ASSESSMENT RESULT SHEET

Centre:

**Centre for Cyber & Cloud**

**Belfast Metropolitan College**

**Castlereagh Campus**

Course:

**Foundation Degree in Software and Cloud Development with Data Analysis [P00989 series]**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Name: George Dornan | | | Mark Awarded (out of 100) | Weighting Within Module:  30% | |
| Module Title:  Data Analysis and Visualization | |  | Date Issued: | Date Due:  Submit by Canvas and/or as instructed by your tutor.  Note the additional instructions in following assignment specification: | |
|
| Assessment No**: Coursework** | |
| Intended Learning Outcomes |  | | | |

**A3**: Demonstrate and evaluate a range of data optimisation solutions  
**B1**: Design and develop appropriate data algorithms and models policies   
**B2** : Diagnose and analyse various data monitoring solutions and recommend, with justification, valid data optimisation techniques.  
**B3** : Evaluate the effectiveness of data security systems  
**C1**: Construct and relay technical information to technical, management, user and academic audiences  
**C2**: Assess using case studies the repercussions of a poorly structured database system  
**C3:** Develop effective solutions to practical problems individually and as a member of Team  
**D1:** Learn effectively in various situations making use of information retrieval skills and learning resources  
**D2:** Communicate effectively, using a range of media and with a variety of audiences  
**D3**: Work effectively under guidance or supervision, independently and as part of a team  
**D4**: Manage one’s own learning and development including time management, organisational skills and awareness of entrepreneurship  
**D5**: Demonstrate continued professional development in recognition of the need for life-long learning

|  |  |  |
| --- | --- | --- |
| Aspects Assessed | Percentage Awarded | Percentage Available |
| **Task 1 – Data Collection**  **Task 2 – Data Cleaning**  **Task 3 – In-Depth Analysis**  **Task 4 – Critique Existing Visualizations**  **Task 5 – Create Visualizations**  **Task 6 – Documentation** |  | **10**  **20**  **25**  **10**  **25**  **10** |

Signature of Assessor Date:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Cross Assessor Signature Date:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

IM Signature Date:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

|  |
| --- |
| ASSESSMENT FEEDBACK |

# Data Analysis and Visualization Assignment

|  |
| --- |
| **Late Submission**    Unauthorised late submission will be marked according to the regulations set out by the College and the Awarding Organisation “Regulations for OU Validated awards at Belfast Met”.    Where coursework is submitted late and there are no accepted extenuating circumstances it will be penalised in line with the following tariff:    Submission within 6 working days: a 10% reduction deducted from the overall marked score for each working day late, down to the 40% pass mark (for UG) and 50% pass Mark (PG awards) and no further. Submission that is late by 7 or more working days: submission refused, mark of 0.    A working day is defined by the partner and submission after the deadline will be assumed to be the next working day.    Students who fail to submit work for assessments or attend examinations shall be deemed to have failed the assessments components concerned and will be marked as **0**.    **Minimum Standards**    Presentation skills are considered essential for Computing and Infrastructure students.  If your assignment document does not conform to the following minimum standards the marks available will be reduced by up to 15%.    All assignments must have the appropriate front sheets attached. If these are not attached the assignment will not be accepted.    Please ensure you submit all attachments, documents etc relating to the coursework/assessment to the VLE (Canvas) and send a backup copy via email to your module lead within the submission window.    **Originality**    All work submitted must be original. Any material duplicated, either in published articles or other students’ assignments will automatically be disqualified. When two or more assignments are found to have similar portions, all parties carry the same penalty. This includes the use of generative AI as defined by the college policy available at [www.belfastmet.ac.uk](http://www.belfastmet.ac.uk/).   **Resubmissions**    A reminder from the handbook:  *A student who, for the first assessment opportunity, has failed to achieve a Pass for that module specification shall be expected to undertake a reassessment. The following rules apply for Resubmission:*     * *Only one opportunity for reassessment of the module will be permitted.* * *Reassessment of course work, project or portfolio-based assessment shall normally involve the reworking of the original task.* * *A student who undertakes a reassessment will have their grades capped at a Pass (40%) for that module.* * *A student will not be entitled to be reassessed in any component of assessment for which a Pass grade or higher has already been awarded.*     *No fee will be charged for resubmission of course work / resit exam where attendance at classes is not requested.*    *Students that are required to re-sit a module with attendance may be required to pay tuition fees.* |

**Scenario: Analysis of Global COVID-19 Impact on Public Health and Economy**

**Project Overview:**

You are a data analyst working for a global non-profit organization focused on public health. Your team is tasked with assessing the global impact of the COVID-19 pandemic on both public health (infection rates, death rates, vaccination rates) and economic indicators (unemployment, GDP growth, inflation). This analysis will help policymakers and healthcare organizations better understand how different countries were affected and what factors influenced their outcomes.

Your final deliverables will be a set of visualizations and insights to guide health officials, policymakers, and economists on how to address future pandemics. You will be using a mixture of **Python**, **R**, and **PowerBI** to collect, analyse, and visualize the data.

**Steps to Complete the Project:**

**1. Data Collection (10 marks):**

* **Task**: Collect two datasets from reliable sources: one for public health data and one for economic indicators. For public health data, you’ll use data from **Johns Hopkins University** (infection rates, death rates, and vaccination rates). For economic data, use **World Bank** datasets (unemployment rates, GDP growth, and inflation rates).
* **Considerations**: Ensure that the datasets cover the same time range (2020-2023) and contain data for multiple countries across continents for a comprehensive, global analysis.

**2. Data Cleaning (20 marks):**

* **Tools**: Use **Python** (pandas) or **R** (dplyr) for data cleaning.
* **Task**: Handle missing values, outliers, and inconsistencies in the dataset. You may need to:
  + Remove irrelevant columns (e.g., columns unrelated to the scope).
  + Handle missing values in vaccination rates using interpolation or imputation.
  + Standardize the economic data to account for differences in currency and inflation adjustments across countries.
* **Documentation**: Clearly document each cleaning step. For example, note how you handled countries that reported inconsistent COVID-19 data and how you aligned timelines for health and economic data.

**3. Data Analysis (25 marks):**

* **Task**: Conduct an in-depth analysis of the following:
  + Trends in infection and vaccination rates over time in different regions.
  + Economic changes during and after major infection waves.
  + The relationship between public health measures (e.g., vaccination rates) and economic recovery (e.g., GDP growth).
* **Tools**: Use statistical techniques like **correlation analysis**, **linear regression**, and **clustering** (e.g., identifying clusters of countries based on how they were impacted). You could use **Python’s SciPy**, **R’s ggplot2**, or **PowerBI's analytical tools**.
* **Insights**: Provide insights such as "Countries with higher vaccination rates saw a faster economic recovery," supported by statistical evidence.

**4. Critique Existing Visualisations (10 marks):**

* **Task**: Identify and critique three visualizations from **public dashboards** (e.g., World Health Organization’s COVID-19 dashboard, or Bloomberg’s economic impact visualizations).
* **Critique**: Provide detailed analysis of each visualization, highlighting strengths (e.g., clear legends, appropriate color schemes) and weaknesses (e.g., lack of interactivity, poor handling of multidimensional data).
* **Recommendations**: Offer recommendations on how to improve these visualizations. For example, suggest using a **heatmap** to visualize the severity of economic downturns by region.

**5. Create Visualisations (25 marks):**

* **Task**: Use **Python (matplotlib, seaborn)**, **R (ggplot2)**, or **PowerBI** to create the following types of visualizations:
  + A **heatmap** showing infection rates per 100,000 people in different countries.
  + A **line chart** comparing GDP growth trends before, during, and after the pandemic for different regions.
  + A **bar chart** comparing vaccination rates and economic recovery (GDP growth) in selected countries.
  + A **dashboard** in PowerBI that integrates health and economic data, allowing users to filter by region, timeframe, or specific metrics (e.g., vaccine efficacy vs. unemployment rates).
* **Audience-specific Visualizations**:
  + For **policymakers**: Simplified, high-level visuals focusing on actionable insights (e.g., "Countries with early vaccine rollouts saw lower death rates").
  + For **data scientists**: More detailed, technical visualizations such as regression plots showing the statistical relationships between variables (e.g., correlation between unemployment rates and infection rates).
  + For the **public**: Interactive, user-friendly dashboards that summarize key points without overwhelming them with too much technical detail.

**6. Documentation (10 marks):**

* **Task**: Document the entire process, including:
  + Code snippets for data cleaning, analysis, and visualizations.
  + A breakdown of the statistical techniques used (e.g., why you chose to use linear regression for economic analysis).
  + Rationales for your visualization choices, especially why certain visual formats were selected for specific audiences.
  + An explanation of any database functionalities used, such as how you queried the NoSQL database to merge health and economic data.

**Final Deliverables:**

1. **Cleaned dataset** with documented steps.
2. **In-depth data analysis** showing key insights.
3. **Critique of three existing visualizations**.
4. **Custom visualizations** tailored for policymakers, data scientists, and the public.
5. **Complete project documentation**, including code, rationale, and technical details.

This scenario will showcase your ability to work with multidimensional data, critique and improve visualizations, and communicate insights to diverse audiences.