: Patents, R\&D and Technological Spillovers for a Panel of Firms
  \item \texttt{Somerville}: Visits to Lake Somerville
  \item \texttt{StrikeNb}: Number of Strikes in Us Manufacturing
  \end{itemize}
\item[d] duration model
  \begin{itemize}
  \item \texttt{Oil}: Oil Investment
  \item \texttt{Strike}: Strike Duration Data
  \item \texttt{StrikeDur}: Strikes Duration
  \item \texttt{UnempDur}: Unemployment Duration
  \item \texttt{Unemployment}: Unemployment Duration
  \end{itemize}
\item[m] multinomial model
  \begin{itemize}
  \item \texttt{Car}: Stated Preferences for Car Choice
  \end{itemize}
\end{description}
- **Catsup**: Choice of Brand for Catsup
- **Cracker**: Choice of Brand for Crackers
- **Fishing**: Choice of Fishing Mode
- **HC**: Heating and Cooling System Choice in Newly Built Houses in California
- **Heating**: Heating System Choice in California Houses
- **Ketchup**: Choice of Brand for Ketchup
- **Mode**: Mode Choice
- **ModeChoice**: Data to Study Travel Mode Choice
- **Tuna**: Choice of Brand for Tuna
- **Yogurt**: Choice of Brand for Yogurts

**• ordered model**
- **Kakadu**: Willingness to Pay for the Preservation of the Kakadu National Park
- **Mathlevel**: Level of Calculus Attained for Students Taking Advanced Micro–economics
- **NaturalPark**: Willingness to Pay for the Preservation of the Alentejo Natural Park

**• panel**
- **Airline**: Cost for U.S. Airlines
- **Cigar**: Cigarette Consumption
- **Cigarette**: The Cigarette Consumption Panel Data Set
- **Crime**: Crime in North Carolina
- **Fatality**: Drunk Driving Laws and Traffic Deaths
- **Gasoline**: Gasoline Consumption
- **Grunfeld**: Grunfeld Investment Data
- **LaborSupply**: Wages and Hours Worked
- **Males**: Wages and Education of Young Males
- **MunExp**: Municipal Expenditure Data
- **Produc**: Us States Production
- **SumHes**: The Penn Table
- **Wages**: Panel Datas of Individual Wages

**• system of equations**
- **BudgetItaly**: Budget Shares for Italian Households
- **BudgetUK**: Budget Shares of British Households
- **Electricity**: Cost Function for Electricity Producers
- **Klein**: Klein's Model I
- **ManufCost**: Manufacturing Costs
- **Nerlove**: Cost Function for Electricity Producers, 1955
- **University**: Provision of University Teaching and Research

**• time–series**
- **CRSPday**: Daily Returns from the CRSP Database
- **CRSPmon**: Monthly Returns from the CRSP Database
- **Capm**: Stock Market Data
- **Consumption**: Quarterly Data on Consumption and Expenditure
- **DM**: DM Dollar Exchange Rate
- **Forward**: Exchange Rates of US Dollar Against Other Currencies
- **Garch**: Daily Observations on Exchange Rates of the US Dollar Against Other Currencies
- **Hstarts**: Housing Starts
- **Icecream**: Ice Cream Consumption
- **IncomeUK**: Seasonally Unadjusted Quarterly Data on Disposable Income and Expenditure
- **Irates**: Monthly Interest Rates
- **LT**: Dollar Sterling Exchange Rate
- **MW**: Growth of Disposable Income and Treasury Bill Rate
- **Macrodat**: Macroeconomic Time Series for the United States
- **Mishkin**: Inflation and Interest Rates
- **MoneyUS**: Macroeconomic Series for the United States
- **Mpyr**: Money, National Product and Interest Rate
- **Orange**: The Orange Juice Data Set
- **PE**: Price and Earnings Index
- **PPP**: Exchange Rates and Price Indices for France and Italy
- **Pound**: Pound-dollar Exchange Rate
- **Pricing**: Returns of Size-based Portfolios
- **Solow**: Solow’s Technological Change Data
- **Tbrate**: Interest Rate, GDP and Inflation
- **Yen**: Yen-dollar Exchange Rate

### Economic fields

---

**Index.Economics**

**Economic fields**

---

**Description**

- consumer behavior
  - **BudgetFood**: Budget Share of Food for Spanish Households
  - **BudgetItaly**: Budget Shares for Italian Households
  - **BudgetUK**: Budget Shares of British Households
  - **Car**: Stated Preferences for Car Choice
  - **Cigar**: Cigarette Consumption
  - **Cigarette**: The Cigarette Consumption Panel Data Set
  - **Doctor**: Number of Doctor Visits
  - **Fishing**: Choice of Fishing Mode
  - **Gasoline**: Gasoline Consumption
  - **HC**: Heating and Cooling System Choice in Newly Built Houses in California
  - **Heating**: Heating System Choice in California Houses
  - **Icecream**: Ice Cream Consumption
• **Mode**: Mode Choice
  
  • **ModeChoice**: Data to Study Travel Mode Choice
  
  • **Somerville**: Visits to Lake Somerville
  
  • **Tobacco**: Households Tobacco Budget Share
  
  • **Train**: Stated Preferences for Train Traveling

• **economics of education**
  
  • **Caschool**: The California Test Score Data Set
  
  • **MCAS**: The Massachusetts Test Score Data Set
  
  • **Mathlevel**: Level of Calculus Attained for Students Taking Advanced Micro-economics
  
  • **Star**: Effects on Learning of Small Class Sizes

• **environmental economics**
  
  • **Airq**: Air Quality for Californian Metropolitan Areas
  
  • **Kakadu**: Willingness to Pay for the Preservation of the Kakadu National Park
  
  • **NaturalPark**: Willingness to Pay for the Preservation of the Alentejo Natural Park

• **finance**
  
  • **CRSPday**: Daily Returns from the CRSP Database
  
  • **CRSPmon**: Monthly Returns from the CRSP Database
  
  • **Capm**: Stock Market Data
  
  • **DM**: DM Dollar Exchange Rate
  
  • **Forward**: Exchange Rates of US Dollar Against Other Currencies
  
  • **Garch**: Daily Observations on Exchange Rates of the US Dollar Against Other Currencies
  
  • **Irates**: Monthly Interest Rates
  
  • **LT**: Dollar Sterling Exchange Rate
  
  • **PPP**: Exchange Rates and Price Indices for France and Italy
  
  • **Pound**: Pound-dollar Exchange Rate
  
  • **Pricing**: Returns of Size-based Portfolios
  
  • **Yen**: Yen-dollar Exchange Rate

• **game theory**
  
  • **FriendFoe**: Data from the Television Game Show Friend Or Foe?

• **health economics**
  
  • **DoctorAUS**: Doctor Visits in Australia
  
  • **DoctorContacts**: Contacts With Medical Doctor
  
  • **MedExp**: Structure of Demand for Medical Care
  
  • **OFP**: Visits to Physician Office
  
  • **VietNamH**: Medical Expenses in Viet-nam (household Level)
  
  • **VietNamI**: Medical Expenses in Viet-nam (individual Level)

• **hedonic prices**
  
  • **Computers**: Prices of Personal Computers
  
  • **Diamond**: Pricing the C’s of Diamond Stones
  
  • **Hedonic**: Hedonic Prices of Census Tracts in Boston
- **Housing**: Sales Prices of Houses in the City of Windsor
- **Journals**: Economic Journals Dat Set

- **labor economics**
  - **Benefits**: Unemployment of Blue Collar Workers
  - **Bwages**: Wages in Belgium
  - **CPSch3**: Earnings from the Current Population Survey
  - **Earnings**: Earnings for Three Age Groups
  - **Griliches**: Wage Data
  - **HI**: Health Insurance and Hours Worked By Wives
  - **LaborSupply**: Wages and Hours Worked
  - **Labour**: Belgian Firms
  - **Males**: Wages and Education of Young Males
  - **Mroz**: Labor Supply Data
  - **PSID**: Panel Survey of Income Dynamics
  - **Participation**: Labor Force Participation
  - **RetSchool**: Return to Schooling
  - **Schooling**: Wages and Schooling
  - **Strike**: Strike Duration Data
  - **StrikeDur**: Strikes Duration
  - **StrikeNb**: Number of Strikes in Us Manufacturing
  - **Treatment**: Evaluating Treatment Effect of Training on Earnings
  - **UnempDur**: Unemployment Duration
  - **Unemployment**: Unemployment Duration
  - **Wages**: Panel Datas of Individual Wages
  - **Wages1**: Wages, Experience and Schooling
  - **Workinghours**: Wife Working Hours

- **macroeconomics**
  - **Consumption**: Quarterly Data on Consumption and Expenditure
  - **Hstarts**: Housing Starts
  - **IncomeUK**: Seasonally Unadjusted Quarterly Data on Disposable Income and Expenditure
  - **Klein**: Klein’s Model I
  - **Longley**: The Longley Data
  - **MW**: Growth of Disposable Income and Treasury Bill Rate
  - **Macrodat**: Macroeconomic Time Series for the United States
  - **Mishkin**: Inflation and Interest Rates
  - **Money**: Money, GDP and Interest Rate in Canada
  - **MoneyUS**: Macroeconomic Series for the United States
  - **Mpyr**: Money, National Product and Interest Rate
  - **PE**: Price and Earnings Index
  - **Produc**: Us States Production
  - **Solow**: Solow’s Technological Change Data
Index.Observations

• SumHes: The Penn Table
• Tbrate: Interest Rate, GDP and Inflation

• marketing
  – Catsup: Choice of Brand for Catsup
  – Cracker: Choice of Brand for Crakers
  – Ketchup: Choice of Brand for Ketchup
  – Tuna: Choice of Brand for Tuna
  – Yogurt: Choice of Brand for Yogurts

• producer behavior
  – Accident: Ship Accidents
  – Airline: Cost for U.S. Airlines
  – Bids: Bids Received By U.S. Firms
  – Clothing: Sales Data of Men’s Fashion Stores
  – Electricity: Cost Function for Electricity Producers
  – Grunfeld: Grunfeld Investment Data
  – Hdma: The Boston HDMA Data Set
  – ManufCost: Manufacturing Costs
  – Metal: Production for SIC 33
  – Mofa: International Expansion of U.S. Mofa’s (majority–owned Foreign Affiliates in Fire (finance, Insurance and Real Estate)
  – Nerlove: Cost Function for Electricity Producers, 1955
  – Oil: Oil Investment
  – Orange: The Orange Juice Data Set
  – PatentsHGH: Dynamic Relation Between Patents and R&D
  – PatentsRD: Patents, R&D and Technological Spillovers for a Panel of Firms
  – TranspEq: Statewide Data on Transportation Equipment Manufacturing
  – University: Provision of University Teaching and Research

• socio–economics
  – Crime: Crime in North Carolina
  – Fair: Extramarital Affairs Data
  – Fatality: Drunk Driving Laws and Traffic Deaths

<table>
<thead>
<tr>
<th>Index.Observations</th>
<th>Observations</th>
</tr>
</thead>
</table>

Description

• country
  – Consumption: Quarterly Data on Consumption and Expenditure
  – DM: DM Dollar Exchange Rate
  – Garch: Daily Observations on Exchange Rates of the US Dollar Against Other Currencies
- **Gasoline**: Gasoline Consumption  
- **Hstarts**: Housing Starts  
- **Icecream**: Ice Cream Consumption  
- **IncomeUK**: Seasonally Unadjusted Quarterly Data on Disposable Income and Expenditure  
- **Irates**: Monthly Interest Rates  
- **Klein**: Klein’s Model I  
- **LT**: Dollar Sterling Exchange Rate  
- **Longley**: The Longley Data  
- **MW**: Growth of Disposable Income and Treasury Bill Rate  
- **Macrodat**: Macroeconomic Time Series for the United States  
- **ManufCost**: Manufacturing Costs  
- **Mishkin**: Inflation and Interest Rates  
- **Mofa**: International Expansion of U.S. Mofa’s (majority–owned Foreign Affiliates in Fire (finance, Insurance and Real Estate)  
- **Money**: Money, GDP and Interest Rate in Canada  
- **Mpyr**: Money, National Product and Interest Rate  
- **Orange**: The Orange Juice Data Set  
- **PE**: Price and Earnings Index  
- **PPP**: Exchange Rates and Price Indices for France and Italy  
- **Pound**: Pound-dollar Exchange Rate  
- **Solow**: Solow’s Technological Change Data  
- **StrikeNb**: Number of Strikes in Us Manufacturing  
- **SumHes**: The Penn Table  
- **Tbrate**: Interest Rate, GDP and Inflation  
- **Yen**: Yen-dollar Exchange Rate

**goods**
- **Computers**: Prices of Personal Computers  
- **Diamond**: Pricing the C’s of Diamond Stones  
- **Housing**: Sales Prices of Houses in the City of Windsor  
- **Journals**: Economic Journals Dat Set

**households**
- **BudgetFood**: Budget Share of Food for Spanish Households  
- **BudgetItaly**: Budget Shares for Italian Households  
- **BudgetUK**: Budget Shares of British Households  
- **HC**: Heating and Cooling System Choice in Newly Built Houses in California  
- **Heating**: Heating System Choice in California Houses  
- **VietNamH**: Medical Expenses in Viet-nam (household Level)

**individuals**
- **Benefits**: Unemployment of Blue Collar Workers  
- **Bwages**: Wages in Belgium  
- **CPSch3**: Earnings from the Current Population Survey
– Car: Stated Preferences for Car Choice
– Catsup: Choice of Brand for Catsup
– Cracker: Choice of Brand for Crackers
– Doctor: Number of Doctor Visits
– DoctorAUS: Doctor Visits in Australia
– Earnings: Earnings for Three Age Groups
– Fair: Extramarital Affairs Data
– Fishing: Choice of Fishing Mode
– FriendFoe: Data from the Television Game Show Friend Or Foe?
– Griliches: Wage Datas
– HI: Health Insurance and Hours Worked By Wives
– Hdma: The Boston HDMA Data Set
– Kakadu: Willingness to Pay for the Preservation of the Kakadu National Park
– Ketchup: Choice of Brand for Ketchup
– Males: Wages and Education of Young Males
– Mathlevel: Level of Calculus Attained for Students Taking Advanced Microeconomics
– Mode: Mode Choice
– ModeChoice: Data to Study Travel Mode Choice
– Mroz: Labor Supply Data
– NaturalPark: Willingness to Pay for the Preservation of the Alentejo Natural Park
– OFP: Visits to Physician Office
– PSID: Panel Survey of Income Dynamics
– Participation: Labor Force Participation
– RetSchool: Return to Schooling
– Schooling: Wages and Schooling
– Somerville: Visits to Lake Somerville
– Star: Effects on Learning of Small Class Sizes
– Tobacco: Households Tobacco Budget Share
– Train: Stated Preferences for Train Traveling
– Tuna: Choice of Brand for Tuna
– Unemployment: Unemployment Duration
– VietNamI: Medical Expenses in Viet–nam (individual Level)
– Wages: Panel Datas of Individual Wages
– Wages1: Wages, Experience and Schooling
– Workinghours: Wife Working Hours
– Yogurt: Choice of Brand for Yogurts

• production units
– Airline: Cost for U.S. Airlines
– Bids: Bids Received By U.S. Firms
– CRSPday: Daily Returns from the CRSP Database
– CRSPmon: Monthly Returns from the CRSP Database
## Index.Source

- **Clothing**: Sales Data of Men’s Fashion Stores
- **Electricity**: Cost Function for Electricity Producers
- **Grunfeld**: Grunfeld Investment Data
- **Labour**: Belgian Firms
- **Nerlove**: Cost Function for Electricity Producers, 1955
- **Oil**: Oil Investment
- **PatentsHGH**: Dynamic Relation Between Patents and R\&D
- **PatentsRD**: Patents, R\&D and Technological Spillovers for a Panel of Firms

### regional
- **Airq**: Air Quality for Californian Metropolitan Areas
- **Cigar**: Cigarette Consumption
- **Cigarette**: The Cigarette Consumption Panel Data Set
- **Crime**: Crime in North Carolina
- **Fatality**: Drunk Driving Laws and Traffic Deaths
- **Hedonic**: Hedonic Prices of Cencus Tracts in Boston
- **Metal**: Production for SIC 33
- **MunExp**: Municipal Expenditure Data
- **Produc**: Us States Production
- **TransEq**: Statewide Data on Transportation Equipment Manufacturing

### schools
- **Caschool**: The California Test Score Data Set
- **MCAS**: The Massashusets Test Score Data Set
- **University**: Provision of University Teaching and Research

### Description

  - **Bids**: Bids Received By U.S. Firms
  - **BudgetFood**: Budget Share of Food for Spanish Households
  - **BudgetItaly**: Budget Shares for Italian Households
  - **BudgetUK**: Budget Shares of British Households
  - **Car**: Stated Preferences for Car Choice
  - **Computers**: Prices of Personal Computers
  - **Crime**: Crime in North Carolina
  - **Doctor**: Number of Doctor Visits
  - **Earnings**: Earnings for Three Age Groups
  - **HI**: Health Insurance and Hours Worked By Wives
  - **Housing**: Sales Prices of Houses in the City of Windsor
  - **Males**: Wages and Education of Young Males
– **Mathlevel**: Level of Calculus Attained for Students Taking Advanced Microeconomics
– **MoneyUS**: Macroeconomic Series for the United States
– **MunExp**: Municipal Expenditure Data
– **OFP**: Visits to Physician Office
– **Oil**: Oil Investment
– **Participation**: Labor Force Participation
– **PatentsRD**: Patents, R&D and Technological Spillovers for a Panel of Firms
– **Train**: Stated Preferences for Train Traveling
– **Unemployment**: Unemployment Duration
– **University**: Provision of University Teaching and Research
– **Workinghours**: Wife Working Hours

  – **Benefits**: Unemployment of Blue Collar Workers
  – **Catsup**: Choice of Brand for Catsup
  – **Cracker**: Choice of Brand for Crakers
  – **Kakadu**: Willingness to Pay for the Preservation of the Kakadu National Park
  – **Ketchup**: Choice of Brand for Ketchup
  – **LaborSupply**: Wages and Hours Worked
  – **Mofa**: International Expansion of U.S. Mofa’s (majority-owned Foreign Affiliates in Fire (finance, Insurance and Real Estate)
  – **Somerville**: Visits to Lake Somerville
  – **Tuna**: Choice of Brand for Tuna
  – **Yogurt**: Choice of Brand for Yogurts

  – **Diamond**: Pricing the C’s of Diamond Stones
  – **FriendFoe**: Data from the Television Game Show Friend Or Foe?

• **Kenneth Train’s home page**: [http://elsa.berkeley.edu/~train/](http://elsa.berkeley.edu/~train/)
  – **HC**: Heating and Cooling System Choice in Newly Built Houses in California
  – **Heating**: Heating System Choice in California Houses
  – **Mode**: Mode Choice

  – **Cigar**: Cigarette Consumption
  – **Crime**: Crime in North Carolina
  – **Gasoline**: Gasoline Consumption
  – **Grunfeld**: Grunfeld Investment Data
  – **Hedonic**: Hedonic Prices of Census Tracts in Boston
  – **Produc**: Us States Production
  – **Wages**: Panel Data of Individual Wages
  - DoctorContacts: Contacts With Medical Doctor
  - Fishing: Choice of Fishing Mode
  - LaborSupply: Wages and Hours Worked
  - MedExp: Structure of Demand for Medical Care
  - PSID: Panel Survey of Income Dynamics
  - PatentsHGH: Dynamic Relation Between Patents and R&D
  - RetSchool: Return to Schooling
  - StrikeDur: Strikes Duration
  - Treatment: Evaluating Treatment Effect of Training on Earnings
  - UnempDur: Unemployment Duration
  - VietNamH: Medical Expenses in Viet–nam (household Level)
  - VietNamI: Medical Expenses in Viet–nam (individual Level)
  - Bids: Bids Received By U.S. Firms
  - DoctorAUS: Doctor Visits in Australia
  - OFP: Visits to Physician Office
  - PatentsHGH: Dynamic Relation Between Patents and R&D
  - Somerville: Visits to Lake Somerville
  - StrikeNb: Number of Strikes in Us Manufacturing
  - CRSPday: Daily Returns from the CRSP Database
  - CRSPmon: Monthly Returns from the CRSP Database
  - Consumption: Quarterly Data on Consumption and Expenditure
  - Doctor: Number of Doctor Visits
  - Earnings: Earnings for Three Age Groups
  - Hstarts: Housing Starts
  - MW: Growth of Disposable Income and Treasury Bill Rate
  - Money: Money, GDP and Interest Rate in Canada
  - Participation: Labor Force Participation
  - Tbrate: Interest Rate, GDP and Inflation
  - Accident: Ship Accidents
  - Airline: Cost for U.S. Airlines
  - Electricity: Cost Function for Electricity Producers
  - Fair: Extramarital Affairs Data
  - Grunfeld: Grunfeld Investment Data
  - Klein: Klein’s Model I
  - Longley: The Longley Data
Index Source

- **ManufCost**: Manufacturing Costs
- **Metal**: Production for SIC 33
- **ModeChoice**: Data to Study Travel Mode Choice
- **Mroz**: Labor Supply Data
- **MunExp**: Municipal Expenditure Data
- **Nerlove**: Cost Function for Electricity Producers, 1955
- **Solow**: Solow’s Technological Change Data
- **Strike**: Strike Duration Data
- **TransEq**: Statewide Data on Transportation Equipment Manufacturing


- **DM**: DM Dollar Exchange Rate
- **Electricity**: Cost Function for Electricity Producers
- **Griliches**: Wage Data
- **LT**: Dollar Sterling Exchange Rate
- **Mishkin**: Inflation and Interest Rates
- **Mpyr**: Money, National Product and Interest Rate
- **Nerlove**: Cost Function for Electricity Producers, 1955
- **Pound**: Pound-dollar Exchange Rate
- **SumHes**: The Penn Table
- **Yen**: Yen-dollar Exchange Rate


- **CPSch3**: Earnings from the Current Population Survey
- **Caschool**: The California Test Score Data Set
- **Cigarette**: The Cigarette Consumption Panel Data Set
- **Fatality**: Drunk Driving Laws and Traffic Deaths
- **Hdma**: The Boston HDMA Data Set
- **Journals**: Economic Journals Data Set
- **MCAS**: The Massachusetts Test Score Data Set
- **Macrodat**: Macroeconomic Time Series for the United States
- **Orange**: The Orange Juice Data Set
- **Star**: Effects on Learning of Small Class Sizes


- **Airq**: Air Quality for Californian Metropolitan Areas
- **Benefits**: Unemployment of Blue Collar Workers
- **Bwages**: Wages in Belgium
- **Capm**: Stock Market Data
- **Clothing**: Sales Data of Men’s Fashion Stores
- **Forward**: Exchange Rates of US Dollar Against Other Currencies
- **Garch**: Daily Observations on Exchange Rates of the US Dollar Against Other Currencies
– **Housing**: Sales Prices of Houses in the City of Windsor
– **Icecream**: Ice Cream Consumption
– **IncomeUK**: Seasonally Unadjusted Quarterly Data on Disposable Income and Expenditure
– **Irates**: Monthly Interest Rates
– **Labour**: Belgian Firms
– **Males**: Wages and Education of Young Males
– **MoneyUS**: Macroeconomic Series for the United States
– **NaturalPark**: Willingness to Pay for the Preservation of the Alentejo Natural Park
– **PE**: Price and Earnings Index
– **PPP**: Exchange Rates and Price Indices for France and Italy
– **PatentsRD**: Patents, R&D and Technological Spillovers for a Panel of Firms
– **Pricing**: Returns of Size-based Portfolios
– **SP500**: Returns on Standard \\& Poor's 500 Index
– **Schooling**: Wages and Schooling
– **Tobacco**: Households Tobacco Budget Share
– **Wages1**: Wages, Experience and Schooling

---

**Description**

- **annual**
  - **Klein**: Klein's Model I
  - **LT**: Dollar Sterling Exchange Rate
  - **Longley**: The Longley Data
  - **ManufCost**: Manufacturing Costs
  - **Mpyr**: Money, National Product and Interest Rate
  - **PE**: Price and Earnings Index
  - **Solow**: Solow's Technological Change Data
- **daily**
  - **CRSPday**: Daily Returns from the CRSP Database
  - **Garch**: Daily Observations on Exchange Rates of the US Dollar Against Other Currencies
  - **SP500**: Returns on Standard \\& Poor's 500 Index
- **four–weekly**
  - **Icecream**: Ice Cream Consumption
- **monthly**
  - **CRSPmon**: Monthly Returns from the CRSP Database
  - **Capm**: Stock Market Data
  - **Forward**: Exchange Rates of US Dollar Against Other Currencies
- **Irates**: Monthly Interest Rates
- **Mishkin**: Inflation and Interest Rates
- **Orange**: The Orange Juice Data Set
- **PPP**: Exchange Rates and Price Indices for France and Italy
- **Pricing**: Returns of Size-based Portfolios
- **StrikeNb**: Number of Strikes in Us Manufacturing

- **quarterly**
  - **Consumption**: Quarterly Data on Consumption and Expenditure
  - **Hstarts**: Housing Starts
  - **IncomeUK**: Seasonally Unadjusted Quarterly Data on Disposable Income and Expenditure
  - **MW**: Growth of Disposable Income and Treasury Bill Rate
  - **Macrod**: Macroeconomic Time Series for the United States
  - **Money**: Money, GDP and Interest Rate in Canada
  - **MoneyUS**: Macroeconomic Series for the United States
  - **Tbrate**: Interest Rate, GDP and Inflation

- **weekly**
  - **DM**: DM Dollar Exchange Rate
  - **Pound**: Pound-dollar Exchange Rate
  - **Yen**: Yen-dollar Exchange Rate

---

**Irates**

*Monthly Interest Rates*

---

**Description**

monthly observations from 1946–12 to 1991–02

_number of observations_: 531

_observation_: country

_country_: United–States

**Usage**

data(Irates)

**Format**

A time serie containing:

- **r1**: interest rate for a maturity of 1 months (% per year).
- **r2**: interest rate for a maturity of 2 months (% per year).
- **r3**: interest rate for a maturity of 3 months (% per year).
- **r5**: interest rate for a maturity of 5 months (% per year).
**r6** interest rate for a maturity of 6 months (% per year).
**r11** interest rate for a maturity of 11 months (% per year).
**r12** interest rate for a maturity of 12 months (% per year).
**r36** interest rate for a maturity of 36 months (% per year).
**r60** interest rate for a maturity of 60 months (% per year).
**r120** interest rate for a maturity of 120 months (% per year).

**Source**

**References**

**See Also**
[Index.Source], [Index.Economics], [Index.Econometrics], [Index.Observations], [Index.Time.Series]

---

**Journals**

*Economic Journals Dat Set*

**Description**

- a cross-section from 2000
  - *number of observations*: 180
  - *observation*: goods

**Usage**

`data(Journals)`

**Format**

- A dataframe containing:
  - **title** journal title
  - **pub** publisher
  - **society** scholarly society
  - **libprice** library subscription price
  - **pages** number of pages
Kakadu

Willingness to Pay for the Preservation of the Kakadu National Park

Description

a cross-section

number of observations : 1827

observation : individuals

country : Australia

Usage

data(Kakadu)

Format

A dataframe containing :

lower  lowerbound of willingness to pay, 0 if observation is left censored

upper  upper bound of willingness to pay, 999 if observation is right censored

answer  an ordered factor with levels nn (respondent answers no, no), ny (respondent answers no,

yes or yes, no), yy (respondent answers yes, yes)

recparks  the greatest value of national parks and nature reserves is in recreation activities (from 1
to 5)

jobs  jobs are the most important thing in deciding how to use our natural ressources (from 1 to 5)

Source

Professor Theodore Bergstrom of the Department of Economics at the University of California, San

Diego.

References

Stock, James H. and Mark W. Watson (2003) Introduction to Econometrics, Addison-Wesley Edu-


See Also

Index.Source, Index.Economics, Index.Econometrics, Index.Observations
Kakadu

lowrisk  development should be allowed to proceed where environmental damage from activities such as mining is possible but very unlikely (from 1 to 5)

wildlife  it’s important to have places where wildlife is preserved (from 1 to 5)

future  it’s important to consider future generations (from 1 to 5)

aboriginal  in deciding how to use areas such as Kakadu national park, their importance to the local aboriginal people should be a major factor (from 1 to 5)

finben  in deciding how to use our natural resources such as mineral deposits and forests, the most important thing is the financial benefits for Australia (from 1 to 5)

mineparks  if areas within natural parks are set aside for development projects such as mining, the value of the parks is greatly reduced (from 1 to 5)

moreparks  there should be more national parks created from state forests (from 1 to 5)

gov  the government pays little attention to the people in making decisions (from 1 to 4)

envcon  the respondent recycles things such as paper or glass and regularly buys unbleached toilet paper or environmentally friendly products?

vparks  the respondent has visited a national park or bushland recreation area in the previous 12 months?

tenv  the respondent watches TV programs about the environment? (from 1 to 9)

conservation  the respondent is member of a conservation organization?

sex  male,female

age  age

schooling  years of schooling

income  respondent’s income in thousands of dollars

major  the respondent received the major–impact scenario of the Kakadu conservation zone survey?

Source


References


See Also

Index.Source, Index.Economics, Index.Econometrics, Index.Observations
Ketchup

Choice of Brand for Ketchup

Description

a cross-section

number of observations: 4956

observation: individuals

country: United States

Usage

data(Ketchup)

Format

A dataframe containing:

bid individuals identifiers

id purchase identifiers

choice one of heinz, hunts, delmonte, stb (store brand)

price.z price of brand z

Source


References


See Also

index.Source, index.Economics, index.Econometrics, index.Observations
Klein | Klein's Model I

**Description**
annual observations from 1920 to 1941

- number of observations : 22
- observation : country
- country : United States

**Usage**

data(Klein)

**Format**
A time serie containing :

- cons consumption
- profit corporate profits
- privwage private wage bill
- inv investment
- lcap previous year's capital stock
- gnp GNP
- pubwage government wage bill
- govspend government spending
- taxe taxes

**Source**

**References**

**See Also**
Description

a panel of 532 observations from 1979 to 1988

number of observations : 5320

Usage

data(LaborSupply)

Format

A dataframe containing :

lnhr  log of annual hours worked
lnwg  log of hourly wage
kids  number of children
age   age
disab bad health
id    id
year  year

Source


References


See Also

Labour  Belgian Firms

Description

a cross-section from 1996

number of observations: 569
observation: production units
country: Belgium

Usage

data(Labour)

Format

A dataframe containing:
capital total fixed assets, end of 1995 (in 1000000 euro)
labour number of workers (employment)
output value added (in 1000000 euro)
wage wage costs per worker (in 1000 euro)

References


See Also

indexSource, indexEconomics, indexEconometrics, indexObservations

Longley  The Longley Data

Description

annual observations from 1947 to 1962

number of observations: 16
observation: country
country: United States
Usage
data(Longley)

Format
A time serie containing:

- **employ**: employment (1,000s)
- **price**: GNP deflator
- **gnp**: nominal GNP (millions)
- **armed**: armed forces

Source

References

See Also

---

**LT**

*Dollar Sterling Exchange Rate*

Description
annual observations from 1791 to 1990

- **number of observations**: 200
- **observation**: country
- **country**: United Kingdom

Usage
data(LT)
Format

A time serie containing:

- `s` US Dollar / Pound exchange rate
- `uswpi` US wholesale price index, normalized to 100 for 1914
- `ukwpi` US wholesale price index, normalized to 100 for 1914

Source


References


See Also


---

Macrodat  
*Macroeconomic Time Series for the United States*

Description

- quarterly observations from 1959-1 to 2000-4
- number of observations: 168
- observation: country
- country: United States

Usage

data(Macrodat)

Format

A time serie containing:

- `lthur` unemployment rate (average of months in quarter)
- `punew` cPl (Average of Months in Quarter)
- `fyff` federal funds interest rate (last month in quarter)
- `fygm3` 3 month treasury bill interest rate (last month in quarter)
- `fygt1` 1 year treasury bond interest rate (last month in quarter)
- `exruk` dollar / Pound exchange rate (last month in quarter)
- `gdpjp` real GDP for Japan
Males

Source


References


See Also


Males  

*Wages and Education of Young Males*

Description

a panel of 545 observations from 1980 to 1987

*number of observations* : 4360

*observation* : individuals

*country* : United States

Usage

data(Males)

Format

A dataframe containing :

- **nr**  identifiant
- **year**  year
- **school**  years of schooling
- **exper**  years of experience (=age-6-school)
- **union**  wage set by collective bargaining  
- **ethn**  a factor with levels (black,hisp,other)
- **maried**  maried  
- **health**  health problem  
- **wage**  log of hourly wage
- **industry**  a factor with 12 levels
- **occupation**  a factor with 9 levels
- **residence**  a factor with levels (rural area, north east, northern central, south)
Source

National Longitudinal Survey (NLS Youth Sample).

References


See Also

*Index.Source, Index.Economics, Index.Econometrics, Index.Observations,*
*Index.Time.Series*

---

**ManufCost**

*Manufacturing Costs*

Description

annual observations from 1947 to 1971

*number of observations*: 25

*observation*: country

*country*: United States

Usage

data(ManufCost)

Format

A time serie containing:

*cost*  cost index

*sk*  capital cost share

*sl*  labor cost share

*se*  energy cost share

*sm*  materials cost share

*pk*  capital price

*pl*  labor price

*pe*  energy price

*pm*  materials price
Mathlevel

Source


References


See Also


---

Mathlevel | Level of Calculus Attained for Students Taking Advanced Micro-economics
---|---

**Description**

- a cross-section from 1983 to 1986
- number of observations: 609
- observation: individuals
- country: United States

**Usage**

data(Mathlevel)

**Format**

A dataframe containing:

- **mathlevel** highest level of math attained, an ordered factor with levels 170, 171a, 172, 171b, 172b, 221a, 221b
- **sat** sat Math score
- **language** foreign language proficiency?
- **sex** male, female
- **major** one of other, eco, oss (other social sciences), ns (natural sciences), hum (humanities)
- **mathcourse** number of courses in advanced math (0 to 3)
- **physicourse** number of courses in physics (0 to 2)
- **chemistcourse** number of courses in chemistry (0 to 2)
Source

References

See Also
Index.Source, Index.Economics, Index.Econometrics, Index.Observations

---

**MCAS**

*The Massashusets Test Score Data Set*

**Description**
a cross-section from 1997-1998

*number of observations*: 220

*observation*: schools

*country*: United States

**Usage**

data(MCAS)

**Format**
A dataframe containing:

- **code**: district code (numerical)
- **municipa**: municipality (name)
- **district**: district name
- **regday**: spending per pupil, regular
- **specneed**: spending per pupil, special needs
- **bilingua**: spending per pupil, bilingual
- **occupday**: spending per pupil, occupational
- **totday**: spending per pupil, total
- **spc**: students per computer
- **speed**: special education students
- **lnchpct**: eligible for free or reduced price lunch
- **tchratio**: students per teacher
percap per capita income

totsc4 4th grade score (math+english+science)

totsc8 8th grade score (math+english+science)

avgsalary average teacher salary

pctel percent english learners

Source

Massachusetts Comprehensive Assessment System (MCAS), Massachusetts Department of Education, 1990 U.S. Census.

References


See Also

Index.Source, Index.Economics, Index.Econometrics, Index.Observations

| MedExp | Structure of Demand for Medical Care |

Description


number of observations: 5574

Usage

data(MedExp)

Format

A time serie containing:

med annual medical expenditures in constant dollars excluding dental and outpatient mental

lc \( \log(\text{coinsrate}+1) \) where coinsurance rate is 0 to 100

idp individual deductible plan?

lpi \( \log(\text{annual participation incentive payment}) \) or 0 if no payment

fmde \( \log(\max(\text{medical deductible expenditure})) \) if IDP=1 and MDE>1 or 0 otherwise

physlim physical limitation?

ndisease number of chronic diseases

health self–rate health (excellent,good,fair,poor)
**Source**


**References**


**See Also**


---

**Description**

a cross-section

*number of observations*: 27

*observation*: regional

*country*: United States

**Usage**

`data(Metal)`

**Format**

A dataframe containing:

*va* output

*labor* labor input

*capital* capital input
Source


References


See Also

`Index.Source`, `Index.Economics`, `Index.Econometrics`, `Index.Observations`

---

**Mishkin**

**Inflation and Interest Rates**

**Description**

monthly observations from 1950-2 to 1990-12

*number of observations*: 491

*observation*: country

*country*: United States

**Usage**

`data(Mishkin)`

**Format**

A time serie containing:

- `pail` one-mounth inflation rate (in percent, annual rate)
- `pai3` three-mounth inflation rate (in percent, annual rate)
- `tb1` one-mounth T-bill rate (in percent, annual rate)
- `tb3` three-mounth T-bill rate (in percent, annual rate)
- `cpi` CPI for urban consumers, all items (the 1982-1984 average is set to 100)

**Source**

References


See Also


---

Mode | Mode Choice
---|---

Description

a cross-section

*number of observations*: 453

*observation*: individuals

Usage

data(Mode)

Format

A dataframe containing:

*choice* one of car, carpool, bus or rail

*cost.z* cost of mode z

*time.z* time of mode z

References

Kenneth Train’s home page: http://elsa.berkeley.edu/~train/.

See Also

Index.Source, Index.Economics, Index.Econometrics, Index.Observations
Description

a cross-section

number of observations: 840

observation: individuals

country: Australia

Usage

data(ModeChoice)

Format

A dataframe containing:

mode choice: air, train, bus or car

ttime terminal waiting cost time, 0 for car

invc in vehicle cost-cost component

invt travel time in vehicle

gc generalized cost measure

hinc household income

psize party size in mode chosen

Source


References


See Also

Index.Source, Index.Economics, Index.Econometrics, Index.Observations
**Mofa**

International Expansion of U.S. Mofa's (majority-owned Foreign Affiliates in Fire (finance, Insurance and Real Estate)

**Description**

a cross-section from 1982

- **number of observations**: 50
- **observation**: country
- **country**: United States

**Usage**

`data(Mofa)`

**Format**

A dataframe containing:

- **gdp**: gross domestic product. Source: "World Bank, World Development Report 1984." Table 3. (This variable is scaled by a factor of 1/100,000)
- **sales**: sales made by the majority owned foreign affiliates of nonbank U.S. parents in finance, insurance and real estate. Source: "U.S. Direct Investment Abroad: 1982 Benchmark Survey Data." Table III.D 3. (This variable is scaled by a factor of 1/100)
- **nbaf**: the number of U.S. affiliates in the host country. Source: "U.S. Direct Investment Abroad: 1982 Benchmark Survey Data." Table 5. (This variable is scaled by a factor of 1/100)
- **netinc**: net income earned by MOFA's of nonbank U.S. corporations operating in the nonbanking financial sector of the host country. Source: "U.S. Direct Investment Abroad: 1982 Benchmark Survey Data." Table III.D 6. (This variable is scaled by a factor of 1/10)

**Source**


**References**


**See Also**

Index.Source, Index.Economics, Index.Econometrics, Index.Observations
Money, GDP and Interest Rate in Canada

Description

quarterly observations from 1967-1 to 1998-4

\textit{number of observations} : 128

\textit{observation} : country

\textit{country} : Canada

Usage

data(Money)

Format

A time serie containing :

\textbf{m} log of the real money supply

\textbf{y} the log of GDP, in 1992 dollars, seasonally adjusted

\textbf{p} the log of the price level

\textbf{r} the 3-month treasury till rate

Source

CANSIM Database of Statistics Canada.

References


See Also

\url{Index.Source}, \url{Index.Economics}, \url{Index.Econometrics}, \url{Index.Observations},
\url{Index.Time.Series}
MoneyUS  Macroeconomic Series for the United States

Description

quarterly observations from 1954–01 to 1994–12

number of observations : 164

country : United States

Usage

data(MoneyUS)

Format

A time serie containing :

m  log of real M1 money stock
infl  quarterly inflation rate (change in log prices), % per year
cpr  commercial paper rate, % per year
y  log real GDP (in billions of 1987 dollars)
tbr  treasury bill rate

Source


References


See Also

Mpyr

Money, National Product and Interest Rate

Description

annual observations from 1900 to 1989

number of observations : 90

observation : country

country : United States

Usage

data(Mpyr)

Format

A time serie containing :

m natural log of M1

p natural log of the net national product price deflator

y natural log of the net national product

r the commercial paper rate in percent at an annual rate

Source


References


See Also

Description

a cross-section

number of observations : 753

observation : individuals
country : United States

Usage

data(Mroz)

Format

A dataframe containing :

work participation in 1975 ?
hoursw wife’s hours of work in 1975
child6 number of children less than 6 years old in household
child618 number of children between ages 6 and 18 in household
agew wife’s age
educw wife’s educational attainment, in years
hearnw wife’s average hourly earnings, in 1975 dollars
wagew wife’s wage reported at the time of the 1976 interview (not= 1975 estimated wage)
hoursh husband’s hours worked in 1975
ageh husband’s age
educu husband’s educational attainment, in years
wageh husband’s wage, in 1975 dollars
income family income, in 1975 dollars
educwm wife’s mother’s educational attainment, in years
educwf wife’s father’s educational attainment, in years
unemprate unemployment rate in county of residence, in percentage points
city lives in large city (SMSA) ?
experience actual years of wife’s previous labor market experience

Source

1976 Panel Study of Income Dynamics.
MunExp

References

See Also
Index.Source, Index.Economics, Index.Econometrics, Index.Observations

---

**MunExp**

**Municipal Expenditure Data**

**Description**
- a panel of 265 observations from 1979 to 1987
  - number of observations: 2385
  - observation: regional
  - country: Sweeden

**Usage**
- `data(MunExp)`

**Format**
- A dataframe containing:
  - id: identification
  - year: date
  - expend: expenditure
  - revenue: revenue from taxes and fees
  - grants: grants from Central Government

**Source**

**References**

**See Also**
- Index.Source, Index.Economics, Index.Econometrics, Index.Observations,
  - Index.Time.Series
Description

quarterly observations from 1963-3 to 1975-4

number of observations : 50

observation : country
country : United States

Usage

data(MW)

Format

A time serie containing :

rdi the rate of growth of real U.S. disposable income, seasonally adjusted
trate the U.S. treasury bill rate

Source


References


See Also

Index.Source, Index.Economics, Index.Econometrics, Index.Observations,
Index.Time.Series
Description

a cross-section from 1987

number of observations : 312
observation : individuals
country : Portugal

Usage

data(NaturalPark)

Format

A dataframe containing :

bid1 initial bid, in euro
bidh higher bid
bidl lower bid
answers a factor with levels (nn,ny,yn,yy)
age age in 6 classes
sex a factor with levels (male,female)
income income in 8 classes

Source


References


See Also

Index.Source, Index.Economics, Index.Econometrics, Index.Observations
Description

a cross-section from 1955 to 1955

*number of observations*: 159

*observation*: production units

*country*: United States

Usage

data(Nerlove)

Format

A dataframe containing:

- **cost**  total cost
- **output**  total output
- **pl**  wage rate
- **sl**  cost share for labor
- **pk**  capital price index
- **sk**  cost share for capital
- **pf**  fuel price
- **sf**  cost share for fuel

Source


References


See Also

[Index.Source], [Index.Economics], [Index.Econometrics], [Index.Observations]
Description

A data frame describing names containing character codes rare or non-existent in standard English text, e.g., with various accent marks that may not be coded consistently in different locales or by different software.

Usage

data(nonEnglishNames)

Format

A data frame with two columns:

- **nonEnglish** a character vector containing names that often have non-standard characters with the non-standard characters replaced by "_"
- **English** a character vector containing a standard English-character translation of nonEnglish

See Also

grepNonStandardCharacters, subNonStandardCharacters

Examples

data(nonEnglishNames)

all.equal(dim(nonEnglishNames), c(17, 2))

Description

a cross-section

number of observations: 4406

observation: individuals

country: United States
Usage
data(OFP)

Format

A dataframe containing:

- **ofp** number of physician office visits
- **ofnp** number of nonphysician office visits
- **opp** number of physician outpatient visits
- **opnp** number of nonphysician outpatient visits
- **emr** number of emergency room visits
- **hosp** number of hospitalizations
- **numchron** number of chronic conditions
- **adldiff** the person has a condition that limits activities of daily living?
- **age** age in years (divided by 10)
- **black** is the person african–american?
- **sex** is the person male?
- **maried** is the person married?
- **school** number of years of education
- **faminc** family income in 10000$
- **employed** is the person employed?
- **privins** is the person covered by private health insurance?
- **medicaid** is the person covered by medicaid?
- **region** the region (noreast, midwest, west)
- **hlth** self-perceived health (excellent, poor, other)

Source


References


See Also

[Index.Source, Index.Economics, Index.Econometrics, Index.Observations]
Oil

Oil Investment

Description

a cross-section from 1969 to 1992

number of observations : 53

observation : production units

country : United Kingdom

Usage

data(Oil)

Format

A dataframe containing:

dur duration of the appraisal lag in months (time span between discovery of an oil field and beginning of development, i.e. approval of annex B).

size size of recoverable reserves in millions of barrels

waterd depth of the sea in metres

gasres size of recoverable gas reserves in billions of cubic feet

operator equity market value (in 1991 million pounds) of the company operating the oil field

p real after–tax oil price measured at time of annex B approval

vardp volatility of the real oil price process measured as the squared recursive standard errors of the regression of pt-pt-1 on a constant

p97 adaptative expectations (with parameter theta=0.97) for the real after–tax oil prices formed at the time of annex B approval

varp97 volatility of the adaptative expectations (with parameter theta=0.97) for real after tax oil prices measured as the squared recursive standard errors of the regression of pt on pte(theta)

p98 adaptative expectations (with parameter theta=0.98) for the real after–tax oil prices formed at the time of annex B approval

varp98 volatility of the adaptative expectations (with parameter theta=0.98) for real after tax oil prices measured as the squared recursive standard errors of the regression of pt on pte(theta)

Source


References

The Orange Juice Data Set

Description

monthly observations from 1948-01 to 2001-06

number of observations : 642

observation : country
country : United States

Usage

data(Orange)

Format

A time serie containing :

priceoj producer price for frozen orange juice

pricefg producer price index for finished goods

fdd freezing degree days (from daily minimum temperature recorded at Orlando area airports)

Source

U.S. Bureau of Labor Statistics for PPIOJ and PWFSA, National Oceanic and Atmospheric Admin-
istration (NOAA) of the U.S Department of Commerce for FDD.

References

Stock, James H. and Mark W. Watson (2003) Introduction to Econometrics, Addison-Wesley Edu-

See Also

Participation  Labor Force Participation

Description

a cross-section

*number of observations*: 872

*observation*: individuals

*country*: Switzerland

Usage

data(Participation)

Format

A dataframe containing:

- `lfp` labour force participation
- `lnnlinc` the log of nonlabour income
- `age` age in years divided by 10
- `educ` years of formal education
- `nyc` the number of young children (younger than 7)
- `noc` number of older children
- `foreign` foreigner

Source


References


See Also

- Index.Source
- Index.Economics
- Index.Econometrics
- Index.Observations
Description

a panel of 346 observations from 1975 to 1979

number of observations : 1730

observation : production units

country : United States

Usage

data(PatentsHGH)

Format

A dataframe containing :

obsno  firm index

year  year

cusip  Compustat’s identifying number for the firm (Committee on Uniform Security Identification Procedures number)

ardsic  a two-digit code for the applied R&D industrial classification (roughly that in Bound, Cummins, Griliches, Hall, and Jaffe, in the Griliches R&D, Patents, and Productivity volume)

scisect  is the firm in the scientific sector ?

logk  the logarithm of the book value of capital in 1972.

sumpat  the sum of patents applied for between 1972-1979.

logr  the logarithm of R&D spending during the year (in 1972 dollars)

logr1  the logarithm of R&D spending (one year lag)

logr2  the logarithm of R&D spending (two years lag)

logr3  the logarithm of R&D spending (three years lag)

logr4  the logarithm of R&D spending (four years lag)

logr5  the logarithm of R&D spending (five years lag)

pat  the number of patents applied for during the year that were eventually granted

pat1  the number of patents (one year lag)

pat2  the number of patents (two years lag)

pat3  the number of patents (three years lag)

pat4  the number of patents (four years lag)
Source


References


See Also


| PatentsRD | Patents, R&D and Technological Spillovers for a Panel of Firms |

Description

a panel of 181 observations from 1983 to 1991

number of observations: 1629

observation: production units

country: world

Usage

data(PatentsRD)

Format

A dataframe containing:

year year
fi firm’s id
sector firm’s main industry sector, one of aero (aerospace), chem (chemistry), comput (computer), drugs, elec (electricity), food, fuel (fuel and mining), glass, instr (instruments), machin (machinery), metals, other, paper, soft (software), motor (motor vehicles)
geo geographic area, one of eu (European Union), japan, usa, rotw (rest of the world)
patent numbers of European patent applications
rdexp log of R&D expenditures
spil log of spillovers
Source

References

See Also

---

**PE**

*Price and Earnings Index*

---

**Description**
annual observations from 1800 to 1931

*number of observations* : 132

*observation* : country
country : United States

**Usage**
data(PE)

**Format**
A time serie containing :

*price*  S\&P composite stock price index

*earnings*  S\&P composite earnings index

**Source**
Robert Shiller.

**References**
politicalknowledge

See Also


---

politicalKnowledge  Political knowledge in the US and Europe

Description

Data from McChesney and Nichols (2010) on domestic and international knowledge in Denmark, Finland, the UK and the US among college graduates, people with some college, and roughly 12th grade only.

Usage

data(politicalknowledge)

Format

A data.frame containing 12 columns and 4 rows.

country  a factor of Denmark, Finland, UK, and US, being the four countries compared in this data set.

DomesticKnowledge.hs, DomesticKnowledge.sc, DomesticKnowledge.c  percent correct answers to calibrated questions regarding knowledge of prominent items in domestic news in a survey of residents of the four countries among college graduates (ending "c"), some college ("sc") and high school ("hs"). Source: McChesney and Nichols (2010, chapter 1, chart 8).

InternationalKnowledge.hs, InternationalKnowledge.sc, InternationalKnowledge.c  percent correct answers to calibrated questions regarding knowledge of prominent items in international news in a survey of residents of the four countries by education level as for DomesticKnowledge. Source: McChesney and Nichols (2010, chapter 1, chart 7).

PoliticalKnowledge.hs, PoliticalKnowledge.sc, PoliticalKnowledge.c  average of domestic and international knowledge

PublicMediaPerCapita  Per capital spending on public media in 2007 in US dollars from McChesney and Nichols (2010, chapter 4, chart 1)

PublicMediaRel2US  Spending on public media relative to the US, being PublicMediaPerCapita / PublicMediaPerCapita

Author(s)

Spencer Graves

Source

Examples

## 1. Combine first 2 rows
```r
data(politicalKnowledge)
pk <- politicalKnowledge[-1,]
pk[1, -1] <- ((politicalKnowledge[1, -1] +
                politicalKnowledge[2, -1])/2)
pk[1, 'country'] <- 'DK-FI'
```

## plot
```r
xlim <- range(pk[, 'PublicMediaPerCapita'])
ylim <- 100*range(pk[2:7])
text.cex <- 2

# to label the lines
(US.UK <- (pk[2, -1]+pk[3, -1])/2)

#png('Knowledge v. public media.png')
op <- par(mar=c(5, 7, 4, 2)+1)
plot(c(0, 110), 100*ylim, type='n', axes=FALSE,
     xlab='Public media $ per capita',
     ylab='Political Knowledge\n(% of standard questions)',
     cex.lab=2)
axis(1, cex.axis=2)
axis(2, las=2, cex.axis=2)

with(pk, text(PublicMediaPerCapita, 100*PoliticalKnowledge hs,
              country, cex=text.cex, xpd=NA,
              col=c('forestgreen', 'orange', 'red')))
with(pk, text(PublicMediaPerCapita, 100*PoliticalKnowledge sc,
              country, cex=text.cex, xpd=NA,
              col=c('forestgreen', 'orange', 'red')))
with(pk, text(PublicMediaPerCapita, 100*PoliticalKnowledge c,
              country, cex=text.cex, xpd=NA,
              col=c('forestgreen', 'orange', 'red')))
with(pk, lines(PublicMediaPerCapita, 100*PoliticalKnowledge hs,
               type='b', pch=' '))
with(pk, lines(PublicMediaPerCapita, 100*PoliticalKnowledge sc,
               type='b', pch=' '))
with(pk, lines(PublicMediaPerCapita, 100*PoliticalKnowledge c,
               type='b', pch=' '))
with(US.UK, text(PublicMediaPerCapita, 100*PoliticalKnowledge hs,
                 'High School\nor less', srt=37, cex=1.5))
with(US.UK, text(PublicMediaPerCapita, 100*PoliticalKnowledge sc,
                 'some\ncollege', srt=10.5, cex=1.5))
with(US.UK, text(PublicMediaPerCapita, 100*PoliticalKnowledge c,
                 'Bachelor\n\nor more', srt=-1, cex=1.5))

par(op)
#dev.off()
```
## Pound-dollar Exchange Rate

### Description

weekly observations from 1975 to 1989

*number of observations*: 778

*observation*: country

*country*: Germany

### Usage

data(Pound)
Format

A dataframe containing:

- **date**: the date of the observation (19850104 is January 4, 1985)
- **s**: the ask price of the dollar in units of Pound in the spot market on Friday of the current week
- **f**: the ask price of the dollar in units of Pound in the 30-day forward market on Friday of the current week
- **s30**: the bid price of the dollar in units of Pound in the spot market on the delivery date on a current forward contract

Source


References


See Also


---

**PPP**  
*Exchange Rates and Price Indices for France and Italy*

Description

- monthly observations from 1981–01 to 1996–06
  
  *number of observations*: 186
  
  *observation*: country
  
  *country*: France and Italy

Usage

`data(PPP)`
**Format**

A time serie containing:

- `lnit` log price index Italy
- `lnfr` log price index France
- `lnx` log exchange rate France/Italy
- `cpiit` consumer price index Italy
- `cpifr` consumer price index France

**Source**

DataStream.

**References**


**See Also**


---

**Pricing**

**Returns of Size-based Portfolios**

**Description**

monthly observations from 1959–02 to 1993–11

*number of observations*: 418

**Usage**

data(Pricing)

**Format**

A time serie containing:

- `r1` monthly return on portfolio 1 (small firms)
- `r2` monthly return on portfolio 2
- `r3` monthly return on portfolio 3
- `r4` monthly return on portfolio 4
- `r5` monthly return on portfolio 5
- `r6` monthly return on portfolio 6
r7  monthly return on portfolio 7
r8  monthly return on portfolio 8
r9  monthly return on portfolio 9
r10 monthly return on portfolio 10 (large firms)
rf  risk free rate (return on 3-month T-bill)
cons real per capita consumption growth based on total US personal consumption expenditures (nondurables and services)

Source
Center for research in security prices.

References

See Also

<table>
<thead>
<tr>
<th>Produc</th>
<th>Us States Production</th>
</tr>
</thead>
</table>

Description
a panel of 48 observations from 1970 to 1986
number of observations : 816
observation : regional
country : United States

Usage
data(Produc)

Format
A dataframe containing :
state  the state
year   the year
pcap   private capital stock
hwy    highway and streets
water  water and sewer facilities
util  other public buildings and structures
pc  public capital
gsp  gross state products
emp  labor input measured by the employment in non-agricultural payrolls
unemp  state unemployment rate

Source

References

See Also

PSID  Panel Survey of Income Dynamics

Description
a cross-section from 1993
number of observations : 4856
observation : individuals
country : United States

Usage
data(PSID)
Format

A dataframe containing:

- `intnum` 1968 interview number
- `persnum` person number
- `age` age of individual
- `educatn` highest grade completed
- `earnings` total labor income
- `hours` annual work hours
- `kids` live births to this individual
- `married` last known marital status (married, never married, windowed, divorced, separated, NA/DF, no histories)

Source

Panel Survey of Income Dynamics.

References


See Also

`Index.Source, Index.Economics, Index.Econometrics, Index.Observations`

---

RetSchool  

*Return to Schooling*

Description

a panel of 48 observations from 1970 to 1986

- `number of observations` : 5225
- `observation` : individuals
- `country` : United States

Usage

data(RetSchool)
Format

A time serie containing:

- **wage76** wage in 1876
- **grade76** grade level in 1976
- **exp76** experience in 1976
- **black** black?
- **south76** lived in south in 1976?
- **smsa76** lived in smsa in 1976?
- **region** region, a factor with levels (un,midatl,enc,wnc,sa,esc,wsc,m,p)
- **smsa66** lived in smsa in 1966?
- **momdad14** lived with both parents at age 14?
- **sinmom14** lived with mother only at age 14?
- **nodaded** father has no formal education?
- **nomomed** mother has no formal education?
- **daded** mean grade level of father
- **momed** mean grade level of mother
- **famed** father's and mother's education, a factor with 9 levels
- **age76** age in 1976
- **col4** is any 4-year college nearby?

Source


References


See Also

Description

a cross-section from 1976

number of observations: 3010

observation: individuals

country: United States

Usage

data(Schooling)

Format

A dataframe containing:

smsa66 lived in smsa in 1966?

smsa76 lived in smsa in 1976?

nearc2 grew up near 2-yr college?

nearc4 grew up near 4-yr college?

nearc4a grew up near 4-year public college?

nearc4b grew up near 4-year private college?

ed76 education in 1976

ed66 education in 1966

age76 age in 1976

daded dads education (imputed avg if missing)

nodaded dads education imputed?

mommed mothers education

nomomed moms education imputed?

momdad14 lived with mom and dad at age 14?

sinmom14 single mom at age 14?

step14 step parent at age 14?

south66 lived in south in 1966?

south76 lived in south in 1976?

lwage76 log wage in 1976 (outliers trimmed)

famed mom-dad education class (1-9)

black black?

wage76 wage in 1976 (raw, cents per hour)
enroll76 enrolled in 1976?
kww the kww score
iqscore a normed IQ score
mar76 married in 1976?
libcrd14 library card in home at age 14?
exp76 experience in 1976

Source
National Longitudinal Survey of Young Men (NLSYM).

References

See Also
Index.Source, Index.Economics, Index.Econometrics, Index.Observations

---

Solow's Technological Change Data

Description
annual observations from 1909 to 1949

number of observations: 41
observation: country
country: United States

Usage
data(Solow)

Format
A time serie containing:
q output
k capital/labor ratio
a index of technology
Source


References


See Also


<table>
<thead>
<tr>
<th>Somerville</th>
<th>Visits to Lake Somerville</th>
</tr>
</thead>
</table>

Description

a cross-section from 1980

number of observations : 659

observation : individuals

country : United States

Usage

data(Somerville)

Format

A dataframe containing :

visits annual number of visits to lake Somerville

quality quality ranking score for lake Somerville

ski engaged in water-skiing at the lake ?

income annual household income

feeSom annual user fee paid at lake Somerville ?

costCon expenditures when visiting lake Conroe

costSom expenditures when visiting lake Somerville

costHoust expenditures when visiting lake Houston
SP500

Source


References


See Also

Index.Source, Index.Economics, Index.Econometrics, Index.Observations

SP500

Returns on Standard \& Poor’s 500 Index

Description

daily observations from 1981–01 to 1991–04

number of observations: 2783

Usage

data(SP500)

Format

A dataframe containing:

r500 daily return S\&P500 (change in log index)

References


See Also

**Star**  
*Effects on Learning of Small Class Sizes*

### Description

- A cross-section from 1985-89
- *number of observations*: 5748
- *observation*: individuals
- *country*: United States

### Usage

```r
data(Star)
```

### Format

A dataframe containing:

- `tmathssk` total math scaled score
- `treadssk` total reading scaled score
- `classk` type of class, a factor with levels (regular, small.class, regular.with.aide)
- `totexpk` years of total teaching experience
- `sex` a factor with levels (boy, girl)
- `freelunk` qualified for free lunch?
- `race` a factor with levels (white, black, other)
- `schidkn` school indicator variable

### Source


### References


### See Also

[Index.Source, Index.Economics, Index.Econometrics, Index.Observations]
### Strike Duration Data

**Description**

a cross-section from 1968 to 1976  
*number of observations*: 62  
*country*: United States

**Usage**

data(Strike)

**Format**

A dataframe containing:

- **duration**: strike duration in days  
- **prod**: unanticipated output

**Source**


**References**


**See Also**

- Index.Source, Index.Economics, Index.Econometrics, Index.Observations

---

### Strikes Duration

**Description**

a cross-section from 1968 to 1976  
*number of observations*: 566  
*country*: United States

**Usage**

data(StrikeDur)
Format

A dataframe containing:

- **dur**: duration of the strike in days
- **gdp**: measure of stage of business cycle (deviation of monthly log industrial production in manufacturing from prediction from OLS on time, time-squared and monthly dummies)

Source


References


See Also

*Index.Source, Index.Economics, Index.Econometrics, Index.Observations*

---

**StrikeNb**

*Number of Strikes in Us Manufacturing*

Description

monthly observations from 1968(1) to 1976 (12)

- **number of observations**: 108
- **observation**: country
- **country**: United States

Usage

`data(StrikeNb)`

Format

A time serie containing:

- **strikes**: number of strikes (number of contract strikes in U.S. manufacturing beginning each month)
- **output**: level of economic activity (measured as cyclical departure of aggregate production from its trend level)
- **time**: a time trend from 1 to 108
Source


References


See Also


---

### SumHes

**The Penn Table**

**Description**

a panel of 125 observations from 1960 to 1985

*number of observations*: 3250

*observation*: country

*country*: World

**Usage**

data(SumHes)

**Format**

A dataframe containing:

*year*  the year

*country*  the country name (factor)

*opec*  OPEC member ?

*com*  communist regime ?

*pop*  country’s population (in thousands)

*gdp*  real GDP per capita (in 1985 US dollars)

*sr*  saving rate (in percent)
Source


References


See Also

indexSource, indexEconomics, indexEconometrics, indexObservations, indexTimeSeries

---

**Tbrate**

*Interest Rate, GDP and Inflation*

**Description**

quarterly observations from 1950-1 to 1996-4

number of observations : 188

observation : country

country : Canada

**Usage**

data(Tbrate)

**Format**

A time serie containing :

- r the 91-day treasury bill rate
- y the log of real GDP
- pi the inflation rate

**Source**

CANSIM database of Statistics Canada.

**References**

Tobacco

See Also


Tobacco  Households Tobacco Budget Share

Description

a cross-section from 1995-96

number of observations : 2724

observation : individuals

country : Belgium

Usage

data(Tobacco)

Format

A dataframe containing :

occupation  a factor with levels (bluecol, whitecol, inactself), the last level being inactive and self-employed

region  a factor with levels (flanders, wallon, brussels)

nkids  number of kids of more than two years old

nkids2  number of kids of less than two years old

nadults  number of adults in household

lnx  log of total expenditures

stobacco  budgetshare of tobacco

salcohol  budgetshare of alcohol

age  age in brackets (0-4)

Source

National Institute of Statistics (NIS), Belgium.

References


See Also

Index.Source, Index.Economics, Index.Econometrics, Index.Observations
Stated Preferences for Train Traveling

Description

- a cross-section from 1987
- number of observations: 2929
- observation: individuals
- country: Netherland

Usage

data(Train)

Format

A dataframe containing:

- id: individual identifier
- choiceid: choice identifier
- choice: one of choice1, choice2
- pricez: price of proposition z (z=1,2) in cents of guilders
- timez: travel time of proposition z (z=1,2) in minutes
- comfortz: comfort of proposition z (z=1,2), 0, 1 or 2 in decreasing comfort order
- changez: number of changes for proposition z (z=1,2)

Source


References


See Also

*Index.Source, Index.Economics, Index.Econometrics, Index.Observations*
Description

A cross-section

number of observations: 25
observation: regional
country: United States

Usage

data(TranspEq)

Format

A dataframe containing:

state state name
va output
capital capital input
labor labor input
nfirm number of firms

Source


References


See Also

Index.Source, Index.Economics, Index.Econometrics, Index.Observations
Treatment

Evaluating Treatment Effect of Training on Earnings

Description

a cross-section from 1974

*number of observations*: 2675
*country*: United States

Usage

data(Treatment)

Format

A dataframe containing:

- **treat** treated?
- **age** age
- **educ** education in years
- **ethn** a factor with levels ("other","black","hispanic")
- **married** married?
- **re74** real annual earnings in 1974 (pre-treatment)
- **re75** real annual earnings in 1975 (pre-treatment)
- **re78** real annual earnings in 1978 (post-treatment)
- **u74** unemployed in 1974?
- **u75** unemployed in 1975?

Source


References


See Also

*Index.Source, Index.Economics, Index.Econometrics, Index.Observations*
Description

a cross-section

*number of observations*: 13705

*observation*: individuals

*country*: United States

Usage

data(Tuna)

Format

A dataframe containing:

- **hid**: individuals identifiers
- **id**: purchase identifiers
- **choice**: one of skw (Starkist water), cosw (Chicken of the sea water), pw (store-specific private label water), sko (Starkist oil), coso (Chicken of the sea oil)
- **price.z**: price of brand z

Source


References


See Also

*Index.Source, Index.Economics, Index.Econometrics, Index.Observations*
UnempDur

Unemployment Duration

Description


number of observations : 3343

Usage

data(UnempDur)

Format

A time serie containing:

- **spell** length of spell in number of two-week intervals
- **censor1** = 1 if re-employed at full-time job
- **censor2** = 1 if re-employed at part-time job
- **censor3** 1 if re-employed but left job: pt-ft status unknown
- **censor4** 1 if still jobless
- **age** age
- **ui** = 1 if filed UI claim
- **reprate** eligible replacement rate
- **disrate** eligible disregard rate
- **logwage** log weekly earnings in lost job (1985\$)
- **tenure** years tenure in lost job

Source


References


See Also

[Index.Source], [Index.Economics], [Index.Econometrics], [Index.Observations], [Index.Time.Series]
Description

a cross-section from 1993

number of observations : 452

observation : individuals

country : United States

Usage

data(Unemployment)

Format

A dataframe containing :

duration duration of first spell of unemployment, t, in weeks

spell 1 if spell is complete

race one of nonwhite, white

sex one of male, female

reason reason for unemployment, one of new (new entrant), lose (job loser), leave (job leaver), reentr (labor force reentrant)

search 'yes' if (1) the unemployment spell is completed between the first and second surveys and number of methods used to search > average number of methods used across all records in the sample, or, (2) for individuals who remain unemployed for consecutive surveys, if the number of methods used is strictly nondecreasing at all survey points, and is strictly increasing at least at one survey point

pubemp 'yes' if an individual used a public employment agency to search for work at any survey points relating to the individuals first unemployment spell

ftp1 1 if an individual is searching for full time work at survey 1

ftp2 1 if an individual is searching for full time work at survey 2

ftp3 1 if an individual is searching for full time work at survey 3

ftp4 1 if an individual is searching for full time work at survey 4

nobs number of observations on the first spell of unemployment for the record

Source

References


See Also

Index.Source, Index.Economics, Index.Econometrics, Index.Observations

<table>
<thead>
<tr>
<th>University</th>
<th>Provision of University Teaching and Research</th>
</tr>
</thead>
</table>

Description

a cross-section from 1988

number of observations: 62

observation: schools

country: United Kingdom

Usage

data(University)

Format

A dataframe containing:

- undstudents undergraduate students
- poststudents postgraduate students
- nassets net assets
- acnumbers academic numbers
- acrelnum academic related numbers
- clerenum clerical numbers
- compop computer operators
- techn technicians
- stfees student fees
- acpay academic pay
- acrelpay academic related pay
- secrpay secretarial pay
- admpay admin pay
- agresrk aggregate research rank
- furneq furniture and equipment
- landbuild land and buildings
- resgr research grants
Source

References

See Also
* Index.Source, Index.Economics, Index.Econometrics, Index.Observations*

---

**USclassifiedDocuments**  
*Official Secrecy of the United States Government*

**Description**
Data on classification activity of the United States government.

Fitzpatrick (2013) notes that the dramatic jump in derivative classification activity (DerivClassActivity) that occurred in 2009 coincided with "New guidance issued to include electronic environment". Apart from the jump in 2009, the DerivClassActivity tended to increase by roughly 12 percent per year (with a standard deviation of the increase in the natural logarithm of DerivClassActivity of 0.18).

**Usage**
data(USclassifiedDocuments)

**Format**
A dataframe containing:

- **year**  
  the calendar year

- **OCAuthority**  
  Number of people in the government designated as Original Classification Authorities for the indicated year.

- **OCActivity**  
  Original classification activity for the indicated year: These are the number of documents created with an original classification, i.e., so designated by an official Original Classification Authority.

- **TenYearDeclass**  
  Percent of OCActivity covered by the 10 year declassification rules.

- **DerivClassActivity**  
  Derivative classification activity for the indicated year: These are the number of documents created that claim another document as the authority for classification.
Details

The lag 1 autocorrelation of the first difference of the logarithms of DerivClassActivity through 2008 is -0.52. However, because there are only 13 numbers (12 differences), this negative correlation is not statistically significant.

Source


Examples

```r
##
## 1. plot DerivClassActivity
##
plot(DerivClassActivity~year, USclassifiedDocuments)
# Exponential growth?

plot(DerivClassActivity~year, USclassifiedDocuments, log='y')
# A jump in 2009 as discussed by Fitzpatrick (2013).
# Otherwise plausibly a straight line.

##
## 2. First difference?
##
plot(diff(log(DerivClassActivity))~year[-1], USclassifiedDocuments)
# Jump in 2009 but otherwise on distribution

##
## 3. autocorrelation?
##
acf(diff(log(USclassifiedDocuments$DerivClassActivity[sel])))
# lag 1 autocorrelation = (-0.52).
# However, with only 12 numbers,
# this is not statistically significant.
```

### USFinanceIndustry

**US Finance Industry Profits**

**Description**

A data.frame giving the profits of the finance industry in the United States as a proportion of total corporate domestic profits.
USFinanceIndustry

Usage
data(USFinanceIndustry)

Format
A data.frame with the following columns:

- **year**: integer year starting with 1929
- **CorporateProfitsAdj**: Corporate profits with inventory valuation and capital consumption adjustments in billions of current (not adjusted for inflation) US dollars
- **Domestic**: Domestic industries profits in billions
- **Financial**: Financial industries profits in billions
- **Nonfinancial**: Nonfinancial industries profits in billions
- **restOfWorld**: Profits of the "Rest of the world" in their contribution to US Gross Domestic Product in billions
- **FinanceProportion**: = Financial/Domestic

Details
This is extracted from Table 6.16 of the National Income and Product Accounts (NIPA) compiled by the Bureau of Economic Analysis of the United States federal government. This table comes in four parts, A (1929-1947), B (1948-1987), C (1987-2000), and D (1998-present). Parts A, B, C and D contain different numbers of data elements, but the first five have the same names and are the only ones used here. The overlap between parts C and D (1998-2000) have a root mean square relative difference of 0.7 percent; there were no differences between the numbers in the overlap period between parts B and C (1987).

This was created using the following command:
```r
demoDir <- system.file('demoFiles', package='Ecdat')
demoCsv <- dir(demoDir, pattern='csv$', full.names=TRUE)
nipa6.16 <- readNIPA(demoCsv)
USFinanceIndustry <- as.data.frame(nipa6.16)
names(USFinanceIndustry)
< c('year', 'CorporateProfitsAdj', 'Domestic', 'Financial', 'Nonfinancial', 'restOfWorld')
USFinanceIndustry$FinanceProportion <- with(USFinanceIndustry, Financial/Domestic)
```

Source
http://www.bea.gov: Under "U.S. Economic Accounts", first select "Corporate Profits" under "National". Then next to "Interactive Tables", select, "National Income and Product Accounts Tables". From there, select "Begin using the data...". Under "Section 6 - income and employment by industry", select each of the tables starting "Table 6.16". As of February 2013, there were 4 such tables available: Table 6.16A, 6.16B, 6.16C and 6.16D. Each of the last three are available in annual and quarterly summaries. The USFinanceIndustry data combined the first 4 rows of the 4 annual summary tables.

See Also
readNIPA
Examples

```r
data(USFinanceIndustry)
plot(FinanceProportion~year, USFinanceIndustry, type='b',
    ylim=c(0, max(FinanceProportion, na.rm=TRUE)),
    xlab='', ylab='', las=1, cex.axis=2, bty='n', lwd=2,
    col='blue')

# Write to a file for Wikimedia Commons
svg('USFinanceIndustry.svg')
plot(FinanceProportion~year, USFinanceIndustry, type='b',
    ylim=c(0, max(FinanceProportion, na.rm=TRUE)),
    xlab='', ylab='', las=1, cex.axis=2, bty='n', lwd=2,
    col='blue')
dev.off()
```

USStateAbbreviations  Standard abbreviations for states of the United States

Description

The object returned by `readUSstateAbbreviations()` on May 20, 2013.

Usage

data(USstateAbbreviations)

Format

A data.frame containing 10 different character vectors of names or codes for 76 different political entities including the United States, the 50 states within the US, plus the District of Columbia, US territories and other political designation, some of which are obsolete but are included for historical reference.

Name  The standard name of the entity.

Status  description of status, e.g., state / commonwealth vs. island, territory, military mail code, etc.

ISO, ANSI.letters, ANSI.digits, USPS, USCG, Old.GPO, AP, Other  Alternative abbreviations used per different standards. The most commonly used among these may be the 2-letter codes officially used by the US Postal Service (USPS).

Details

This was read from the Wikipedia article on "List of U.S. state abbreviations"

Source

the Wikipedia article on "List of U.S. state abbreviations"
See Also

readUSstateAbbreviations showNonASCII grepNonStandardCharacters subNonStandardCharacters

Examples

```r
##
## to use
##
data(UStaxWords)

##
## to update
##

USstateAbb2 <- readUSstateAbbreviations()
```

<table>
<thead>
<tr>
<th>UStaxWords</th>
<th>Number of Words in US Tax Law</th>
</tr>
</thead>
</table>

**Description**

Thousands of words in US tax law for 1995 to 2005 in 10 year intervals. This includes income taxes and all taxes in the code itself (written by congress) and regulations (written by government administrators).

**Usage**

```r
data(UStaxWords)
```

**Format**

A `data.frame` containing:

- `year` tax year
- `IncomeTaxCode` number of words in thousands in the US income tax code
- `otherTaxCode` number of words in thousands in US tax code other than income tax
- `EntireTaxCode` number of words in thousands in the US tax code
- `IncomeTaxRegulations` number of words in thousands in US income tax regulations
- `otherTaxRegulations` number of words in thousands in US tax regulations other than income tax
- `IncomeTaxCodeAndRegs` number of words in thousands in both the code and regulations for the US income tax
- `otherTaxCodeAndRegs` number of words in thousands in both code and regulations for US taxes apart from income taxes.
- `EntireTaxCodeAndRegs` number of words in thousands in US tax code and regulations
Details

Thousands of words in the US tax code and federal tax regulations, 1955-2005. This is based on data from the Tax Foundation (taxfoundation.org), adjusted to eliminate an obvious questionable observation in otherTaxRegulations for 1965. This series was not reported directly by the Tax Foundation but is easily computed as the difference between their Income and Entire tax numbers. This series shows the numbers falling by 48 percent between 1965 and 1975 and by 1.5 percent between 1995 and 2005. These are the only declines seen in these numbers and seem inconsistent with the common concern (expressed e.g., in Moody, Warcholik and Hodge, 2005) about the difficulties of simplifying any governmental program, because vested interest appear to defend almost anything.

The decline of 48 percent seems more curious for two additional reasons: First, it was preceded by a tripling of otherTaxRegulations between 1955 and 1965. Second, it was NOT accompanied by any comparable behavior of otherTaxCode. Instead, the latter grew each decade by between 17 and 53 percent, similar to but slower than the growth in IncomeTaxCode and IncomeTaxRegulations.

Accordingly, otherTaxRegulations for 1965 is replaced by the average of the numbers for 1955 and 1975, and EntireTaxRegulations for 1965 is comparably adjusted. This replaces (1322, 2960) for those two variables for 1965 with (565, 2203). In addition, otherTaxCodeAndRegs and EntireTaxCodeAndRegulations are also changed from (1626, 3507) to (870, 2751).

Independent of whether this adjustment is correct or not, it’s clear that there have been roughly 3 words of regulations for each word in the tax code. Most of these are income tax regulations, which have recently contained 4.5 words for every word in code. The income tax code currently includes roughly 50 percent more words than other tax code.

Author(s)

Spencer Graves

Source

Tax Foundation: Number of Words in Internal Revenue Code and Federal Tax Regulations, 1955-2005

References


Examples

data(UStaxWords)
plot(EntireTaxCodeAndRegs/1000 ~ year, UStaxWords, type='b',
    ylab='Millions of words in US tax code & regs')

# Write to a file for Wikimedia Commons
svg('UStaxWords.svg')

matplot(UStaxWords$year, UStaxWords[c(2:3, 5:6)]/1000,
    type='b', bty='n', ylab='',
    ylim=c(0, max(UStaxWords$EntireTaxCodeAndRegs)/1000),
    las=1, xlab='''', cex.axis=2)
Medical Expenses in Vietnam (household Level)

Description

a cross-section from 1997

number of observations : 5999
observation : households
country : Vietnam

Usage

data(VietNamH)

Format

A dataframe containing :

sex  gender of household head (male,female)
age  age of household head
educyr  schooling year of household head
farm  farm household ?
urban  urban household ?
hhsize  household size
intotal log household total expenditure
lnmed log household medical expenditure
lnrlfood log household food expenditure
lnexp12m log of total household health care expenditure for 12 months
commune commune

Source

References

See Also
Index.Source, Index.Economics, Index.Econometrics, Index.Observations

---

VietNamI Medical Expenses in Viet–nam (individual Level)

Description

a cross-section from 1997

*number of observations*: 27765

*observation*: individuals

*country*: Vietnam

Usage
data(VietNamI)

Format

A dataframe containing:

pharvis number of direct pharmacy visits
lnhhexp log of total medical expenditure
age age of household head
sex gender (male,female)
married married ?
educ completed diploma level ?
illness number of of illnesses experiences in past 12 months
Wages

**Source**

**References**

**See Also**
*Index.Source, Index.Economics, Index.Econometrics, Index.Observations*

---

<table>
<thead>
<tr>
<th>Wages</th>
<th>Panel Datas of Individual Wages</th>
</tr>
</thead>
</table>

**Description**
a panel of 595 observations from 1976 to 1982

- number of observations: 4165
- observation: individuals
- country: United States

**Usage**
data(Wages)

**Format**
A dataframe containing:

- exp years of full-time work experience
- wks weeks worked
- bluecol blue collar?
- ind works in a manufacturing industry?
- south resides in the south?
- smsa resides in a standard metropolitan statistical area?
- married married?
sex  a factor with levels (male,female)
union individual’s wage set by a union contract?
ed years of education
black is the individual black?
lwage logarithm of wage

Source
Panel study of income dynamics.

References

See Also

<table>
<thead>
<tr>
<th>Wages1</th>
<th>Wages, Experience and Schooling</th>
</tr>
</thead>
</table>

Description
a panel of 595 observations from 1976 to 1982

*number of observations*: 3294
*observation*: individuals
*country*: United States

Usage
data(Wages1)

Format
A time serie containing :

**exper** experience in years
**sex** a factor with levels (male,female)
**school** years of schooling
**wage** wage (in 1980 \$) per hour
References


See Also


---

<table>
<thead>
<tr>
<th>Workinghours</th>
<th>Wife Working Hours</th>
</tr>
</thead>
</table>

**Description**

- a cross-section from 1987
- *number of observations*: 3382
- *observation*: individuals
- *country*: United States

**Usage**

`data(Workinghours)`

**Format**

A dataframe containing:

- **hours** wife working hours per year
- **income** the other household income in hundreds of dollars
- **age** age of the wife
- **education** education years of the wife
- **child5** number of children for ages 0 to 5
- **child13** number of children for ages 6 to 13
- **child17** number of children for ages 14 to 17
- **nonwhite** non–white?
- **owned** is the home owned by the household?
- **mortgage** is the home on mortgage?
- **occupation** occupation of the husband, one of mp (manager or
- **unemp** local unemployment rate in %
Source

References

See Also
Index.Source, Index.Economics, Index.Econometrics, Index.Observations

<table>
<thead>
<tr>
<th>Yen</th>
<th>Yen-dollar Exchange Rate</th>
</tr>
</thead>
</table>

Description
weekly observations from 1975 to 1989
number of observations : 778
observation : country
country : Japan

Usage
data(Yen)

Format
A dataframe containing :

date the date of the observation (19850104 is January, 4, 1985)
s the ask price of the dollar in units of Yen in the spot market on friday of the current week
f the ask price of the dollar in units of Yen in the 30-day forward market on friday of the current week
s30 the bid price of the dollar in units of Yen in the spot market on the delivery date on a current forward contract

Source

References
Yogurt

See Also


<table>
<thead>
<tr>
<th>Yogurt</th>
<th>Choice of Brand for Yogurts</th>
</tr>
</thead>
</table>

**Description**

- a cross-section
- number of observations: 2412
- observation: individuals
- country: United States

**Usage**

data(Yogurt)

**Format**

A dataframe containing:

- id  individuals identifiers
- choice one of yoplait, dannon, hiland, weight (weight watcher)
- feat.z is there a newspaper feature advertisement for brand z?
- price.z price of brand z

**Source**


**References**


**See Also**

Index.Source, Index.Economics, Index.Econometrics, Index.Observations
Index

*Topic datasets
  Accident, 4
  Airline, 5
  Airq, 6
  bankingCrises, 7
  Benefits, 8
  Bids, 9
  BudgetFood, 10
  BudgetItaly, 11
  BudgetUK, 12
  Bwages, 13
  Capm, 14
  Car, 15
  Caschool, 16
  Catsup, 17
  Cigar, 18
  Cigarette, 19
  Clothing, 20
  Computers, 21
  Consumption, 22
  CPSch3, 23
  Cracker, 24
  CRANpackages, 25
  Crime, 26
  CRSPday, 27
  CRSPmon, 28
  Diamond, 29
  DM, 30
  Doctor, 31
  DoctorAUS, 32
  DoctorContacts, 33
  Earnings, 34
  Electricity, 35
  Fair, 36
  Fatality, 37
  FinancialCrisisFiles, 38
  Fishing, 39
  Forward, 40
  FriendFoe, 41

  Garch, 42
  Gasoline, 43
  Griliches, 44
  Grunfeld, 45
  HC, 46
  Hdma, 47
  Heating, 48
  Hedonic, 49
  HI, 50
  Housing, 51
  Hstarts, 52
  Icecream, 53
  incomeInequality, 54
  IncomeUK, 59
  Irates, 73
  Journals, 74
  Kakadu, 75
  Ketchup, 77
  Klein, 78
  LaborSupply, 79
  Labour, 80
  Longley, 80
  LT, 81
  Macrodat, 82
  Males, 83
  ManufCost, 84
  Mathlevel, 85
  MCAS, 86
  MedExp, 87
  Metal, 88
  Mishkin, 89
  Mode, 90
  ModeChoice, 91
  Mofa, 92
  Money, 93
  MoneyUS, 94
  Mpyr, 95
  Mroz, 96
  MunExp, 97
INDEX

MW, 98
NaturalPark, 99
Nerlove, 100
nonEnglishNames, 101
OFP, 101
Oil, 103
Orange, 104
Participation, 105
PatentsSHG, 106
PatentsRD, 107
PE, 108
politicalKnowledge, 109
Pound, 111
PPP, 112
Pricing, 113
Produc, 114
PSID, 115
RetSchool, 116
Schooling, 118
Solow, 119
Somerville, 120
SPS00, 121
Star, 122
Strike, 123
StrikeDur, 123
StrikeNb, 124
SumHes, 125
Tbrate, 126
Tobacco, 127
Train, 128
TranspEq, 129
Treatment, 130
Tuna, 131
UnempDur, 132
Unemployment, 133
University, 134
USClassifiedDocuments, 135
USFinanceIndustry, 136
USStateAbbreviations, 138
UStaxWords, 139
VietNamH, 141
VietNamI, 142
Wages, 143
Wages1, 144
Workinghours, 145
Yen, 146
Yogurt, 147

*Topic documentation
FriendFoe, 41, 63, 67, 69
Garch, 42, 62, 63, 65, 71, 72
Gasoline, 43, 61, 62, 66, 69
grepNonStandardCharacters, 101, 139
Griliches, 44, 64, 67, 71
Grunfeld, 45, 61, 65, 68–70
HC, 46, 61, 62, 66, 69
Hdma, 47, 60, 63, 67, 71
Heating, 48, 61, 62, 66, 69
Hedonic, 49, 63, 68, 69
HI, 50, 60, 64, 67, 68
Housing, 51, 64, 66, 68, 72
Hstarts, 52, 62, 64, 66, 70, 73
Icecream, 53, 62, 66, 72
incomeInequality, 54
IncomeUK, 59, 62, 64, 66, 72, 73
Irates, 62, 63, 66, 72, 73, 73
Journals, 64, 66, 71, 74
Kacadu, 61, 63, 67, 69, 75
Ketchup, 61, 65, 67, 69, 77
Klein, 61, 64, 66, 70, 72, 78
LaborSupply, 61, 64, 69, 70, 79
Labour, 64, 68, 72, 80
Longley, 64, 66, 70, 72, 80
LT, 62, 63, 66, 71, 72, 81
Macrodat, 62, 64, 66, 71, 73, 82
Males, 61, 64, 67, 68, 72, 83
ManufCost, 61, 65, 66, 71, 72, 84
Mathlevel, 61, 63, 67, 69, 85
MCAS, 63, 68, 71, 86
MedExp, 63, 70, 87
Metal, 65, 68, 71, 88
Mishkin, 62, 64, 66, 71, 73, 89
Mode, 61, 63, 67, 69, 90
ModeChoice, 61, 63, 67, 71, 91
Mofa, 60, 65, 66, 69, 92
Money, 64, 66, 70, 73, 93
MoneyUS, 62, 64, 69, 72, 73, 94
Mpyr, 62, 64, 66, 71, 72, 95
Mroz, 60, 64, 67, 71, 96
MunExp, 61, 68, 69, 71, 97
MW, 62, 64, 66, 70, 73, 98
NaturalPark, 61, 63, 67, 72, 99
Nerlove, 61, 65, 68, 71, 100
nonEnglishNames, 101
OFP, 60, 63, 67, 69, 70, 101
Oil, 60, 65, 68, 69, 103
Orange, 62, 65, 66, 71, 73, 104
Participation, 60, 64, 67, 69, 70, 105
PatentsGH, 60, 65, 68, 70, 106
PatentsRD, 60, 65, 68, 69, 72, 107
PE, 62, 64, 66, 72, 108
politicalKnowledge, 109
Pound, 62, 63, 66, 71, 73, 111
PPP, 62, 63, 66, 72, 73, 112
Pricing, 62, 63, 72, 73, 113
Produc, 61, 64, 68, 69, 114
PSID, 64, 67, 70, 115
read.xls, 39
readFinancialCrisisFiles, 7
readNIPA, 137
readUSSstateAbbreviations, 139
RetSchool, 64, 67, 70, 116
Schooling, 64, 67, 72, 118
showNonASCII, 139
Solow, 62, 64, 66, 71, 72, 119
Somerville, 60, 63, 67, 69, 70, 120
SP500, 72, 121
Star, 63, 67, 71, 122
Strike, 60, 64, 71, 123
StrikeDur, 60, 64, 70, 123
StrikeNb, 60, 64, 66, 70, 73, 124
subNonStandardCharacters, 101, 139
SumHes, 61, 65, 66, 71, 125
Tbrate, 62, 65, 66, 70, 73, 126
Tobacco, 60, 63, 67, 72, 127
Train, 60, 63, 67, 69, 128
TranspEq, 65, 68, 71, 129
Treatment, 64, 70, 130
Tuna, 61, 65, 67, 69, 131
UnempDur, 60, 64, 70, 132
Unemployment, 60, 64, 67, 69, 133
University, 61, 65, 68, 69, 134
USclassifiedDocuments, 135
USFinanceIndustry, 136
USstateAbbreviations, 138
USTaxWords, 139
VietNamH, 63, 66, 70, 141
VietNamI, 63, 67, 70, 142
Wages, 61, 64, 67, 69, 143
Wages1, 64, 67, 72, 144
Workinghours, 60, 64, 67, 69, 145
Yen, 62, 63, 66, 71, 73, 146
Yogurt, 61, 65, 67, 69, 147