

|  |
| --- |
| **COURSEWORK ASSESSMENT SPECIFICATION** |

|  |  |
| --- | --- |
| Module Title: | Big Data Analytics |
| Module Number: | LD7186 |
| Module Tutor Name(s): | Nitsa Herzog |
| Academic Year: | 2022-2023 |
| % Weighting (to the overall module): | 100% |
| Coursework Title: | Assignment |
| Average Study Time Required by Student: | 60 study hours |

Dates and Mechanisms for Assessment Submission and Feedback

|  |
| --- |
| Date of Handout to Students:  During week 1 |
| Mechanism for Handout to Students:  Via eLP, discussed during the Lecture |
| Date and Time of Submission by Student: **15 May 2023, 16:00**  If for personal reasons, you cannot submit by that date and you feel you have grounds for requesting an extension, you should contact [nu.london@northumbria.ac.uk](mailto:nu.london@northumbria.ac.uk) |
| Mechanism for Submission of Work by Student:  Electronic submission on Turnitin via Blackboard. The file you will be submitting should be renamed in the following format: **ID\_Surname\_Firstname** |
| Date by which Work, Feedback and Marks will be returned to Students:  Within 20 working days after submission of this assignment |
| Mechanism for return of assignment work, feedback and marks to students:  Formal feedback will take place following the completion of all reviews and internal moderation of results. |

**Learning Outcomes Assessed in this assessment**

This assignment will assess the following learning outcomes:

1. Demonstrate a critical understanding of big data management, manipulation and modelling methods and techniques.
2. Critical appraisal of statistical and big data science methods, techniques and tools applied for business intelligence.
3. Systematically identify and critically analyse big data-related problems and develop robust solutions using state-of-the-art big data analytics techniques and tools in a problem scenario.
4. Demonstrate a critical awareness of using big data and analytics for sustainability.

**General Information**

This assignment constitutes 100% towards the final mark for this module. Any queries relating to this assignment should be discussed with the module tutor:

**Type of the submission required**

This is an INDIVIDUAL piece of work contributing towards the module assessment. Deliverables should be assembled into a single report document, which includes (critical appraisal, research, and snapshot evidence of tasks carried out and justification of technologies used). Submission will be in the form of an MS word report (4000 words).

**Section 1 – Big Data Analytics (Python)**

**Case study:**

You are required to analyze the dataset “***Billionaires.csv***” using Python Tools for Big Data Analysis (The dataset can be found in the Assessment folder). Perform data analysis according to the regions and cities. As an outcome of your analysis, you have to answer the following research questions:

What are the top 10 countries with the highest number of billionaires?

What industries/sectors are most successful?

What are the main industries with the highest number of women billionaires?

What age range represents the highest and lowest number of billionaires?

Describe and analyse the additional factors that might influence wealth, e.g. inheritance.

Individual tasks which are required to be discussed in detail are presented below.

**Task 1: Problem Domain, Data Description, and Research Question (15 Marks).**

* Research state-of-the-art literature and describe the application domain to which the dataset belongs and provide a detailed description of the dataset.
* Formulate an appropriate hypothesis (NULL and alternative) based on correlation or other metrics, which can be tested using different statistical methods.

**Task 2: Solution Exploration (15 Marks)**

* Evaluate a range of approaches and technologies for developing big data applications.
* Describe the solutions and techniques applied to a similar problem.
* State your chosen methodological approach with justification.
* Support your research with relevant references.

**Task 3: Solution Development (35 Marks)**

* Perform data pre-processing if it is required.
* Provide a general analysis of your dataset (descriptive statistics) and answer the research questions. Include graphs and other types of visual representation that give essential information about the data components.
* Describe in detail the steps you have taken to reach your solution to the selected problem. Support your analysis with data visualization. Provide screenshots of the codding parts where appropriate.
* Present the results clearly and explicitly.
* Perform statistical significance testing on your data to accept or reject the NULL hypothesis.

**Task 4: Evaluation and Future Development (15 Marks)**

* Describe any potential impact of your results and how they can be used within the application domain.
* Include considerations for any limitations of your work.
* Detail a future work with an explanation of how the solution could be enhanced and developed further in terms of the latest technologies.

**Section 2 – Business Intelligence (Tableau)**

As a Data Analyst of the R&D of Amazon, you were asked to analyse the dataset obtained from Walmart, one of the leading supermarkets in the USA. The Dataset named “***Electronic Sales***” contains the product name, price, and order details (The dataset can be found in the Assessment folder).

You must use Tableau for the dataset analysis and for the completion of the following tasks.

**Task 1:** Using the parameter function, display a chart showing ten top-selling and ten low-selling products. **[3 marks]**

**Task 2:** Using an appropriate chart, display the Total/Average sales made in each city. The chart should allow the user to switch between the Sum and Average Sales. Place the result in descending order. **[3 marks]**

**Task 3:** You are asked to provide the stores’ performance on a weekly basis. These charts should demonstrate the average and total sales. **[4 marks]**

**Task 4:** The ordered items come with 6 months warranty from the date of purchase. Use the Date function to calculate the Warranty End date. **[4 marks]**

**Task 5:** Create an interactive dashboard with at least four sheets. The interactivity means that if you click on any sheet of the dashboard, the other sheets should simultaneously display the related information. **[6 marks]**

**Please note that each task should be supported by screenshots that demonstrate the stages of your process.**

**Grading Guidance**

**Distinction:**

Excellent in-depth application and critical research on the processes and user requirements. Provide in-depth knowledge of how to analyse and provide insight using Big Data.

Excellent in-depth understanding and demonstration of data analytics using Tableau.

**Commendation:**

Shows good research on the processes and user requirements. Provide in-depth knowledge of how to analyse and provide insight using Big Data.

**Pass:**

Provides a basic understanding of the deliverables. Provides end-to-end design, and all requirements are met. The report has some errors and lacks adequate explanation. The robustness and correctness of Database and data analytics are not explained thoroughly. Evidence of design is shown but inadequate explanation using Tableau.

**Fail:**

Provides incomplete attempts or lacks substantial parts of the deliverables. Fails to demonstrate understanding of the concepts required to implement deliverables. Work lacks serious clarity and detail. There are several errors in the report.

**Academic Integrity Statement:**

You must adhere to the university regulations on academic conduct. Formal inquiry proceedings will be instigated if there is any suspicion of plagiarism or any other form of misconduct in your work. Refer to the University’s Assessment Regulations for Northumbria Awards if you are unclear about the meaning of these terms. The latest copy is available on the University website. [Assessment Regulations and Polices](https://northumbria-cdn.azureedge.net/about-us/university-services/student-library-and-academic-services/quality-and-teaching-excellence/assessment/guidance-for-students/) (last accessed on the 18th of August 2022).

**Formative Feedback**

There will be an opportunity for formative feedback during the semester. You are advised to start working on this assignment as early as possible to seek clarification from the module tutor regarding any questions you might have during the semester. Note that tutors will not predict your grade, and you should not comment on any aspect of your work as indicating that it is correct. It would help if you made every effort to take advantage of formative feedback as tutors will not comment on draft work at other times. Remember that you will get more useful feedback from us by asking specific questions than just presenting us with your documentation and asking, ‘Is this right?’

**Penalties for Exceeding Word Limits:**

The following penalties will be applied after any reductions in marks due to late submission have been made; penalties will be applied as defined in the University Policy on Word Limits

Policy [click here](https://northumbria-cdn.azureedge.net/about-us/university-services/student-library-and-academic-services/quality-and-teaching-excellence/assessment/guidance-for-students/) (Last accessed on the 18th of August 2022).

The actual word count is to be declared on the front of the assessment submission.

|  |
| --- |
|  |
|  | |

Late Submission Policy:

For coursework submitted up to 1 working day (24 hours) after the published hand-in deadline without approval, 10% of the total marks available for the assessment (i.e.100%) shall be deducted from the assessment mark. Penalties will be applied as defined in the University Policy on the Late submission work [click here](https://northumbria-cdn.azureedge.net/about-us/university-services/student-library-and-academic-services/quality-and-teaching-excellence/assessment/guidance-for-students/) (Last accessed on the 18th of August 2019).

For clarity: a late piece of work that would have scored 65%, 55%, or 45% had it been handed in on time will be awarded 55%, 45% or 35%, respectively as 10% of the total available marks will have been deducted.

Failure to submit: The University requires all students to submit assessed coursework by the deadline stated in the assessment brief.  Where coursework is submitted without approval after the published hand-in deadline, penalties will be applied as defined in the University Policy on the Late Submission of Work [click here](https://northumbria-cdn.azureedge.net/about-us/university-services/student-library-and-academic-services/quality-and-teaching-excellence/assessment/guidance-for-students/) (Last accessed on the 18th of August 2).