Roman Matkovskyy

Project list for the course "Financial Modelling and Analysis", MSc Business Analytics

You have to present the results of the 6 projects.

The projects consist of the two parts: technical realization and interpretations. Both sides shall be presented. More precisely: the projects should include a short presentation and discussion of the methodology used, description of data, a realization in R and interpretations/conclusions/implications. It can be done in R notebook.

Projects

Team projects – 30% of the final grade

Project 1 – Warm up your creativity

Estimate probability of having cryptocurrency price from Coinbase exchange lower by at least 9% at the close on at least one of the next 5 trading days

(Hint: cryptocurrencies are traded 7 days a week)

Project 2 - Exploratory analysis

Select any 2 financial time series of interest, e.g., stocks, bonds, cryptocurrencies etc. Min 5 years of observations (the longer – the better) from YahooFinance, and estimate the following tests:

- General statistics, i.e., mean, median, variance, quantiles, kurtosis and skewness.
- Check a potential presence of outliers.
- Univariate tests for each series: Shapiro-Wilk, Kolmogorov-Smirnov, Jarque-Bera, D'Agostino test of skewness, Anscombe-Glynn test of kurtosis, Bonett-Seier test of kurtosis, Anderson-Darling goodness of fit test.
- Multivariate tests: Pearson's product moment correlation coefficient t-test, Spearman rank correlation test, Kendall's tau correlation coefficient test
- Two sample t-test for the difference in sample means

Estimate granger causality and volatility.

Visualize. Interpret the results. Provide recommendations .

Project 3: Regression specification and transformation

Use the data set USMacroG in R's AER package.

- This data set contains quarterly times series on 12 U.S. macroeconomic variables for the period 1950–2000.
- Or any other source of data to collect/use the following variables:

– consumption = real consumption expenditures,

– dpi = real disposable personal income,

– government = real government expenditures, and

– unemp = unemployment rate.

– Investment = investment

– Cpi = inflation

– interest = interest rate

– Predict changes in consumption from changes in the other variables.

1. Identify any features, for instance outliers, mean, variance, skewness, kurtosis etc.

2. Which variables can be used to get a better prediction?

3. Which variables seem useful for predicting changes in consumption?

4. What does ANOVA table advise?

5. Which variables are better to remove from the model, and in what order? (use stepAIC function)

6. Do you want to add other variables? If yes – which ones and why?

7. Was the improvement large/significant?

8. Check the leverage.

9. Test which transformation can derive the better results.

10. Conclusions /interpretations

Individual projects – 70% of the final grade

Project 4: Binary regression

Use a dataset *default of credit card clients.xls*.

This research employed a binary variable, default payment (Yes = 1, No = 0), as the response variable. This study reviewed the literature and used the following 23 variables as explanatory variables:

X1: Amount of the given credit (dollar): it includes both the individual consumer credit and his/her family (supplementary) credit.

X2: Gender (1 = male; 2 = female).

X3: Education (1 = graduate school; 2 = university; 3 = high school; 4 = others).

X4: Marital status (1 = married; 2 = single; 3 = others).

X5: Age (year).

X6 - X11: History of past payment. We tracked the past monthly payment records (from April to September, 2005) as follows: X6 = the repayment status in September, 2005; X7 = the repayment

status in August, 2005; . . .;X11 = the repayment status in April, 2005. The measurement scale for the repayment status is: -1 = pay duly; 1 = payment delay for one month; 2 = payment delay for two months; . . .; 8 = payment delay for eight months; 9 = payment delay for nine months and above.

X12-X17: Amount of bill statement (dollar). X12 = amount of bill statement in September, 2005; X13 = amount of bill statement in August, 2005; . . .; X17 = amount of bill statement in April, 2005.

X18-X23: Amount of previous payment (dollar). X18 = amount paid in September, 2005; X19 = amount paid in August, 2005; . . .;X23 = amount paid in April, 2005.

Estimate the probability of default.

Estimate the probability that it is a Male client, if it has the following characteristics:

LIMIT_BAL	170000
EDUCATION	1
MARRIAGE	1
AGE	35
PAY_0	0
PAY_2	0
PAY_3	-1
PAY_4	-1
PAY_5	-1
PAY_6	-1
BILL_AMT1	86239
BILL_AMT2	75600
BILL_AMT3	2800
BILL_AMT4	21881
BILL_AMT5	0
BILL_AMT6	6780
PAY_AMT1	1512
PAY_AMT2	2800
PAY_AMT3	21881
PAY_AMT4	0
PAY_AMT5	6780
PAY_AMT6	530
default payment next month	0

Or a shorter version of characteristics, for instance :

LIMIT_BAL	170000
EDUCATION	1
MARRIAGE	1
AGE	35
PAY_0	0
PAY_2	0
PAY_3	-1
PAY_4	-1
PAY_5	-1

PAY_6	-1
BILL_AMT1	86239
default payment next month	0

Interpret and visualize results.

Project 5 – Herding.

- Estimate herding in the market of interest. Interpret the results.
- Estimate herding on up/down market days. Interpret the results.

You can apply different models (TV, Markov switching, Bayesian etc.)

Project 6 – Cryptocurrency.

Download closing prices of min 10 financial assets (50% cryptocurrencies, 50% of any traditional assets, e.g., bonds, metals, indexes etc).

- Calculate log return, correlation.
- Estimate an economic factor model. What are the factor loadings? What are the variances of the unique risks? Is the model fitted well? Interpretation.
- Estimate a statistical factor model. What is the pattern of factor influence?
- Comparison/Interpretations/conclusions/implications.