MIS772

Predictive Analytics

Workshop: Time Series and Forecasting

Time series data pre-processing, exploration and decomposition, subsequently forecasting

Deakin University CRICOS Provider Code: 00113B

Workshop Plan

Objectives:
Your task is to create a time series model to predict future trends in real-estate market in Ames, USA.

Data Set:
Use file "AmesHousingPast.csv"

Acknowledgements:
Dean De Cock, Truman State University, 2011.

Original Data from Kaggle:
https://www.kaggle.com/c/house-prices-advanced-regression-techniques

Method:
Attend the seminar, follow the tutor's demo and instructions, take notes. Note that the class and online seminar will be recorded and their videos linked to the CloudDeakin topic for later access and study.

Acquire data for time-series analysis

- Load Ames housing data and unzip
- Read the "past" data set, and store
- Chart and explore past data

Prepare time series data

- Select attributes and replace missing values
- Create a time-series index
- Create aggregates of price and number of sales
- Run, explore and save

Create a simple moving average model

- Extend previous process
- Add moving average simple filter
- Explore time series smoothing, save

Create a forecasting model

- Modify previous process
- Investigate classic time series decomposition
- Train a forecast model
- Apply the model, analyse and save

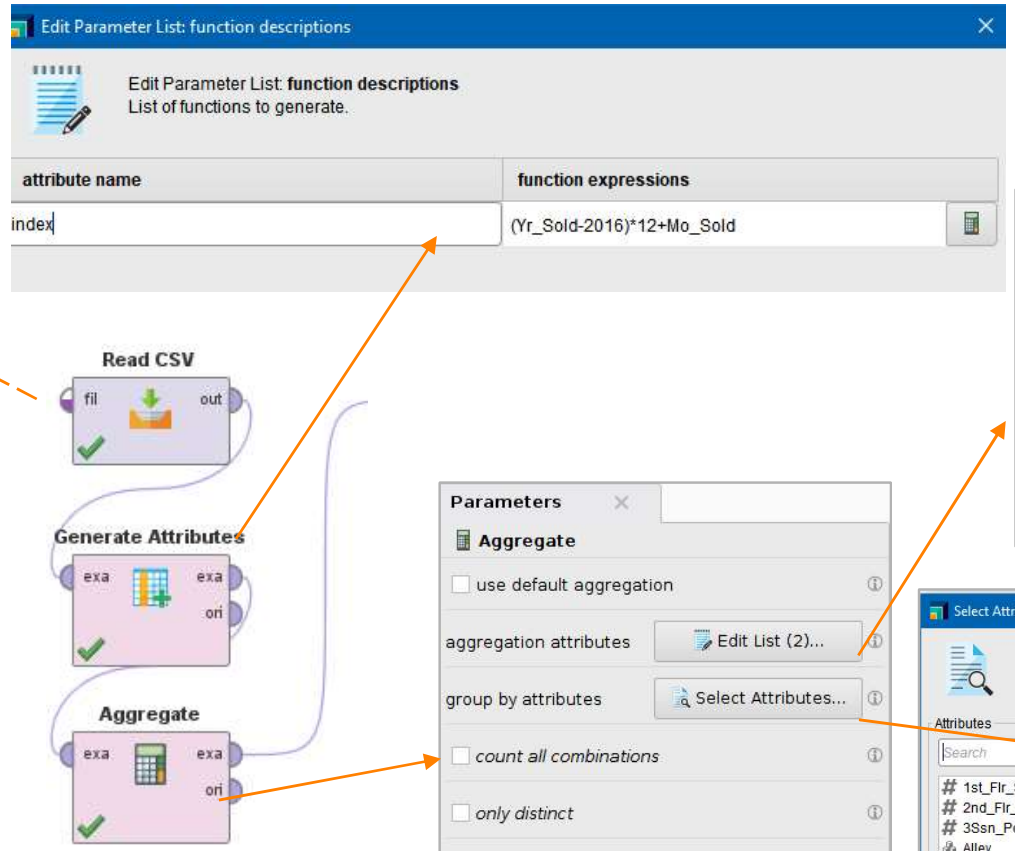
DEAKIN BUSINESS SCHOOL

Time Series Pre-Processing

Load AmesHousing (past) data, select numeric attributes only, replace all missing values with average, generate a new time series index in months, and aggregate a median sale price and a count of sales per month. Run and save.

index	count(PID)	median(SalePrice)
1	18	176700
2	24	188050
3	51	149500
4	48	141576
5	75	157000
6	97	155000
7	122	156750
8	45	187750
9	41	197000
10	49	142500
11	31	209000
12	24	161000

Check the time series with an index in monthly increments, median monthly sale price as well as the count of monthly sales.

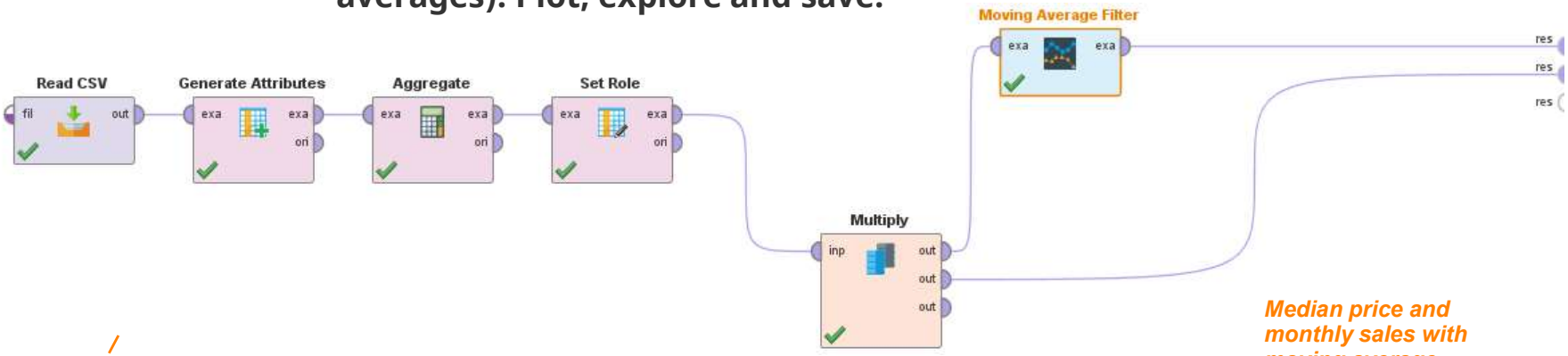


DEAKIN BUSINESS SCHOOL

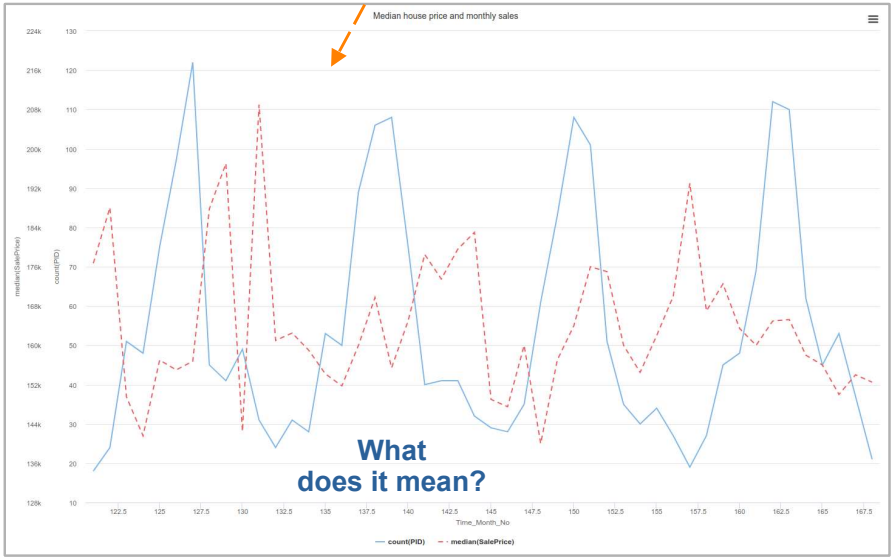
Explore Moving Averages

Extend the previous process. Set the time index role as ID. Multiply the result (for comparison). Add moving average simple filter. Produce raw data series and compare with smoothed data series (with moving averages). Plot, explore and save.

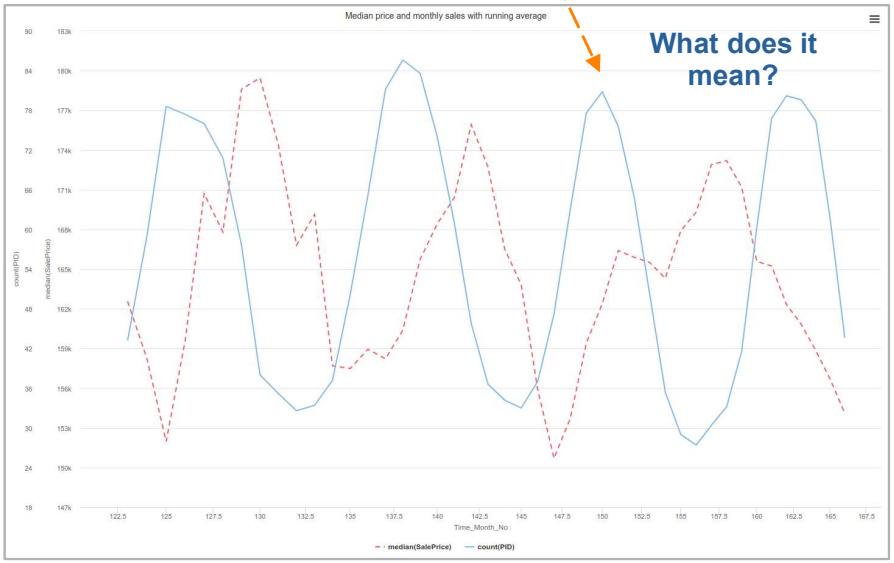
Challenge: use RM visualisation to plot two charts in a single pane.



Median price and monthly sales – raw data



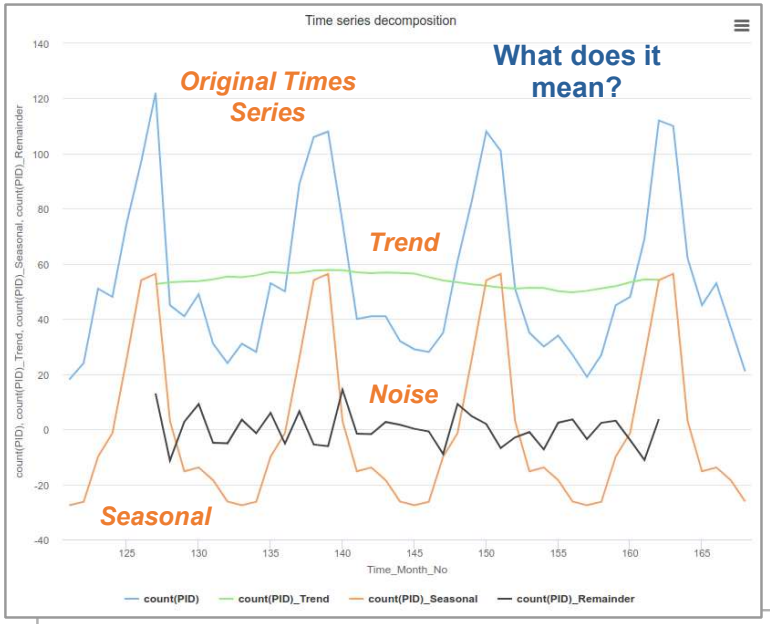
Median price and monthly sales with moving average

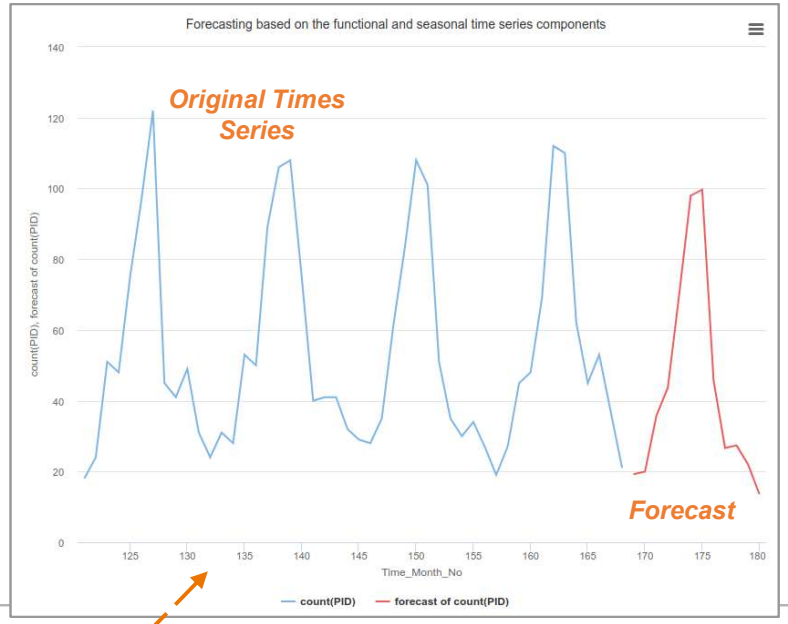


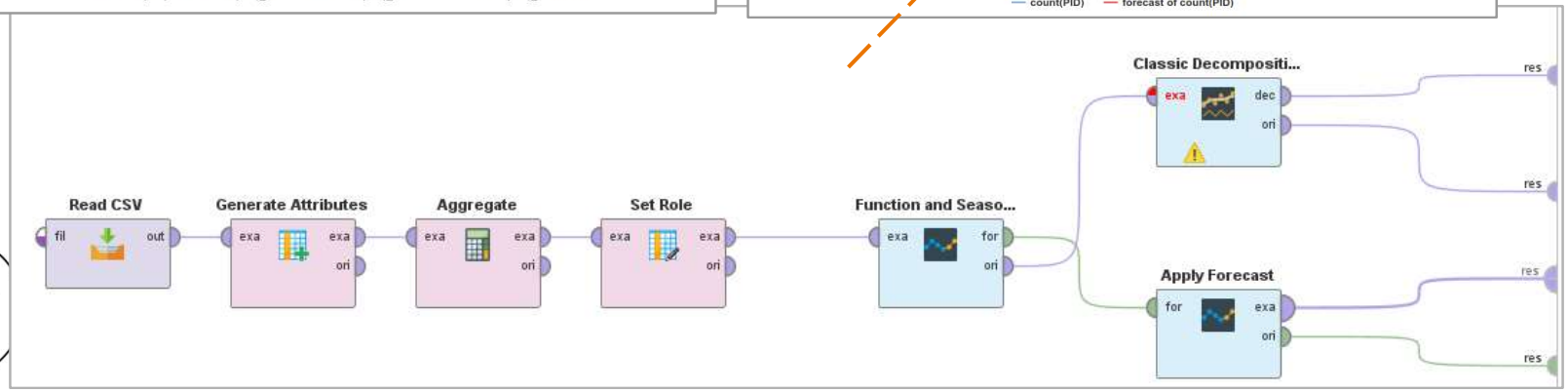
DEAKIN BUSINESS SCHOOL

Explore Moving Averages, Forecast

Extend the previous process. Investigate classic time series decomposition. Train a functional and seasonal forecast model. Apply the model, analyse and save.







Parameters

Function and Seasonal Component Forecast

time series attribute	count(PID)
indices attribute	index
decomposition method	classic decomposition
function fitting method	polynomial function
degree	1
decomposition mode	additive
ignore invalid values	
seasonality	12

Parameters

Apply Forecast

forecast horizon	12
add original time series	
add combined time series	

Parameters

Classic Decomposition

attribute filter type	single
attribute	count(PID)
invert selection	
include special attributes	
has indices	
indices attribute	index
decomposition mode	additive
ignore invalid values	
seasonality	12

DEAKIN BUSINESS SCHOOL