BSBXBD403

Analyse big data

Assessment 3 of 4

Project

Version 1.1

Assessment Instructions

Task overview

This assessment task is divided into five parts having eight (8) demonstration activities. Read each question carefully before documenting the demonstration task evidence in the spaces provided.

To complete this assessment, you will need the following:

Required software applications

- Microsoft Excel.

Additional resources and supporting documents

BSBXBD403_03_Project_Scenario documents (zipped folder) - This folder contains the following sub-folders, scenario documents and templates required for reference and use when performing the tasks in this assessment.

- AUS Retail_Datasets (folder)
    - AUS Retail_Sales (.xlsx)
    - AUS Retail_Products (.xlsx)
- AUS Retail_Data exploration policy.pdf
- AUS Retail_Data dictionary reference.xlsx
- AUS Retail_Data exploration_template.xlsx
- AUS Retail_ Data flow and dataset schemas.pdf

Assessment Information

Submission

You are entitled to three (3) attempts to complete this assessment satisfactorily. Incomplete assessments will not be marked and will count as one of your three attempts.

All questions must be responded to correctly to be assessed as satisfactory for this assessment.

Answers must be typed into the space provided and submitted electronically via the LMS. Hand-written assessments will not be accepted unless previously arranged with your assessor.

Reasonable adjustment

Students may request a reasonable adjustment for assessment tasks.

Reasonable adjustment usually involves varying:

- the processes for conducting the assessment (e.g. allowing additional time)
- the evidence gathering techniques (e.g. oral rather than written questioning, use of a scribe, modifications to equipment)

However, the evidence collected must allow the student to demonstrate all requirements of the unit.

Refer to the Student Handbook or contact your Trainer for further information.

Please consider the environment before printing this assessment.

Part A: Project scenario

All tasks in this assessment refer to a simulated environment where conditions are typical of a work environment that uses big data related to a fictitious retail business organisation called **'AUS Retail'**.

Read the project scenario carefully before doing the demonstration tasks in Part B.

A1. Company Background

AUS Retail started as a single retail store based in Sydney, NSW. They now have retail store locations across several other states and territories in Australia and continue to grow with the goal of eventually setting up stores across all states in Australia. As the business is growing rapidly, the management requires a more accurate and efficient way to gain insights into their retail sales, stores and products.

Current project

To achieve the organisation's big data analytic requirements, the company has set up a separate team to analyse the organisation's transactional and non-transactional data. The team will be led by the **Chief Data Officer (CDO)**, Mia Gonzales.

Previously, a sample of raw big data from 2018-2021 was tested and validated by the testing team. The two validated datasets (transactional and non-transactional) are now made available for analysts to carry out data exploration tasks. Mia had prepared a series of policy and procedure documents with up-to-date information on the process that should be followed by the analyst team when further exploring the datasets.

A2. Your role

You have recently joined AUS Retail as a trainee analyst and have been given the opportunity to work on the new big data analysis project. Your supervisor is Mia Gonzales (CDO). You must comply with any legislative requirements and follow any standard operating procedures as outlined in AUS Retail's policy and procedure documents when carrying out big data exploration tasks.

A3. Standards, legislative requirements and procedures

You are provided with the following organisational documents and data files related to the fictitious organisation AUS Retail to assist with the big data analysis process.

- **AUS Retail_Data exploration policy.pdf** – This includes organisational procedures, legislative requirements and industry standards for exploring datasets to better understand the data.
- **AUS Retail_Data exploration_template.xlsx** – This is an Excel worksheet template to be used when recording the outcomes of data exploration activities performed on each dataset.
- **AUS Retail_ Data flow and dataset schemas.pdf** – This contains the details of internal organisational systems from which various types of data flows, their relationships and the recommended dataset schemas to be used for categorising data in each dataset.
- **AUS Retail_Data dictionary reference.xlsx** – This contains additional details of the data type, format and data structure specifications for each field included in the given datasets.

A4. Big datasets for analysis

The data analyst team is provided with access to the 'AUS Retail_Datasets' folder which contains historical retail sales data and product data, extracted from the organisation's internal systems for the years 2018-2021.

The *AUS Retail_Datasets* folder contains the following two datasets that need to be further explored.

1. **AUS Retail_Sales (.xlsx)** – transactional dataset, **'Sales'**
2. **AUS Retail_Products (.xlsx)** – non-transactional dataset, **'Products'**

Part B: Identify internal sources of big data to be analysed

To complete this part of the assessment, you are required to:

- carefully read the scenario details outlined in Part A and within this section
- follow the organisational policy, procedures and legislative requirements provided.

Scenario continued:

You have received the following email from your supervisor Mia Gonzales, to brief you about the task at hand.

From: Gonzales, Mia
To: Student Lastname, Student Firstname
Attached documents:

- AUS Retail_Data exploration policy.pdf
- AUS Retail_ Data flow and dataset schemas.pdf
- AUS Retail_Data dictionary reference.xlsx

Subject: Conduct data exploration

Hi <student name>,

Welcome to the big data analysis project team. I would like to brief you on the first task in the analysis project.

First things first! Before conducting any analysis, you must first explore the datasets to better understand their structure. This information must then be documented in a data dictionary as it will be a useful reference throughout the analysis process.

Please refer to the following sections in the *AUS Retail_ Data exploration policy.pdf* document that will provide you with the necessary information and guidelines on how to prepare a data dictionary and the document templates that should be used.

- 4. Data Exploration
- 4.1 Procedure for documenting data exploration activities
- 4.2 Procedure to prepare a data dictionary

You must also consider any legislative requirements that may be applicable to the data obtained from AUS Retail's internal systems. Refer to the:

- *AUS Retail_ Data flow and dataset schemas.pdf* to understand the types of data that flow within each part of the organisation and the general business logic
- *AUS Retail_Data dictionary reference.xlsx* to understand the data types, formats and data structure specifications of the key fields/columns obtained from AUS Retail's internal systems.

All the best on your first task and kind regards,

Mia Gonzales
Chief Data Officer (CDO)
Mia.Gonzales@ausretail.com.au

B1. Identify the structure of the transactional dataset, Sales

In this task, you are required to identify the structure of the transactional dataset 'Sales' by interpreting various sources of information and following organisational policies, procedures and legislative requirements.

Task:

a. Identify and interpret the information from the given scenario, the transactional dataset 'Sales' and the policy and procedure document.

b. Document the required information relevant to the transactional dataset 'AUS Retail_Sales (.xlsx)' by:
   - using recommended documentation templates according to organisational policy and procedures
   - removing any data columns that may not comply with privacy and data protection regulations according to the organisation's legislative requirements
   - using clear, specific, and industry-related terminology to prepare a 'Data Dictionary' for the 'Sales' dataset.

c. Provide a screenshot of the completed 'Data Dictionary' for the 'Sales' dataset in the space provided under 'Screenshot evidence:'

   Note: The screenshot should clearly show the name of the Excel worksheet file displayed on the title bar containing your name initials and current date.

Screenshot evidence:

Assessor instructions: The Data Dictionary's **'Field/Column name'** list should not include **'Customer Name'** details, as this should be removed as part of complying with the legislative requirements. All other fields (as shown in the sample screenshot below) should be identified by the student from the internal data source, **'AUS Retail_Sales.xslx'.** A sample screenshot is provided below.

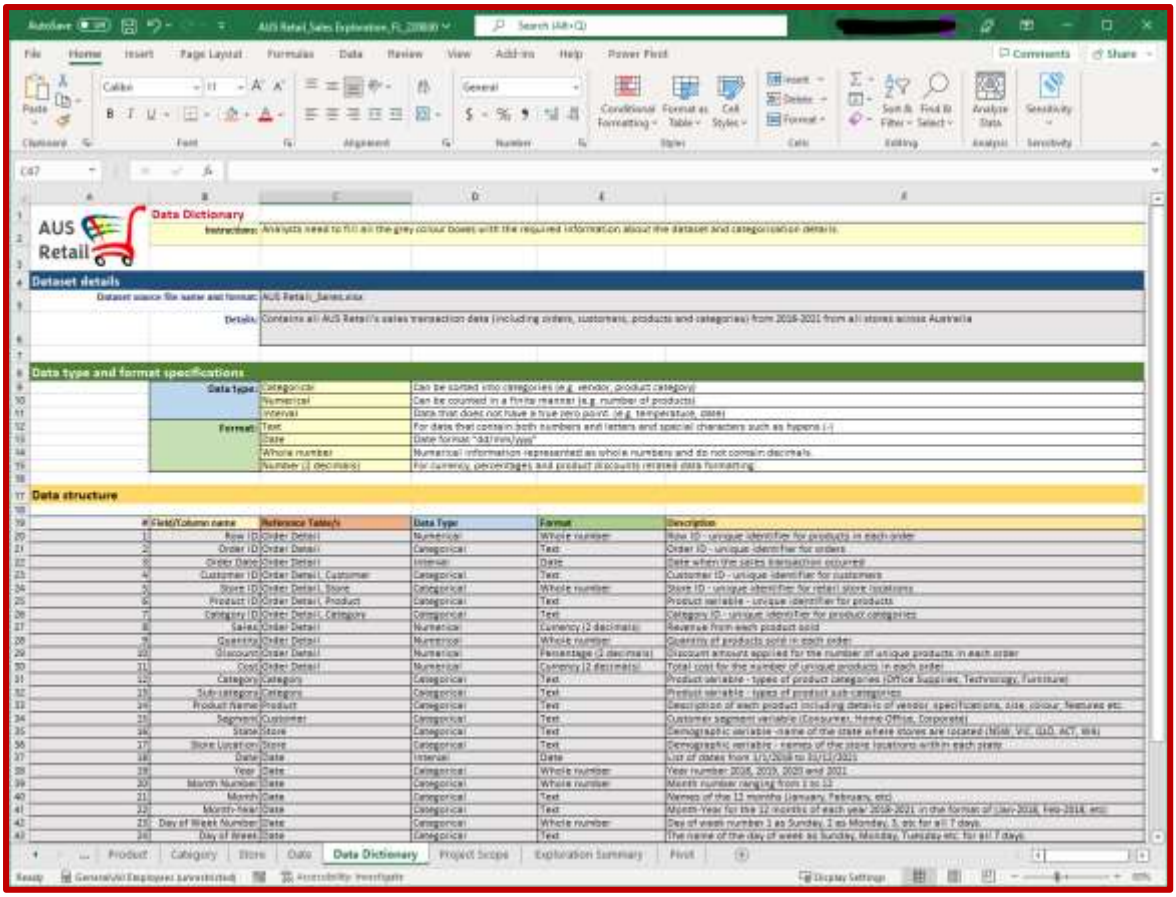


Figure 1 - Screenshot evidence of the completed 'Data Dictionary' tab for the 'Sales' dataset

Important:

- Once the 'Data Dictionary' tab is completed, save and close the Excel worksheet file.
- You will continue to work on this file again in Part C and Part D to complete the rest of the tabs.
- Always retain an up-to-date copy of this Excel worksheet file as you work through the tasks.
- You will be asked to submit a completed version of this file in Part E of this assessment.

B2. Identify the structure of the non-transactional dataset, Products

In this task, you are required to identify the structure of the non-transactional dataset 'Products' by interpreting various sources of information and following organisational policies and procedures.

Task:

a. Identify and interpret the information from the given scenario, the non-transactional dataset 'Products' and the policy and procedure document.

b. Document the required information relevant to the transactional dataset 'AUS Retail_Products (.xlsx)' by:
   - using recommended documentation templates according to organisational policy and procedures
   - using clear, specific, and industry-related terminology to prepare a 'Data Dictionary' for the 'Products' dataset.

c. Provide a screenshot of the completed 'Data Dictionary' for the 'Products' dataset in the space provided under 'Screenshot evidence:'.

Note: The screenshot should clearly show the name of the Excel worksheet file displayed on the title bar containing your name initials and current date.

Screenshot evidence:

Assessor instructions: All fields (as shown in the sample screenshot below) should be identified by the student from the internal data source, **'AUS Retail_Products.xslx'.** A sample screenshot is provided below.

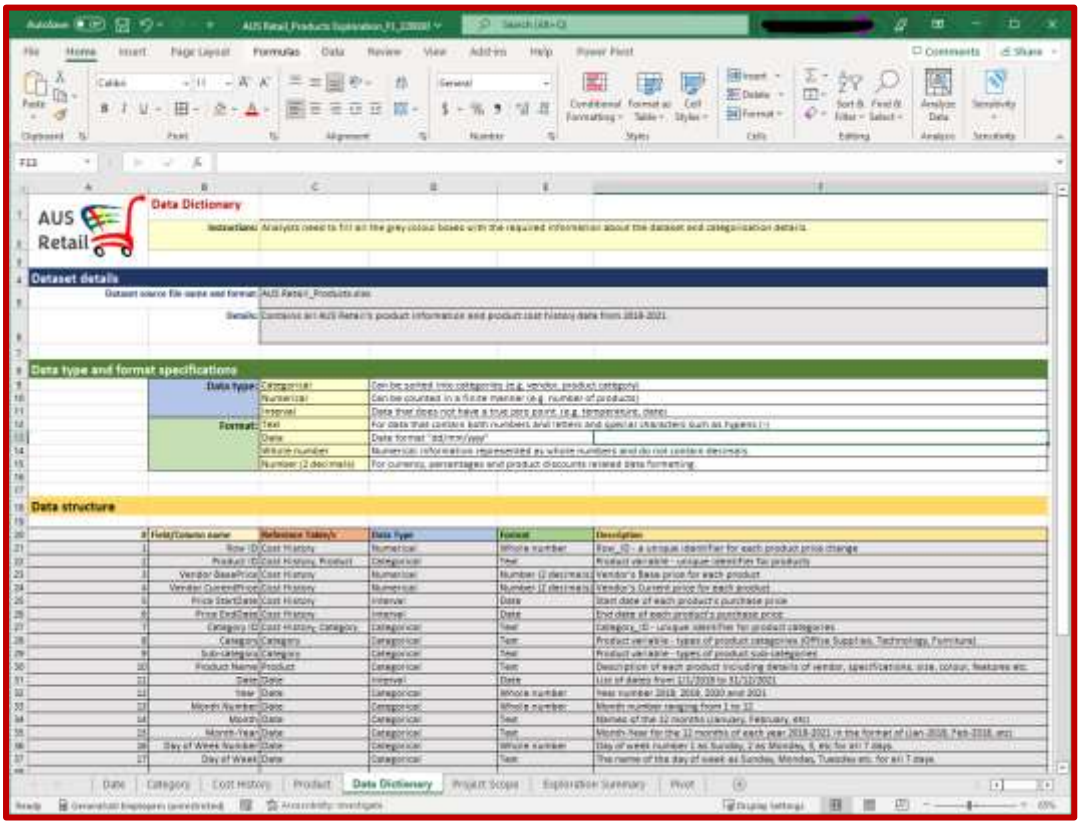


Figure 2 - Screenshot evidence of the completed 'Data Dictionary' tab for the 'Products' dataset

Important:

- Once the 'Data Dictionary' tab is completed, save and close the Excel worksheet file.
- You will continue to work on this file again in Part C and Part D to complete the rest of the tabs.
- Always retain an up-to-date copy of this Excel worksheet file as you work through the tasks.
- You will be asked to submit a completed version of this file in Part E of this assessment.

Part C: Determine the project scope

To complete this part of the assessment, you are required to:

- carefully read the scenario details outlined within this section
- follow sub-section '4.3 Procedure for determining project scope' outlined in the *AUS Retail_ Data exploration policy.pdf* document
- use Microsoft Excel's functions and features to perform the required demonstration tasks
- use the same template documents you have created in Tasks B1 and B2 to record your answers.
  - *AUS Retail_Sales Exploration _NameInitials_ddmmyyyy.xlsx*
  - *AUS Retail_Products Exploration _NameInitials_ddmmyyyy.xlsx*

C1. Outline the scope for analysing the Sales dataset

In this task, you are required to carefully read the scenario details, interpret business requirements and document the scope for analysing the **'Sales'** dataset according to the organisational policies and procedures.

Scenario continued:

During the project kick-off meeting your supervisor Mia Gonzales, outlined the following details regarding AUS Retail's specific requirements for analysing the **'Sales'** dataset.

Considering the **AUS Retail_Sales (.xlsx)** transactional dataset, the following sales, cost and profit summary report have been provided by AUS Retail's management as a reference for the analysis.

Sales, Costs and Profits Summary (2018-2021)

Year	Total Cost	Total Sales	Gross Profit	Gross Profit Margin %
2018	$433,426.06	$482,575.77	$49,149.71	10.18%
2019	$408,307.96	$469,550.09	$61,242.13	13.04%
2020	$526,337.35	$607,759.33	$81,421.98	13.40%
2021	$638,356.38	$731,232.48	$92,876.10	12.70%
Total	**$2,006,427.75**	**$2,291,117.67**	**$284,689.92**	**12.43%**

The purpose of the analysis is to further investigate these values recorded in the report to derive insights into trends and relationships based on the four (4] years of data provided in the dataset. Specifically, the following needs to be included in the sales dataset analysis.

Trends in:

- ✓ 'Total Sales', 'Total Cost', 'Gross Profit' and 'Gross Profit Margin %' by 'Month-Year' per 'Store Location'.
- ✓ 'Total Sales, 'Gross Profit' and 'Gross Profit Margin %' by 'Store Location' per 'Month-Year'.

Relationships between:

- ✓ 'Total Sales' and 'Total Costs' by 'Month-Year' per 'Store Location'
- ✓ 'Total Sales' and 'Gross Profit' by 'Store Location' per 'Month-Year'

Task:

a. Determine organisational requirement for analysing the transactional dataset 'Sales' and record the following information under the 'Project Scope' tab of the template document:
   - Requirements for the analysis
   - Shape and size of the data – use the appropriate Excel functions to complete the required mathematical calculations.

b. Provide a screenshot of the completed 'Project Scope' tab for the 'Sales' dataset in the space provided under 'Screenshot evidence:'.

   Note: The screenshot should clearly show the name of the Excel worksheet file displayed on the title bar containing your name initials and current date.

Evidence of performing the task:

Assessor instructions: The 'Project Scope' must be completed by the student as indicated in the sample screenshot provided below.

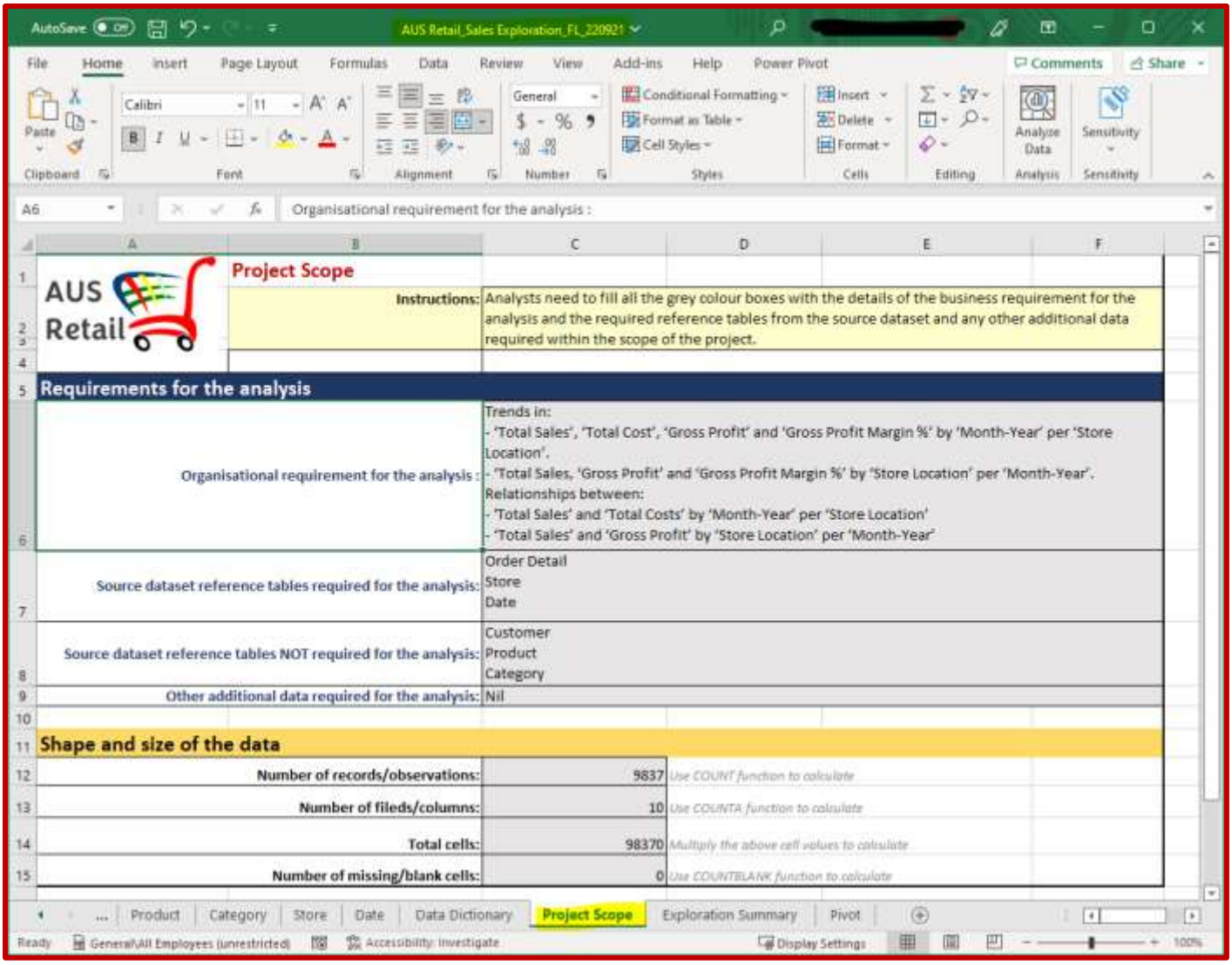


Figure 3 - Screenshot evidence of the completed 'Project Scope' tab for the 'Sales' dataset

Important:

- Once the 'Project Scope' tab is completed, save and close the Excel worksheet file.
- You will continue to work on this file again in Part D to complete the last tab.
- Always retain an up-to-date copy of this Excel worksheet file as you work through the tasks.
- You will be asked to submit a completed version of this file in Part E of this assessment.

C2. Outline the scope for analysing the Products dataset

In this task, you are required to carefully read the scenario details, interpret business requirements and document the scope for analysing the **Products** dataset according to the organisational policies and procedures.

Scenario continued:

During the project kick-off meeting your supervisor Mia Gonzales, outlined the following details regarding AUS Retail's specific requirements for analysing the **Products** dataset.

Considering the **AUS Retail_Products (.xlsx)** non-transactional dataset, the following product purchase costs summary report has been provided by AUS Retail's management as a reference for the analysis.

AUS Retail

Product purchase cost Summary (2018-2021)

Year	Total Vendor price changes	Total unique Products with price changes
2018	1984	1222
2019	2055	1241
2020	2565	1363
2021	3238	1514
Total	**9842**	**1857**

The purpose of the analysis is to further investigate these values recorded in the report to derive insights into trends and relationships based on the four (4) years of data provided in the dataset. Specifically, the following needs to be included in the product dataset analysis.

Trends in:

- ✓ 'Total Vendor price changes', 'Total unique Products with price changes' and 'Total unique Vendors that changed prices' by 'Week Number for each Year' per product 'Sub-category'.
- ✓ 'Total Vendor price changes', 'Total unique Products with price changes' and 'Total unique Vendors that changed prices' by 'Sub-category' per 'Week Number of each Year'.

Relationships between:

- ✓ 'Total unique Vendors that changed prices' and 'Total Vendor Price Changes' by 'Week Number of each Year' for each 'Sub-Category'.
- ✓ 'Total unique Products with price changes' and 'Total unique Vendors that changed prices', by 'Sub-category' for each 'Week Number of each Year'.

Task:

a. Determine organisational requirement for analysing the non-transactional dataset 'Products' and record the following information under the 'Project Scope' tab of the template document:
   - Requirements for the analysis
   - Shape and size of the data – use the appropriate Excel functions to complete the required mathematical calculations.

b. Provide a screenshot of the completed 'Project Scope' tab for the 'Products' dataset in the space provided under 'Screenshot evidence:'.

Note: The screenshot should clearly show the name of the Excel worksheet file displayed on the title bar containing your name initials and current date.

Evidence of performing the task:

Assessor instructions: The 'Project Scope' must be completed by the student as indicated in the sample screenshot provided below.

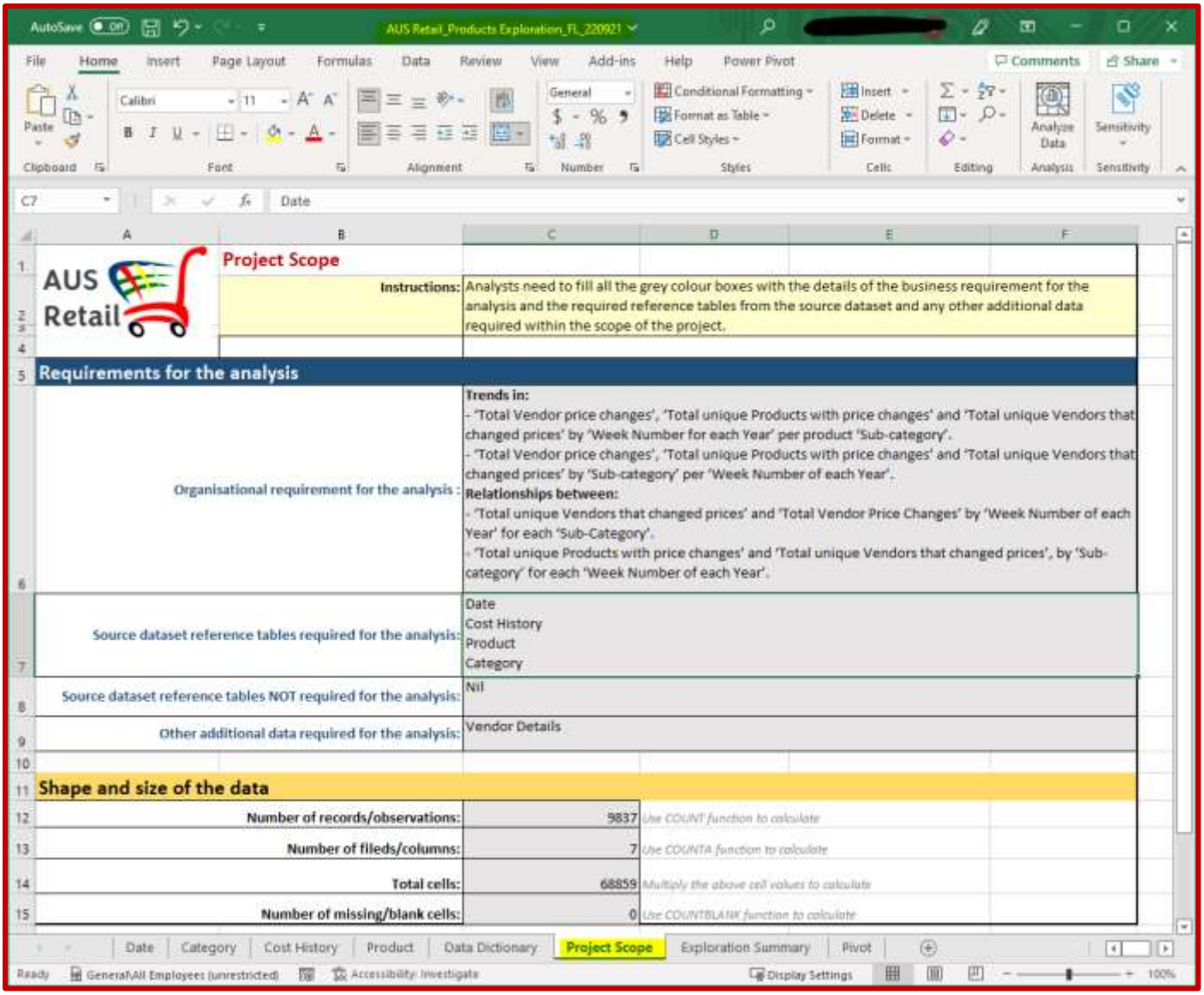


Figure 4 - Screenshot evidence of the completed 'Project Scope' tab for the 'Products' dataset

Important:

- Once the 'Project Scope' tab is completed, save and close the Excel worksheet file.
- You will continue to work on this file again in Part D to complete the last tab.
- Always retain an up-to-date copy of this Excel worksheet file as you work through the tasks.
- You will be asked to submit a completed version of this file in Part E of this assessment.

Part D: Establish parameters to be applied in the analysis

To complete this part of the assessment, you are required to:

- carefully read the scenario details outlined within this section
- follow the organisational policy and procedures provided
- use Microsoft Excel's functions and features to perform the required demonstration tasks.
- use the same template documents you've created in Tasks C1 and C2 to record your answers.
    - *AUS Retail_Sales Exploration _NameInitials_ddmmyyyy.xlsx*
    - *AUS Retail_Products Exploration _NameInitials_ddmmyyyy.xlsx*

D1. Establish parameters for transactional dataset, Sales

In this task, you are required to determine the most important parameters (variables, key measures, data categories and other specifications) to be applied when analysing the 'Sales' dataset in line with the business requirements identified in Task C1.

Task:

Establish the parameters required for analysing the transactional dataset 'Sales' by doing the following.

a. Read through and follow the guidelines in section '4.4 Procedure for establishing parameters' of the *AUS Retail_ Data exploration policy.pdf* document.

b. Outline the details of the required parameters (variables and key measures) and their specifications. You must:
    - identify three (3) variables and three (3) key measures
    - use the 'Exploration Summary' tab of the Excel worksheet file you've created in Task C1 (e.g. *AUS Retail_Sales Exploration _NameInitials_ddmmyyyy.xlsx*) to record the details of the variables and key measures identified.
    - provide a screenshot of the completed 'Exploration Summary' tab in **'Table 1: D1b'** under 'Screenshot evidence:'.

      Note: The screenshot must clearly show the following details:
        - **A full view of the 'Exploration Summary' tab completed with all required information.**
        - **The name of the Excel worksheet file displayed on the title bar containing your name initials and current date.**

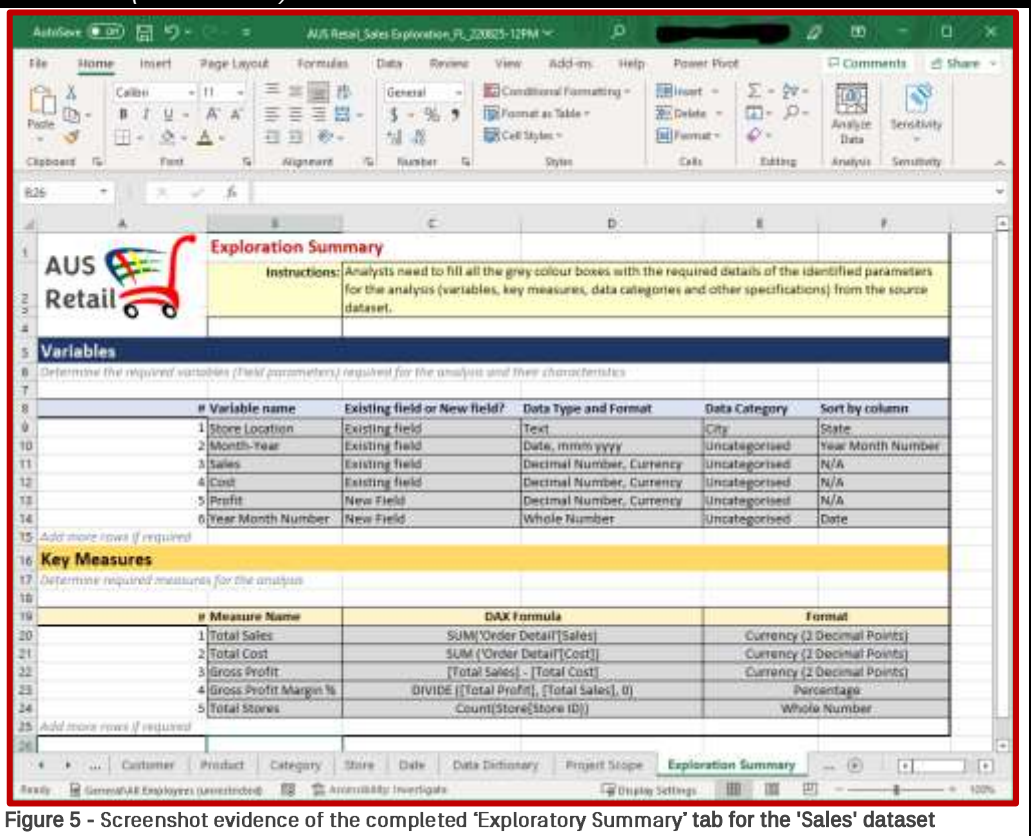
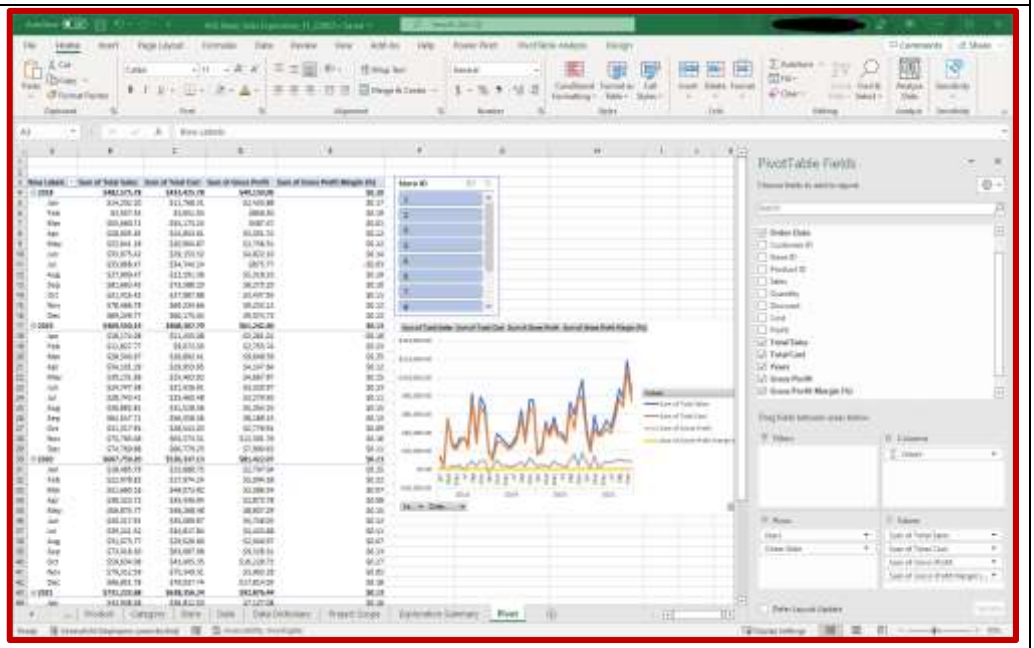
c. Create a PivotTable that summarises and validates the established parameters. You must:
    - use correct mathematical calculations, Excel functions and features to include the identified variables and key measures in the PivotTable
    - accurately record the numerical data in the 'Pivot' tab of the template document
      provide a screenshot of the completed 'Pivot' tab in **'Table 1: D1c'** under 'Screenshot evidence:'.

      Note: The screenshot must clearly show the following details:
        - **A full view of the 'Pivot' tab completed with the summarised information of the dataset, an expanded view of the PivotTables Fields column showing any filters, columns, rows, and values used to create the PivotTable, including the use of slicers and a PivotChart.**
        - **The name of the Excel worksheet file displayed on the title bar containing your name initials and current date.**

Screenshot evidence:

Assessor instructions: The students must identify three (3) variables and three (3) key measures for the analysis. The screenshots below show an additional number of possible variables and key measures that are applicable for the given scenario.

Table 1 - Evidence of establishing parameters for Sales (transactional dataset)

Demonstration tasks:	Evidence *(Screenshot)*:
D1b. **Completed 'Exploration Summary' for Sales (transactional dataset).** **Note:** The screenshot must clearly show the following details: ☐ A full view of the 'Exploration Summary' tab completed with all required information. ☐ The name of the Excel worksheet file displayed on the title bar containing your name initials and current date.	 Figure 5 - Screenshot evidence of the completed 'Exploratory Summary' tab for the 'Sales' dataset
D1c **Completed 'Pivot' tab for Sales (transactional dataset).** **Note:** The screenshot must clearly show the following details: ☐ A full view of the 'Pivot' tab completed with the summarised information of the dataset, an expanded view of the PivotTables Fields column showing any filters, columns, rows, and values used to create the PivotTable, including the use of slicers and a PivotChart. ☐ The name of the Excel worksheet file displayed on the title bar containing your name initials and current date.	 Figure 6 - Screenshot evidence of the completed 'Pivot' tab for the 'Sales' dataset

Important:

- Once the 'Exploratory Summary' and 'Pivot' tabs are completed, save and retain an up-to-date copy of this Excel worksheet file.
- You will be asked to submit the completed version of this file in Part E of this assessment.

D2. Establish parameters for non-transactional dataset, Products

In this task, you are required to determine the most important parameters (variables, key measures, data categories and other specifications) to be applied when analysing the 'Products' dataset in line with the business requirements identified in Task C2.

Task:

Establish the parameters required for analysing the non-transactional dataset 'Products' by doing the following.

a. Read through and follow the guidelines in section '4.4 Procedure for establishing parameters' of the *AUS Retail_ Data exploration policy.pdf* document.

b. Outline the details of the required parameters (variables and key measures) and their specifications. You must:
   - identify three (3) variables and three (3) key measures
   - use the 'Exploration Summary' tab of the Excel worksheet file you've created in Task C2 (e.g. *AUS Retail_Products Exploration _NameInitials_ddmmyyyy.xlsx*) to record the details of the variables and key measures identified
   - provide a screenshot of the completed 'Exploration Summary' tab in **'Table 2: D2b'** under 'Screenshot evidence'.

     Note: The screenshot must clearly show the following details:
     - **A full view of the 'Exploration Summary' tab completed with all required information.**
     - **The name of the Excel worksheet file displayed on the title bar containing your name initials and current date.**

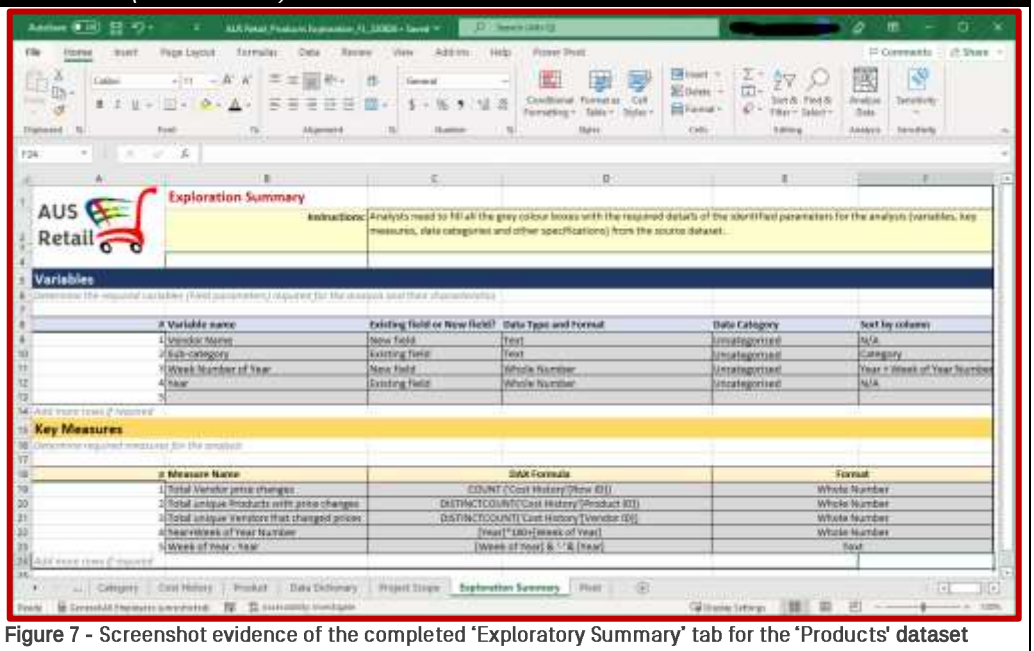
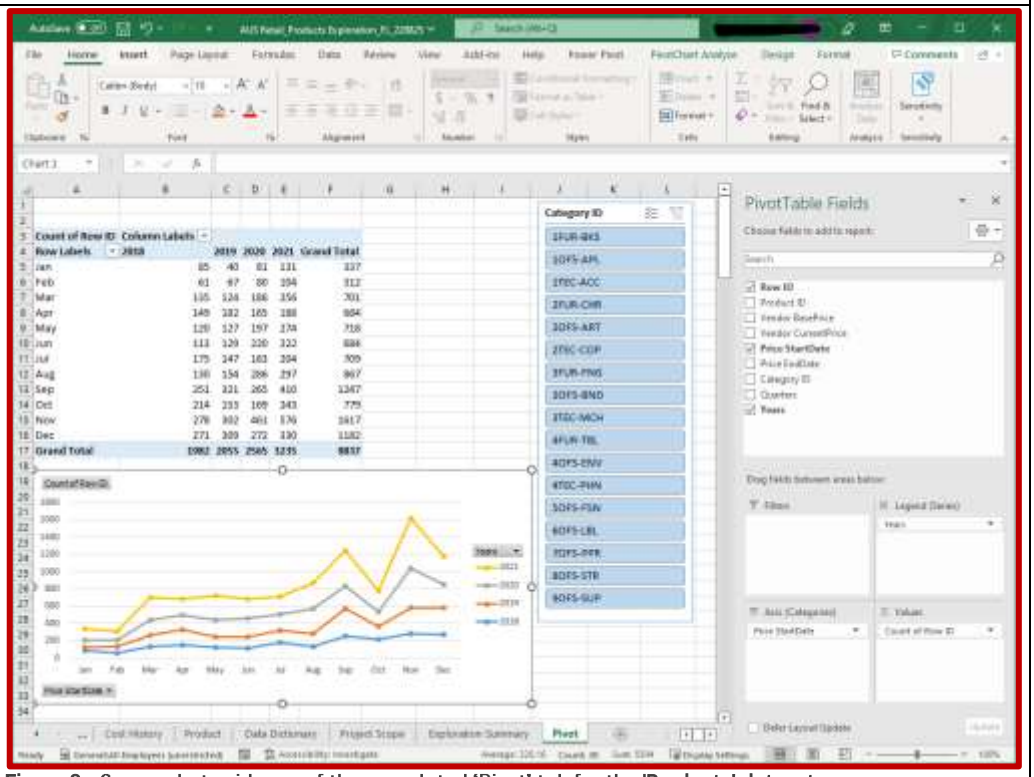
c. creating a PivotTable that summarises and validates the established parameters. You must:
   - use appropriate mathematical calculations, Excel functions and features to include the identified variables and key measures in the PivotTable
   - accurately record the numerical data in the 'Pivot' tab of the template document
   - provide a screenshot of the completed 'Pivot' tab in **'Table 2: D2c'** under 'Screenshot evidence:'.

     Note: The screenshot must clearly show the following details:
     - **A full view of the 'Pivot' tab completed with the summarised information of the dataset, an expanded view of the PivotTables Fields column showing any filters, columns, rows, and values used to create the PivotTable, including the use of slicers and a PivotChart.**
     - **The name of the Excel worksheet file displayed on the title bar containing your name initials and current date.**

Screenshot evidence:

Assessor instructions: The students must identify three (3) variables and three (3) key measures for the analysis. The screenshots below show an additional number of possible variables and key measures that are applicable for the given scenario.

Table 2 - Evidence of establishing parameters for Products (non-transactional dataset)

Demonstration tasks:	Evidence *(Screenshot)*:
D2b. **Completed 'Exploration Summary' for Products (non-transactional dataset).** Note: The screenshot must clearly show the following details: ☐ A full view of the 'Exploration Summary' tab completed with all required information. ☐ The name of the Excel worksheet file displayed on the title bar containing your name initials and current date.	 Figure 7 - Screenshot evidence of the completed 'Exploratory Summary' tab for the 'Products' dataset
D2c. **Completed 'Pivot' tab for Products (non-transactional dataset).** Note: The screenshot must clearly show the following details: ☐ A full view of the 'Pivot' tab completed with the summarised information of the dataset, an expanded view of the PivotTables Fields column showing any filters, columns, rows, and values used to create the PivotTable, including the use of slicers and a PivotChart. ☐ The name of the Excel worksheet file displayed on the title bar containing your name initials and current date.	 Figure 8 - Screenshot evidence of the completed 'Pivot' tab for the 'Products' dataset

Important:

- Once the 'Exploratory Summary' and 'Pivot' tabs are completed, save and retain an up-to-date copy of this Excel worksheet file.
- You will be asked to submit the completed version of this file in Part E of this assessment.

Part E: Store supporting evidence and confirm parameters

To complete this part of the assessment, you are required to

- carefully read the scenario details outlined within this section
- follow the relevant sections outlined in the *AUS Retail_ Data exploration policy.pdf* document.

Scenario:

You have received the following email from your supervisor Mia Gonzales, to brief you about the task at hand.

From: Gonzales, Mia
To: Student Lastname, Student Firstname
Attached documents: AUS Retail_Data exploration policy.pdf

Subject: Advice on storing supporting evidence and confirming parameters

---

Hi <student name>,

Please store all the relevant supporting evidence related to the data exploration tasks you have carried out for both Sales and Product datasets according to our organisation's standard procedure and legislative requirements.

Refer to the following sections of the *AUS Retail_ Data exploration policy.pdf* document for more information.

- ✓ 5. Storing associated supporting evidence
- ✓ 5.1 Legislative requirements
- ✓ 5.2 Procedure for storing results of data exploration activities

Once the documents are completed and stored in the correct location, please notify me via email, as I need to confirm them before you move forward with the analysis. Refer to sub-section '4.5 Procedure for confirming parameters' outlined in the *AUS Retail_ Data exploration policy.pdf* for more details regarding the process.

All the best and kind regards,

Mia Gonzales
Chief Data Officer (CDO)
Mia.Gonzales@ausretail.com.au

Before printing this email please consider the environment.
This message may contain privileged information or confidential information or both and is intended for the recipient named. If you are not the intended addressee, please delete it and notify the sender.

E1. Store supporting evidence for both datasets

In this task, you are required to store the associated supporting evidence you've created when exploring both datasets (Sales and Products) in the previous tasks according to AUS Retail's procedures and legislative requirements.

Read through the scenario (Email received from your supervisor) carefully to understand the relevant sections of the *AUS Retail_ Data exploration policy.pdf* document that you must refer to when doing this task.

Preparation:

As preparation for this task create a new folder in your local computer with your name details called 'BSBXBD403_03_Firstname_Lastname'. For example, a folder created by John Smith should have the name: 'BSBXBD403_03_John_Smith'.

Note: All the folders and files you are required to store as part of this task, should be saved inside this folder.

Task:

a. Create the required folder for storing documents related to data exploration activities according to AUS Retail's standard policy and procedure.

b. Ensure legislative requirements have been met.
Note: If there are any columns/fields in the dataset work files, that do not meet the legislative requirements, these need to be removed before storing the files in the appropriate folder location.

Assessor instructions: The students are expected to remove the data in the **'Customer Name'** field of the 'Customer' tab from the *AUS Retail_Sales Exploration_NameInitials_ddmmyyyy.xlsx* file before storing it in the folder according to the legislative requirement outlined in the policy document.

c. Store the supporting evidence for each dataset, in this folder.
Assessor instructions: The students should store the dataset exploration work files in the **'Supporting documents'** folder according to the policy and procedure document.

d. Provide a screenshot of the folder structure you have created and evidence of storing the relevant documents within this folder, in the space provided under 'Screenshot evidence:'.

e. Create a zipped file of the 'BSBXBD403_03_Firstname_Lastname' folder. You must ensure that this folder includes completed versions of the associated supporting documents for each dataset stored within the correct folder structure before compressing(zipping) it to create the zipped file.

Important: You must submit the 'BSBXBD403_03_Firstname_Lastname' zipped file along with this completed assessment document for marking. Refer to assessment submission instructions for more information.

Screenshot evidence:

Assessor instructions: The folder submitted by the student **'BSBXBD403_03_Firstname_Lastname'** (once downloaded and extracted by assessors) must include the 'Supporting documents' folder. The associated supporting documents for both datasets must be stored within the 'Supporting documents' folder.

The screenshot provided by the student must show the correct folder structure and relevant files are stored within these folders. The sample screenshot provided below indicates the correct folder names, file names and their structure.

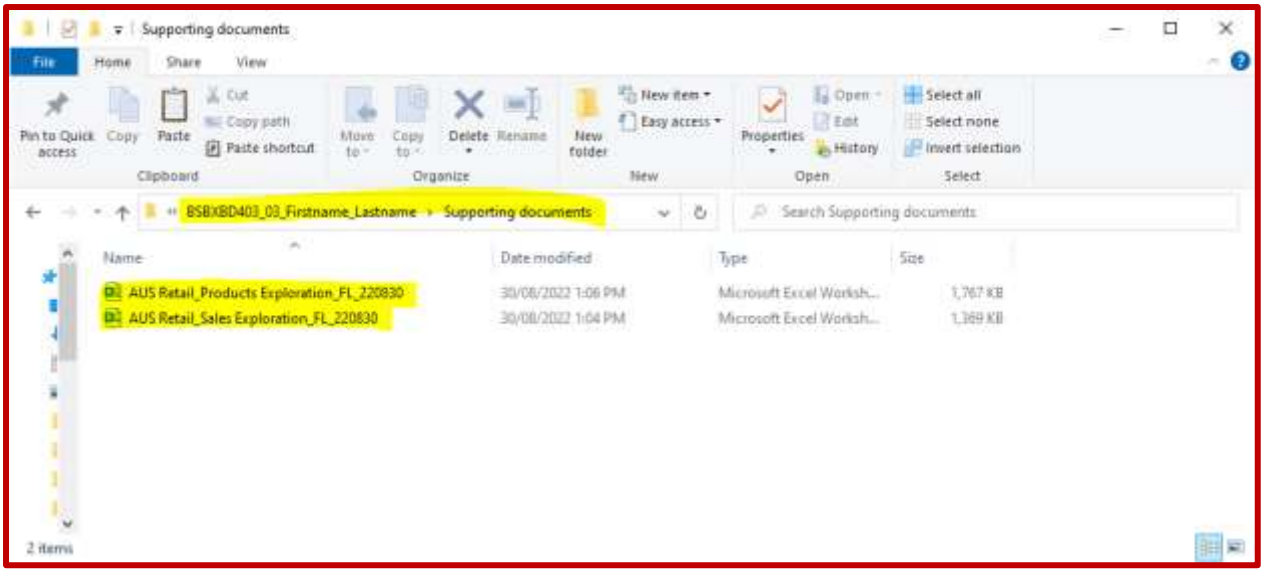


Figure 9 - Screenshot evidence of storing associated supporting evidence.

E2. Confirm parameters to be applied in the analysis of both datasets

In this task, you are required to confirm the parameters by sending this information to your supervisor for confirmation.

Read through the scenario (Email received from your supervisor) carefully to understand the relevant sections of the *AUS Retail_ Data exploration policy.pdf* document that you must refer to when doing this task.

Task:

Write a draft email addressed to your supervisor to request confirmation for the established parameters for both Sales and Product related datasets. When drafting the email, you must:

a. briefly outline the purpose of the email using clear, specific and industry-related terminology
b. mention any supporting evidence or attachments that the email recipient needs to refer to
c. request a response from the recipient, to confirm the established parameters for both datasets
d. use AUS Retail's standard email template to draft the email.

(Word count: 65 – 90 words in the email body).

Answer: Drafted email to Supervisor

Draft your email in the space given below.

Assessor instructions: Student responses are likely to include different wording than the sample answer provided. However, the acceptable responses must:

- be within the specified word limit (for the email body)
- reflect the characteristics described in the exemplar answer.

A sample answer is provided below.

Lastname, Firstname

From: Lastname, Firstname
Sent: Monday, 23 September 2022 1:35 PM
To: Gonzales, Mia

Attached documents: Supporting documents (folder): AUS Retail_Sales Exploration_FL_ddmmyyyy.xlsx, AUS Retail_Products Exploration_FL_ddmmyyyy.xlsx

Subject: Request confirmation for parameters to be applied for the analysis (Sales, Products)

Hi Mia,

I have established the parameters for analysing the Sales and Products datasets and have documented the outcomes of the data exploration activities according to the organisational policy and procedures.

The associated supporting evidence for both datasets is stored in the **'Supporting documents'** folder (which I've attached as a zipped file) for your reference.

Please confirm if the established parameters are correct according to the requirement for analysis, or if there are other parameters that I would need to include.

Thanks, and kind regards

Firstname Lastname

Trainee analyst

Firstname.Lastname@ausretail.com.au

Assessment checklist:

Students must have completed all activities within this assessment before submitting. This includes:

Part B: Determine the purpose and scope of big data analysis		
B1	Screenshot – Evidence of the completed 'Data Dictionary' tab for the Sales dataset.	☐
B2	Screenshot – Evidence of the completed 'Data Dictionary' tab for the Products dataset.	☐
Part C: Determine the project scope		
C1	Screenshot – Evidence of the completed 'Project Scope' tab for the Sales dataset.	☐
C2	Screenshot – Evidence of the completed 'Project Scope' tab for the Products dataset.	☐
Part D: Establish parameters to be applied in the analysis		
D1	Table 1 – Evidence of establishing parameters for the Sales dataset • D1a – Screenshot of the complete 'Exploration Summary' tab • D1b – Screenshot of the completed 'Pivot' tab	☐
D2	Table 2 – Evidence of establishing parameters for the Products dataset • D2a – Screenshot of the complete 'Exploration Summary' tab • D2b – Screenshot of the completed 'Pivot' tab	☐
Part E: Store associated supporting evidence and confirm parameters to be applied in the analysis		
E1	Screenshot – Evidence of the storing supporting evidence for both datasets. File submission – **BSBXBD403_03_Firstname_Lastname** (zipped file) which includes the data exploration work files for each dataset, within the appropriate folder structure. • AUS Retail_Sales Exploration _NameInitials_ddmmyyyy.xlsx • AUS Retail_Products Exploration _NameInitials_ddmmyyyy.xlsx	☐
E2	Email to Supervisor – email draft for confirming parameters.	☐

Congratulations you have reached the end of Assessment [3]!

References:

Learning Container. 2020. *Sample sales data excel xls*. [online] Available at: https://www.learningcontainer.com/download/sample-sales-data-excel-xls/ [Accessed 04 April 2022].