# University of Westminster

# Department of Computer Science

|  |  |
| --- | --- |
| **6BUIS017W CRM & CM With Business Intelligence** | |
| Module leader | Panagiotis Chountas |
| Unit | Coursework 2 |
| Weighting: | 50% |
| Qualifying mark | 30% |
| Description | Students will be given a large data set to perform predictive modelling using linear regression techniques in Python. The students will be expected to produce a word project report on their analysis of the data set resulting from applying their own implementation and discuss how predictive modelling is used in the Oil and Gas Industry for Implementing Change Management Strategies. |
| Learning Outcomes Covered in this Assignment: | 1. This assignment contributes towards the following Learning Outcomes (LOs): Describe the roles of customer relationship management (CRM), change management (CM) and data mining approaches in today’s business world; 2. Define customer relationship management, change management problems, select and apply appropriate Business Intelligence (BI) methodologies and evaluate BI solutions to these problems; 3. Apply CRM knowledge and CM to support change and improve operational processes of service organizations; |
| Handed Out: | 28th November 2024 |
| Due Date | 09th January 2025  Submission by 13:00 hours |
| Expected deliverables | Submit on Blackboard a zip file containing the required documentation (either in docx or pdf format). All implemented codes should be included in your documentation together with the results/analysis. |
| Method of Submission: | Electronic submission on BB via a provided link close to the submission time. |
| Type of Feedback and Due Date: | Feedback will be provided on BB, on 25th January 2025 |
| BCS CRITERIA MEETING IN THIS ASSIGNMENT | * 7.1.6 Use appropriate processes * 7.1.7 Investigate and define a problem * 7.1.8 Apply principles of supporting disciplines * 8.1.1 Systematic understanding of knowledge of the domain with depth in particular areas * 8.1.2 Comprehensive understanding of essential principles and practices * 8.2.2 Tackling a significant technical problem * 10.1.2 Comprehensive understanding of the scientific techniques |

Assessment regulations

Refer to section 4 of the “How you study” guide for undergraduate students for a clarification of how you are assessed, penalties and late submissions, what constitutes plagiarism etc.

Penalty for Late Submission

If you submit your coursework late but within 24 hours or one working day of the specified deadline, 10 marks will be deducted from the final mark, as a penalty for late submission, except for work which obtains a mark in the range 40 – 49%, in which case the mark will be capped at the pass mark (40%). If you submit your coursework more than 24 hours or more than one working day after the specified deadline you will be given a mark of zero for the work in question unless a claim of Mitigating Circumstances has been submitted and accepted as valid.

It is recognised that on occasion, illness or a personal crisis can mean that you fail to submit a piece of work on time. In such cases you must inform the Campus Office in writing on a mitigating circumstances form, giving the reason for your late or non-submission. You must provide relevant documentary evidence with the form. This information will be reported to the relevant Assessment Board that will decide whether the mark of zero shall stand. For more detailed information regarding University Assessment Regulations, please refer to the following website:<http://www.westminster.ac.uk/study/current-students/resources/academic-regulations>

**Coursework Tasks**

Predict the price of Natural Gas Using Linear Regression

By observing the markets you learn everything about people, and most importantly this knowledge is provided in such a way where it is perfect for data scientists to put their hands on it. The main problem that tends to appear with markets, is that they are often unpredictable.

The mathematical models might all show that the value of a certain commodity will go up and then something unpredictable happens (e.x. COVID-19) and everything changes. Thus, it becomes obvious that the markets are extremely prone to external influence and factors. Nevertheless, your task is to attempt, by using linear regression, to predict the price of Brent oil.

We now have all Brent oil price information for the last 20 years under “BrentOilPrices.csv” on BB. It is now time to see how our data looks in order to determine what features we want to keep and which ones we want to discard. It appears that we are quite lucky! The dataset only contains the date and the price, thus there is no sort of tampering needed on our part.

Task A

1. Data Visualisation: Define simple line chart to give an idea of the stock price change Brent oil price information for a period of 30 years;

**[8 Marks]**

1. Build explanatory variables — the features we are going to use to predict the price of Brent oil. The variables we will be using at this stage, are the **moving averages** for the past **three (MA3)** and nine **days (MA9), based on input from the oil stock market.**

**[14 Marks]**

A moving average is a technical indicator that market analysts and investors may use to determine the direction of a trend. It sums up the data points of a financial security over a specific time period and divides the total by the number of data points to arrive at an average. It is called a “moving” average because it is continually recalculated based on the latest price data.

Analysts use the moving average to examine support and resistance by evaluating the movements of an asset’s price. A moving average reflects the previous price action/movement of a security. Analysts or investors then use the information to determine the potential direction of the asset price. It is known as a lagging indicator because it trails the price action of the underlying asset to produce a signal or show the direction of a given trend.

1. Define the Train and Test Data: This step covers the preparation of the train data and the test data;

**[10 Marks]**

1. Build a Linear Regression Model (LR) using the **moving averages** for the past for the past for the past **three (MA3)** and nine **days (MA9), as inputs**;

[10 Marks]

1. Prediction Function and Result: In this step, run the model using the test data we defined in step four. Visualise the predicted versus the actual stock values for the specific time period and calculate the model’s accuracy;

[10 Marks]

1. Calculate the alpha and betas values in Python: Define the linear regression equation using the alpha and betas values;

**[8 Marks]**

**[60 Marks]**

**How Oil and Gas Industry is Implementing Change Management Strategies Using Predictive Modelling**

**Task B**

**Define as part of a report “How Oil and Gas Industry is Implementing Data Driven Change Management Strategies Using Predictive Modelling and Analytics”.** More specifically focus on implementing technology platforms that can eliminate inefficiencies and improve day-to-day operations that cut time, generate revenue and improve working conditions for their employees.

**Present your findings for the above questions in the form of a research paper.  
The paper must express your own conclusions and findings. The paper size  
should be between [900-1000] words, excluding references.  
Papers violating the lower limit or exceeding the upper limit of allowable words  
will be subject to a penalty of 10%, (2 Marks out of 20)**

**[40 Marks]**

# Marking Scheme

Due to the nature of the assessment candidates may come up with more than one equally, good solutions. Thus, marks will be allocated as follows

Predict the price of Natural Gas Using Linear Regression

Task A

1. Data Visualisation: Plot and comment on a simple line chart to give an idea of the stock price change Brent oil price information for a period of 30 years;

* *Plot and comment on a simple line chart using Python [4 Marks]*
* *Comment on line chart behaviour [4 Marks]*

**[8 Marks]**

1. Build explanatory variables — the features we are going to use to predict the price of gas. The variables we will be using at this stage, are the **moving averages** for the past **three (MA3)** and nine **days (MA9), based on input from the oil stock market;**
   * Build explanatory variables (MA3) *using Python [7 Marks]*
   * Build explanatory variables (MA9) *using Python [7 Marks]*

[14 Marks]

A moving average is a technical indicator that market analysts and investors may use to determine the direction of a trend. It sums up the data points of a financial security over a specific time period and divides the total by the number of data points to arrive at an average. It is called a “moving” average because it is continually recalculated based on the latest price data.

Analysts use the moving average to examine support and resistance by evaluating the movements of an asset’s price. A moving average reflects the previous price action/movement of a security. Analysts or investors then use the information to determine the potential direction of the asset price. It is known as a lagging indicator because it trails the price action of the underlying asset to produce a signal or show the direction of a given trend.

1. Define the Train and Test Data: This step covers the preparation of the train data and the test data;
   * Define the Train Data set *using Python [4 Marks]*
   * Define the Test Data set *using Python [4 Marks]*
   * Define Reproducible Seed *using Python [2 Marks]*

**[10 Marks]**

1. Build a Linear Regression Model (LR) using the ****moving averages**** for the past for the past for the past **three (MA3)** and nine **days (MA9), **as inputs****;

* Define the Linear Regression Model *using Python [5 Marks]*
* Train the Linear Regression Model *using Python [5 Marks]*

**[10 Marks]**

1. Prediction Function and Result: In this step, run the model using the test data we defined in step four. Visualise the predicted versus the actual stock values for the specific time period and calculate the model’s accuracy;

* Test the Linear Regression Model *using Python [5 Marks]*
* Visualise the Linear Regression Model *using Python [5 Marks]*

**[10 Marks]**

1. Calculate the alpha and betas values in Python: Define the linear regression equation using the alpha and betas values;

* Extract alpha and betas values *using Python [4 Marks]*
* Formulate the Linear Regression Model *[4 Marks]*

**[8 Marks]**

**[60 Marks]**

**Task B**

**Define as part of a report “How Oil and Gas Industry is Implementing Data Driven Change Management Strategies Using Predictive Modelling and Analytics”.** More specifically focus on implementing technology platforms that can eliminate inefficiencies and improve day-to-day operations that cut time, generate revenue and improve working conditions for their employees.

**Present your findings for the above questions in the form of a research paper.  
The paper must express your own conclusions and findings. The paper size  
should be between [900-1000] words, excluding references.  
Papers violating the lower limit or exceeding the upper limit of allowable words  
will be subject to a penalty of 10%, (2 Marks out of 20).**

Marks will be allocated for:

* ***Originality of the report***

***[8 Marks]***

* ***Critical analysis***

***[8 Marks]***

* ***Technical content***

***[8 Marks]***

* ***Clarity of the paper***

***[8 Marks]***

* ***Use of references***

***[8 Marks]***

**[40 Marks]**