## ASSESSMENT AND INTERNAL VERIFICATION FRONT SHEET (Individual Criteria)

| Course Title | B.Sc. Software Development, B.Sc. Business Analytics | | Lecturer Name & Surname | Frankie Inguanez | |
|---|---|---|---|---|---|
| Unit Number & Title | ITBSI-506-1601 Business Intelligence & Reporting | | | | |
| Assignment Number, Title / Type | **02** | | | | |
| Date Set | | | **Deadline Date** | **18/05/2020** | |
| Student Name | | | **ID Number** | | **Class / Group** |

☐ *Student's declaration prior to handing-in of assignment:*
  ❖ *I certify that the work submitted for this assignment is my own and that I have read and understood the respective Plagiarism Policy*

☐ ***Student's declaration on assessment special arrangements (Tick only if applicable)***
  ❖ *I certify that adequate support was given to me during the assignment through the Institute and/or the Inclusive Education Unit.*

☐ ❖ *I declare that I refused the special support offered by the Institute.*

| Student Signature: | | **Date :** | **08/04/2020** |
|---|---|---|---|

| Assessment Criteria | Maximum Mark | Mark Achieved |
|---|---|---|
| *KU2 Identify a number of diverse scenarios in which BI techniques can be applied and highlight the benefits that can be achieved from this application* | 7 | |
| *AA2 Use appropriate database tools to physically create the relational data warehouse structure* | 7 | |
| *KU4 Identify the various data sources needed to populate a data warehouse and script the data integration operations required to extract and merge this data* | 5 | |
| *KU5 Construct a data profiling exercise to identify and script the required transformation operations that need to be carried out on a given data set prior to loading within a data warehouse* | 5 | |
| *AA3 Use appropriate software tools to set up an ETL process for the population and continuous maintenance of a data warehouse* | 7 | |
| *AA4 Explain a number of OLAP operations through examples on a given data cube* | 7 | |
| *SE2 Assess the use of varying data analysis techniques to extract knowledge from data held within multi-dimensional cubes* | 10 | |
| *SE3 Compare different report formats and explain their utility in varying circumstances* | 10 | |
| Total Mark | **56** | |

**Assessor's feedback to student**

*(If necessary, use reverse side of page for IV feedback on assignment brief / sample of assessment decisions)*

| | **Name & Surname** | **Signature** | **Date** |
|---|---|---|---|
| **Internal Verifier :** Approval of a*ssignment brief* | | For approval signature, please refer to electronic audit trail | |
| **Lecturer / Assessor :** Issue of results and feedback to student | | For approval signature, please refer to electronic audit trail | |
| **Internal Verifier :** Approval of *assessment decisions (Sample)* | | For approval signature, please refer to electronic audit trail | |
| **Learner's signature upon collection of corrected assignment.** | | | |

# Business Intelligence & Reporting

# Assignment 02

**Instructions to Students**

Read the following instructions carefully before you start the assignment. If you do not understand any of them, ask your lecturer.

- This assignment is a home assignment and should be completed within the communicated deadline.

- This assignment caries a total of 56% from the final module mark.

- You are required to upload your work to Moodle as specified by your lecturer.

- Copying is strictly prohibited, and any students caught will be subject to the respective MCAST Disciplinary Procedures.

## Scenario 01

You are requested to head over to the World Air Quality Index website (https://waqi.info/) which displays the air quality readings from land sensors. Use the map to locate all local sensors (4 in total) and 4 of another major European city. Download the data for the chosen sensors for the last 4 years 2016-2020.

Save all the assignment SQL code in *Script.sql* making sure to add a comment to show which section you're attempting.

### Section A (KU5)

You are required to create a schema named *csv* and load each provided csv data file into a separate table. Perform any necessary data cleaning required. Such as proper country name, adding longitude and latitude coordinates.

### Section B (AA3, KU4)

Design a relational 3NF model based on the attributes of the data loaded in the *csv* schema. Save this model B.PNG. vImplement the required tables in a schema named *oltp*, and populate the tables using the data from the *csv* schema.

### Section C (AA2)

Design an OLAP cube using the star schema with the intention of analyzing air quality readings. Save the model as *Section C.PNG.* Implement all the required tables in a schema named olap, and load data from the oltp schema.

### Section D (AA4)

Prepare the following queries:

   a.  For the chosen city and for all of Malta display the average readings for every month of every year.
   b.  Pivot the data such that in the first column you have region (Malta is a region, whilst a European city such as London is another region), then every month of the year 2020 as columns. Show the average NO2 readings.

### Section E (SE2)

Using Power BI, connect to the olap schema and prepare 1 report with three pages that represent the data stored on your cube. The pages must be designed to showcase the readings such that the user can choose the region/country and date period.

**Section F (SE3)**

Prepare a time-series visualization and a forecasting visualization based on the data with the necessary filters.


**Section G (KU2)**

Consider the reports you have created in *Section E* and *Section F*, comment about the effect of the COVID-19 pandemic induced lockdowns on the air quality for the chosen region and that of Malta.

**Marking Scheme**

| Section A – KU5 | Max | Awarded |
|---|---|---|
| Tables created as needed. | 2 | |
| Data loaded from csv to database. | 1 | |
| Data cleaned to a satisfactory level. | 2 | |
| **Total** | **5** | |

| Section B – KU4 | Max | Awarded |
|---|---|---|
| Design OLTP model. | 2 | |
| Correct creation of database based on OLTP model. | 3 | |
| **Total** | **5** | |

| Section B – AA3 | Max | Awarded |
|---|---|---|
| Data loaded from csv schema to oltp schema. Marks are deducted per incorrectly loaded data. | 7 | |
| **Total** | **7** | |

| Section C – AA2 | Max | Awarded |
|---|---|---|
| Design OLAP model. | 2 | |
| Database created based on OLAP model. | 2 | |
| Data loaded from oltp schema to olap schema | 3 | |
| **Total** | **7** | |

| Section D – AA4 | Max | Awarded |
|---|---|---|
| Query a produces desired result | 3 | |
| Query b produces desired result | 4 | |
| **Total** | **7** | |

| Section E – SE2 | Max | Awarded |
|---|---|---|
| 3 Pages with enough interactivity and applied use created (2.5 marks each) | 7.5 | |
| Attention to detail. | 2.5 | |
| **Total** | **10** | |

| Section F – SE3 | Max | Awarded |
|---|---|---|
| Time-series visualization and comparison | 5 | |
| Forecasting visualization and comparison | 5 | |
| **Total** | **10** | |

| Section G (KU2) | Max | Awarded |
|---|---|---|
| 2 points on usefulness of report (2.5 marks each) | 5 | |
| **Total** | **5** | |