Assignment Project Exam Help

Predictive Analytics

Week 11: Time Series Forecasting

https://powcoder.com

Semester 2, 2018

Add WeChat powcoder

Discipline of Business Analytics, The University of Sydney Business School

QBUS2820 content structure

Assignment Project Exam Help

1. Statistical and Machine Learning foundations and applications.
2. Advanced regression methods.
3. Classification methods.
4. **Time series forecasting**

https://powcoder.com

Add WeChat powcoder

Week 11: Time Series Forecasting

1. Problem definition
2. Time series patterns
3. Simple forecasting methods
4. Model diagnostics
5. Model validation
6. Random walk model

Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

Time series

A **time series** is a set of observations $y_1, y_2, \ldots, y_t$ ordered in time.

Assignment Project Exam Help

Examples:

- Weekly unit sales of a product.

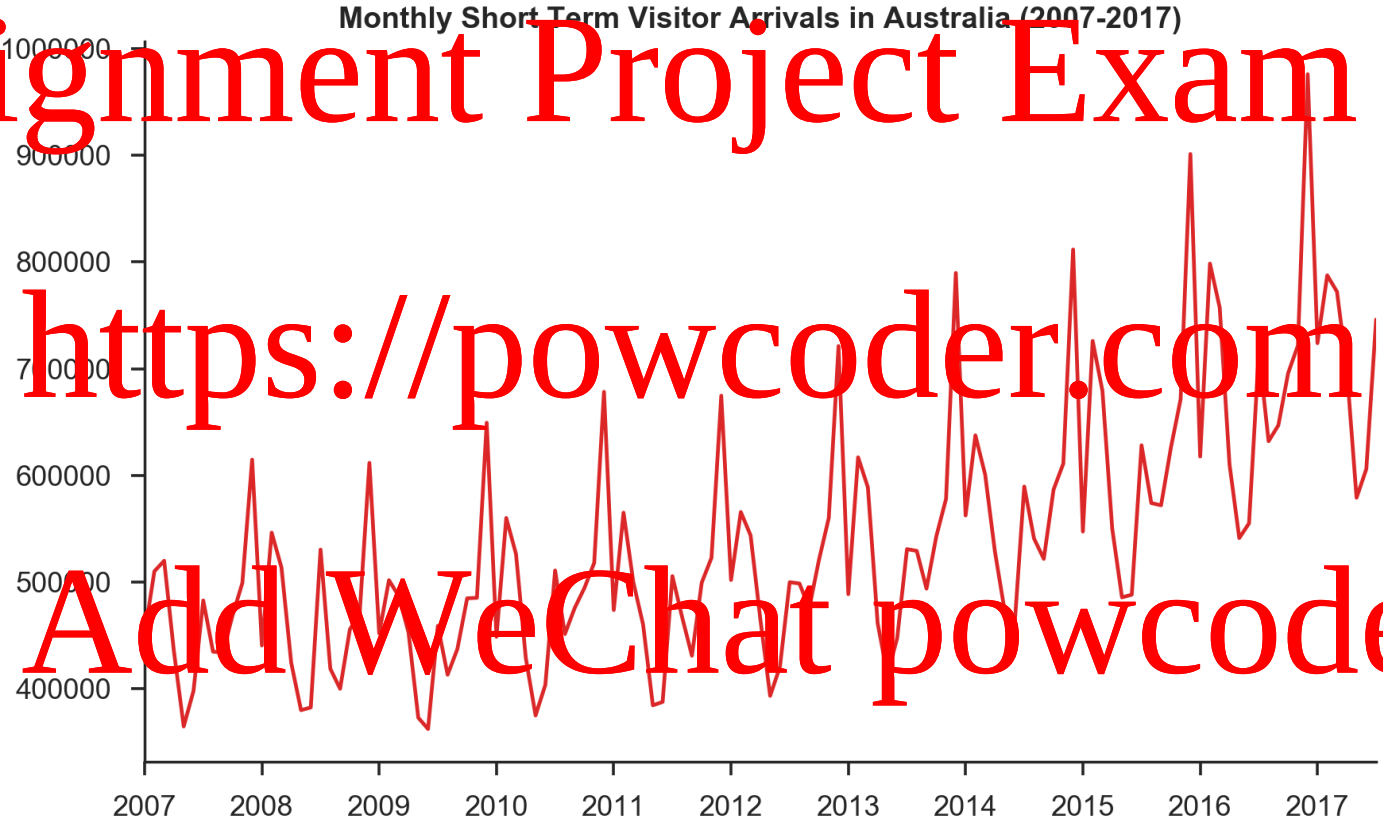
https://powcoder.com

- Unemployment rate in Australia each quarter.
- Daily production levels of a product.

Add WeChat powcoder

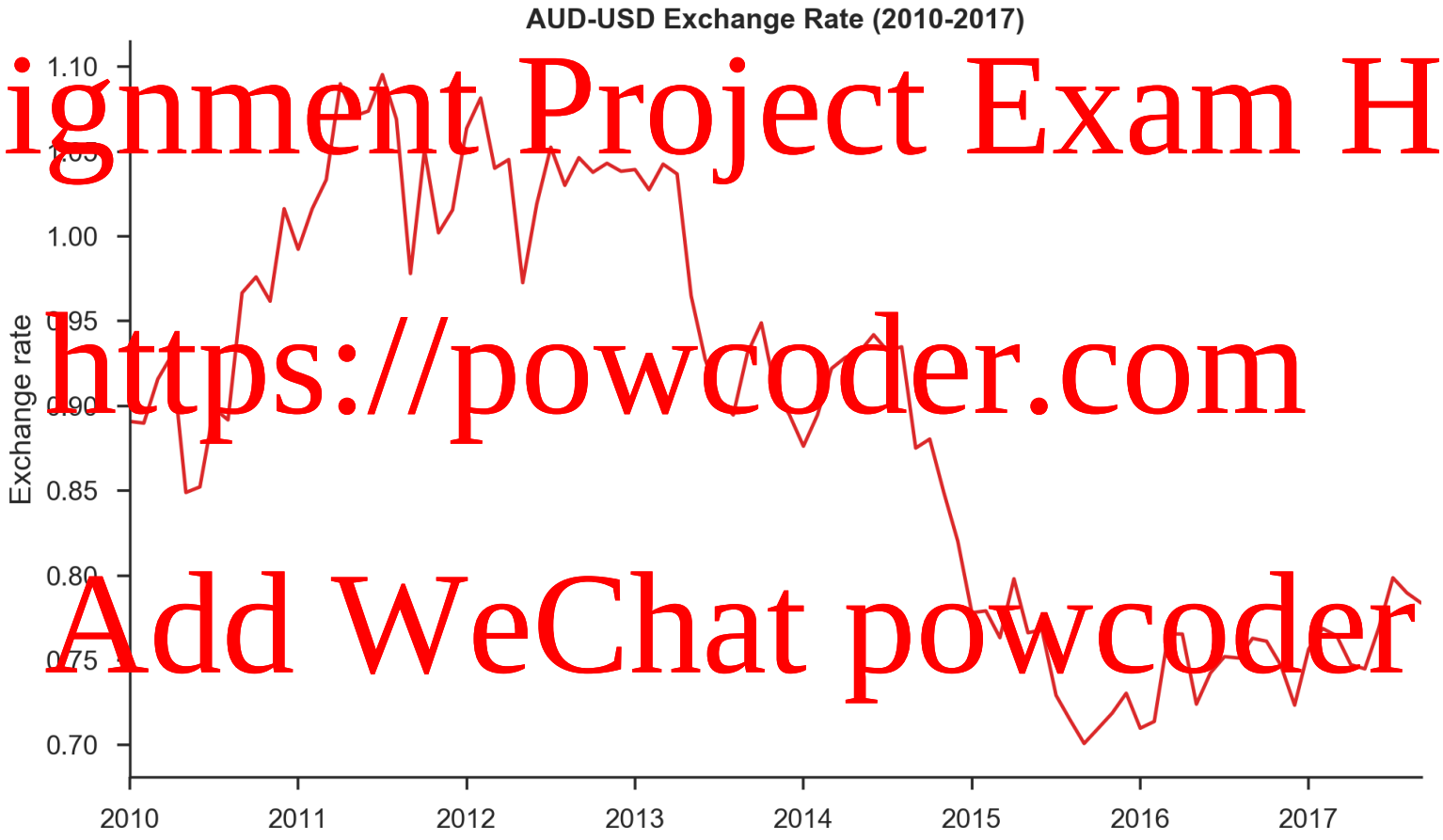
- Average annual temperature in Sydney.
- 5 minute prices for CBA stock on the ASX.

Example: visitor Arrivals in Australia



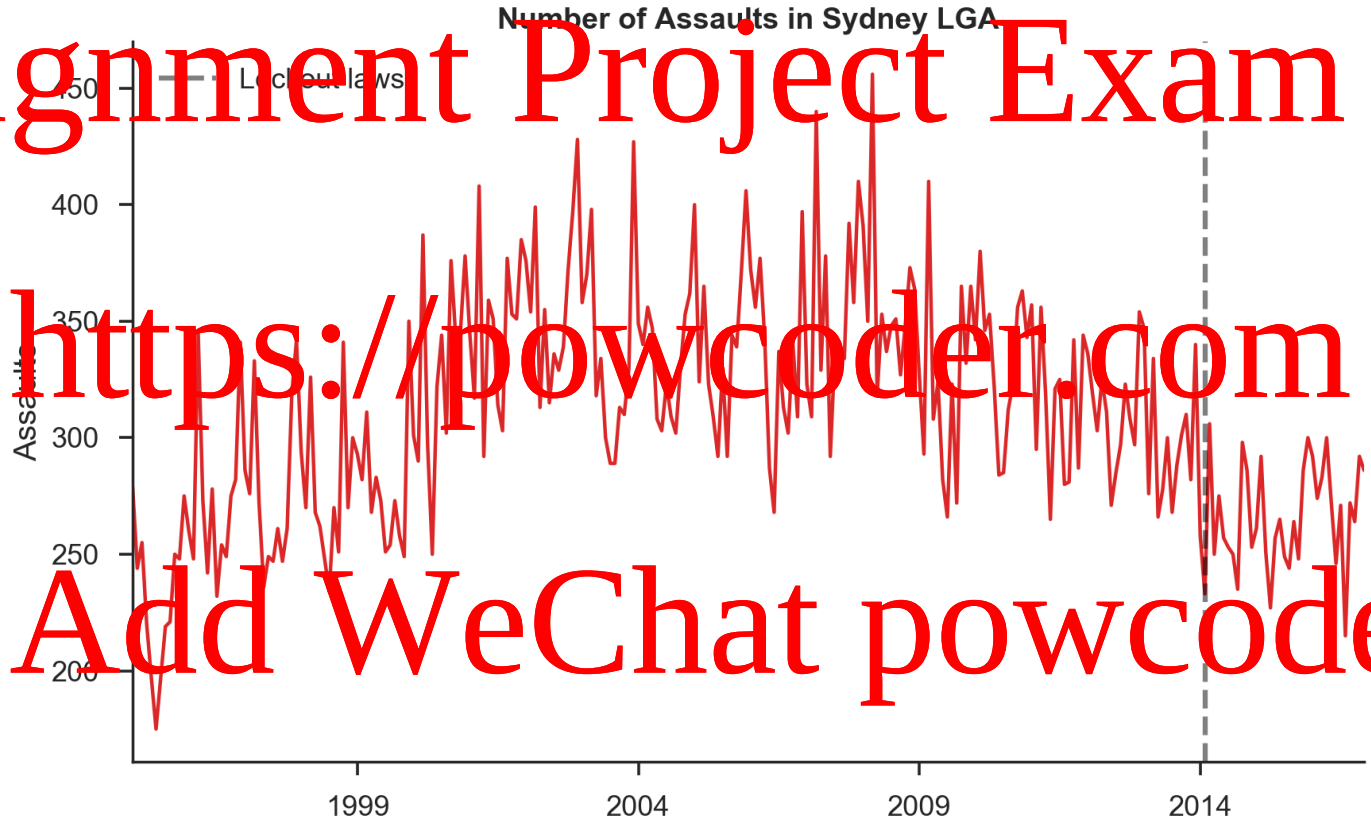
Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

Example: AUD/USD exchange rate



Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

Example: assaults in Sydney



Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

Forecasting

Assignment Project Exam Help

A **forecast** is a prediction about future events and conditions given all current information, including historical data and knowledge of any future events that might impact these events.

https://powcoder.com

The act of making such predictions is called **forecasting**.

Forecasting informs business and economic decision making, planning, government policy, etc.

Add WeChat powcoder

Examples

- Governments need to forecast unemployment, interest rates, expected revenues from income taxes to formulate policies.
- Retail stores need to forecast demand to control inventory levels, hire employees and provide training.
- Banks/Investors/financial analysts need to forecast financial returns, risk or volatility, market 'timing'.
- University administrators need to forecast enrollments to plan for facilities and for faculty recruitment.
- Sports organisations need to project sports performance, crowd figures, club gear sales, revenues, etc. in the coming season.

Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

Forecasting in business

Different problems lead to different approaches under the umbrella of forecasting.

Assignment Project Exam Help

- Quantitative (data based) forecasting (our focus in this unit).

https://powcoder.com

- Qualitative (judgmental) forecasting.

Add WeChat powcoder

- Prediction markets.
- Common approach: judgmentally adjusted statistical forecasting.

Assignment Project Exam Help

Problem definition

https://powcoder.com

Add WeChat powcoder

Forecasting

Assignment Project Exam Help

Our objective is to predict the value of a time indexed response variable at a future point $t + h$, given the observed series until the present point t. That is, we want to predict Y_{t+h} given $y_1, y_2, \ldots, y_t$, where h is the **forecast horizon**.

https://powcoder.com

We can extend this setting to allow for the presence of predictors $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_t$, leading to a **dynamic regression** problem.

Add WeChat powcoder

Decision theory

Assignment Project Exam Help

We denote a **point forecast** as $\widehat{Y}_t = f(Y_{1:t})$. As before, we assume a squared error loss function:

https://powcoder.com

$$L(Y_{t+h}, f(Y_{1:t}) = (Y_{t+h} - f(Y_{1:t}))^2$$

Add WeChat powcoder

We use the slice notation $Y_{1:t}$ as a compact way to write $Y_1, \ldots, Y_t$.

Point forecasting (key concept)

Assignment Project Exam Help

Using the arguments from earlier earlier in the unit, the optimal point forecast under the squared error loss is the conditional expectation:

https://powcoder.com

$$f(Y_{1:t}) = E(Y_{t+h}|Y_{1:t})$$

Add WeChat powcoder

Our objective is therefore to approximate the conditional expectation of Y_{t+h} given the historical data, possible for multiple values of h.

Interval forecasting (key concept)

Assignment Project Exam Help

Uncertainty quantification is an essential for business forecasting.

https://powcoder.com

A **density forecast** $\widehat{p}(Y_{t+h}|y_1,\ldots,y_t)$ is an estimate of the entire conditional density $p(Y_{t+h}|y_1,\ldots,y_t)$.

Add WeChat powcoder

An **interval forecast** is an interval $(\widehat{y}_{t+h,L}, \widehat{y}_{t+h,U})$ such that $\widehat{P}(\widehat{y}_{t+h,L} < Y_{t+h} < y_{t+h,U}) = 1-\alpha$.

Fan chart (key concept)

Assignment Project Exam Help

- For consecutive forecast horizons, construct prediction intervals for different probability levels (say, 75%, 90%, and 99%) and plot the using different shades.

https://powcoder.com

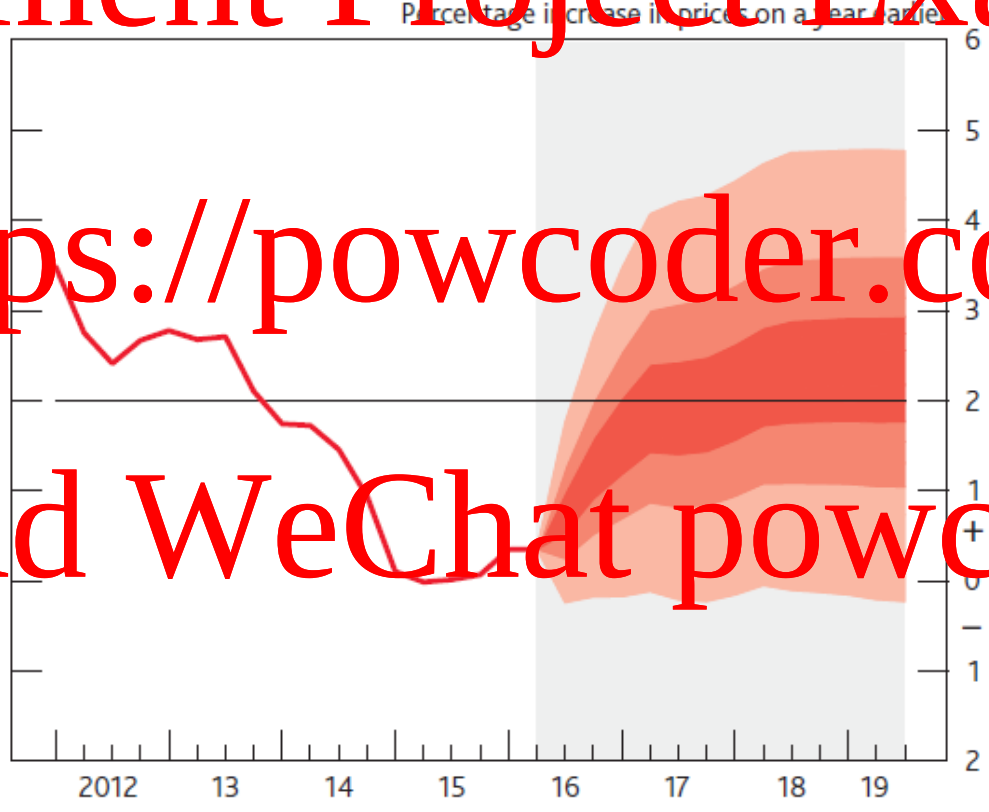
- The intervals typically get wider with the horizon, representing increasing uncertainty about future values.

Add WeChat powcoder

- Fan charts are useful tools for presenting forecasts.

Example: fan chart

Chart 5.12 CPI inflation projection based on constant nominal interest rates at 0.25%, other policy actions as announced



Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

Assignment Project Exam Help

Time series patterns

https://powcoder.com

Add WeChat powcoder

Time series patterns (key concept)

We interpret a time series as

$$Y_t = f(T_t, S_t, C_t, E_t),$$

where T_t is the trend component, S_t is the seasonal component, C_t is the cyclic component, and E_t is an irregular or error component.

Trend. The systematic long term increase or decrease in the series.

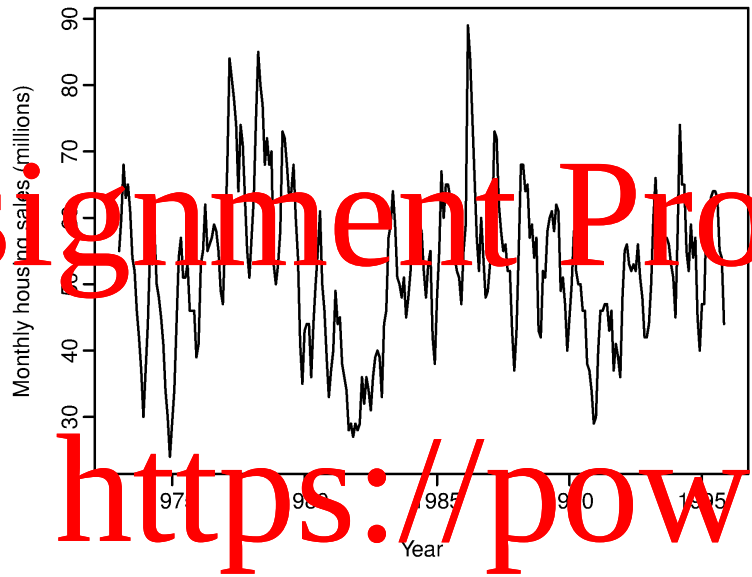
Seasonal. A systematic change in the mean of the series due to seasonal factors (month, day of the week, etc).

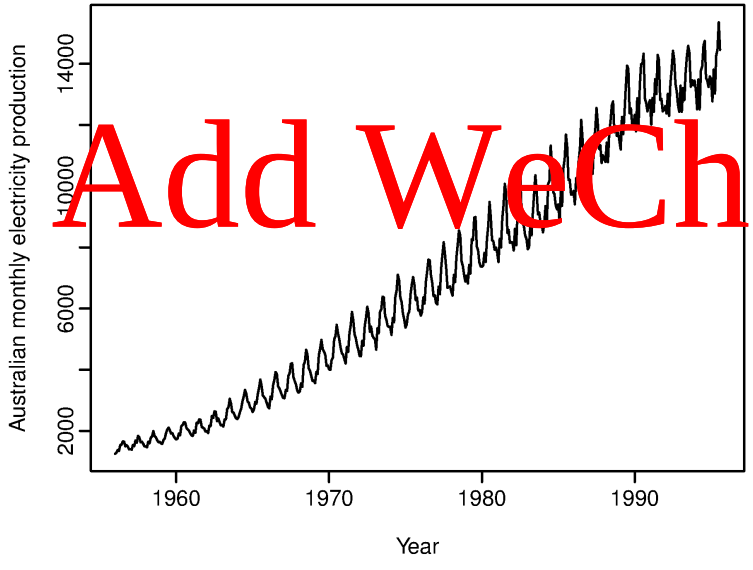
Cyclic. A cyclic pattern exists when there are medium or long run fluctuations in the time series that are not of a fixed period.

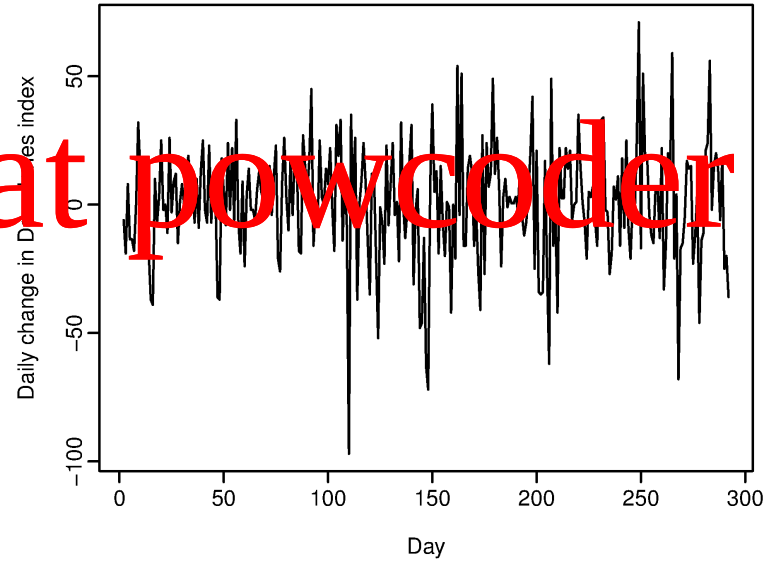
Irregular. Short term fluctuations and noise.

Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

Examples: time series patterns



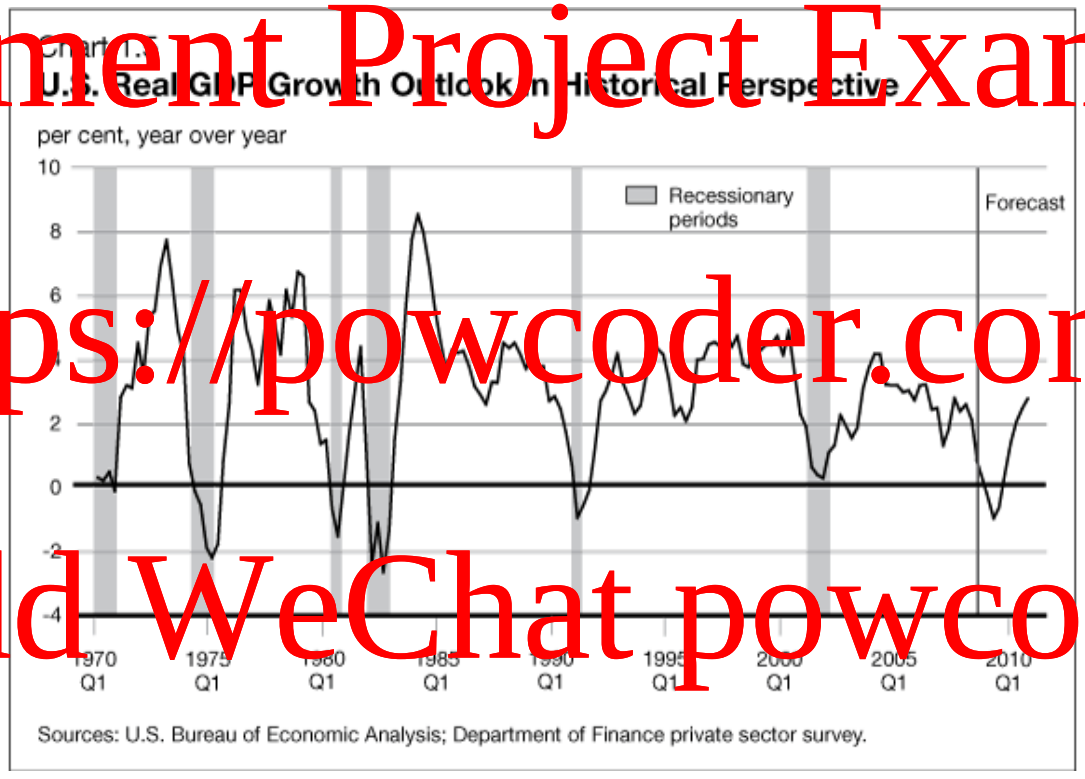




Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

Example: cyclic series

Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder



Time series models

Assignment Project Exam Help

Time series models can be additive or multiplicative.

https://powcoder.com

Additive: $Y_t = T_t + S_t + E_t$

Add WeChat powcoder

Multiplicative: $Y_t = T_t \times S_t \times E_t$

Log transformation

Assignment Project Exam Help

When the time series displays a multiplicative behaviour

$$Y_t = T_t \times S_t \times E_t,$$

https://powcoder.com

we usually apply to log-transformation to obtain a more convenient additive specification

Add WeChat powcoder

$$\log Y_t = \log T_t + \log S_t + \log E_t.$$

Choosing an additive or multiplicative specification

Assignment Project Exam Help

Is the seasonal variation proportional to the trend?

https://powcoder.com

- If yes, a multiplicative model is more adequate.
- If not, we use an additive model.

Add WeChat powcoder

Time series decomposition

Assignment Project Exam Help

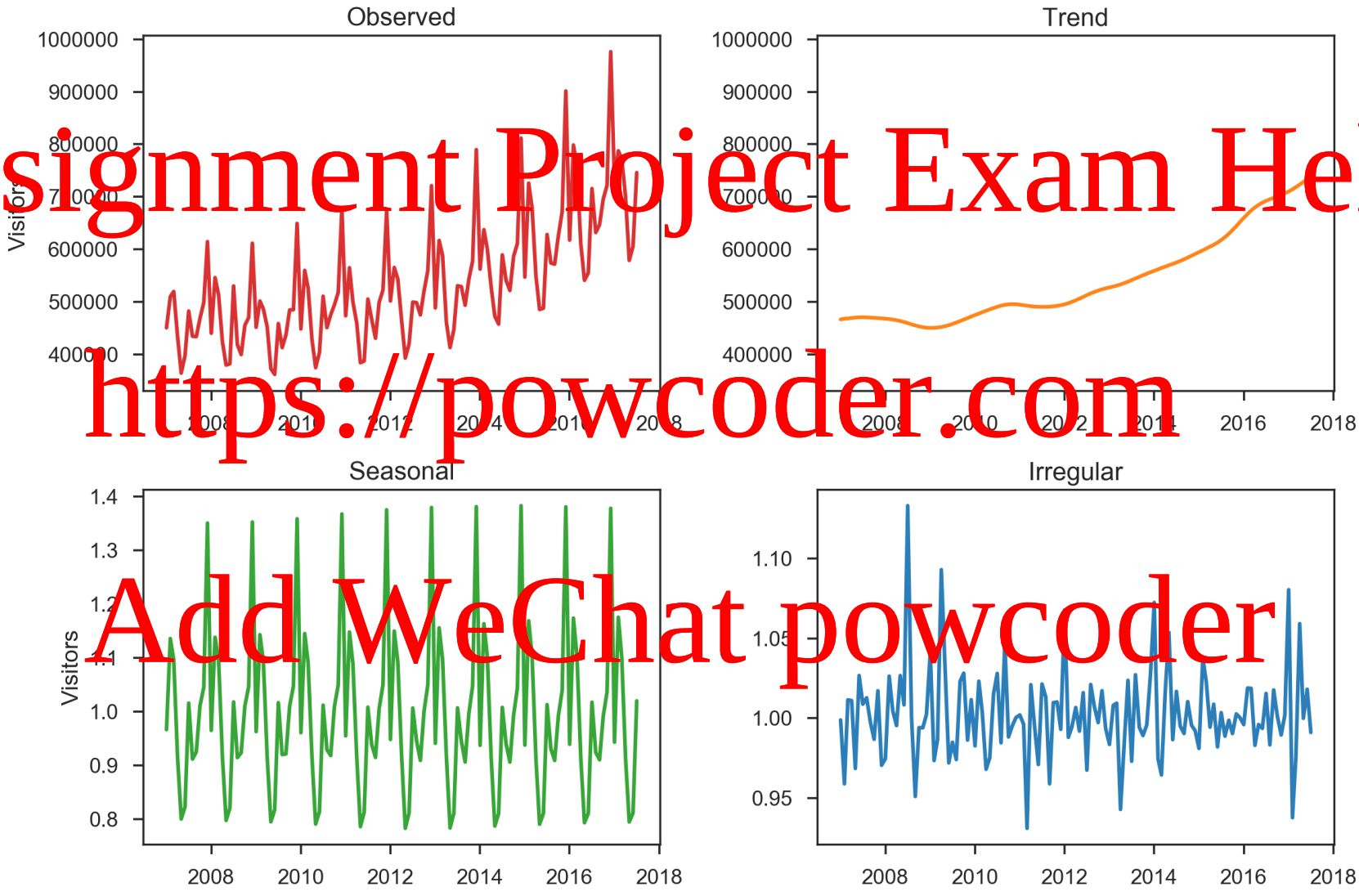
Time series decomposition methods are algorithms for splitting a time series into different components, typically for purposes of seasonal adjustment.

https://powcoder.com

In the context of forecasting, decomposition methods are useful tools for exploratory data analysis, allowing us to visualise patterns in the data.

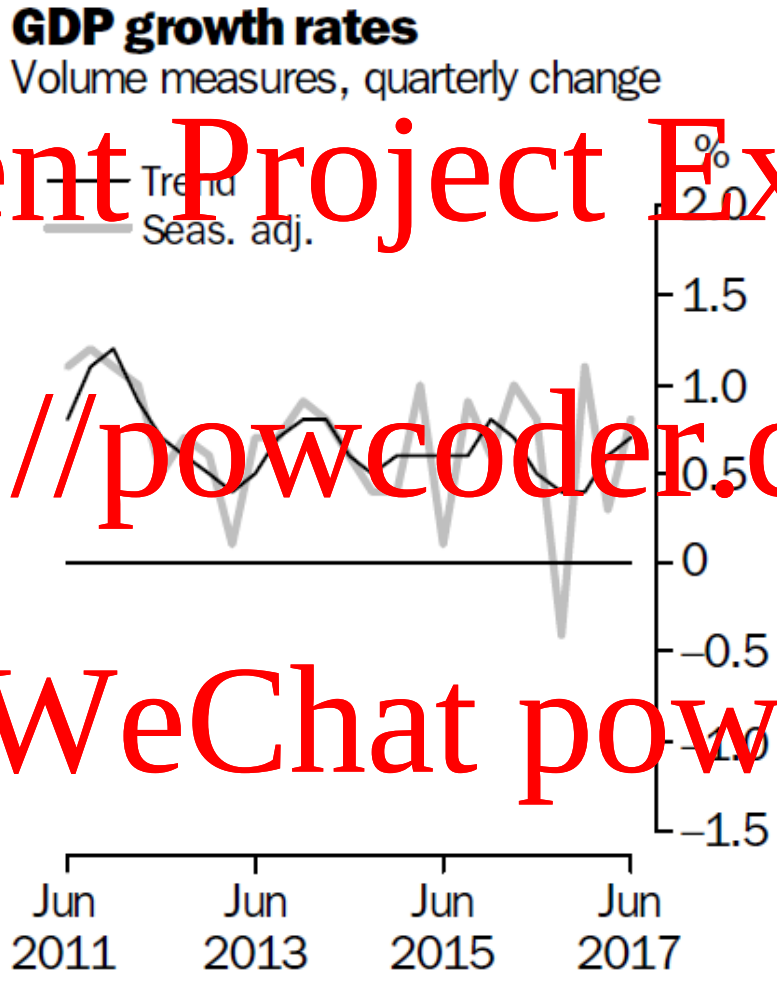
Add WeChat powcoder

Time series decomposition: visitor arrivals



Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

Example: seasonal adjustment and trend extraction



Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

Assignment Project Exam Help

Simple forecasting methods

https://powcoder.com

Add WeChat powcoder

Random walk

Assignment Project Exam Help

The **random walk method** (called the naïve method in the book) forecasts the series using the value of the last available observation:

https://powcoder.com

$$\widehat{y}_{t+h} = y_t$$

Add WeChat powcoder

Seasonal random walk

For time series with seasonal patterns, we can extend the random walk method by forecasting the series with the value of the last available observation *in the same season*:

$$\widehat{y}_{t+h} = y_{t+h-m} \qquad (\text{if } h \leq m),$$

where m is the seasonal period. For example, $m = 12$ and $m = 4$ for monthly and quarterly data respectively.

The general formula is

$$\widehat{y}_{t+h} = y_{t+h-km}, \qquad k = \lfloor (h-1)/m + 1 \rfloor.$$

Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

Drift method

Assignment Project Exam Help

The **drift method** forecasts the series as the sum of the most recent value (as in the naive method) and the average change over time:

https://powcoder.com

$$\widehat{y}_{t+1} = y_t + \sum_{i=2}^{t} \frac{y_i - y_{i-1}}{t-1}$$

Add WeChat powcoder

$$\widehat{y}_{t+h} = y_t + h \times \sum_{i=2}^{t} \frac{y_i - y_{i-1}}{t-1}$$

Assignment Project Exam Help

Model diagnostics

https://powcoder.com

Add WeChat powcoder

Autocorrelation (key concept)

The **autocorrelation** of a time series process is

$$\rho_k = \frac{E\left[(Y_t - \mu)(Y_{t-k} - \mu)\right]}{\sigma^2} = \text{Corr}(Y_t, Y_{t+k}),$$

where k is the lag, and μ and σ^2 are the mean and variance of the time series (assuming that they do not depend on t).

The **sample autocorrelation** is

$$r_k = \frac{\sum_{t=1}^{T-k}(y_{t+k} - \bar{y})(y_t - \bar{y})}{\sum_{t=1}^{T}(y_t - \bar{y})^2}.$$

The autocorrelation function (**ACF**) plot displays the autocorrelation for a range of lags.

Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

White noise process (key concept)

Assignment Project Exam Help

A **white noise process** is a sequence of independently and identically distributed random variables with mean 0 and finite variance σ^2.

https://powcoder.com

If a time series model is well specified, we expect the residual series of the fitted model to behave like a white noise process.

Add WeChat powcoder

Model diagnostics (key concept)

Residual plot. The presence of patterns in the time series of residuals (such as non-constant variance over time) may suggest assumption violations and the need for alternative models.

Residual ACF plot. Well specified models should lead to small and insignificant sample autocorrelations, consistent with a white noise process.

Residual distribution plots (histogram, KDE, Q-Q plots, etc). Inspecting the distribution of the residuals will suggest the appropriate assumptions for interval forecasting.

Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

Assignment Project Exam Help

Model validation

https://powcoder.com

Add WeChat powcoder

Training and validation sets

- We incorporate model validation into the forecasting process by setting aside a validation sample for estimating and comparing the performance of different models.
- We allocate the last part of the sample (typically 20-50% of the data) to the validation set.
- In time series, the training set is called "in-sample data" and the validation set the "out-of-sample data".
- Due to the dynamic nature of forecasting, there is no test set (though we may sometimes refer to model validation as forecast evaluation).

Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

Real time forecasts (key concept)

We validate forecasts by following the "real time" approach: at every period t, we use all the available data at present to estimate the model and predict the future value of the series.

Assignment Project Exam Help

1. Starting at $t = n$, use the observations at times $1, 2, \ldots, t$ to estimate the forecasting model. Use the estimated model to forecast the observation at time $t + 1$.

https://powcoder.com

2. Repeat the above step for $t = n + 1, \ldots, T - 1$.

Add WeChat powcoder

3. Compute forecast accuracy measures based on the prediction errors $y_{n+1} - \widehat{y}_{n+1}, \ldots, y_T - \widehat{y}_T$.

We follow a similar procedure for multi-step forecasts.

Expanding and rolling windows

Assignment Project Exam Help

We can consider two schemes for updating the estimation sample at each validation period:

https://powcoder.com

Expanding window. At each step, add the latest observation to the the estimation sample.

Rolling window. At each step, use only the most recent n observations for estimation. The rolling window scheme implicitly assumes that the dynamics of the series has a time changing nature, so that data far in the past are less relevant for estimation.

Add WeChat powcoder

Measuring forecast accuracy

Assignment Project Exam Help

We typically assume the squared error loss and compute the out-of-sample MSE to measure forecast accuracy.

https://powcoder.com

However, it is useful to be familiar with other measures that are common in business forecasting:

Add WeChat powcoder

- Percentage errors.
- Scaled errors.

Percentage errors

- The percentage error is given by $p_t = 100 \times ((y_t - \widehat{y}_t)/y_t)$. It has the advantage of being scale-independent.
- The most commonly used measure is mean absolute percentage error

  $$\text{MAPE} = \text{mean}(|p_t|).$$

- Measures based on percentage errors have the disadvantage of being infinite or undefined if $y_t = 0$ for any t in the period of interest, and having extreme values when any $y_t = 0$ is close to zero.

- Percentage errors are only valid under a meaningful zero.

Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

Scaled errors

- Hyndman and Koehler (2006) proposed scaling the errors based on the training MAE (or MSE) from a benchmark method (typically a simple model).
- For a non-seasonal time series, a useful way to define a scaled error uses naïve forecasts:

$$q_t = \frac{y_t - \widehat{y}_t}{\frac{1}{T-1}\sum_{i=2}^{T} |y_i - y_{i-1}|}.$$

- Because the numerator and denominator both involve values on the scale of the original data, q_j is independent of the scale of the data.

Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

Mean absolute scaled error

Assignment Project Exam Help

The **mean absolute scaled error** is

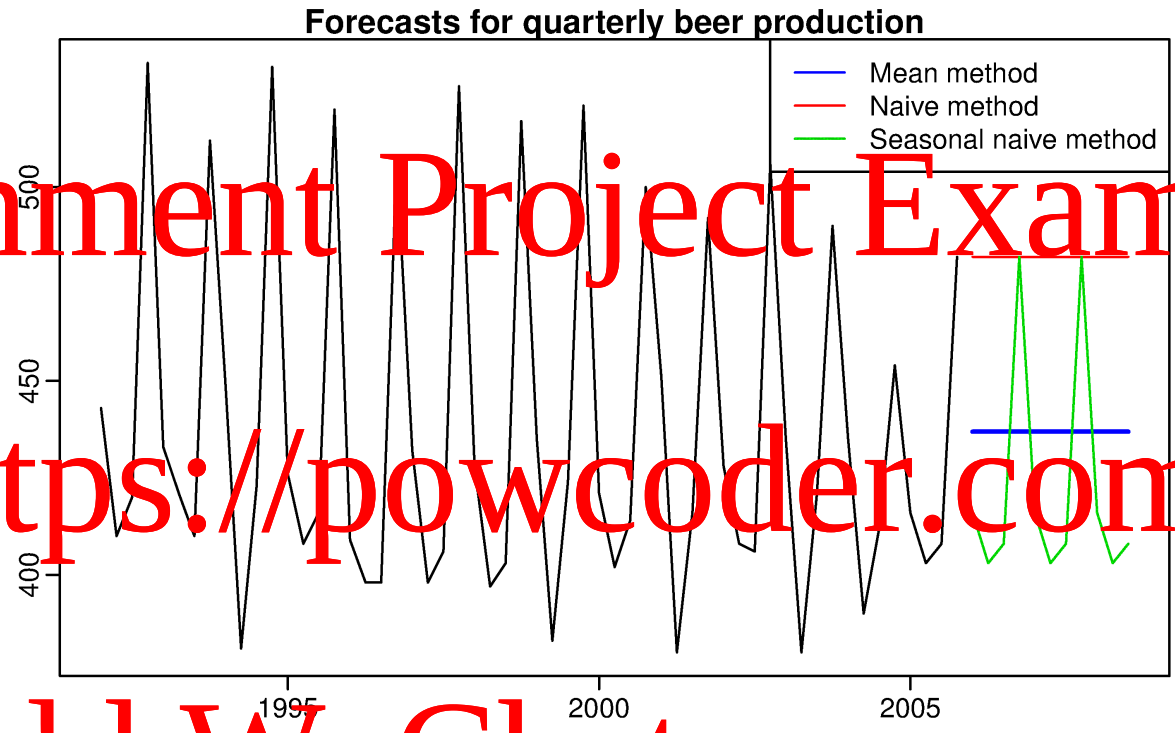
$$\text{MASE} = \text{mean}(|q_t|).$$

https://powcoder.com

A scaled error is less than one if it arises from a better set of forecasts than the random walk method evaluated on the training data.

Add WeChat powcoder

Example: Quarterly Australian Beer Production



Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

The figure shows shows three forecasting methods applied to the quarterly Australian beer production using data to the end of 2005. We compute the forecast accuracy measures for 2006-2008.

Example: Quarterly Australian Beer Production

Assignment Project Exam Help

Method	RMSE	MAE	MAPE	MASE
Mean method	38.01	33.78	8.17	2.30
Naïve method	70.91	63.91	15.88	4.35
Seasonal naïve method	**12.97**	**11.27**	**2.73**	**0.77**

https://powcoder.com

Add WeChat powcoder

It is clear from the graph that the seasonal naive method is best for the data, although it can still be improved.

Assignment Project Exam Help

Random walk model

https://powcoder.com

Add WeChat powcoder

Random walk model (key example)

Assignment Project Exam Help

In this section, we use the random walk method to illustrate how to obtain point and interval forecasts for multiple horizons based on a time series model.

https://powcoder.com

We assume the model

$$Y_t = Y_{t-1} + \varepsilon_t,$$

Add WeChat powcoder

where ε_t is i.i.d with constant variance σ^2.

Random walk model

Since $Y_t = Y_{t-1} + \varepsilon_t$, we can use back substitution to show that

Assignment Project Exam Help

$$Y_{t+1} = Y_t + \varepsilon_{t+1}$$

https://powcoder.com

$$\begin{aligned} Y_{t+2} &= Y_{t+1} + \varepsilon_{t+2} \\ &= Y_t + \varepsilon_{t+1} + \varepsilon_{t+2} \end{aligned}$$

Add WeChat powcoder

$$\begin{aligned} Y_{t+h} &= Y_{t+h-1} + \varepsilon_{t+h} \\ &= Y_t + \varepsilon_{t+1} + \ldots + \varepsilon_{t+h} \end{aligned}$$

Point forecast

Assignment Project Exam Help

$$Y_{t+h} = Y_t + \sum_{i=1}^{h} \varepsilon_{t+i}$$

https://powcoder.com

Therefore, we obtain the point forecast for any horizon as

Add WeChat powcoder

$$\begin{aligned}
\widehat{y}_{t+h} &= E(Y_{t+h}|y_{1:t}) \\
&= E\left(Y_t + \sum_{i=1}^{h} \varepsilon_{t+i} \middle| y_{1:t}\right) \\
&= y_t
\end{aligned}$$

Forecast

The conditional variance is

$$\begin{aligned}\text{Var}(Y_{t+h}|y_{1:t}) &= \text{Var}(y_t + \sum_{i=1}^{h} \varepsilon_{t+i}|y_{1:t}) \\ &= h\sigma^2.\end{aligned}$$

For density forecasting, we need to make further assumptions about the errors. If we assume that $\varepsilon_t \sim N(0, \sigma^2)$,

$$Y_{t+h}|y_{1:t} \sim N\left(y_t, h\sigma^2\right).$$

Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

Forecast interval

Under the Gaussian assumption,

$$Y_{t+h}|y_{1:t} \sim N\left(y_t, h\sigma^2\right),$$

leading to the forecast interval

$$y_t \pm z_{\alpha/2} \times \sqrt{h\widehat{\sigma}^2},$$

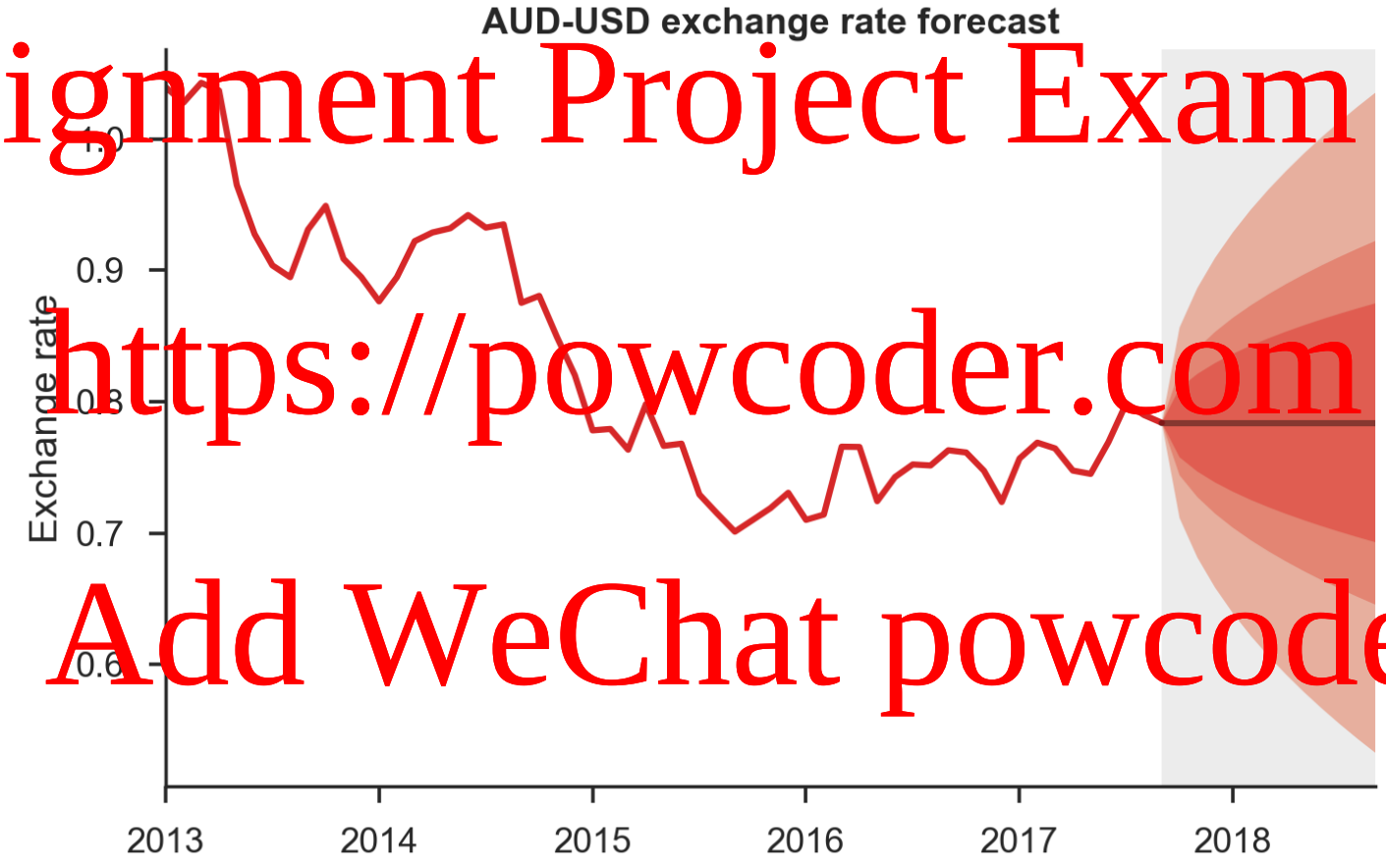
where

$$\widehat{\sigma}^2 = \frac{\sum_{t=2}^{T}(y_t - y_{t-1})^2}{T-1},$$

and $z_{\alpha/2}$ is the appropriate critical value from the normal distribution.

Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

Example: USD/AUD exchange rate



Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

Forecast interval

Forecast interval based on the assumption of normal errors:

$$y_t \pm z_\alpha \times \sqrt{h\widehat{\sigma}^2}$$

Assignment Project Exam Help

- This forecast interval is based on the **plug-in method**, as we replace the unknown σ^2 with an estimate.

https://powcoder.com

- The plug in method is a standard approach, but you should be aware that it ignores parameter uncertainty, leading to prediction intervals that are too narrow.

Add WeChat powcoder

- If the errors are not Gaussian, you should use other methods such as the Bootstrap algorithm (not in the scope of our unit).

Review questions

Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

- What is point and interval forecasting?
- What are the four time series components?
- Which diagnostics do we use for univariate time series models, and why?
- How to we conduct model validation for forecasting?
- How do we compute forecasts and prediction intervals for the random walk model?