

**Cardiff Metropolitan University**  
**B.Sc. (Hons) in Business Information Systems**  
**Assignment Cover Sheet**

Student Details ( Student should fill the content)				
Name				
Student ID				
Scheduled unit details				
Unit code	CIS6008			
Unit title	Analytics and Business Intelligence			
Unit enrolment details	Year	3		
	Study period	2022		
Lecturer	Mr. Induneth De Silva			
Mode of delivery	Full Time			
Assignment Details				
Nature of the Assessment	Coursework (a report)			
Topic of the Case Study	Coursework; Application of statistical and geospatial business analytics tools, techniques and methodologies to generate business intelligence those are essential for informed decision making <b>in power and renewable energy sectors</b> in Sri Lanka.			
Learning Outcomes covered	LO2, LO3, LO4			
Word count	3200(report)			
Due date / Time	03 <sup>rd</sup> June 2022			
Extension granted?	Yes	No	Extension Date	
Is this a resubmission?	Yes	No	Resubmission Date	
Declaration				
I certify that the attached material is my original work. No other person's work or ideas have been used without acknowledgement. Except where I have clearly stated that I have used some of this material elsewhere, I have not presented it for examination / assessment in any other course or unit at this or any other institution				
Name/Signature			Date	
Submission				
Return to:				
Result				
Marks by 1 <sup>st</sup> Assessor		Name & Signature of the 1 <sup>st</sup> Assessor		Agreed Mark
Marks by 2 <sup>nd</sup> Assessor		Name & Signature of the 2 <sup>nd</sup> Assessor		
Comments on the Agreed mark				

**CMU B.Sc. (HONS) BIS - ASSIGNMENT FEEDBACK SHEET -ICBT CAMPUS**

<b>STUDENT NAME:</b>		<b>STUDENT NUMBER:</b>
<b>Module Number &amp; Title:</b> Analytics and Business Intelligence		<b>Semester: II</b>
<b>Assignment Type &amp; Title:</b> Coursework: Application of statistical and geospatial business analytics tools, techniques and methodologies to generate business intelligence those are essential for informed decision making <b>in power and renewable energy sectors</b> in Sri Lanka.		
<b>For student use:</b> <i>Critical feedback on the individual progression towards achieving the assignment outcomes</i>		
<b>For the Assessors' feedback</b> <b>Indicate the Task number strength and Weaknesses and the marks for each task</b>		
<b>Task No/Question No</b>	<b>Strengths (1st Assessor)</b>	<b>Strengths (2<sup>nd</sup> Assessor)</b>

Task No / Question No	Weaknesses (1 <sup>st</sup> Assessor)			Weaknesses (2 <sup>nd</sup> Assessor)	
Areas for future improvement					
Comments by 1 <sup>st</sup> Assessor			Comments by 2 <sup>nd</sup> Assessor		
Marks					
Task /Question No	Marks by 1 <sup>st</sup> Assessor	Marks by 2 <sup>nd</sup> Assessor	Marks by IV (if any)	IV comments (If Any)	
Total Marks					
Name and the Signature of the 1st Assessor				Date:	
Name & Signature of the 2nd Assessor :				Date :	
Name & Signature of the IV: (If any)				Date :	

## Assignment Brief

### Purpose:

This assignment is to assess student's ability to perform business analysis using Statistical and Geographic Information Systems(GIS) related tools, techniques and methodologies to find out applicable and useful intelligence for taking informed decision making in private and government sector institutions in the island and different parts of the countries in the world. The relevant higher level administrative officials of the those institutions can use GIS to generate maximum efficiency and benefits on informed business decision making while eliminating discrimination, ambiguity and uncertainty.

### Tasks Introduction

Understand the given tasks based on Sri Lanka **power and renewable energy sectors** using data science domain associate with non-geospatial and geospatial data models. Apply relevant tools, techniques and methodologies found in business analytics subject relevant to the module scope and conduct analysis on different subject matters with the support of source data provided in shape files (.shp), raster files(.asc),comma separated (.csv),text and excel file formats. The data analysis and visual demonstration required to be done using standard software tools recommended for the module (**R, R-studio, R-commander, QGIS, PostgreSQL** and **Google Earth** etc.).

### Task 1 – Report (100 Marks)

The student required to do the following data analysis visualizations based on datasets provided with this assignment using R, R-Studio, R-commander, QGIS, PostgreSQL, GOOGLE EARTH and other related supportive tools. All required datasets have been included within the “**Data Sets**” folder as separate subfolders per each question.

- a) Application of business analytics to analysis how effectively various factors contribute electricity generation and how to manage such factors for taking informed decision making in electricity generation on US (United States) survey. A dataset named ***Electricity\_Generation\_1955.CSV*** has been provided for you to uncover possible **associations** available with *output-Cost, output-labor, output-capital, output-capital-fuel* using correlation analysis supported by suitable hypothesis with the aid of suitable business analytics tools such as R, R-studio and R-commander. Analysis should be followed by **critical discussions** of the findings.

(15 Marks)

- b) As per the findings of **question a)** you are required to develop a well explained statistical model(s) to predict precisely the electricity output generating factors using statistical **regression analysis** followed by explanation of the resulting model(s) with the aid of suitable business analytics tools such as R, R-studio, R-commander. Analysis should be followed by **critical discussions** of the findings.

(15 Marks)

- c) Develop a Sri Lanka map that visualizes poverty information by **poverty headcount index** based on **districts** during **1995/96** and **2016** time period using **table 2.13** in **page no 62** Poverty indicators by sector ,province and district chapter of **Economic and Social Statistics of Sri Lanka** report publishes by Central Bank of Sri Lanka (Provided with the dataset). The map processing should be done using provided **vector** data set. The map should be **critically** explained based on visualized information by **comparing** statistical figures during aforementioned time periods. You can further refer all possible local and international reliable information sources to support your justifications. (CRS: Kandawala / Sri Lanka Grid - EPSG:5234)

(10 Marks)

- d) **Clearly recognize** the area explained and do a crime analysis of the area. Develop a heat map and explain the importance of maintaining the law and order of the area for its residents' economic and social development. The map should be well developed and the discussion should be analytically and critically done using all possible information sources. (CRS: Kandawala / Sri Lanka Grid - EPSG: 5234)

(10 Marks)

- e) Develop a digitized informative map of **Kelanitissa Thermal Power Plant Complex** and its suburbs with the support of suitable information provided with the dataset. The digitization should be carried out with the support of the power plant complex image after completing **geo referencing** properly. **Digitization** and **geo referencing** should be done in association with QGIS open layer plugins. Every vector layer attribute table should contain **id**, **name** and **type** fields and associated data. The developed map should be well explained considering digitized entities and infrastructure facilities with the support of reliable local and international information sources. (For map development Use coordinates reference system as WGS84-EPSG4326).

(10 Marks)

- f) Develop a **geospatial database** called "**SL\_Renewable\_Energy**" to include all the information available in the Sri Lanka map: **Cumulative Solar SPP Capacity Addition 2019** in the web page <http://www.energy.gov.lk/index.php/en/renewable-energy/technologies/solar-energy> by Sri Lanka Sustainable Energy Authority (<http://www.energy.gov.lk/index.php/en/>). The database should contain the vector files used to develop the below mentioned map. Using **SL\_Renewable\_Energy** geospatial database develop a Sri Lanka map to visualize following information. The classification of the map should be done by Cumulative Solar Energy Generated in Mega Watts. The map should be **analytically** explained considering how well sustainable energy can be used to solve electricity generation problems in the island with the aid of all possible local and international reliable information sources. (CRS: Kandawala / Sri Lanka Grid - EPSG: 5234)

- 1) District Name
- 2) Cumulative Solar Energy Generated in Mega Watts(MW)

(10 Marks)

- g) Develop a Sri Lanka map to visualize **Power Plants** belongs to **Ceylon Electricity Board(CEB)** with geo spatial information such as address, latitude and longitude. The accurate GPS locations should be retrieved via **Google Earth** as a **KML/KMZ** file. The map should be well described considering entities indicated and the advantages provided by those entities for the national economy. Supportive official information can be collected from the CEB's official web presence: <https://ceb.lk/>. Suitable and supportive vector data can be collected from provided dataset.

(10 Marks)

- h) Using the provided vector data set about Peradeniya city area in Sri Lanka, develop an area **suitability map** that clearly indicate the most suitable area for a new **Sustainable Energy Generation Research Center** to develop within the air distance range of **1Km** away from **Power House** and **2Km** away from **Urawala Primary School**. Find out the most suitable area for constructing aforesaid new government building while finding following details through geo-processing and analysis tools of QGIS. The map should be followed by a full description of the feasibility of the found area. (CRS: Kandawala / Sri Lanka Grid - EPSG:5234)

- i) Total number of buildings situated.
- ii) Names of buildings situated.
- iii) Total area of the land occupied by the buildings.
- iv) Total area available within the suitability region.

(20 Marks)

**Guidelines for the report format**

- Paper A4
- Margins 1.5" left, 1" right, top and bottom
- Page numbers – bottom, right
- Line spacing 1.5
- Font
  - Headings 14pt, Bold
  - Normal 12pt
  - Font face- Times New Roman
- Referencing and in-text citation should be done **strictly using Harvard Referencing System.**

## **Marking Scheme**

### **Learning Outcomes:**

1. Demonstrate understanding of the leading technologies relating to business intelligence, data analysis, predictive and other analytical technologies (e.g. geospatial, social), and be able to apply them appropriately in real world scenarios.
2. Demonstrate understanding of and application of specialist technologies used to harvest, analyses and visualize business data in an intelligent way.
3. Critically evaluate, design, prototype and implement business intelligence from data harvesting, processing visualizations to business analysis and storytelling.
4. Explore the latest visualization techniques, business-IT project governance and related industry certifications.

### **Learning Outcomes covered from the course work.**

#### **LO2, LO3, LO4**

- a) LO2, LO3
- b) LO2, LO3
- c) LO2, LO3, LO4
- d) LO2, LO3, LO4
- e) LO2, LO3, LO4
- f) LO2, LO3, LO4
- g) LO2, LO3, LO4
- h) LO2, LO3, LO4

### Marking criteria – for Task 1(Coursework)

Marks	Description of the criteria
<p><b>Part a)</b> Application of business analytics to analysis how effectively various factors contribute electricity generation and how to manage such factors for taking informed decision making in electricity generation on US (United States) survey. A dataset named <i>Electricity_Generation_1955.CSV</i> has been provided for you to uncover possible <b>associations</b> available with <i>output-Cost, output-labor, output-capital, output-capital-fuel</i> using correlation analysis supported by suitable hypothesis with the aid of suitable business analytics tools such as R, R-studio and R-commander. Analysis should be followed by <b>critical discussions</b> of the findings.</p> <p style="text-align: right;">(15 Marks)</p>	
0-1	No or Very poor reporting and statistical testing has been done based on subject matter.
1-5	Basic reporting with <b>hypothesis based correlation</b> testing has been done for the subject matter while selecting most suitable variables.
5-8	Very good reporting with <b>hypothesis based correlation</b> testing has been done for the subject matter while selecting most suitable variables. Box plot graphical simulation has been supported with the findings.
8-15	Excellent reporting with <b>full scale of hypothesis based correlation</b> tests have been done for the subject matter. <b>Scatter plot</b> graphical simulations have been supported with the findings. A clear justification of the findings has been done.
<p><b>Part b)</b> As per the findings of <b>question a)</b> you are required to develop a well explained statistical model(s) to predict precisely the electricity output generating factors using statistical <b>regression analysis</b> followed by explanation of the resulting model(s) with the aid of suitable business analytics tools such as R, R-studio, R-commander. Analysis should be followed by <b>critical discussions</b> of the findings.</p> <p style="text-align: right;">(15 Marks)</p>	
0-1	No or Very poor reporting and statistical testing has been done based on subject matter.
1-4	Basic reporting with full scale of <b>regression</b> analysis has been done for the subject matter.
4-7	Very good reporting with full scale of <b>regression</b> analysis has been done for the subject matter. Scatter plot graphical simulation has been supported with the findings. Precise <b>statistical model</b> has been developed.
7-15	Excellent reporting with full scale of <b>regression</b> analysis have been done for the subject matter. Scatter plot graphical simulation has been supported with the findings. Precise <b>statistical model</b> has been developed. A clear justification of the findings has been done.
<p><b>Part c)</b> Develop a Sri Lanka map that visualizes poverty information by <b>poverty headcount index</b> based on <b>districts</b> during <b>1995/96</b> and <b>2016</b> time period using <i>table 2.13</i> in <i>page no 62</i> Poverty indicators by sector ,province and district chapter of <b>Economic and Social Infrastructure</b> of Sri Lanka report publishes by Central Bank of Sri Lanka (Provided with the dataset). The map processing should be done using provided <b>vector</b> data set. The map should be <b>critically</b> explained based on visualized information by <b>comparing</b> statistical figures during aforementioned time periods. You can further refer all possible local and international reliable information sources to support your justifications. (CRS: Kandawala / Sri Lanka Grid - EPSG:5234)</p> <p style="text-align: right;">(10 Marks)</p>	
0-2	No or Very poor ordinary map has been included.
3-5	Basic map with <b>some</b> required information has been included. No vector data layer used for the map to visualize information clearly.
5-7	Very good map with <b>all</b> required information has been included. All standard map elements ( <b>North Arrow, Map Scale-Graphic, Map Scale- numeric, Map title, Map</b>



	<b>legends)</b> have been included to easily read the map. A vector data layer used for the map to visualize information clearly. A suitable <b>base map</b> has been included. The map has been generally explained.
7-10	Excellent map with all required information has been included. All standard map elements ( <b>North Arrow, Map Scale-Graphic, Map Scale- numeric, Map title, Map legends)</b> have been included to easily read the map. A suitable <b>base map</b> has been included. The map has been properly <b>captioned</b> . The map has been critically and well explained.
<b>Parts d) Clearly recognize</b> the area explained and do a crime analysis of the area. Develop a heat map and explain the importance of maintaining the law and order of the area for its residents' economic and social development. The map should be well developed and the discussion should be analytically and critically done using all possible information sources. (CRS: Kandawala / Sri Lanka Grid - EPSG:5234) <b>(10 Marks)</b>	
0-3	No or Very poor ordinary map has been included.
3-6	Basic map with <b>some</b> required information has been included.
6-8	Very good map with <b>all</b> required information has been included. All standard map elements ( <b>North Arrow, Map Scale-Graphic, Map Scale- numeric, Map title, Map legends)</b> have been included to easily read the map. A suitable <b>base map</b> has been included and area found. The map explained in general.
8-10	Excellent map with all required information has been included. All standard map elements ( <b>North Arrow, Map Scale-Graphic, Map Scale- numeric, Map title, Map legends)</b> have been included to easily read the map. A suitable <b>base map</b> has been included area found. The map has been properly <b>captioned</b> . The map well explained with precise and critical justifications.
<b>Part e)</b> Develop a digitized informative map of <b>Kelanitissa Thermal Power Plant Complex</b> and its suburbs with the support of suitable information provided with the dataset. The digitization should be carried out with the support of the power plant complex image after completing <b>geo referencing</b> properly. <b>Digitization</b> and <b>geo referencing</b> should be done in association with QGIS open layer plugins. Every vector layer attribute table should contain <b>id, name</b> and <b>type</b> fields and associated data. The developed map should be well explained considering digitized entities and infrastructure facilities with the support of reliable local and international information sources. (For map development Use coordinates reference system as WGS84-EPSG4326). <b>(10 Marks)</b>	
0-3	No or Very poor ordinary map has been included.
3-5	Basic map with <b>some</b> required digitized information has been included.
5-7	Very good map with <b>all</b> required information has been included. All standard map elements ( <b>North Arrow, Map Scale-Graphic, Map Scale- numeric, Map title, Map legends)</b> have been included to easily read the map. A suitable <b>base map</b> has been included. The map has been explained in general.
7-10	Excellent map with all required information has been included. All standard map elements ( <b>North Arrow, Map Scale-Graphic, Map Scale- numeric, Map title, Map legends)</b> have been included to easily read the map. A suitable <b>base map</b> has been included. The map has been properly <b>captioned</b> . The map has been critically and well explained.
<b>Part f)</b> Develop a <b>geospatial database</b> called “ <b>SL_Renewable_Energy</b> ” to include all the information	

<p>available in <i>the</i> Sri Lanka map: <b>Cumulative Solar SPP Capacity Addition 2019</b> in the web page <a href="http://www.energy.gov.lk/index.php/en/renewable-energy/technologies/solar-energy">http://www.energy.gov.lk/index.php/en/renewable-energy/technologies/solar-energy</a> by Sri Lanka Sustainable Energy Authority (<a href="http://www.energy.gov.lk/index.php/en/">http://www.energy.gov.lk/index.php/en/</a>). The database should contain the vector files used to develop the below mentioned map. Using <b>SL_Renewable_Energy</b> geospatial database develop a Sri Lanka map to visualize following information. The classification of the map should be done by Cumulative Solar Energy Generated in Mega Watts. The map should be <b>analytically</b> explained considering how well sustainable energy can be used to solve electricity generation problems in the island with the aid of all possible local and international reliable information sources. (CRS: Kandawala / Sri Lanka Grid - EPSG:5234)</p> <ul style="list-style-type: none"> <li>• District Name</li> <li>• Cumulative Solar Energy Generated in Mega Watts(MW)</li> </ul> <p style="text-align: right;"><b>(10 Marks)</b></p>	
0-1	No or Very poor ordinary map has been included.
1-3	Basic map with <b>some</b> required information has been included.
3-5	Very good map with <b>all</b> required information has been included. All standard map elements ( <b>North Arrow, Map Scale-Graphic, Map Scale- numeric, Map title, Map legends</b> ) have been included to easily read the map. A suitable <b>base map</b> has been included. Geo database has been developed using PostGIS spatial DBMS. Suitable screen shots of the work have been included in appendix.
5-10	Excellent map with all required information has been included. All standard map elements ( <b>North Arrow, Map Scale-Graphic, Map Scale- numeric, Map title, Map legends</b> ) have been included to easily read the map. A suitable <b>base map</b> has been included. The map has been properly <b>captioned</b> . Geo database has been developed using PostGIS spatial DBMS. The map has been well explained. Suitable screen shots of the work have been included in appendix.
<p><b>Part g)</b> Develop a Sri Lanka map to visualize <b>Power Plants</b> belongs to <b>Ceylon Electricity Board(CEB)</b> with geo spatial information such as address, latitude and longitude. The accurate GPS locations should be retrieved via <b>Google Earth</b> as a <b>KML/KMZ</b> file. The map should be well described considering entities indicated and the advantages provided by those entities for the national economy. Supportive official information can be collected from the CEB's official web presence: <a href="https://ceb.lk/">https://ceb.lk/</a>. Suitable and supportive vector data can be collected from provided dataset.</p> <p style="text-align: right;"><b>(10 Marks)</b></p>	
0-1	No or Very poor ordinary map has been included.
1-3	Basic map with <b>some</b> required information has been included.
3-5	Very good map with <b>all</b> required information has been included. All standard map elements ( <b>North Arrow, Map Scale-Graphic, Map Scale- numeric, Map title, Map legends</b> ) have been included to easily read the map. A suitable <b>base map</b> has been included. <b>KML/KMZ</b> file(s) included using <b>Google Earth</b> . Suitable screen shots of the work have been included in appendix. The map has been explained in general.
5-10	Excellent map with all required information has been included. All standard map elements ( <b>North Arrow, Map Scale-Graphic, Map Scale- numeric, Map title, Map legends</b> ) have been included to easily read the map. A suitable <b>base map</b> has been included. The map has been properly <b>captioned</b> . <b>KML/KMZ</b> file(s) included using <b>Google Earth</b> . The map has been well explained with critical justifications. Suitable screen shots of the work have been included in appendix.
<p><b>Part h)</b> Using the provided vector data set about Peradeniya city area in Sri Lanka, develop an area <b>suitability map</b> that clearly indicate the most suitable area for a new <b>Sustainable Energy Generation Research Center</b> to develop within the air distance range of <b>1Km</b> away from <b>Power House</b> and <b>2Km</b></p>	

away from **Urawala Primary School**. Find out the most suitable area for constructing aforesaid new government building while finding following details through geo-processing and analysis tools of QGIS. The map should be followed by a full description of the feasibility of the found area.

- 1) Total number of buildings situated.
- 2) Names of buildings situated.
- 3) Total area of the land occupied by the buildings.
- 4) Total area available within the suitability region.

**(20 Marks)**

0-1	No or Very poor ordinary map has been included.
1-5	Basic map with <b>some</b> required information has been included.
5-12	Very good map with <b>all</b> required information has been included. All standard map elements ( <b>North Arrow, Map Scale-Graphic, Map Scale- numeric, Map title, Map legends</b> ) have been included to easily read the map. A suitable <b>geo-processing tools</b> such as <b>buffering, clipping and intersection etc.</b> has been used. Suitable screen shots of the work have been included in appendix. The map has been explained in general.
12-20	Excellent map with all required information has been included. All standard map elements ( <b>North Arrow, Map Scale-Graphic, Map Scale- numeric, Map title, Map legends</b> ) have been included to easily read the map. A suitable <b>geo-processing tools</b> such as <b>buffering, clipping and intersection etc.</b> has been used. <b>Sub questions</b> have been answered. The map has been well explained. Suitable screen shots of the work have been included in appendix.

#### **Final Grading criteria for the coursework**

Marks	Final Grade
$\geq 70$	1
69-60	2:1
59-50	2:2
49-40	3
$< 40$	fail