# 6CS030 Worksheet One – 10%

# Hand-out: Week 3. Due: Week 4

|  |  |
| --- | --- |
| **Student Number:** | 1828421 |
| **Student Name:** | Prakash Dahal |

This worksheet is based on Workbooks 1 and 2.

1. Download the *Employment Rate & Qualifications Profile of Adults* Excel spreadsheet from Canvas

This spreadsheet consists of 3 worksheets: E*mployment Rate, Degree-Level Qualification* and *No Qualifications.*

You need to examine just one of the worksheets.

First take your student number and divide it by 3. Use the ***remainder*** value to pick one of the following worksheets:

|  |  |
| --- | --- |
| **Remainder Value** | **Worksheet to use** |
| 0 | Employment Rate |
| 1 | Degree-Level Qualification |
| 2 | No Qualifications |

|  |  |
| --- | --- |
| **My Remainder Value:** | 2 |

1. Examine your worksheet and carry out the following tasks:

|  |  |  |
| --- | --- | --- |
| **Task no** | **Task** | **Marks** |
| a | Identify any cleaning issues in the worksheet and take steps to rectify the issues.  Document what changes have been made.  Include:   * Evidence of issues in the data, e.g., include screen shots as evidence (Snipping tool can be useful for this); * Include a visualisation to show issues; * Explanation of steps taken to rectify the problems and why you chose the methods taken; * The updated data | 40 |

|  |  |  |
| --- | --- | --- |
| b | Convert the worksheet so that it is in a suitable format for a relational table and import into Oracle.  Choose appropriate names for the table and column names.  Document what changes you have made.  Include:   * Documentation of what you did to make the data suitable for a relational database; * Screen shot of the table structure in Oracle (e.g., use the Column tab in SQL Developer) * Screen shot of sample data (you do not need to show it all, e.g., use the Data tab in SQL Developer) | 20 |

|  |  |  |
| --- | --- | --- |
| c | Write two SQL commands that demonstrate an OLAP query.  Include:   * 2 SQL Queries and their results * A short sentence to say what the query represents | 20 |

|  |  |  |
| --- | --- | --- |
| d | Name one advantage to using this approach for handing Big Data and include brief explanation of why you think this is an advantage.  Include:   * Advantage * Short explanation | 10 |

|  |  |  |
| --- | --- | --- |
| e | Name one disadvantage to using this approach for handing Big Data and include brief explanation of why you think this is a disadvantage.  Include:   * Disadvantage * Short explanation | 10 |

Remember:

* This is an individual assessment
* You must use the correct worksheet – 0 marks will be given if you pick the wrong one!
* Upload this document on Canvas when complete.