Make the SLR model for each dataset provided.

Explain all the steps that you performed.

Play with the proportions of training and test sets. How do different proportions affect on results?

Check the documentation on LinearRegression class on sklearn libraries website. Are there any other parameters which you can tune?

Below are the explanations to the datasets (the last columns are outputs which you need to predict):

AirPassengers R Documentation

Monthly Airline Passenger Numbers 1949-1960

Description

The classic Box & Jenkins airline data. Monthly totals of international airline passengers, 1949 to 1960.

Usage

```
AirPassengers
```

Format

A monthly time series, in thousands.

Source

Box, G. E. P., Jenkins, G. M. and Reinsel, G. C. (1976) *Time Series Analysis, Forecasting and Control.* Third Edition. Holden-Day. Series G.

---

bigcity R Documentation

Population of U.S. Cities

Description

The `bigcity` data frame has 49 rows and 2 columns.

The `city` data frame has 10 rows and 2 columns.

The measurements are the population (in 1000's) of 49 U.S. cities in 1920 and 1930.
The 49 cities are a random sample taken from the 196 largest cities in 1920.
The `city` data frame consists of the first 10 observations in `bigcity`.

Usage

```
bigcity
```

Format

This data frame contains the following columns:

`u`

: The 1920 population.

`x`

: The 1930 population.

Source

The data were obtained from

Cochran, W.G. (1977) *Sampling Techniques*. Third edition. John Wiley

References

Davison, A.C. and Hinkley, D.V. (1997) *Bootstrap Methods and Their Application*. Cambridge University Press.

CanPop R Documentation

Canadian Population Data

Description

The `CanPop` data frame has 16 rows and 2 columns. Decennial time-series of Canadian population, 1851–2001.

Usage

```
CanPop
```

Format

This data frame contains the following columns:

year

census year.

population

Population, in millions

Source

Urquhart, M. C. and Buckley, K. A. H. (Eds.) (1965) *Historical Statistics of Canada*. Macmillan, p. 1369.

Canada (1994) *Canada Year Book*. Statistics Canada, Table 3.2.

Statistics Canada: http://www12.statcan.ca/english/census01/products/standard/popdwell/Table-PR.cfm.

References

Fox, J. (2008) *Applied Regression Analysis and Generalized Linear Models*, Second Edition. Sage.

USPop R Documentation

Population of the United States

Description

The `USPop` data frame has 22 rows and 1 columns. This is a decennial time-series, from 1790 to 2000.

Usage

```
USPop
```

Format

This data frame contains the following columns:

year

census year.

population

Population in millions.

Source

U.S.~Census Bureau: http://www.census-charts.com/Population/pop-us-1790-2000.html, downloaded 1 May 2008.

References

Fox, J. (2008) *Applied Regression Analysis and Generalized Linear Models*, Second Edition. Sage.