

# MN-M585: Applied Data Analytics

## Coursework Assignment

Submission Deadline: Friday, 30<sup>th</sup> April 2021 by 3pm

This coursework assignment is an individual coursework assignment worth 60% of the overall module mark. Please save your assignment file using your student ID number.

### Section 1: Linear Regression Analysis

The *capm.xlsx* file contains daily close prices of Microsoft and S&P500 along with the risk-free rate, SMB and HML. Use a subsample of these data running from your birthday in 2010 to your birthday in 2013 to reply to the following questions:

1. Estimate the Capital Asset Pricing model (CAPM). Provide the model in equation form. Interpret the coefficients. Is the stock aggressive or defensive?  
(4 marks)
2. Test the significance of the coefficients at the 5% significance level.  
(4 marks)
3. Test, at the 5% level, the hypothesis that Microsoft stock has the same volatility as the market portfolio against the alternative that the Microsoft stock does not have the same volatility as the market. What do you conclude?  
(7 marks)
4. Test the residuals of the regression for violation of the CLRM assumptions and report any violations.  
(8 marks)
5. Add 2 factors to the CAPM model and estimate the new model – the so called Fama-French 3-factor model. Provide the model in equation form. Test the addition of these factors to the original model.  
(4 marks)
6. Select a final model and report on it, justifying the inclusion of the selected variables.  
(6 marks)
7. Report the goodness of fit statistic of the final model and comment on the goodness of fit of each regression.  
(7 marks)

**(Overall: 40 marks)**

## Section 2: Time Series Analysis

1. Select a financial time series of daily close prices for an international equity index. The time series should start in your birthday in 2014 and end in your birthday in 2019.  
(4 marks)
2. Use the appropriate plots and test to evaluate whether an ARMA model fits the daily close prices series. If not, to which series would you fit an ARMA model? Why?  
(13 marks)
3. Using both the Akaike (AIC) and Schwarz (SBIC) Information Criteria, select the ARMA (p,q) model that describes the data best and provide it in equation form. How did you decide the model order? Report AIC and SBIC scores for all models tested and the chosen model.  
(10 marks)
4. Use the appropriate visualizations and test to evaluate whether a GARCH-type model fits the data.  
(10 marks)
5. Fit the GARCH(1,1), TGARCH(1,1) and EGARCH(1,1) models to the first 900 days of your data set. Comment on it. In your answer you should provide the models in equation form, describe which features of financial data each model captures and discuss the constraints related to each model and if they are satisfied.  
(13 marks)
6. Forecast one-day-ahead volatility for the remaining days of the data set. Plot these forecasts along with the true volatility, report the forecast errors and comment.  
(10 marks)

**(Overall: 60 marks)**

### Tips for writing your report

- The report should be split into 2 sections as above. The maximum word limit is 2,000 words.
- You should concisely explain your methodology and justify your decisions for each point in all sections based on the STATA output.
- Please include all commands from the STATA in an appendix to the report.