Name	Assessment 1 – Data Analytics Notebook
Due	Fri 2 Sept 11:59pm
Weight	40% (indicative weighting)
Submit	Jupyter Notebook to Blackboard (Formative Part A via JupyterHub autograder)

Rationale and Description

Foundational to addressing business concerns with data analytics is an understanding of potential data sources, the kinds of techniques that may be used to process and analyse those data, and an ability to present the final analytics in a way that is meaningful for the stakeholders.

This assessment will involve the completion of a number of quiz notebooks (Part A) and the creation of a complete Jupyter notebook (Part B), which together demonstrate your understanding of the technical process required to ask good questions and obtain meaningful answers using data analytics.

You will use your knowledge from the studio sessions together with the techniques practiced in the tutorial sessions, and apply both to completing a basic skills notebook, and creating another notebook that uses data analytics to address a business concern. You will not only perform the necessary analysis steps, but also provide an explanation of your decision process.

Learning Outcomes

Successful completion of this task will demonstrate:

1. An understanding of how a variety of analysis techniques can be used to take data from different sources and analyse it in a way that is meaningful to a specific question.
2. An ability to shape your data analytics decisions to suit a question within an area of concern.
3. An ability to select, prepare, and use appropriate data, analysis techniques, and visualisations.

Essential Elements

Success in this task will require you to:

1. Select and apply analytics techniques appropriately
2. Demonstrate each phase of the data analytics cycle
3. Document your reasoning and decision-making
4. Provide a cohesive and context appropriate analysis
5. Create a notebook that demonstrates:
    a. foundational data analytics techniques
    b. address 2 questions with data analytics

Detailed Instructions

Using the provided notebook templates, complete Part A quiz notebooks to demonstrate your ability to use foundational techniques, and complete Part B to address 2 questions (provided in the template) by analysing the provided data.

- Question 1 is focused on structured data and will require numerical analysis techniques.
- Question 2 is focused on unstructured or semi-structured data and will require text analysis techniques.

For each question, you will complete a full data analytics cycle, and drawn on appropriate techniques from the first 5 weeks of the unit. *You should avoid using techniques that have not been presented in class, but you may like to vary those techniques if you have the technical knowledge to do so (and if those variations are beneficial to answering the question).*

It is critical that you document your thinking behind the selection of techniques as you will be assessed on your thinking and your ability to make decisions appropriate to the question and area of concern. You need to demonstrate that you *understand* the techniques that you are using, and unnecessary complexity will detract rather than add to the quality of your work.

You will NOT be assessed on the quality of your code apart from its utility to perform the necessary tasks.

The data analytics cycle for each question should tell a story (narrative) that starts with the question and data preparation, moves through the analysis, visualisation of the analytics, and concludes with connecting insights to the question under consideration. The story should clearly address the business concern.

For each step, you must document your decision making and explain why you did what you did. This description of thinking should align with the overall narrative.

1. Question: State the question and describe its significance to the business concern. A description of how you interpret your question should be provided with the question itself.

2. Data: Write the necessary code to obtain the data and make it available for analysis in your notebook. The data may need to be processed prior to analysis, including cleaning in certain cases depending on the analysis techniques being employed.

3. Analysis: Select analysis techniques that are appropriate to each question and implement them appropriately, documenting decisions as you make them.

4. Visualisation: Create visualisations that are appropriate to both the results of your analysis, and for answering the question. You should have at least one visualisation for each question, but you may choose to include more if you decide that it is important to provide the necessary insight.

5. Insight: Answer the question in a way that is meaningful to the business concern. This may involve providing additional descriptive text that explains how the analytics and accompanying visualisation/s address the question.

Prior to submission of your notebooks, reset all cells and restart the kernel, then run the notebook completely, checking that all cells have run without error. Download the resulting notebook/s and submit following the submission instructions for this task.
Important: ensure that you name your notebook with your student number and name.

Submission

The assessment will be completed in 2 parts, and these have different notebook templates, and will be submitted separately:

- **PART A** – Complete the provided Jupyter notebooks (5 quiz notebooks) to demonstrate your basic skills. These notebooks will be released after the relevant content has been covered in class. Part A will be computer graded (autograder), and you will be able to submit multiple times for grading until you are successful with all tasks. Computer marking runs will be initiated regularly through till the assignment due date. Please note that there are no extensions on the autograding of Part A. Updates on when autograding has been completed will be posted to Slack. You will be able to use the error messages in the autograding process to identify bugs and fix your work before resubmission. **Successful completion of all 5 quiz notebooks (full marks) together with a reasonable attempt at Part B will guarantee you a minimum of a 4 for this assessment**. It is advisable that you aim for early success with Part A, so that you can focus most of your time on Part B of the assignment. Note, if Part A is not completed successfully, the grade for this assignment will be based on all notebooks submitted according to the criteria sheet (mostly Part B).

- **PART B** – Using the provided template, follow the instructions to undertake a complete data analytics cycle for each provided question. Download your completed notebook (ensuring no errors and all data and visualisations visible), and submit to the Blackboard assignment upload link before the due date.

Marking Criteria

This assessment is criteria referenced, meaning that your grade for the assessment will be given based on your ability to satisfy key criteria. Refer to the attached Criteria Sheet and ensure that you understand the detailed criteria.

It is important to realise that the assessment not only requires that you know or understand, but also that you *demonstrate* or *provide evidence* of your understanding. This means that you are making your knowledge and understanding clear to the person marking your assignment. It is not the marker's job to try and guess what is in your head, they will mark based on the evidence you provide.

You will *not* receive individual marks or percentages for this assessment. You will receive an overall grade (e.g. pass - 4, high distinction - 7) based on the extent to which you meet the criteria. In general, the criteria is listed in order of importance, so if there is ambiguity in assigning an overall grade, the top criteria will carry more weight.

In general, completion of criteria successfully in order A – D will enable you to achieve higher grades for the assessment. This is indicated clearly in the attached criteria sheet.

IAB303 Assessment 1: Data Analytics Notebook – Criteria Feedback Sheet

Name				**Stud No.**			
Criteria	**1**	**2**	**3**	**4**	**5**	**6**	**7**
[D] Making a Meaningful Connection between Data Analytics and Business Concern	There **is little or no evidence** of a **meaningful connection** between the data analytics and a business concern.			**There is some connection between** the data analytics to a business concern but the overall narrative **lacks meaning and/or clarity**.	**Mostly connects** the data analytics to a business concern in a meaningful way but the overall narrative **could be improved**.	**Makes a meaningful connection** between the data analytics and a business concern through the use of **an appropriate narrative.**	**Makes a meaningful connection** between data analytics and a business concern through an **engaging and consistently clear narrative**.
[C] Applying the Data Analytics Cycle appropriately to a Business Scenario	The **application of the data analytics cycle** to the scenario is either **inappropriate and/or is missing important elements**. There is **little or no** demonstration of **technique selection or use**, and/or technique use is inappropriate for the scenario.			The **application of the data analytics cycle** to the scenario is **complete and mostly appropriate**. Some techniques could be improved.		The **application of the data analytics cycle** to the scenario is **complete** and all techniques are **appropriate** and **used effectively.**	
[B] Effective use of Unstructured Data Analysis Techniques	There is **little or no** evidence of understanding of analysis techniques for **unstructured data**.		Provides evidence of **mostly** understanding analysis techniques for **unstructured data**.	Provides **solid evidence** of understanding analysis techniques for **unstructured data**.			
[A] Effective use of Structured Data Analysis Techniques	There is **little or no** evidence of understanding of analysis techniques for **structured data**.	Provides evidence of **mostly** understanding analysis techniques for **structured data**.	Provides **solid evidence** of understanding analysis techniques for **structured data**.				
PART A Completion	Provides **limited or no evidence** of understanding foundational analysis techniques.			Provides **satisfactory evidence** of understanding foundational analysis techniques. (Minimum grade of 4 with reasonable part B attempt)			
TASK GRADE:	1	2	3	4	5	6	7
Additional comments:							