Assignment Cover Sheet

Qualification		Module Number and Title	
Higher Diploma in Computing & Software Engineering		CSE5014- Business Analytics	
Student Name & No.		**Assessor**	
		.Induranga De Silva	
Hand out date		**Submission Date**	
Assessment type Coursework-Individual (3000 words equivalent)	**Duration/Length of Assessment Type**	**Weighting of Assessment** 100 %	

Learner declaration
I, …………………………………………<name of the student and registration number>, certify that the work submitted for this assignment is my own and research sources are fully acknowledged.

Marks Awarded	
First assessor	
IV marks	
Agreed grade	
Signature of the assessor		Date	

FEEDBACK FORM

INTERNATIONAL COLLEGE OF BUSINESS & TECHNOLOGY

Module:

Student:

Assessor:

Assignment:

Strong features of your work:

Areas for improvement:

Marks Awarded:

Learning outcomes covered

- Explain Business Analytics methodologies, tools and the techniques.
- Evaluate business advantages produced by business analytics.

- Develop propose solution for a business problem or creating opportunity using appropriate business analytics methodologies, tools and the techniques.

Scenario and the Task

Introduction

The Business Analytics subject domain is considered to be one of the major area where most of the companies and various profitable and non- profitable institutions consider for achieving the best decision support in their respective various operations life cycles. Because of the economy is growing rapidly to achieve tangible and intangible as well as financial and non-financial benefits all most all the government and private organizations required to consider precision and accuracy of their management decisions in all major three managerial levels; operational, tactical and strategic. The amount of information generated in modern agile economic environment is very high and maintaining the consistency also a challenge. Because of the fact that utilization of big data analysis considered as one of the prime concern to deal with credible and valuable informed decision making. During the process of conversion data in to decision supportive information, it is very important to use different data analysis methods, techniques and tools that are comprehensively explained in data analytics. The inclusion of data analytics subject elements in modern management information systems and decision support information systems for enabling effective and efficient online analytical processing incorporated with connected operational databases, data marts and data warehouses considered to be vital.

In performing big data analysis, it is very important to use good statistical software. At present there are many such products available under generic or bespoke software category considering open source or closed source. Usually open source products are financially feasible for many organizations compared to closed source products. At present big data analysis rapid growth

identifiable in open source category with frequently released updated versions and many feature extensions compared to closed source. On the other hand side much reliable many software products have been released by industry pioneer solution providers. Therefore, section of the best product for data analysis is also need to be done wisely by relevant authorities of organizations for their objectives to be precisely achieved.

SCENARIO

Prestige of an occupation considered to be vital factor in planning human resource needed for national development planning in short term and long term. **Identification of factors contribute for prestige** of an occupation is very crucial in deciding what industries and professions need to be given priority when national development planning is done by Sri Lanka government with the support of ministries such as industry and commerce.

The dataset comprises of information about various types occupations in Canada. The information includes education, income, women, prestige, census, occupation type. Any changes in those directly influence to have prestige score variations.

The ministry of industry and commerce is considering getting full-scale data science support to develop essential business intelligence required for further elaboration of the suggestions given by the consultants with the support of national and international real world examples in human resource management. The company has appointed a panel of experts in the field of **data science** to facilitate the aforementioned with the aid of local and foreign intelligence. Assume that you are one of the data scientists in the panel assigned with a special set of tasks to accomplish and provide a report of what factors contribute in prestige of occupation. You are required to complete the below mentioned tasks and prepare a report based on the findings.

The updated dataset has been provided to you as a separate data file labeled "**Prestigue_New.csv**" which was taken from Canada (1971) *Census of Canada*. Vol. 3, Part 6. Statistics Canada [pp. 19-1–19-21].

Survey Data Dictionary

	Data Field (Vector/Variable)	Description
1	Occupation	Name of the occupation
2	education	Average education of occupational incumbents, years, in 1971.
3	income	Average income of incumbents, dollars, in 1971.
4	women	Percentage of incumbents who are women.
5	prestige	Pineo-Porter prestige score for occupation, from a social survey conducted in the mid-1960s.
6	census	Canadian Census occupational code.
7	type	Type of occupation. A factor with levels; bc -Blue Collar; prof - Professional, Managerial, and Technical; wc - White Collar.
8	NA	Data is **N**ot **A**vailable

Tasks

1. Describe with credible examples of possible advantages generated by analytics and business intelligence found in data science for various levels of officials in ministry of **ministry of industry and commerce** of Sri Lanka in associating effective, efficient and suitable informed decisions making. **(5 Marks)**

2. Explain tools, techniques and methodologies going to use for this case study based analysis. . **(6 Marks)**

3. Find out minimum, maximum, mean, median, mode of **income** of the incumbents. . **(6 Marks)**

4. Find out summary statistics of **prestige, education**, **income** of the incumbents. The answers should be followed by descriptive justifications. **(6 Marks)**

5. Conduct central tendency analysis for **prestige, education**, **income** of incumbents**.** Represent finding graphically using bell curves. The answers should be followed by descriptive justifications. . **(10 Marks)**

6. Conduct statistical analysis for determining whether or not **prestige** of incumbents changed **significantly** based on the **type** of occupation. Statistical analysis should be comprised of numerical findings followed by graphical analysis. Discuss the advantages generated by the analysis for informed decision making. . **(12 Marks)**

7. Using statistical hypothetical testing, prove whether there is a statistically significant relationship exists with **prestige** and **education of** incumbents. The answers should be followed by descriptive justifications. **(10 Marks)**

8. Using statistical hypothetical testing, prove whether there is a statistically significant relationship exists with **prestige** and **income of** incumbents**.** The answers should be followed by descriptive justifications. **(10Marks)**

.

9. Using statistical hypothetical testing, prove whether there is a statistically significant relationship exists with **prestige** and percentage of **women of** incumbents**.** The answers should be followed by descriptive justifications.
   (10 Marks)

10. Write a conclusion based on the findings of the data analysis and suggest necessary recommendations as the solution for the problems identified. The answers should be followed by descriptive justifications.   **(15 Marks)**

   • Proper report format carries **10 Marks** separately.

   • Question no 7,8 and 9 should be incorporated with hypothesis based **normality testing**

   • Note: The conclusion can include the findings of suitable **regression analysis** as well.

   • Students are stirckly required to submit their ***R project file*** , ***R script file*** and suitable evidences of own work along with their report. Fialure to do so would disqualify their report for marking.

Marking Scheme

Task-1 contains 5 marks

Criteria	Marks
	Out of 5
Fail Did not Explain the scenario and not included the benefits provided by the analysis for the selected company/institution.	0-1
Pass Explained the scenario.	1-2
Good Explained the scenario and benefits provided by the analysis for the selected company/institution.	2-3
Excellent Well explained the scenario and the benefits provided by the analysis for the selected company/institution.	3-5

Task-2 contains 6 marks

Criteria	Marks
	Out of 6
Fail Did not state tools, techniques and methodologies going to be used for analysis	0-1
Pass Stated tools, techniques and methodologies going to be used for analysis	1-2
Good Stated tools, techniques and methodologies going to be used for analysis. Explained the mentioned tools, techniques and methodologies	2-4
Excellent Stated tools, techniques and methodologies going to be used for analysis. Well explained the mentioned tools, techniques and methodologies.	4-6

Task-3 contains 6 marks

Criteria	Marks
	Out of 6
Fail Did not use minimum, maximum, mean, median, mode functions to get respective statistics of provided data	0-2
Pass Used minimum, maximum, mean, median, mode functions to get respective statistics of provided data	2-4
Good Used minimum, maximum, mean, median, mode functions to get respective statistics of provided data and got an accurate set of results and **explained briefly** about what is obtained.	4-5
Excellent Used minimum, maximum, mean, median, mode functions to get respective statistics of provided data and got an accurate set of results and **explained descriptively** about what is obtained.	5-6

Task-4 contains 6 marks

Criteria	Marks
	Out of 6
Fail Did not include summery statistical data using summary statistical function.	0
Pass Included summery statistical data using summary statistical function and explained	6

Task-5 contains 10 marks

Criteria	Marks
	Out of 10
Fail Did not include bell curves based on data provided	0-1
	1-5

Pass Included bell curves based on data provided		
Good Included bell curves based on data provided and Described information represented in the charts briefly	5-7	
Excellent Included bell curves based on data provided and Well described information represented in the charts briefly	7-10	

Task-6 contains 12 marks

Criteria	Marks
	Out of 12
Fail Did not include hypothetical statement considering null hypothesis and alternative hypothesis	0-1
Pass Included hypothetical statement considering null hypothesis and alternative hypothesis and statistical justification of result based on **mean / variance** analysis tests.	1-6
Good Included hypothetical statement considering null hypothesis and alternative hypothesis and statistical justification of result based on **mean / variance** analysis tests supported by basic discussion and graphical analysis.	6-10
Excellent Included hypothetical statement considering null hypothesis and alternative hypothesis and statistical justification of result based on **mean / variance** analysis tests supported by **extensive** discussion and graphical analysis.	10-12

Task-7 contains 10 marks

Criteria	Marks
	Out of 10
Fail Did not include of hypothetical statement considering null hypothesis and alternative hypothesis	0-1
Pass Included hypothetical statement considering null hypothesis and alternative hypothesis and statistical justification of result based on correlation analysis test	1-5
Good Included hypothetical statement considering null hypothesis and alternative hypothesis and statistical justification of result based on correlation analysis test supported by normality test	6-7
Excellent Included hypothetical statement considering null hypothesis and alternative hypothesis and statistical justification of result based on correlation analysis test supported by **normality** test.	7-10

Task-8 contains 10 marks

Criteria	Marks
	Out of 10
Fail Did not include hypothetical statement considering null hypothesis and alternative hypothesis	0-1
Pass Included hypothetical statement considering null hypothesis and alternative hypothesis and statistical justification of result based on correlation analysis test	1-5
Good Included hypothetical statement considering null hypothesis and alternative hypothesis and statistical justification of result based on correlation analysis test supported by normality test	6-7

Criteria	Marks
Excellent Included hypothetical statement considering null hypothesis and alternative hypothesis and statistical justification of result based on correlation analysis test supported by **normality** test.	7-10

Task-9 contains 10 marks

Criteria	Marks Out of 10
Fail Did not include hypothetical statement considering null hypothesis and alternative hypothesis	0-1
Pass Included hypothetical statement considering null hypothesis and alternative hypothesis and statistical justification of result based on correlation analysis test	1-5
Good Included hypothetical statement considering null hypothesis and alternative hypothesis and statistical justification of result based on correlation analysis test supported by normality test	6-7
Excellent Included hypothetical statement considering null hypothesis and alternative hypothesis and statistical justification of result based on correlation analysis test supported by **normality** test.	7-10

Task-10 contains 15 marks

Criteria	Marks Out of 15
Fail Did not provide conclusion related to statistical findings done based on hypothetical testing.	0-1
Pass	1-5

Provided conclusion related to statistical findings done based on hypothetical testing.	
Good Provided conclusion related to statistical findings done based on hypothetical testing. The conclusion should comprise of sub conclusions of question no 07, 08 and 09.	5-10
Excellent Provided conclusion related to statistical findings done based on hypothetical testing. The conclusion should comprise of sub conclusions of question no 07, 08 and 09. The student provided recommendation to support improvements of the organization. Applications of suitable *regression analysis* findings are highly anticipated.	10-15

Report structure contains 10 marks

Criteria	Marks
	Out of 10
Fail No or weak report structure adhered.	0-1
Pass Standard report structure included but no scientific reporting style used to explain the findings.	1-4
Good Standard report structure included and scientific reporting style used to explain the findings. References included but have some errors.	4-7
Excellent Standard report structure included and scientific reporting style used to explain the findings. Proper citation with matching references included.	7-10

__Submission Guidelines__

Report Structure:

- Executive Summery
- Table of contents, Table of Figures, Table of Tables
- Introduction of the Organization & its operational environment /Scenario
- Data Analysis and Discussion
- Conclusion
- Future Recommendation
- Gantt chart & its Description
- Referencing
- Appendix (Appendix A, Appendix B, etc.) for Group meetings, Samples of Questionnaire

Report Format:

▪ Submission format: Report	▪ Header and Footer: 1 Inch	
▪ Paper Size: A4	▪ Basic Font Size:12	
▪ Words: 3000 words	▪ Line Spacing: 1.5	
▪ Printing Margins: LHS; RHS: 1 Inch	▪ Font Style: Times New Roman	
▪ Binding Margin: ½ Inch	▪ Referencing should be done **strictly using Harvard system**	