



ASSESSOR GUIDE

BSBXBD403

Analyse big data

Assessment 4 of 4

Project



Assessment Instructions

Task overview

This assessment task is divided into six parts having 12 demonstration activities. Read each question carefully and provide screenshot evidence of performing each demonstration task in the space provided.

To complete this assessment, you will need the following:

Information and telecommunications equipment

- A computer installed with the Windows operating system.
- Microsoft Power BI Desktop App - Download and install the free 'Power BI Desktop' app from Microsoft Store: [Downloads | Microsoft Power BI](https://powerbi.microsoft.com/en-au/downloads/) (Long URL: <https://powerbi.microsoft.com/en-au/downloads/>)

Additional resources and supporting documents

BSBXBD403_04_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_Approved Vendors List (.pdf)
- AUS Retail_Data analysis policy (.pdf)
- AUS Retail_Data transformation specifications [.pdf]
- AUS Retail_Data categorisation specifications [.pdf]
- 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:

1. the processes for conducting the assessment (e.g. allowing additional time)
2. 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 are conducted in a simulated environment where conditions are typical of a work environment that uses big data as it relates 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 off 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, data exploration tasks have been carried out on two big datasets (Sales and Products) to establish the required parameters. Such as variables, key measures, data categories and other specifications to be applied in the analysis. Analysts have documented these parameters clearly using the **AUS Retail's preferred data exploration template** and have requested **Mia Gonzales** to review and confirm the parameters. Mia further discussed **these parameters with AUS Retail's** key stakeholders and management to obtain their feedback and confirmation on the established parameters.
- Now, the analysts are required to carry out further processing and investigations on each dataset to find meaningful relationships and trends based on the confirmed parameters. Mia had prepared a series of policy and procedure documents, technical specifications and task checklists with up-to-date information on the process that should be followed by the analyst team to analyse the big 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 the **AUS Retail's policy and procedure documents** when carrying out big data analysis tasks.

In your role in this new project, you are required to:

- identify trends and relationships within transactional and non-transactional big datasets
- establish data acceptability within the analysed transactional and non-transactional big datasets
- form recommendations based on the analysis
- report on analysis findings.

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 analysis policy.pdf** – This includes organisational procedures, legislative requirements and industry standards for analysing big data.
- **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.

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 datasets that need to be analysed.

- **AUS Retail_Sales (.xlsx)** – transactional dataset, 'Sales'
- **AUS Retail_Products (.xlsx)** – non-transactional dataset, 'Products'
- **AUS Retail_Approved Vendors List (.pdf)** – an additional data source required for the analysis which includes unstructured data.

A5. Specification documents

The following technical specification documents contain details of the confirmed parameters for both 'Sales' and 'Products' datasets that must be referred to when performing data transformations and categorisations in preparation for analysis.

- **AUS Retail_Data transformation specifications.pdf**
- **AUS Retail_Data categorisation specifications.pdf**

Part B: Extract and transform big data in preparation for analysis

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 and procedures provided
- access 'Power Query Editor' from the Microsoft Power BI technology platform to perform data extraction and transformation tasks.

As preparation for this task, do the following.

- Open the 'Power BI Desktop' application.
 - Save two new Power BI work files for each dataset (Sales and Products) using the following file name format with your name initials and current date information:
 - AUS Retail_Sales Analysis_NameInitials_ddmmyyyy
 - AUS Retail_Products Analysis_NameInitials_ddmmyyyy
- E.g. A file saved for the 'Sales' dataset on the 21st September 2022 by John Smith should have the name: 'AUS Retail_Sales Analysis_JS_21092022'

Note: You will be using these Power BI work files to perform the demonstration tasks in this section.

Scenario continued:

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

From: Gonzales, Mia

To: Student Lastname, Student Firstname

Attached document: AUS Retail_Data transformation specifications (.pdf)

Subject: Extract and transform datasets

Hi <student name>,

As we have confirmed all the parameters required for the analysis, you can now extract and transform the **Sales** and **Product** related datasets in preparation for analysis. Please find attached the *AUS Retail_Data transformation specifications.pdf* document that outlines the detailed specifications for the confirmed parameters of the 'Sales' and 'Product' datasets.

Important: Please read through the following sections of the *AUS Retail_Big data analysis policy.pdf* document as they will provide details of the recommended procedures to be used when extracting and transforming the required data from the dataset files.

- 4.1 Procedure to extract data into a new Power BI work file
- 4.2 Procedure to transform data
- 4.3 Procedure to extract and transform unstructured data into a Power BI work file

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.

B1. Extract and transform the structured dataset, Sales

In this task, you will be extracting and transforming the structured, transactional dataset 'Sales' in preparation for analysis.

Task:

Do the following in the Power BI file *AUS Retail_Sales Analysis_NamelInitials_ddmmyyyy (.pbix)*.

- a. Extract the reference tables required from the 'Sales' dataset *AUS Retail_Sales (.xlsx)*, into Power BI.
- b. Transform the 'Sales' data by referring to the relevant specifications provided in the *AUS Retail_Data transformation specifications.pdf* document. Make sure that you:
 - i. edit each new transformation step listed under 'Query Settings' > 'Applied Steps' to reflect the details of the changes made
 - ii. check whether the data in each column is 100% valid, and there are 0% errors and 0% empty cells by using data profiling tools to view the column distribution, column quality and column profile details
 - iii. perform the required table transformations.

Note: The 'Order Detail' table must be grouped by 'Order ID', 'Order Date' and 'Store ID' having summarised values for 'Sales' and 'Cost' columns.

Assessor instructions: See the steps and options selected when grouping the 'Order Detail' table in the screenshot below. This screenshot is for assessor reference only.

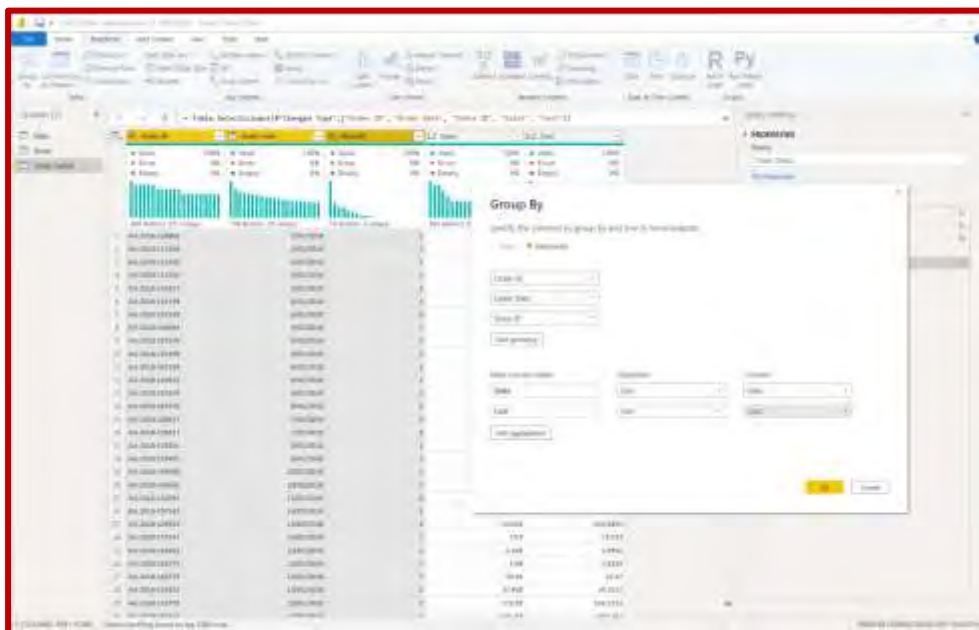


Figure 1 - Screenshot showing the steps used when grouping Power BI Desktop © Microsoft (for Assessor reference only)

- iv. add new columns required for the analysis using M query statements and mathematical formulas.
- c. Provide two (2) screenshots of the 'Power Query Editor' window as evidence of completing the tasks in 'Table 1' under 'Screenshot evidence:'. Read the 'Note' sections carefully to understand the details that must be visible in each screenshot.

Screenshot evidence:

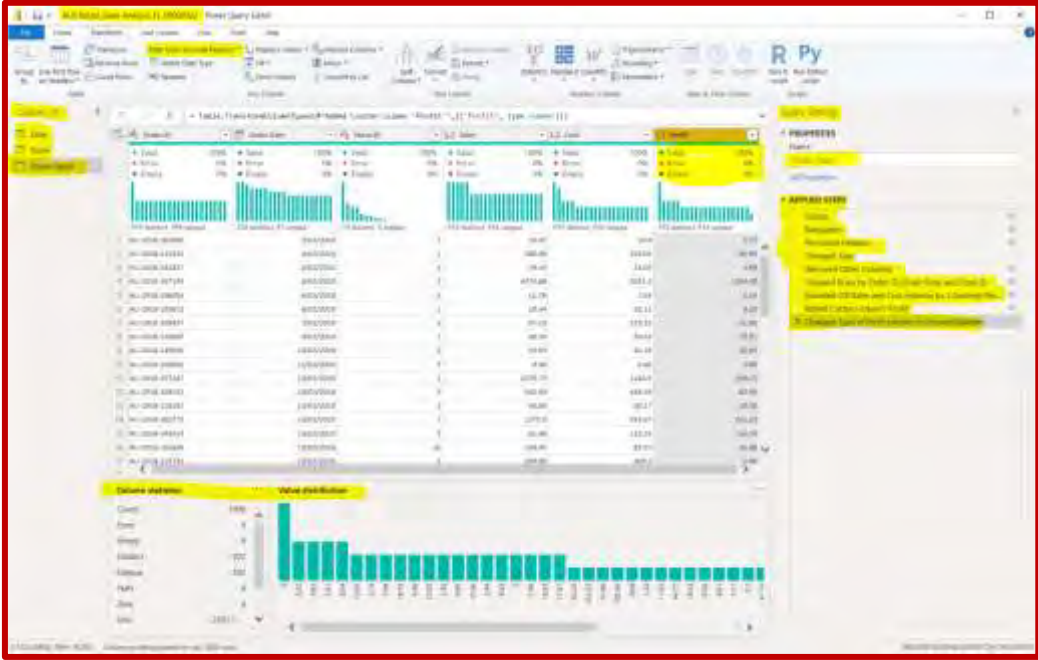
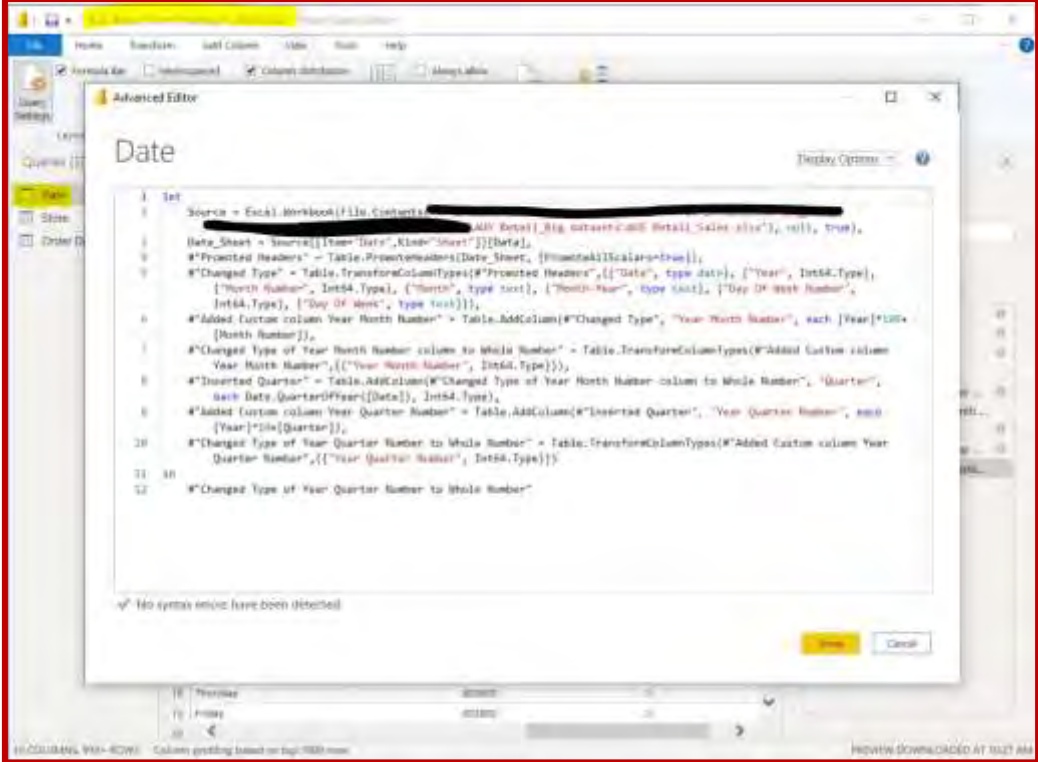
Assessor instructions: The following transformations must be performed by the student after extracting the transactional dataset 'Sales' as indicated in the screenshots provided.

- There are three (3) tables are listed under the 'Queries' pane (e.g. Order Detail, Store, Date).

- The first row is converted as the header row.
- The data type and format of the Profit column is 'Decimal Number', rounded off to 2 decimal points.
- The transformation steps are displayed on the 'Query Settings' > 'Applied Steps'.
- The column profile, quality and distribution details indicate that the column data is 'Valid 100%', 'Errors 0%', and 'Empty 0%'.
- The 'Advanced Editor' window displays the M query statements used to transform the 'Date' table.

Sample screenshots are provided below.

Table 1 - Evidence of extracting and transforming structured transactional dataset, Sales.

Required evidence:	Screenshot
<p>Screenshot of the 'Power Query Editor' window with the 'Profit' column selected.</p> <p>Note: The screenshot must clearly show the following details:</p> <ul style="list-style-type: none"> □ 'Queries' pane. □ 'Query Settings' > 'Applied Steps'. □ 'Transform' tab selected on the 'Power Query Editor' ribbon. □ 'Column statistics' and 'Value distribution' for the selected 'Profit' column. □ Column data validation (Valid, Error, Empty) % values. □ Title bar showing the file name that includes your name initials and current date. 	 <p>Figure 2 - Screenshot 1 for task B1 using Power BI Desktop © Microsoft</p>
<p>Screenshot of the 'Power Query Editor' with the 'Advanced Editor' opened for the 'Date' table.</p> <p>Note: The screenshot must clearly show the following details:</p> <ul style="list-style-type: none"> □ 'Queries' pane with 'Date' table selected. □ 'Advanced Editor' window expanded to show all M query statements used to perform the transformations in the selected table. □ Title bar showing the file name that includes your name initials and current date. 	 <pre> 1 let 2 Source = Excel.Workbook(File.Contents(3 [Redacted Path]), [Redacted Workbook Name], true), 4 Data_Sheet = Source[[Item="Date",Kind="Sheet"]][Data], 5 #Projected Headers = Table.PromoteHeaders(Data_Sheet, {FromTextAsScalars=true}), 6 #Changed Type = Table.TransformColumnTypes(#Projected Headers,{{"Date", type date}, {"Year", Int64.Type}, {"Month Number", Int64.Type}, {"Month", type text}, {"Month-Year", type text}, {"Day Of Week Number", Int64.Type}, {"Day Of Week", type text}}), 7 #Added Custom column Year Month Number = Table.AddColumn(#Changed Type, "Year-Month Number", each [Year]*100 + [Month Number]), 8 #Changed Type of Year Month Number column to Whole Number = Table.TransformColumnTypes(#Added Custom column Year Month Number,{{"Year-Month Number", Int64.Type}}), 9 #Inserted Quarter = Table.AddColumn(#Changed Type of Year Month Number-column to Whole Number, "Quarter", each Date.QuarterOfYear([Date]), Int64.Type), 10 #Added Custom column Year Quarter Number = Table.AddColumn(#Inserted Quarter, "Year-Quarter Number", each [Year]*100 + [Quarter]), 11 #Changed Type of Year Quarter Number to Whole Number = Table.TransformColumnTypes(#Added Custom column Year Quarter Number,{{"Year-Quarter Number", Int64.Type}}), 12 #Change Type of Year Quarter Number to Whole Number </pre> <p>Figure 3 - Screenshot 2 for task B1 using Power BI Desktop © Microsoft</p>

Note: Once the transformations are done and you have recorded the screenshot evidence,

- continue to load the data into Power BI
- save and close your Power BI work file as you will continue to work on this file again in Part C.

B2. Extract and transform the structured dataset, Products

In this task, you will be extracting and transforming the structured, non-transactional dataset 'Products' in preparation for analysis.

Task:

Do the following in the Power BI file *AUS Retail_Products Analysis_NamelInitials_ddmmyyyy (.pbix)*.

- Extract the reference tables required from the 'Products' dataset 'AUS Retail_Products (.xlsx)' into Power BI.
- Transform the 'Products' data by referring to relevant specifications provided in the *AUS Retail_Data transformation specifications.pdf* document. Make sure that you:
 - edit each new transformation step listed under 'Query Settings' > 'Applied Steps' to reflect the details of the changes made
 - check whether the data in each column is 100% valid, and there are 0% errors and 0% empty cells by using data profiling tools to view the column distribution, column quality and column profile details
 - perform the required table transformations
 - add new columns required for the analysis using M query statements and mathematical formulas.
- Provide two [2] screenshots of the 'Power Query Editor' window as evidence of completing the tasks in 'Table 2' under 'Screenshot evidence:'. Read the 'Note' sections carefully to understand the details that must be visible in each screenshot.

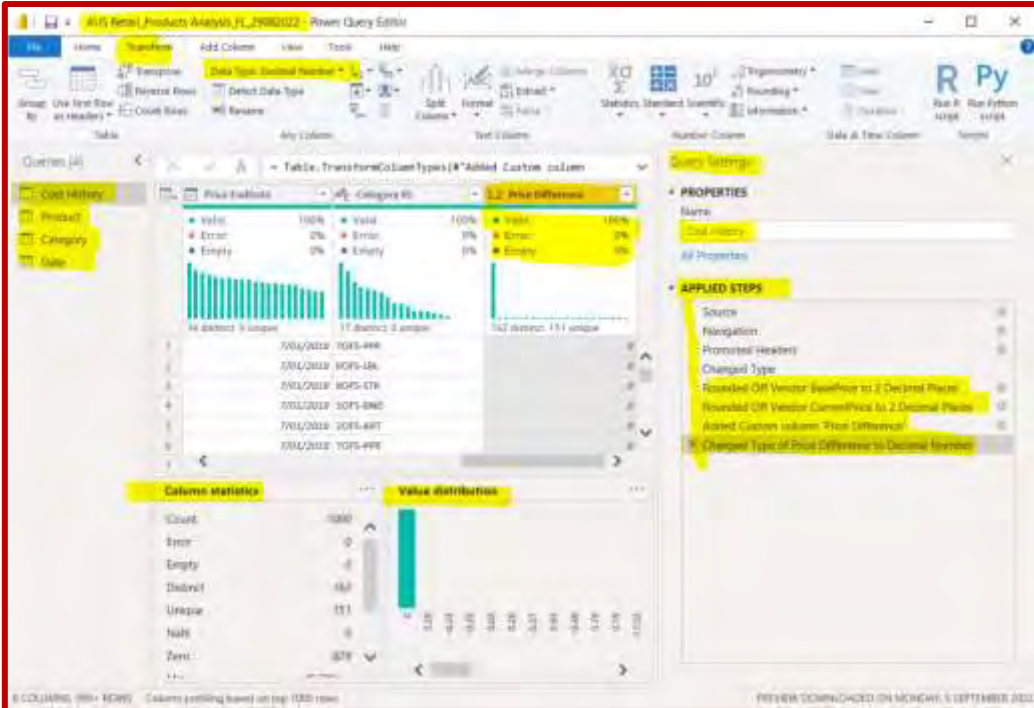
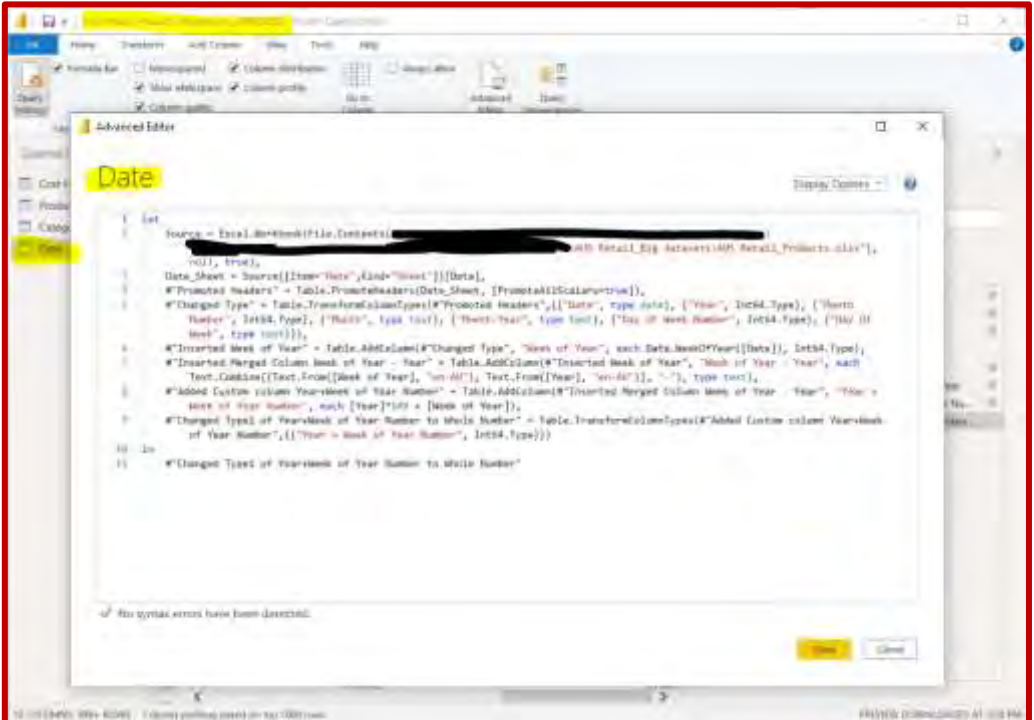
Screenshot evidence:

Assessor instructions: The following transformations must be performed by the student after extracting the non-transactional dataset 'Products' as indicated in the screenshots provided.

- Four tables are listed under the 'Queries' pane (e.g. Cost History, Date, Category, Product).
- The first row is converted as the header row.
- The data type and format of the 'Price Difference' column is 'Decimal Number'.
- The transformation steps should be displayed on the 'Query Settings' > 'Applied Steps'.
- The column profile, quality and distribution details indicate that the column data is 'Valid 100%', 'Errors 0%', and 'Empty 0%'.
- The 'Advanced Editor' window displays the M query statements used to transform the 'Date' table.

Sample screenshots are provided below.

Table 2 - Evidence of extracting and transforming the structured non-transactional dataset, Products.

Required evidence:	Evidence (Screenshot):
<p>Screenshot of the 'Power Query Editor' window with the 'Price Difference' column selected.</p> <p>Note: The screenshot must clearly show the following details:</p> <ul style="list-style-type: none"> <input type="checkbox"/> 'Queries' pane. <input type="checkbox"/> 'Query Settings' > 'Applied Steps'. <input type="checkbox"/> 'Transform' tab selected on the 'Power Query Editor' ribbon. <input type="checkbox"/> 'Column statistics' and 'Value distribution' for the selected 'Price Difference' column. <input type="checkbox"/> Column data validation (Valid, Error, Empty) % values. <input type="checkbox"/> Title bar showing the file name that includes your name initials and current date. 	 <p>Figure 4 - Screenshot 1 for task B2 using Power BI Desktop © Microsoft</p>
<p>Screenshot of the 'Power Query Editor' with the 'Advanced Editor' opened for the 'Date' table.</p> <p>Note: The screenshot must clearly show the following details:</p> <ul style="list-style-type: none"> <input type="checkbox"/> 'Queries' pane with 'Date' table selected. <input type="checkbox"/> 'Advanced Editor' window expanded to show all M query statements used to perform the transformations in the selected table. <input type="checkbox"/> Title bar showing the file name that includes your name initials and current date. 	 <p>Figure 5 - Screenshot 2 for task B2 using Power BI Desktop © Microsoft</p>

Note: Once the transformations are done and you have recorded the screenshot evidence,

- continue to **Load** the data into Power BI
- save and close your Power BI work file as you will continue to work on this file again in **Task B3**.

B3. Extract and transform an unstructured dataset

In this section, you will be extracting and transforming unstructured data from a PDF document.

Scenario continued:

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

From: Gonzales, Mia

To: Student Lastname, Student Firstname

Subject: Extract and transform unstructured data

Hi <student name>,

Good job on extracting and transforming the structured datasets. Your next task involves working with unstructured data.

We have a problem **that needs to be solved before moving forward with the 'Products' data analysis and I do have a solution for it.**

Current issue: To analyse the Products dataset, it is necessary to include the Vendor details (Vendor ID, Vendor Name) **within the Products dataset. Although the vendor's name is mentioned within the product name column for each product, this information is incomplete, inaccurate and ambiguous. Therefore, Vendor details are not currently accessible within the Product dataset in a structured format that can be used for the analysis.**

Solution: I have requested a list of AUS Retail's approved Vendors from the Production department to be used in our project. As a result of this request, a PDF document *AUS Retail Approved Vendor List.pdf* containing a complete list of Vendors (including Vendor ID, Vendor Name and associated Product IDs) have been made available to the analyst team. I have placed a copy of this PDF file in the 'AUS Retail_Datasets' folder for your reference.

Therefore, your next task is to extract and transform Vendor details from the attached PDF file into the existing Power BI data model of the 'Products' dataset to be used in the analysis.

This task needs to be done carefully and in a logical sequence to ensure the unstructured data from the PDF file is transformed into a structured format and linked with the products-related dataset.

Please refer to the sub-section '4.3 Procedure to extract and transform unstructured data into a Power BI work file' outlined in the *AUS Retail Data analysis policy.pdf* document for more information.

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.

Task:

Use the same file you have worked on in **Task B2** (AUS Retail_Products Analysis_NamelInitials_ddmmyyyy.pbix) to do the following.

- a. Extract the unstructured data from the *AUS Retail Approved Vendor List.pdf* document into the Power BI work file for the 'Products' dataset. Make sure that you:
 - expand the extracted source table's 'Data' column to show that they contain the required data for the analysis (i.e. 'Vendor ID', 'Vendor Name' and 'Product ID' information)

- provide a screenshot of the expanded 'Data' table columns in 'Power Query Editor' window in 'Table 3: B3a' under 'Screenshot evidence:'. Read the 'Note' section carefully to understand the details that must be visible in the screenshot.

Assessor instructions: Students should follow steps 1 and 2 of sub-section 4.3 Procedure to extract and transform unstructured data into a Power BI work file outlined in the [AUS Retail_Data analysis policy.pdf](#) document. The detailed steps are outlined as follows for assessor reference:

- **Step 1** - In the *AUS Retail_Products Analysis_NamelInitials_ddmmyyyy (.pbix)*, Power BI work file, go to 'Transform data'.
 - **Step 2** – Use the 'New Source' option to get data from the *AUS Retail_Approved Vendor List.pdf* into the Power Query Editor.
Note: There may be multiple tables/pages in the Navigator window, you would only need to select one page/table to be loaded to the platform at this stage. From this single table, you can obtain the entire set of data from the source during the transformation stage.
 - **Step 3** – Expand the Table to include the required data fields.
- b. Transform the unstructured data into a structured table called 'Vendors&Products' that contains data in separate columns: 'Vendor ID', 'Vendor Name' and 'Product ID'. Make sure that you:
- edit each new transformation step listed under 'Query Settings' > 'Applied Steps' to reflect the details of the changes made.
 - check that the columns in the 'Vendors&Products' table are free of any errors or empty cells.
 - provide a screenshot of the transformed 'Vendors&Products' data table in the 'Power Query Editor' window in 'Table 3: B3b' under 'Screenshot evidence:'. Read the 'Note' sections carefully to understand the details that must be visible in each screenshot.

Assessor instructions: The students are provided with general guidelines in step 3 of sub-section 4.2 Procedure to extract and transform unstructured data into a Power BI work file outlined in the [AUS Retail_Data analysis policy.pdf](#).

The detailed steps followed when performing this task are as follows. Please note that this task can also be done in a different order of steps than what is outlined below.

- Step 1** – Used the table transformation option: 'Remove Top Rows', to remove the top 6 rows of the table.
- Step 2** – Used the column transformation option: 'Fill-down' to fill-in the null columns with the required vendor ID and vendor name details.
- Step 3** – Used the table transformation option: 'Filter by values in a column' to ensure that the column that contains Product ID data is free of any unwanted cell values (remove null, repeating characters and phrases)
- Step 4** – Used the table transformation option: 'Filter by values in a column' > 'Text Filter' > 'Does not begin with' > 'Page' to ensure that the column that contains Product ID data does not contain any page numbering information.
- Step 5** – Use the 'Split column by Delimiter' – function to separate the Vendor ID and Vendor name data values in individual columns.
- Step 6** – Rename the columns as 'Vendor ID', 'Vendor Name' and 'Product ID'
- Step 7** – Remove any unwanted column/s only leaving those columns that include 'Vendor ID', 'Vendor Name' and 'Product ID' details.
- Step 8** - Ensure the columns are formatted with the correct data type.
- Step 9**- Rename the transformed data table as 'Vendors&Products'

- c. Prepare the vendor data to be used in the analysis by using the 'Vendors&Products' table to do the following.
- Combine the 'Vendor ID' column into the 'Cost History' fact table.

- Provide a screenshot of the 'Cost History' table in 'Power Query Editor' window (with the 'Vendor ID' column selected) in 'Table 3: B3c' under 'Screenshot evidence:'. Read the 'Note' sections carefully to understand the details that must be visible in the screenshot.

Assessor instructions: The detailed steps followed when performing this task are as follows:

- **Step 1** – Perform a merge of the **Cost History** Table with the **Vendors&Products** table to include Vendor ID for each column matched against the Product ID column in both tables.
- **Step 2** - Expand the merged Vendors&Products column (untick the option to **Use original column name as prefix**) and only select the Vendor ID column.

d. Prepare a separate 'Vendor' table to be used in the analysis by doing the following.

- Create a table called 'Vendor' by referencing the 'Vendor&Products' table.
- Transform the Vendor table to only include the 'Vendor ID' and 'Vendor Name' columns. Ensure that duplicates are removed and the table contains 314 distinct records.
- Ensure that the 'Vendors&Products' table does not load into the Power BI data model.
- Provide a screenshot of the 'Vendor' table in 'Power Query Editor' window in 'Table 3: B3d' under 'Screenshot evidence:'. Read the 'Note' sections carefully to understand the details that must be visible in the screenshot.

Assessor instructions: The detailed steps followed when performing this task are as follows:

- **Step 1**– Right click on the '**Vendors&Products**' table and select '**Reference**'.
- **Step 2** - Rename the new referenced table as '**Vendor**'
- **Step 3** – Remove the Product ID column and ensure it only contains '**Vendor ID**' and '**Vendor Name**' columns.
- **Step 4** - Remove duplicates. Ensure that altogether 314 vendors are recorded in this table.
- **Step 5** - Right-click on the '**Vendors&Products**' table, then untick the option to 'Enable Load'.

Important:

- Once the transformations are done, continue to 'Load' the data into Power BI.
- Save and close your Power BI work file as you will continue to work on this file again in **Part C**.

Screenshot evidence:

Assessor instructions: Students should provide the required screenshots to demonstrate that the unstructured dataset **AUS Retail_Approved Vendors List.pdf** has been extracted and transformed using 'Power Query Editor'.

Sample screenshots are provided below.

Table 3 - Evidence of extracting and transforming unstructured data.

Demonstration tasks: Evidence (Screenshot):

B3a. Screenshot of the extracted data tables from the unstructured big data source.

Note: The screenshot must clearly show the following details:

- 'Queries' pane.
- 'Query Settings' > 'Applied Steps'.
- 'View' tab selected on the 'Power Query Editor' ribbon.
- 'Column distribution', 'Column quality' and 'Column profile' details.
- Title bar showing the file name that includes your name initials and current date.

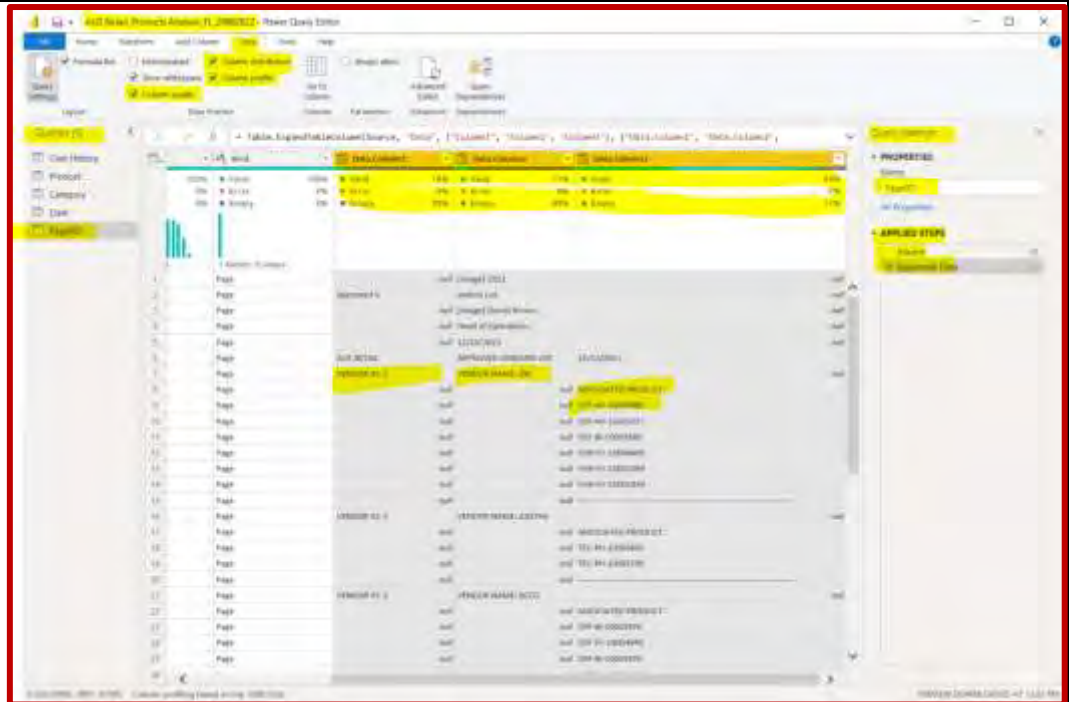


Figure 6 - Screenshot for task B3a using Power BI Desktop © Microsoft

B3b. Screenshot of the transformed Vendors&Products table.

Note: The screenshot must clearly show the following details:

- 'Queries' pane.
- 'Query Settings' > 'Applied Steps'.
- 'Transform' tab selected on the 'Power Query Editor' ribbon.
- Column data validation (Valid, Error, Empty) % values.
- Title bar showing the file name that includes your name initials and current date.

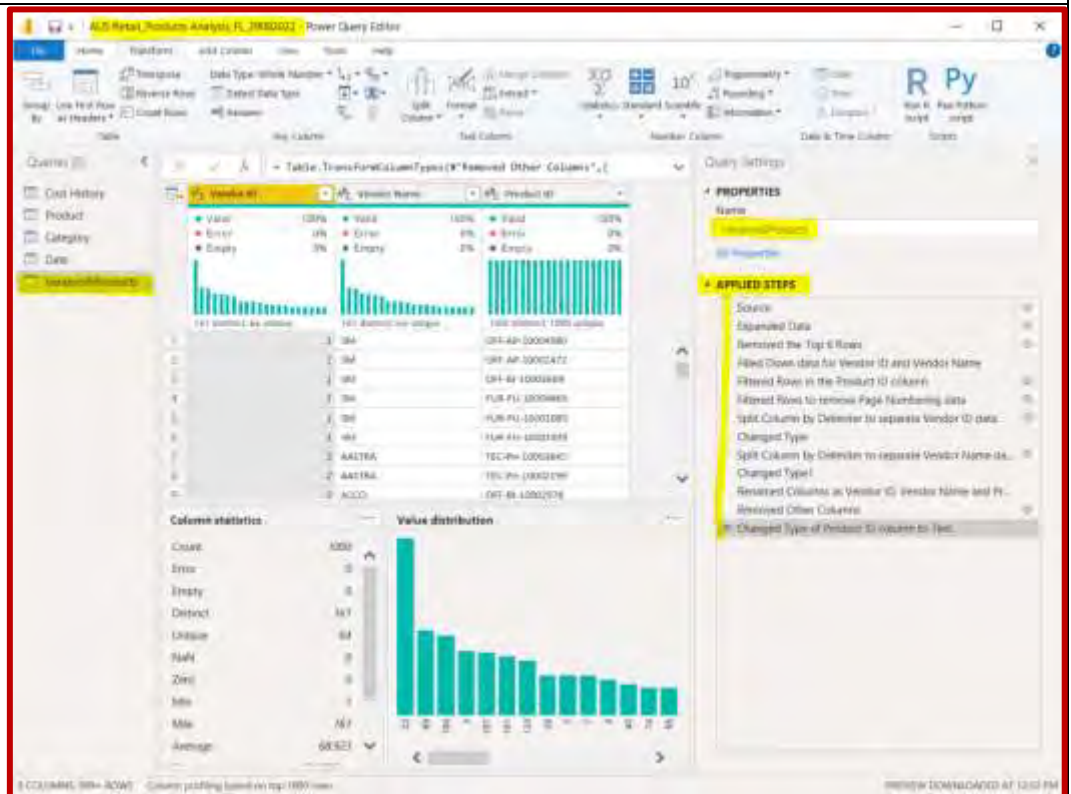


Figure 7 - Screenshot for task B3b using Power BI Desktop © Microsoft

Demonstration tasks: Evidence (Screenshot):

B3c. Screenshot of the **Cost History** table with the merged **Vendor ID** column selected.

Note: The screenshot must clearly show the following details:

- 'Queries' pane.
- 'Query Settings' > 'Applied Steps'.
- 'Transform' tab selected on the 'Power Query Editor' ribbon.
- Column data validation (Valid, Error, Empty) % values.
- Title bar showing the file name that includes your name initials and current date.

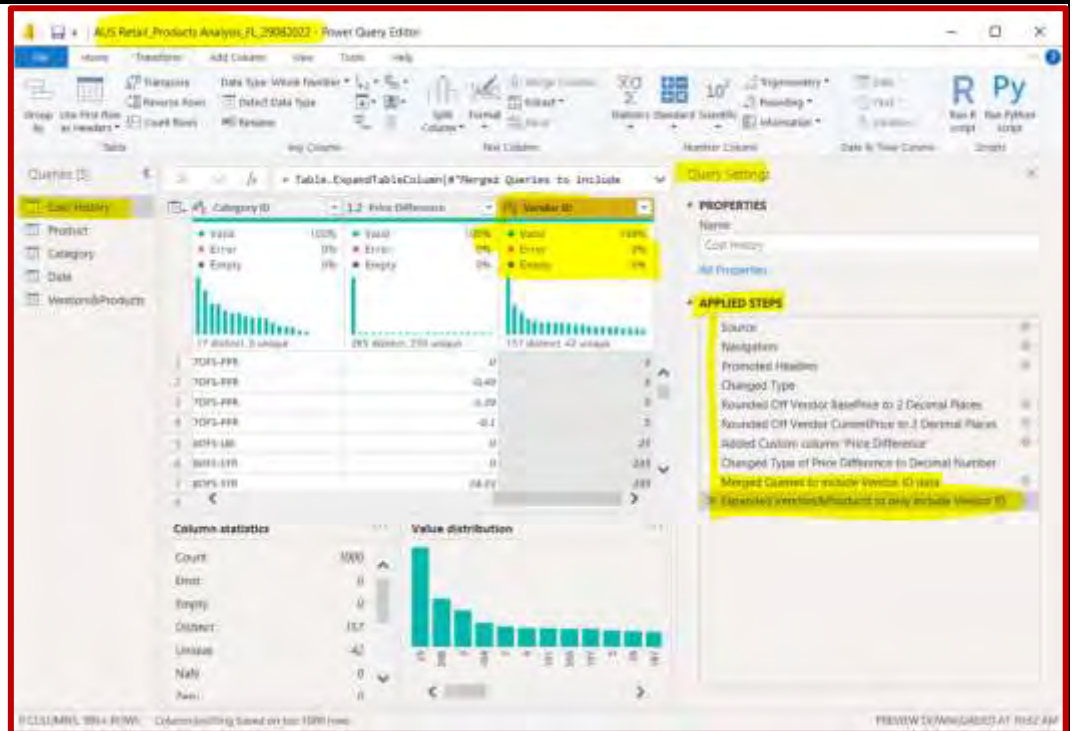


Figure 8 - Screenshot for task B3c using Power BI Desktop © Microsoft

B3d. Screenshot of the **Vendor** table created and transformed.

Note: The screenshot must clearly show the following details:

- 'Queries' pane.
- 'Query Settings' > 'Applied Steps'.
- Column data validation (Valid, Error, Empty) % values.
- Title bar showing the file name that includes your name initials and current date.

Assessor instructions:

The **'Vendors&Products'** table name should display as italicised in 'Power Query Editor' which indicates that it will not load into Power BI.

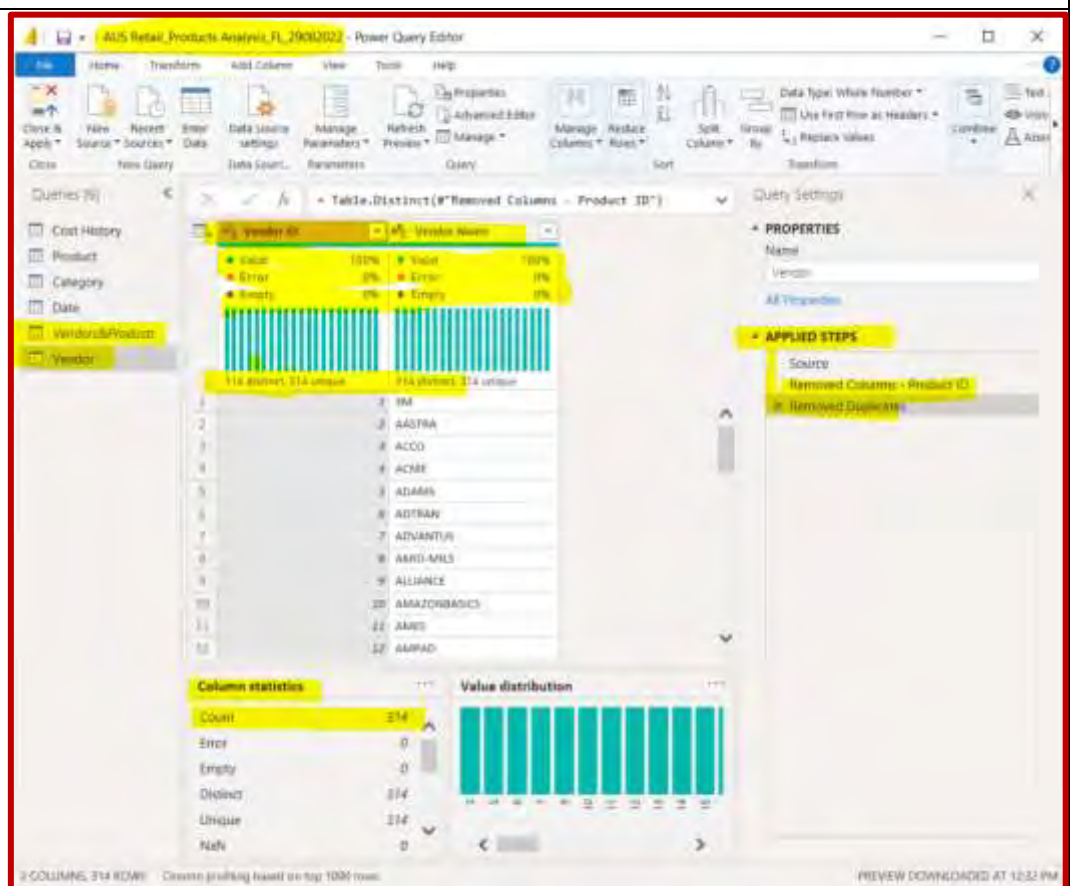


Figure 9 - Screenshot for task B3d using Power BI Desktop © Microsoft

Part C: Categorise and prepare the captured big data for analysis

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

- carefully read the scenario details outlined within this section
- access the 'Microsoft Power BI' technology platform to perform the categorising and preparation tasks
- use the Power BI files that you previously worked on in Part B.

Scenario continued:

You have received the following email from your supervisor Mia Gonzales, briefing you about the next task.

From: Gonzales, Mia

To: Student Lastname, Student Firstname

Attached document: AUS Retail_Data categorisation specifications.pdf

Subject: Categorise and prepare the datasets for analysis

Hi <student name>,

Well done on extracting and transforming both datasets successfully. Your next task is to categorise and prepare the datasets for analysis. **When I say 'categorise' and 'prepare' it means the following.**

- **Categorise:** When categorising the data in the dataset, you first need to create the required parameters (measures and fields) required for the analysis and secondly categorise these parameters using the correct data type, format, category, sorting criteria and summarisations.
- **Prepare:** Once all the parameters are created, categorised, formatted and sorted, you will need to create a dashboard to visualise the data obtained from these parameters. This is a crucial step to further prepare for the analysis. A dashboard helps to provide an overview of the important data points in the dataset and helps to verify how effectively the established parameters can be visualised.

Categorising and preparing each dataset must be done correctly, efficiently and in a logical sequence so that accurate results from the dataset can be obtained during the analysis. Therefore, I have put together four checklists that you must complete to effectively categorise and prepare each dataset.

- Refer to **Table 4 & Table 5** Checklists for categorising and preparing the 'Sales' dataset.
- Refer to **Table 6 & Table 7** Checklists for categorising and preparing the 'Products' dataset.

Please tick/check each task in the checklists as you complete them.

Important:

- When categorising each dataset, please refer to the attached *AUS Retail_Data categorisation specifications.pdf* document that outlines the confirmed parameter specifications.
- When creating the dashboard in preparation for the analysis, please refer to sub-section '**5.1 Dashboard format and structure**' of the *AUS Retail_Data analysis policy.pdf* document to ensure that the recommended format and structure is followed.

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.

C1. Categorise and prepare the Sales dataset for analysis

In this task, you are required to categorise and prepare the captured transactional dataset 'Sales', for analysis.

Task:

Use the same file you have worked on in **Task B1** (AUS Retail_Sales Analysis_NameInitials_ddmmyyyy.pbix) to do the following.

- a. Categorise the required tables, fields and new measures for analysing the 'Sales' dataset by:
 - referring to the relevant confirmed parameter specifications outlined in the *AUS Retail_Data categorisation specifications.pdf* document
 - completing each task in the checklist.

Table 4 - Checklist for categorising the Sales dataset.

#	Task details	Tick/check <input checked="" type="checkbox"/> once completed.
	Do the following in the 'Data' view of Power BI Desktop.	
1	Mark the Date table as 'date table' in the Sales dataset.	<input type="checkbox"/>
2	Create the new measures required for the analysis and categorise each measure.	<input type="checkbox"/>
3	Create the new calculated column required for the analysis and categorise it.	<input type="checkbox"/>
4	Categorise each field required for the analysis.	<input type="checkbox"/>
	Do the following in the 'Model' view of Power BI Desktop.	
5	Create relationships/links between the fact and dimension tables according to the dataset model as outlined in the <i>AUS Retail_Data flow and dataset schemas.pdf</i> document.	<input type="checkbox"/>
6	Hide all fields in each table in the data model, except for those parameters (measures and fields) required for the analysis.	<input type="checkbox"/>

- b. Prepare the **Sales** data for analysis, by:
 - creating a dashboard using the parameters you have previously categorised
 - referring to the organisation's recommended format and structure for the dashboard
 - completing each task in the checklist.

Note: The term 'Field Parameters' when used in the checklist, refers to a method of visualising fields and not as an input to a report visualisation.

Table 5 - Checklist for preparing the Sales dataset.

#	Task details	Tick/check <input checked="" type="checkbox"/> once completed.
	Do the following in the 'Report' view of Power BI Desktop.	
1	Rename the report page as 'Dashboard: Sales Overview'	<input type="checkbox"/>
2	Add the company logo and dashboard title to the dashboard according to the organisation's recommended format and structure. Refer sub-section '5.1 Dashboard format and structure' of <i>AUS Retail_Data analysis policy.pdf</i> .	<input type="checkbox"/>
3	Add 'Card' visuals for each measure (created in Task C1a) to display its value.	<input type="checkbox"/>

#	Task details	Tick/check <input checked="" type="checkbox"/> once completed.
4	Add a 'Slicer' visual to provide the option to select the 'Year' range from 2018-2021.	<input type="checkbox"/>
5	Add a new 'Field parameter' called 'Parameter-Measures' with the measures: <ul style="list-style-type: none"> • 'Total Cost' • 'Total Sales' • 'Gross Profit' Ensure that a 'Slicer' visual is automatically added to the dashboard. Format the 'Slicer Settings' > 'Orientation' to 'Horizontal'.	<input type="checkbox"/>
6	Add a new 'Field parameter' called 'Parameter-Time' with the measures: <ul style="list-style-type: none"> • 'Month-Year' • 'Quarter-Year' • 'Year' Ensure that a 'Slicer' visual is automatically added to the dashboard. Format the 'Slicer Settings' > 'Orientation' to 'Horizontal'.	<input type="checkbox"/>
7	Add an 'Area chart' visual to display 'Total Cost' (Y-axis) by 'Quarter' (X-axis) for each 'Year' (Legend).	<input type="checkbox"/>
8	Add a 'Tree map' visual to display 'Gross Profit' (Values) by 'Store Location' (Category). Format the visual by turning on 'Data labels'.	<input type="checkbox"/>
9	Add a 'Matrix' visual to display the 'Gross Profit Margin %' (Values) for each 'Store Location' (Rows). Ensure that the: <ul style="list-style-type: none"> • Title is set to 'Store Profit Margin %' • column/row totals are turned off • table is expanded to show the full details without having to scroll • 'Gross Profit Margin %' value is conditionally formatted to display data bars • 'Gross Profit Margin %' column is sorted in descending order (high to low % value). 	<input type="checkbox"/>
10	Add a 'Line and clustered column chart' visual to display the 'Parameter-Measure' (Column y-axis) and 'Gross Profit Margin %' (Line-y axis) by 'Parameter-Time' (X-axis). Format the visual by turning on 'Data labels'.	<input type="checkbox"/>
11	Ensure that all visualisations in the dashboard are placed according to the organisation's recommended format and structure. Refer sub-section '5.1 Dashboard format and structure' of <i>AUS Retail_Data analysis policy.pdf</i> .	<input type="checkbox"/>

c. Provide a screenshot of the completed 'Dashboard: Sales Overview' under 'Screenshot evidence:'

Note: The screenshot must clearly show the following details:

- 'Report' view of the Power BI work file.
- 'Fields' pane with all tables expanded.
- Title bar showing the file name that includes your name initials and current date.
- The following should be selected on the dashboard:
 - 'Month-Year' selected in 'Parameter-Time'
 - 'Total Sales' selected in 'Parameter-Measures'
 - The year slicer should have the year range between 2018-2021.
- The 'Line and clustered column chart' visual selected.
- 'Visualizations' pane showing the criteria for the selected visual.

Screenshot evidence:

Assessor instructions: The screenshot provided by the student must indicate that,

- all required measures have been created and only the required parameters are visible under the 'Fields' pane.
- the values and visualisations match what is displayed in the sample screenshot provided.

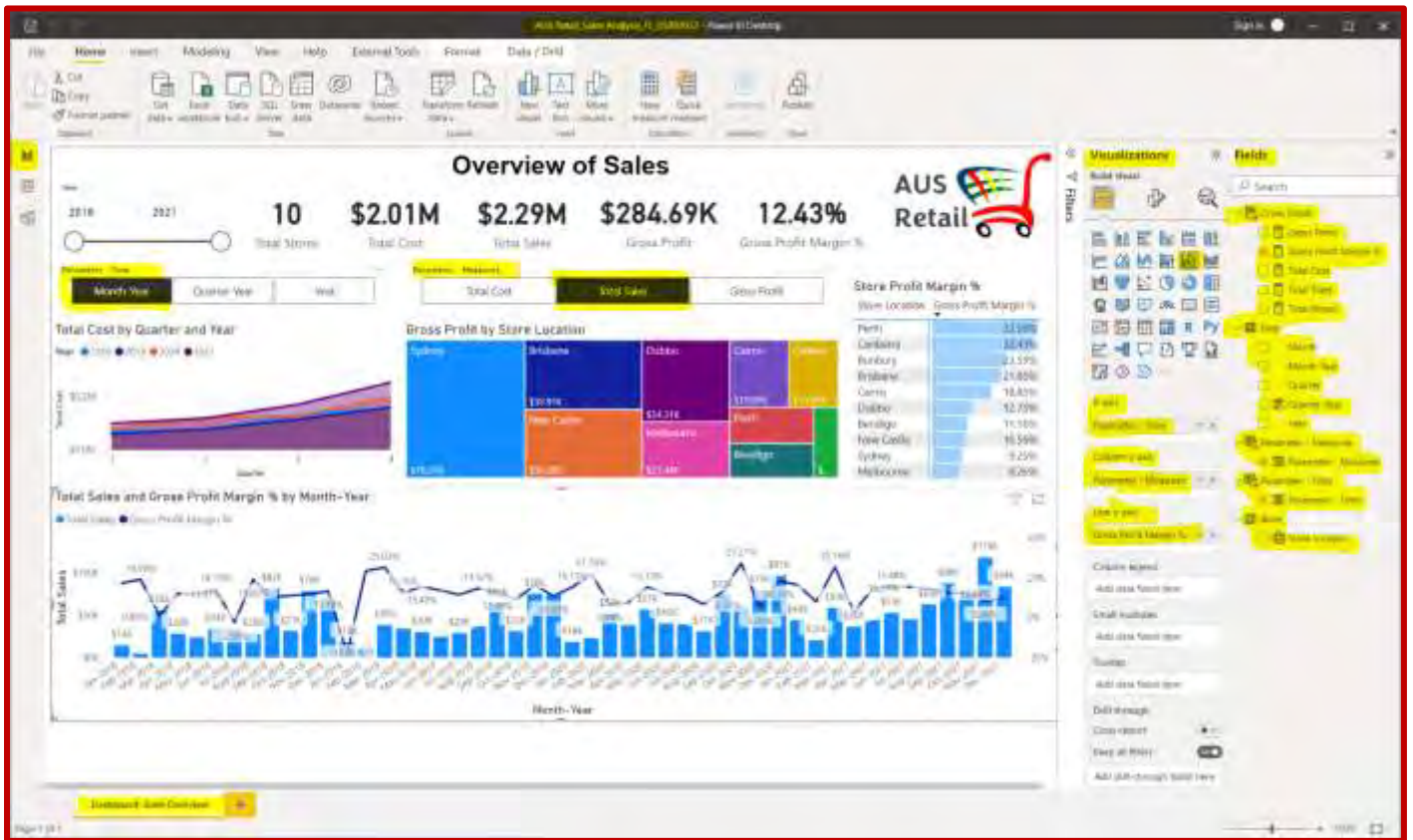


Figure 10 - Screenshot for task C1 using Power BI Desktop © Microsoft

Note: Once the parameters are created, categorised and the dashboard is created, save and close your Power BI work file as you will continue to work on this file again in **Part D**.

C2. Categorise and prepare the Products dataset for analysis

In this task, you are required to categorise and prepare the captured non-transactional dataset **Products**, for analysis.

Task:

Use the same file you have worked on in **Task B3** (AUS Retail_Products Analysis_NamelInitials_ddmmyyyy.pbix) to do the following.

- Categorise the required tables, fields and new measures for analysing the 'Products' dataset by:
 - referring to the relevant confirmed parameter specifications outlined in the *AUS Retail_Data categorisation specifications.pdf* document
 - completing each task in the checklist.

Table 6 - Checklist for categorising the Products dataset.

#	Task details	Tick/check <input checked="" type="checkbox"/> once completed.
1	Do the following in the 'Data' view of Power BI Desktop.	
	Mark the 'Date' table as 'date table' in the 'Products' dataset.	<input type="checkbox"/>

#	Task details	Tick/check <input checked="" type="checkbox"/> once completed.
2	Create the new measures required for the analysis and categorise each measure.	<input type="checkbox"/>
3	Categorise each field required for the analysis.	<input type="checkbox"/>
	Do the following in the 'Model' view of Power BI Desktop.	
4	Create relationships/links between the fact and dimension tables according to the dataset model as outlined in the <i>AUS Retail_Data flow and dataset schemas.pdf</i> document.	<input type="checkbox"/>
5	Hide all fields in each table in the data model, except for those parameters (measures and fields) required for the analysis.	<input type="checkbox"/>

b. Prepare the 'Product' data for analysis, by:

- creating a dashboard using the parameters you have previously categorised
- referring to the organisation's recommended format and structure for the dashboard
- completing each task in the checklist.

Note: The term 'Field Parameters' when used in the checklist, refers to a method of visualising fields and not as an input to a report visualisation.

Table 7 - Checklist for preparing the Products dataset.

#	Task details	Tick/check <input checked="" type="checkbox"/> once completed.
	Do the following in 'Report' view of Power BI Desktop.	
1	Rename the report page as 'Dashboard: Vendor Price Changes Overview'	<input type="checkbox"/>
2	Add the company logo and dashboard title to the dashboard according to the organisation's recommended format and structure. Refer sub-section '5.1 Dashboard format and structure' of <i>AUS Retail_Data analysis policy.pdf</i>	<input type="checkbox"/>
3	Add 'Card' visuals for each measure created to display its value.	<input type="checkbox"/>
4	Add a 'Slicer' visual to provide the option to select the 'Year' range from 2018-2021.	<input type="checkbox"/>
5	Add a new 'Field parameter' called 'Parameter-Measures' with the following measures: <ul style="list-style-type: none"> • 'No. of distinct Product impacted' • 'No. of distinct Vendors contributed' • 'Total Vendor Price changes' Ensure that a 'Slicer' visual is automatically added to the dashboard. Format the 'Slicer Settings' > 'Orientation' to 'Horizontal'.	<input type="checkbox"/>
6	Add a new field parameter called 'Parameter-Time' with the following measures: <ul style="list-style-type: none"> • 'Week of Year' • 'Week of Year – Year' • 'Year' Ensure that a 'Slicer' visual is automatically added to the dashboard. Format the 'Slicer Settings' > 'Orientation' to 'Horizontal'.	<input type="checkbox"/>
7	Add a 'Stacked area chart' visual to display 'No. of distinct Products impacted' (Y-axis) by 'Week of Year' (X-axis) for each 'Year' (Legend).	<input type="checkbox"/>

#	Task details	Tick/check <input checked="" type="checkbox"/> once completed.
8	Add a 'Tree map' visual to display 'Total Vendor Price Changes' (Values) by 'Vendor Name' (Category). Format the visual by turning on 'Data labels'. Filter the visual to only show the top five (5) vendor names by value 'Total Vendor Price Changes'.	<input type="checkbox"/>
9	Add a 'Clustered bar chart' visual to display the 'No. of distinct Vendors contributed' (X-axis) for each 'Sub-category' (Y-axis). Format the visual by turning on 'Data labels'. Sort the visual by 'No. of distinct Vendors contributed' in descending order.	<input type="checkbox"/>
10	Add a 'Stacked column chart' visual to display the 'Parameter-Measure' (Y-axis) by 'Parameter-Time' (X-axis) and product 'Category' [Legend]. Format the visual by turning on 'Total labels'.	<input type="checkbox"/>
11	Ensure that all visualisations in the dashboard are placed according to the organisation's recommended format and structure. Refer sub-section '5.1 Dashboard format and structure' of <i>AUS Retail_Data analysis policy.pdf</i>	<input type="checkbox"/>

c. Provide a screenshot of the completed 'Dashboard: Vendor Price Changes Overview' under 'Screenshot evidence:'.

Note: The screenshot must clearly show the following details:

- 'Report' view of the Power BI work file.
- 'Fields' pane with all tables expanded
- Title bar showing the file name that includes your name initials and current date.
- The following should be selected on the dashboard:
 - 'Week of Year' selected in 'Parameter-Time'
 - 'Total Vendor Price changes' selected in 'Parameter-Measures'
 - The 'Year' slicer visual should have the year range between 2018-2021
- The 'Line and clustered column chart' visual selected
- 'Visualizations' pane showing the criteria for the selected visual.

Screenshot evidence:

Assessor instructions: The screenshot provided by the student must indicate that,

- all required measures have been created and only the required parameters are visible under the 'Fields' pane.
- the values and visualisations match what is displayed in the sample screenshot provided.

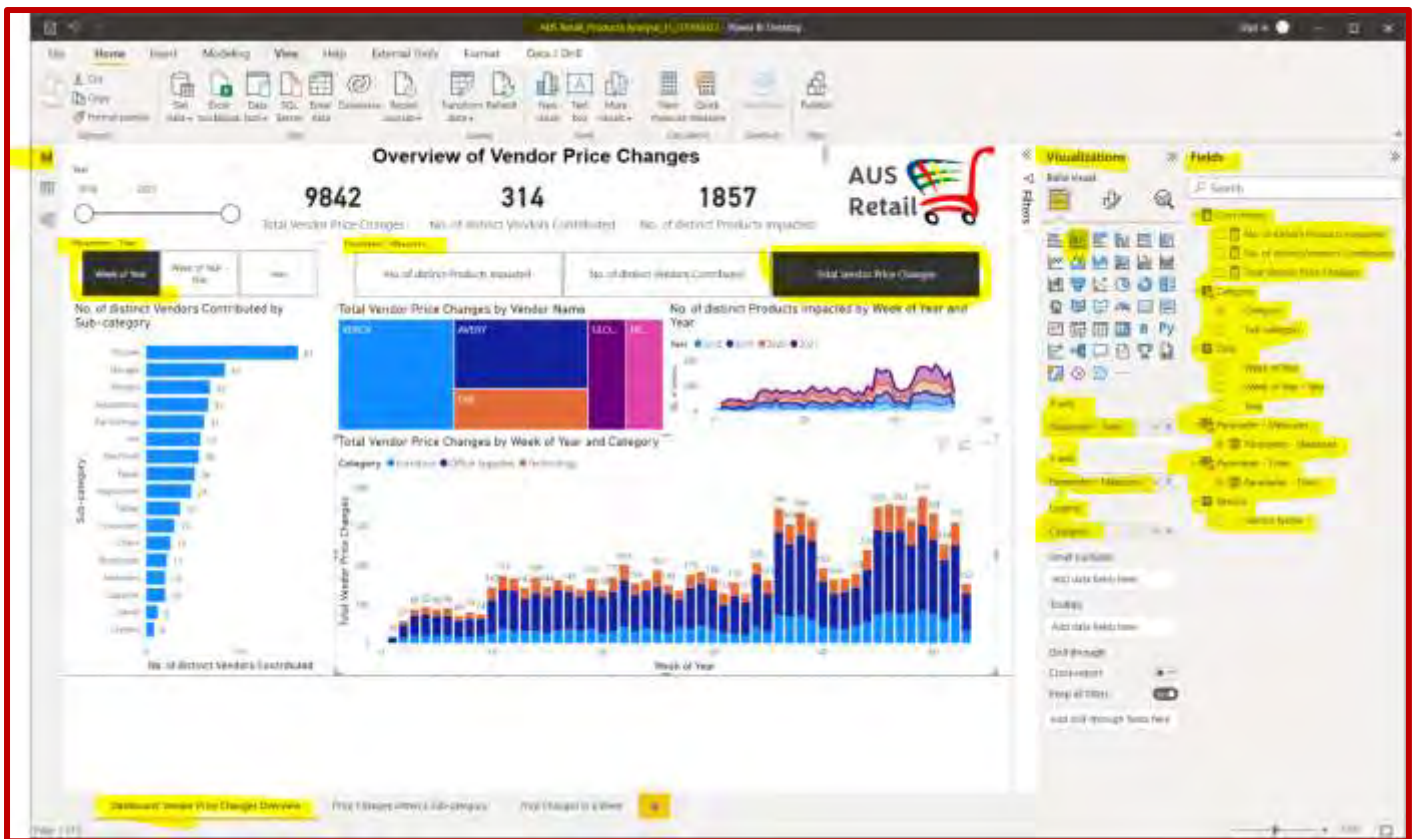


Figure 11 - Screenshot for task C2 using Power BI Desktop © Microsoft

Note: Once the parameters are created, categorised and the dashboard is created, save and close your Power BI work file as you will continue to work on this file again in **Part D**.

Part D: Analyse trends and relationships in big data

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 the required tools and dashboards in Microsoft Power BI technology platform to perform the data analysis tasks
- use the Power BI files that you previously worked on in Part C.

Scenario continued:

You have received the following email from your supervisor Mia Gonzales, briefing you about the next task.

From: Gonzales, Mia

To: Student Lastname, Student Firstname

Subject: Analyse trends and relationships

Hi <student name>,

As you have already categorised and prepared the datasets, now you need to analyse the datasets for trends and relationships using 'drill-through reports'.

Each analysis report must include the following elements as a standard.

- **Company Logo** – Include as per standard report format and structure. The format, structure and placement of these elements are provided in sub-section '6.1 Analysis report format and structure' of the *AUS Retail_Data analysis policy.pdf* document.
- **Report Title** – Include as per standard report format and structure.
- **Summary of key outcomes, results and insights:** Report on the key outcomes, results and insights of the analysis in a short paragraph using clear, specific and industry-related terminology.
- **Visualisations** – The types of visualisations used may vary depending on the requirements of the analysis.
- **Chart Labels** and **Chart Titles** should have meaningful names

To take you through the sequential process of creating the drill-through reports, I have put together four checklists that you must complete to effectively create the drill-through report and perform the analysis.

- Refer to **Table 8 & Table 9** Checklists to create drill-through reports for the 'Sales' dataset analysis.
- Refer to **Table 11 & Table 12** Checklists to create drill-through reports for the 'Products' dataset analysis.

Please tick/check each task in the checklists as you complete them.

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.

D1. Analyse trends and relationships in the Sales dataset

In this task, you will be analysing trends and relationships in the transactional dataset 'Sales' by creating drill-through report pages in Power BI.

Task:

Use the same file you have worked on in **Task C1** (AUS Retail_Sales Analysis_NameInitials_ddmmyyyy.pbix) to do the following.

- a. Create a report page to analyse an individual store's performance. To do this, you must complete the sub-tasks in the checklist given below.

Table 8 - Checklist to perform Task D1a

#	Sub-task details	Tick/check <input checked="" type="checkbox"/> once completed.
1	Create a new report page and rename it 'Store Performance'.	<input type="checkbox"/>
2	Add the company logo and report title as 'Store Performance Analysis' according to the organisational specifications.	<input type="checkbox"/>
3	Add 'Store Location' as a drill-through field to this report page.	<input type="checkbox"/>
4	Add 'Card' visuals for the following parameters to display values within the context of the analysis. <ul style="list-style-type: none"> • 'Total Sales' • 'Total Cost' • 'Gross Profit' • 'Gross Profit Margin %' • 'Store Location' - Format this visual as follows: <ul style="list-style-type: none"> ○ Category label: Off ○ Title: On ○ Title text: Store Location ○ Effects > Background colour: #CBFFD7 	<input type="checkbox"/>
5	Add a 'Slicer' visual to provide the option to select the 'Year' range from 2018-2021.	<input type="checkbox"/>
6	Add a 'Line and clustered column chart' as the primary visual to display the following parameters: <ul style="list-style-type: none"> • Column Y-axis: 'Total Sales', 'Total Costs' • Line Y-axis: 'Gross Profit Margin %' • X-axis: 'Month-Year' Format the visual by turning on 'Data labels'.	<input type="checkbox"/>
7	Generate a summary of the data presented in the 'Line and clustered column chart' visual. Ensure the summary text is customised as follows: <ul style="list-style-type: none"> • Title: On • Title Text: Summary of key outcomes, results and insights • Effects > Background Color: #CBFFD7 	<input type="checkbox"/>
8	Add a 'Waterfall chart' as a secondary visual to display 'Gross Profit' (Y-axis) by 'Year' (Category) and 'Quarter' (Breakdown). Sort the visual by 'Year' in ascending order.	<input type="checkbox"/>
9	Analyse trends and relationships using this report page for the following context using the suggested method. Context: Analyse the monthly performance of the store location that reported the lowest 'Gross Profit Margin %' over 2018-2021. Method: From the 'Dashboard: Sales Overview' page, right-click on the 'Store Location' of interest from a visual (according to the given context) and select 'Drill through' > 'Store Performance' to go to the analysis report page for further details.	<input type="checkbox"/>
10	Provide a screenshot as evidence of the 'Store Performance' report, in 'Table 10: D1a' under 'Screenshot evidence:'. Read the 'Note' sections carefully to understand the details that must be visible in each screenshot.	<input type="checkbox"/>

- b. Create a report page to analyse the sales, costs, profit and margin figures of all stores in a specific month. To do this, you must complete the sub-tasks in the checklist given below.

Table 9 - Checklist to perform Task D1b

#	Sub-task details	Tick/check <input checked="" type="checkbox"/> once completed.
1	Create a new report page and rename it as 'All Stores in Month'.	<input type="checkbox"/>
2	Add the company logo and report title as 'All Stores in Month Analysis' according to the organisational specifications.	<input type="checkbox"/>
3	Add 'Month-Year' as a drill-through field to this report page.	<input type="checkbox"/>
4	Add 'Card' visuals for the following parameters to display values within the context of the analysis. <ul style="list-style-type: none"> • 'Total Sales' • 'Total Cost' • 'Gross Profit' • 'Gross Profit Margin %' • 'Month-Year' -Format this visual as follows: <ul style="list-style-type: none"> ○ Category label: Off ○ Title: On ○ Title text: Month-Year ○ Effects > Background colour: #CBFFD7 	<input type="checkbox"/>
5	Add a 'Line and clustered column chart' as the primary visual to display the following parameters: <ul style="list-style-type: none"> • Column Y-axis: 'Total Sales', 'Total Costs' • Line Y-axis: 'Gross Profit Margin %' • X-axis: 'Store Location' Format the visual by turning on 'Data labels'.	<input type="checkbox"/>
6	Generate a summary of the data presented in the 'Line and clustered column chart' visual. Ensure the summary text is customised as follows: <ul style="list-style-type: none"> • Title: On • Title Text: Summary of key outcomes, results and insights • Effects: Background Colour: #CBFFD7 	<input type="checkbox"/>
7	Add a 'Tree map' as a secondary visual to display 'Gross Profit' (Values) by 'Store Location' (Category). Format the visual by turning on 'Data labels'.	<input type="checkbox"/>
8	Add a 'Matrix' visual as a secondary visual to display the 'Gross Profit Margin %' (Values) for each 'Store Location' (Rows). Ensure that the: <ul style="list-style-type: none"> • Title is set to 'Store Profit Margin %' • 'Column subtotals' and 'Row subtotals' are turned off • table is expanded to show the full details without having to scroll • 'Gross Profit Margin %' value is conditionally formatted to display 'Data bars' • 'Gross Profit Margin %' column is sorted in 'descending' order (high to low % value). 	<input type="checkbox"/>
9	Analyse trends and relationships using this report page for the following context and suggested method. Context: Analyse the sales, costs and profit figures for all store locations for the 'Month-Year' that reported the lowest 'Gross Profit Margin %' over 2018-2021. Method: From the 'Dashboard: Sales Overview' page, right-click on the 'Month-Year' of interest from a visual (according to the given context) and select 'Drill through' > 'All Stores in Month' to go to the analysis report page for further details.	<input type="checkbox"/>



#	Sub-task details	Tick/check <input checked="" type="checkbox"/> once completed.
10	Provide a screenshot as evidence of the 'All Stores in Month' report, in 'Table 10: D1b' under 'Screenshot evidence'. Read the 'Note' sections carefully to understand the details that must be visible in each screenshot.	<input type="checkbox"/>

Screenshot evidence:

Assessor instructions: Each report screenshot must include a summary of the analysis in the form of a short paragraph that reports on the results and insights from each analysis.

Sample screenshots are provided below.

Table 10 - Evidence of analysing the transactional dataset, Sales

Demonstration tasks:	Evidence (Screenshot):
<p>D1a. Report page 1 – Store Performance Analysis</p> <p>Note: The screenshot of the 'Report' must clearly show the following details:</p> <ul style="list-style-type: none"> <input type="checkbox"/> All required elements of the report page (Company Logo, Report title, chart titles, and correct visualisation types) <input type="checkbox"/> 'Card' visuals indicating the context selected from the dashboard page. <input type="checkbox"/> 'Summary of key outcomes, results and insights' included. <input type="checkbox"/> 'Visualisations' pane showing filters based on the context. <input type="checkbox"/> Title bar showing the file name with your name initials and current date. 	 <p>Figure 12 - Screenshot for task D1a using Power BI Desktop © Microsoft</p>
<p>D1b. Report page 2 – All Stores in Month Analysis</p> <p>Note: The screenshot of the 'Report' must clearly show the following details:</p> <ul style="list-style-type: none"> <input type="checkbox"/> All required elements of the report page (Company Logo, Report title, chart titles, and correct visualisation types) <input type="checkbox"/> 'Card' visuals indicating the context selected from the dashboard page. <input type="checkbox"/> 'Summary of key outcomes, results and insights' included. <input type="checkbox"/> 'Visualisations' pane showing filters based on the context. <input type="checkbox"/> Title bar showing the file name with your name initials and current date. 	 <p>Figure 13 - Screenshot for task D1b using Power BI Desktop © Microsoft</p>

Note: Once the analysis reports are created save and close your Power BI work file as you will continue to work on this file again in **Part E1**.

D2. Analyse trends and relationships in the Products dataset

In this task, you will be analysing trends and relationships in the transactional dataset 'Products' by creating drill-through report pages in Power BI.

Task:

Use the same file you have worked on in **Task C2** (AUS Retail_Products Analysis_NamelInitials_ddmmyyyy.pbix) to do the following.

- a. Create a report page to analyse vendor product price changes for a specific product sub-category. To do this, you must complete the sub-tasks in the checklist given below.

Table 11 - Checklist to perform Task D2a

#	Sub-task details	Tick/check <input type="checkbox"/> once completed.
1	Create a new report page and rename it as 'Price Changes within a Sub-category'	<input type="checkbox"/>
2	Add the company logo and report title as 'Analysis of Vendor Price Changes within a Sub-category' according to the organisational specifications.	<input type="checkbox"/>
3	Add 'Sub-category' as a drill-through field to this report page.	<input type="checkbox"/>
4	Add 'Card' visuals for the following parameters to display values within the context of the analysis. <ul style="list-style-type: none"> • 'Total Vendor Price Changes' • 'No. of distinct Vendors Contributed' • 'Sub-category' - Format this visual as follows: <ul style="list-style-type: none"> ○ Category label: Off ○ Title: On ○ Title text: Product Sub-category ○ Effects > Background colour: #C9E1F7 	<input type="checkbox"/>
5	Add a 'Slicer' visual to provide the option to select the 'Year' range from 2018-2021.	<input type="checkbox"/>
6	Add a 'Stacked column chart' as the primary visual to display the following parameters: <ul style="list-style-type: none"> • Y-axis: 'Total Vendor Price Changes' and 'No. of distinct Vendors Contributed' • X-axis: 'Week of Year – Year' Format the visual by turning on 'Total labels'. Sort the visual by 'Week of Year – Year' in ascending order.	<input type="checkbox"/>
7	Generate a summary of the data presented in the 'Line and stacked column' chart visual. Ensure the summary text is customised as follows: <ul style="list-style-type: none"> • Title: On • Title Text: Summary of key outcomes, results and insights • Effects > Background Color: #C9E1F7 	<input type="checkbox"/>
8	Add a 'Tree map' as a secondary visual to display 'Total Vendor Price Changes' (Values) by 'Vendor Name' (Category). Format the visual by turning on 'Data labels'. Filter the visual to only show the top five (5) vendor names by value 'Total Vendor Price Changes'.	<input type="checkbox"/>

#	Sub-task details	Tick/check <input checked="" type="checkbox"/> once completed.
9	<p>Add a 'Matrix' visual as a secondary visual to display the 'No. of distinct Products Impacted' (Values) for each 'Vendor Name' (Rows). Format the visual by;</p> <ul style="list-style-type: none"> • setting the 'Title' as 'Distinct Products Impacted by Vendor Name' • conditionally formatting the 'No. of distinct Products Impacted' value to display 'Data bars' • sorting the 'No. of distinct Products Impacted' column in descending order (high to low % value). 	<input type="checkbox"/>
10	<p>Analyse trends and relationships using this report page for the following context using the suggested method.</p> <p>Context: Analyse the product 'Sub-category' that reported the highest 'No. of distinct Vendors contribution' over 2018-2021.</p> <p>Method: From the 'Dashboard: Vendor Price Changes Overview' page, right-click on the product sub-category of interest from a visual (according to the given context) and select 'Drill through' > 'Price Changes within a Sub-category' to go to the analysis report page for further details.</p>	<input type="checkbox"/>
11	<p>Provide a screenshot as evidence of the 'Store Performance' report, in 'Table 13: D2a' under 'Screenshot evidence:'. Read the 'Note' sections carefully to understand the details that must be visible in each screenshot.</p>	<input type="checkbox"/>

- b. Create a report page to analyse the vendor product price changes for all product sub-categories in a specific week. To do this, you must complete the sub-tasks in the checklist given below.

Table 12 - Checklist to perform Task D2b

#	Sub-task details	Tick/check <input checked="" type="checkbox"/> once completed.
1	Create a new report page and rename it as 'Price Changes in a Week'	<input type="checkbox"/>
2	Add the company logo and report title as 'Analysis of Vendor Price Changes in a Week' according to the organisational specifications.	<input type="checkbox"/>
3	Add 'Week of Year' as a drill-through field to this report page.	<input type="checkbox"/>
4	<p>Add 'Card' visuals for the following parameters to display values within the context of the analysis.</p> <ul style="list-style-type: none"> • 'Total Vendor Price Changes' • 'No. of distinct Vendors Contributed' • 'No. of distinct Products Impacted' • 'Week of Year – Year' - Format this variable as follows: <ul style="list-style-type: none"> ○ Category label: Off ○ Title: On ○ Title text: Week of Year – Year ○ Effects > Background colour: #C9E1F7 	<input type="checkbox"/>
5	<p>Add a 'Line and clustered column' chart as the primary visual to display the following parameters:</p> <ul style="list-style-type: none"> • Column y-axis: 'No. of distinct Vendors Contributed' and 'No. of distinct Products Impacted' • Line y-axis: 'Total Vendor Price Changes' • X-axis: 'Sub-category' <p>Format the visual by turning on 'Data labels'.</p>	<input type="checkbox"/>
6	<p>Generate a summary of the data presented in the 'Line and clustered column chart' visual. Format the visual by;</p> <ul style="list-style-type: none"> • setting the 'Title' as 'Summary of key outcomes, results and insights' 	<input type="checkbox"/>

#	Sub-task details	Tick/check <input checked="" type="checkbox"/> once completed.
	<ul style="list-style-type: none"> changing the 'Effects' > 'Background' > 'Color' to #CBFFD7. 	
7	<p>Add a 'Tree map' as a secondary visual to display 'Total Vendor Price Changes' (Values) by 'Vendor Name' (Category). Format the visual by turning on 'Data labels'.</p>	<input type="checkbox"/>
8	<p>Add a <i>Matrix</i> visual as a secondary visual to display the 'No. of distinct Products Impacted' (Values) for each 'Vendor Name' (Rows). Format the visual by;</p> <ul style="list-style-type: none"> setting the 'Title' as 'Distinct Products Impacted by Vendor Name' turning off 'Column subtotals' turning on 'Row subtotals' conditionally formatting the 'No. of distinct Products Impacted' value to display 'Data bars' sorting the 'No. of distinct Products Impacted' column in descending order (high to low % value). 	<input type="checkbox"/>
9	<p>Analyse trends and relationships using this report page for the following context using the suggested method.</p> <p>Context: Analyse the 'Week of Year' that reported the highest 'Total Vendor Price Changes' over 2018-2021.</p> <p>Method: From the 'Dashboard: Vendor Price Changes Overview' page, right-click on the specific week of interest from a visual (according to the given context) and select 'Drill through' > 'Price Changes' in a Week to go to the analysis report page for further details.</p>	<input type="checkbox"/>
10	<p>Provide a screenshot as evidence of the 'All Stores in Month' report, in 'Table 13: D2b' under 'Screenshot evidence:'. Read the 'Note' sections carefully to understand the details that must be visible in each screenshot.</p>	<input type="checkbox"/>

Screenshot evidence:

Assessor Instructions: Each report screenshot must include a summary of the analysis in the form of a short paragraph that reports on the results and insights from each analysis.

Sample screenshots are provided below.

Table 13 - Evidence of analysing the non-transactional dataset, Products

Demonstration tasks: Evidence (Screenshot):

D2a. Report page 1 – Price Change within a Sub-category

Note: The screenshot of the 'Report' must clearly show the following details:

- All required elements of the report page (Company Logo, Report title, chart titles, and correct visualisation types)
- 'Card' visuals indicating the context selected from the dashboard page.
- 'Summary of key outcomes, results and insights' included.
- 'Visualisations' pane showing filters based on the context.
- Title bar showing the file name with your name initials and current date.



Figure 14 - Screenshot for task D2a using Power BI Desktop © Microsoft

D2b. Report page 2 – Price Changes in a Week

Note: The screenshot of the 'Report' must clearly show the following details:

- All required elements of the report page (Company Logo, Report title, chart titles, and correct visualisation types)
- 'Card' visuals indicating the context selected from the dashboard page.
- 'Summary of key outcomes, results and insights' included.
- 'Visualisations' pane showing filters based on the context.
- Title bar showing the file name with your name initials and current date.

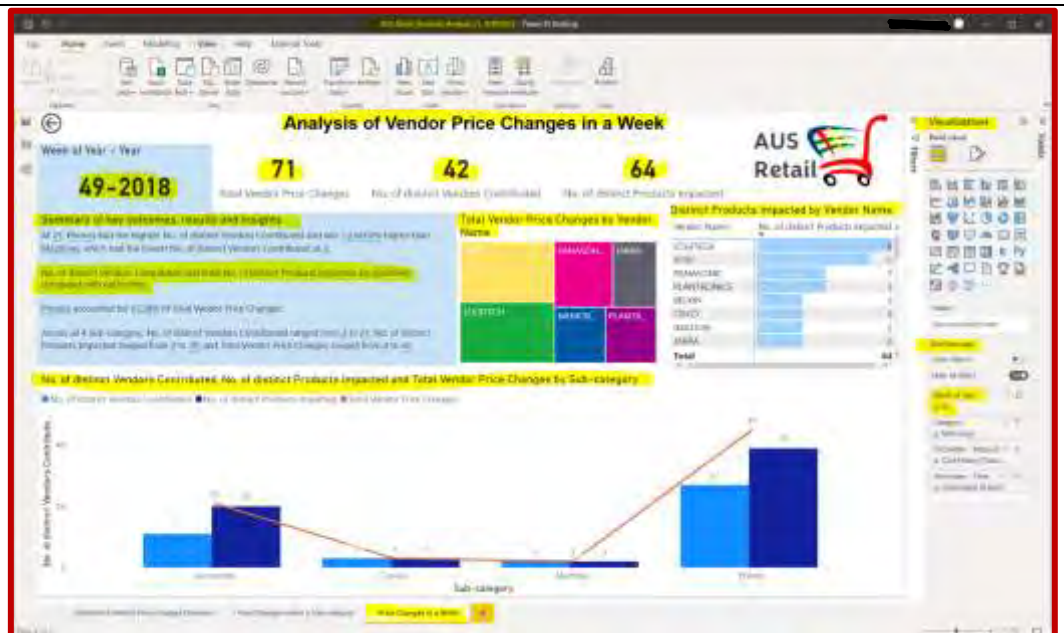


Figure 15 - Screenshot for task D2b using Power BI Desktop © Microsoft

Note: Once the analysis reports are created save and close your Power BI work file as you will continue to work on this file again in Part E2.

Part E: Conduct statistical analysis to confirm accuracy of big data 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 the required tools and dashboards in Microsoft Power BI technology platform to conduct the statistical analysis tasks
- use clear, specific and industry-related terminology to represent outcomes of the analysis in reports
- use the Power BI files that you previously worked on in Part D
- complete at times complex calculations to create measures required for the statistical analysis.

Scenario continued:

You have received the following email from your supervisor Mia Gonzales, briefing you about the next task.

From: Gonzales, Mia

To: Student Lastname, Student Firstname

Subject: Conduct statistical analysis to confirm the accuracy of the analysis

Hi <student name>,

Thank you for preparing the dashboard and the preliminary analysis you've done in the analysis report pages. Please also conduct a statistical analysis to ensure that your analysis data is correct. You must specifically:

- check if the data includes any extreme values that may be providing an inaccurate/biased result in the analysis
- check the strength of the relationships between variables in the analysis and prove that the correlation type is statistically correct.

As you are new to the job, please follow the checklist I've prepared that will guide you through the process. Please tick/check each task in the checklists as you complete them.

- Refer to **Table 14 & Table 15 Checklists** to confirm the accuracy of the 'Sales' dataset analysis.
- Refer to **Table 18 & Table 19 Checklists** to confirm the accuracy of the 'Products' dataset analysis.

Refer to the sample DAX formulas provided in the following sections of the *AUS Retail_Data analysis policy.pdf* document.

- 7. Conducting statistical analysis (univariate)
 - 7.1 Calculate Skewness
 - 7.2 Determine the shape of the data
- 8. Conducting statistical analysis (bivariate)
 - 8.1 Calculate the Correlation Coefficient
 - 8.2 Determine the strength of the relationship between variables
 - 8.3 Determine the type of relationship between variables

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. Confirm the accuracy of Sales data analysis

In this task, you will be conducting statistical analysis to confirm the accuracy of the Sales data analysis you have previously completed in Part D1.

Task:

Conduct the statistical analysis for the **Sales** dataset by doing the following tasks using the same file you have worked on in **Task D1** (AUS Retail_Sales Analysis_Namelnitials_ddmmyyyy.pbix).

- a. Create the necessary statistical measures and use histograms to analyse the distribution of the data in the 'Sales' and 'Profit' fields.

To do this, you must complete the following sub-tasks in the checklist.

Table 14 - Checklist to perform Task E1a

#	Sub-task details	Tick/check <input checked="" type="checkbox"/> once completed.
	Do the following in the 'Data' view of Power BI Desktop.	
1.	<p>Create the required bins to group the 'Sales' data and 'Profit' data in the 'Order Detail' table according to the following specification.</p> <ul style="list-style-type: none"> Configure the bin size of the 'Sales (bins)' to '2'. Configure the bin size of the 'Profit (bins)' to '80'. <p>Notes for assessors: The procedure to do this is as follows:</p> <ul style="list-style-type: none"> Right-click on the 'Sales' field. Select 'New group' to configure the bin size to '2'. Right-click on the 'Profit' field. Select 'New group' to configure the bin size to '80'. 	<input type="checkbox"/>
2.	<p>Create a new empty table called 'Statistics'.</p> <p>Notes for assessors: The procedure to do this is as follows: 'Data view' > 'Table tools' > 'New Table' > Type 'Statistics =' > Press 'Enter' key.</p>	<input type="checkbox"/>
	Within the 'Statistics' table, create the following calculated measures.	
3.	<p>Create new measures to calculate the frequency (count) of 'Sales' and 'Profit' fields called:</p> <ul style="list-style-type: none"> Frequency of Sales Frequency of Profit 	<input type="checkbox"/>
4.	<p>Create new measures to calculate the arithmetic mean of 'Sales' and 'Profit' fields called:</p> <ul style="list-style-type: none"> Sales Average Profit Average 	<input type="checkbox"/>
5.	<p>Create new measures to calculate the median of 'Sales' and 'Profit' fields called:</p> <ul style="list-style-type: none"> Sales Median Profit Median 	<input type="checkbox"/>
6.	<p>Create new measures to calculate the standard deviation of 'Sales' and 'Profit' fields called:</p> <ul style="list-style-type: none"> Sales Standard Deviation Profit Standard Deviation 	<input type="checkbox"/>
7.	<p>Create new measures to calculate the minimum value of 'Sales' and 'Profit' fields called:</p> <ul style="list-style-type: none"> Sales Min Profit Min 	<input type="checkbox"/>
8.	<p>Create new measures to calculate the maximum value of 'Sales' and 'Profit' fields called:</p> <ul style="list-style-type: none"> Sales Max Profit Max 	<input type="checkbox"/>
9.	<p>Create new measures to calculate the range of 'Sales' and 'Profit' fields called:</p> <ul style="list-style-type: none"> Sales Range 	<input type="checkbox"/>

#	Sub-task details	Tick/check <input checked="" type="checkbox"/> once completed.
	<ul style="list-style-type: none"> Profit Range 	
10.	<p>Create new measures to calculate the first quartile of 'Sales' and 'Profit' fields called:</p> <ul style="list-style-type: none"> Sales Q1 Profit Q1 	<input type="checkbox"/>
11.	<p>Create new measures to calculate the third quartile of 'Sales' and 'Profit' fields called:</p> <ul style="list-style-type: none"> Sales Q3 Profit Q3 	<input type="checkbox"/>
12.	<p>Create new measures to calculate the interquartile range (IQR) of 'Sales' and 'Profit' fields called:</p> <ul style="list-style-type: none"> Sales IQR Profit IQR 	<input type="checkbox"/>
13.	<p>Create new measures to calculate the 'Pearson's Median Skewness' for the 'Sales' and 'Profit' fields called:</p> <ul style="list-style-type: none"> Sales Median Skewness Profit Median Skewness <p>Refer to the equation provided in sub-section '7.1 Calculate skewness' in the <i>AUS Retail_Data analysis policy.pdf</i> document.</p>	<input type="checkbox"/>
	Do the following in the 'Report' view of Power BI Desktop.	
14.	Create a new report page and rename it 'Statistical Analysis: Sales, Profit Distribution'.	<input type="checkbox"/>
15.	<p>Add 'Multi-row Card' visual to display the values from the following calculated measures for 'Sales' data.</p> <ul style="list-style-type: none"> Sales Average Sales Median Sales Standard Deviation Sales Min Sales Max Sales Range Sales Q1 Sales Q3 Sales IQR Sales Median Skewness <p>Configure the following format options for this visual:</p> <ul style="list-style-type: none"> Title: On Title Text: Sales Statistics 	<input type="checkbox"/>
16.	<p>Add another 'Multi-row Card' visual to display the values from the following calculated measures for 'Profit' data.</p> <ul style="list-style-type: none"> Profit Average Profit Median Profit Standard Deviation Profit Min Profit Max Profit Range Profit Q1 Profit Q3 Profit IQR Profit Median Skewness <p>Configure the following format options for this visual:</p> <ul style="list-style-type: none"> Title: On Title Text: Profit Statistics 	<input type="checkbox"/>

#	Sub-task details	Tick/check <input checked="" type="checkbox"/> once completed.
17.	<p>Create calculated measures to determine the shape of the 'Sales' and 'Profit' data using the recommended DAX formula.</p> <p>Name these measures as:</p> <ul style="list-style-type: none"> • 'Shape of the Sales data' • 'Shape of the Profit data' <p>Refer to the sub-section '7.2 Determine the shape of the data' in the <i>AUS Retail_Data analysis policy.pdf</i> document to find the recommended DAX formula.</p>	
18.	<p>Create new calculated columns in the 'Order Detail' table using the 'NORM.DIST' function for both 'Sales' and 'Profit' fields called:</p> <ul style="list-style-type: none"> • 'Sales Distribution Curve' • 'Profit Distribution Curve' <p>Note: When using the NORM.DIST calculation formula, you will need to type in the exact values obtained for the associated arithmetic mean and standard deviation values for 'Sales' and 'Profit' data as displayed in the 'Multi-row Card' visuals in the report.</p>	<input type="checkbox"/>
19.	<p>Add a 'Line and clustered column chart' visual to display the distribution of the 'Sales' data with the following criteria.</p> <ul style="list-style-type: none"> • Column Y-axis: 'Frequency of Sales' • Line Y-axis: Average of 'Sales Distribution Curve' • X-axis: 'Sales (bins)' <p>Ensure that the 'Zoom Slider' is turned on to help zoom-in to further investigate the shape of the distribution.</p>	<input type="checkbox"/>
20.	<p>Generate a summary of the data presented in the previously created 'Line and clustered column chart' visual for 'Sales' data with the following settings:</p> <ul style="list-style-type: none"> • Title: On • Title Text: Statistical Analysis of the Sales data • Effects > Background Color: #CBFFD7 	<input type="checkbox"/>
21.	<p>Add another 'Line and clustered column chart' visual to display the distribution of the 'Profit' data with the following criteria.</p> <ul style="list-style-type: none"> • Column Y-axis: 'Frequency of Profit' • Line Y-axis: Average of 'Profit Distribution Curve' • X-axis: 'Profit (bins)' <p>Ensure that the 'Zoom Slider' is turned on to help zoom-in to further investigate the shape of the distribution.</p>	<input type="checkbox"/>
22.	<p>Generate a summary of the data presented in the previously created 'Line and clustered column chart' visual for 'Profit' data with the following settings:</p> <ul style="list-style-type: none"> • Title: On • Title Text: Statistical Analysis of the Profit data • Effects > Background Color: #CBFFD7 	<input type="checkbox"/>
23.	<p>Customise the summary text created for 'Sales' and 'Profit' data to only include information related to the statistical analysis. Such as the following.</p> <ul style="list-style-type: none"> • Shape of the data: Indicate whether it is symmetrical, positively skewed or negatively skewed. Add here a dynamic value to display the relevant information from the previously created measures 'Shape of the Sales data' or 'Shape of the Profit data' as appropriate. • Central tendency: Indicate where the data is centered in terms of mean, median and mode? Are there any observations that are radically different from the rest which may have an effect on the arithmetic mean? 	<input type="checkbox"/>

#	Sub-task details	Tick/check <input checked="" type="checkbox"/> once completed.
	<ul style="list-style-type: none"> Spread of the data: How broadly is the data dispersed? Range (Max value – Min value) and IQR. Are there huge differences between the range and the IQR indicating that extreme values may be present in the data? <p>Note: Delete the automatically generated summary text and only include a summary of your statistical analysis.</p>	
24.	Provide a screenshot as evidence of the 'Statistical Analysis: Sales, Profit Distribution' report, in 'Table 17, E1a' under 'Screenshot evidence:'. Read the 'Note' sections carefully to understand the details that must be visible in each screenshot.	<input type="checkbox"/>

b. Create the necessary statistical measures and visuals to analyse the relationship between 'Sales' and 'Cost' variables in the analysis.

To do this, you must complete the following sub-tasks in the checklist.

Table 15 - Checklist to perform Task E1b

#	Sub-task details	Tick/check <input checked="" type="checkbox"/> once completed.
	<p>Within the 'Statistics' table, create the following calculated measures for the X and Y variables.</p> <p>Note: The variables are defined as follows.</p> <ul style="list-style-type: none"> Variable X = 'Total Sales' Variable Y = 'Total Cost' <p>Refer to the section '8. Conducting statistical analysis (bivariate)' in the <i>AUS Retail_Data analysis policy.pdf</i> document for the recommended DAX formulas.</p>	
1.	Create a new measure to calculate the correlation coefficient between the X and Y variables, by 'Month-Year'. Name this measure as 'Correlation Coefficient for Total Sales & Total Cost'.	<input type="checkbox"/>
2.	Create a calculated measure to determine the strength of the relationship between the X and Y variables, using the recommended DAX formula. Name this measure as 'Correlation strength of Sales & Cost'.	<input type="checkbox"/>
3.	Create a calculated measure to determine the type of relationship between the X and Y variables, using the recommended DAX formula. Name this measure as 'Correlation type of Sales & Cost'.	<input type="checkbox"/>
	Do the following in the 'Report' view of Power BI Desktop.	
4.	Create a new report page and rename it 'Statistical Analysis: Sales, Cost Relationship'.	<input type="checkbox"/>
5.	Add a 'Scatter chart' visual as to display the associated X and Y variables by 'Month-Year'. Ensure that the 'Trend line' is turned on.	
6.	Add a 'Card' visual to display the value for the calculated measure 'Correlation Coefficient for Total Sales & Total Cost'.	<input type="checkbox"/>
7.	Add a 'Slicer' visual to enable customisation of the report visuals based on 'Store Location'. Format the 'Slicer Settings' > 'Orientation' to 'Horizontal'.	<input type="checkbox"/>

#	Sub-task details	Tick/check <input checked="" type="checkbox"/> once completed.
8.	<p>Add a 'Smart Narrative' visual and customise it to include information related to the statistical analysis of the relationship between Sales and Cost such as the following.</p> <ul style="list-style-type: none"> • Correlation strength: Add here a dynamic value to display the relevant information from the measure 'Correlation strength of Sales & Cost'. • Correlation type: Add here a dynamic value to display the relevant information from the measure 'Correlation type of Sales & Cost'. <p>Format the 'Smart Narrative' visual using the following settings:</p> <ul style="list-style-type: none"> • Title: On • Title Text: Relationship Analysis: Sales & Cost • Effects > Background Color: #CBFFD7 	<input type="checkbox"/>
9.	<p>Conduct statistical analysis to confirm the accuracy of the relationship between 'Sales' and 'Cost' variables by using this report. To do this, you must customise this report according to the following context, using the suggested method.</p> <p>Context: The previously created 'Store Performance' drill-through report (in Task D1a) analysed the monthly performance of the store location that reported the lowest 'Gross Profit Margin %' over 2018-2021. The drill-through report for this specific store location identified the type of correlation between 'Total Sales' and 'Total Cost'. This correlation type needs to be confirmed by conducting statistical analysis.</p> <p>Method: In the 'Relationship Analysis: Sales, Profit' report page, select the store location of interest from the 'Slicer' visual according to the context. Ensure other visuals in the report ('Scatter Chart', 'Card', 'Smart Narrative') display the relevant statistical information that confirms the correlation strength and type.</p>	<input type="checkbox"/>
10.	<p>Provide a screenshot as evidence of the customised 'Relationship Analysis: Sales, Profit' report, in 'Table 17 > E1b' under 'Screenshot evidence:'. Read the 'Note' sections carefully to understand the details that must be visible in each screenshot.</p>	<input type="checkbox"/>

- c. Create the necessary statistical measures and visuals to analyse the relationship between 'Sales' and 'Profit' variables in the analysis.

To do this, you must complete the following sub-tasks in the checklist.

Table 16 - Checklist to perform Task E1c

#	Sub-task details	Tick/check <input checked="" type="checkbox"/> once completed.
	<p>Within the 'Statistics' table, create the following calculated measures for the X and Y variables.</p> <p>Note: The variables are defined as follows.</p> <ul style="list-style-type: none"> • Variable X = 'Total Sales' • Variable Y = 'Gross Profit' <p>Refer to the section '8. Conducting statistical analysis (bivariate)' in the <i>AUS Retail Data analysis policy.pdf document</i> for the recommended DAX formulas.</p>	
1.	<p>Create a new measure to calculate the correlation coefficient the X and Y variables, by 'Store Location'.</p> <p>Name this measure as 'Correlation Coefficient for Total Sales & Gross Profit'.</p>	<input type="checkbox"/>
2.	<p>Create a calculated measure to determine the strength of the relationship between the X and Y variables, using the recommended DAX formula.</p> <p>Name this measure as 'Correlation strength of Sales & Profit'.</p>	<input type="checkbox"/>

#	Sub-task details	Tick/check <input checked="" type="checkbox"/> once completed.
3.	Create a calculated measure to determine the type of relationship between the X and Y variables using the recommended DAX formula. Name this measure as 'Correlation type of Sales & Profit'.	<input type="checkbox"/>
	Do the following in the 'Report' view of Power BI Desktop.	
4.	Create a new report page and rename it 'Statistical Analysis: Sales, Profit Relationship'.	<input type="checkbox"/>
5.	Add a 'Scatter chart' visual as to display the associated X and Y variables by 'Store Location'. Ensure that the 'Trend line' is turned on.	<input type="checkbox"/>
6.	Add a 'Card' visual to display the value for the calculated measure 'Correlation Coefficient for Total Sales & Gross Profit'.	<input type="checkbox"/>
7.	Add a 'Slicer' visual to provide the option to select the 'Year' range from 2018-2021. Format the 'Slicer Settings' > 'Selection' > 'Single Select'.	<input type="checkbox"/>
8.	Add another 'Slicer' visual to provide the option to select a specific 'Month'. Format the 'Slicer Settings' > 'Orientation' to 'Horizontal'.	<input type="checkbox"/>
9.	Add a 'Smart Narrative' visual and customise it to include information related to the statistical analysis of the relationship between Sales and Cost such as the following. <ul style="list-style-type: none"> • Correlation strength: Add here a dynamic value to display the relevant information from the measure 'Correlation strength of Sales & Cost'. • Correlation type: Add here a dynamic value to display the relevant information from the measure 'Correlation type of Sales & Cost'. Format the 'Smart Narrative' visual using the following settings: <ul style="list-style-type: none"> • Title: On • Title Text: Relationship Analysis: Sales & Cost • Effects > Background Color: #CBFFD7 	<input type="checkbox"/>
10.	Conduct statistical analysis to confirm the accuracy of the relationship between 'Sales' and 'Profit' variables using this report. To do this, you must customise this report according to the following context, using the suggested method. Context: The previously created 'All Stores in Month' drill-through report (in Task D1b) analysed the sales, costs and profit figures for all store locations for the specific month-year that reported the lowest 'Gross Profit Margin %' over 2018-2021. The drill-through report for this specific month-year identified the type of correlation between 'Total Sales' and 'Gross Profit'. This correlation type needs to be confirmed by conducting statistical analysis. Method: In the 'Relationship Analysis: Sales, Profit' report page, select the 'Year' and 'Month' of interest from the 'Slicer' visuals according to the context. Ensure other visuals in the report ('Scatter Chart', 'Card', 'Smart Narrative') display the relevant statistical information that confirms the correlation strength and type.	<input type="checkbox"/>
11.	Provide a screenshot as evidence of the customised 'Relationship Analysis: Sales, Profit' report, in 'Table 17 > E1c' under 'Screenshot evidence:'. Read the 'Note' sections carefully to understand the details that must be visible in each screenshot.	<input type="checkbox"/>

Screenshot evidence:

Assessor instructions: Each report screenshot provided by the student must include a summary of the analysis in the form of a short paragraph that reports on the results and insights from the statistical analysis.

Sample screenshots are provided below.

Table 17 - Evidence of statistically analysing the transactional dataset, Sales

Demonstration tasks:

Evidence (Screenshot):

E1a. Statistical Analysis: Sales, Profit Distribution

Note: The screenshot of the report page must clearly show the following details:

- 'Multi-row card' visuals indicating the context selected from the dashboard page.
- Summary of the statistical analysis of 'Sales' and 'Profit' data included.
- 'Fields' pane showing all the new statistical measures created for the analysis.
- Title bar showing the file name with your name initials and current date.

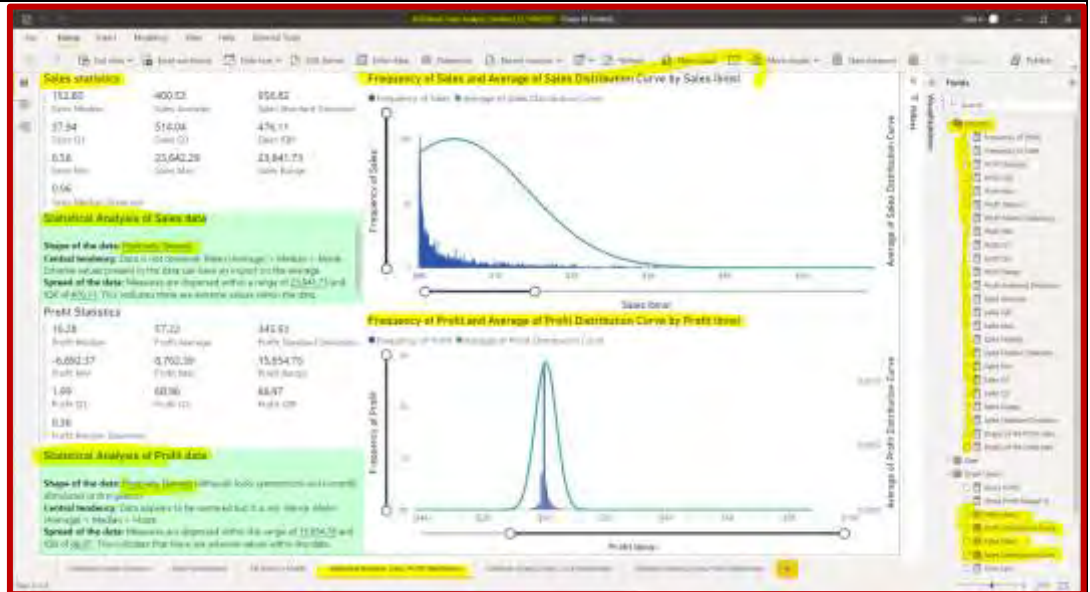


Figure 16 - Screenshot for task E1a using Power BI Desktop © Microsoft

E1b. Statistical Analysis: Sales, Cost Relationship

Note: The screenshot of the report page must clearly show the following details:

- 'Slider' visual with the specific 'Store Location' selected.
- 'Card' visual indicating the coefficient of correlation value for the selected 'Store Location' according to the context.
- Summary of the 'Relationship Analysis: Sales, Cost' indicating correlation strength and type.
- 'Fields' pane showing the relevant statistical measures created for the relationship analysis.
- Title bar showing the file name with your name initials and current date.

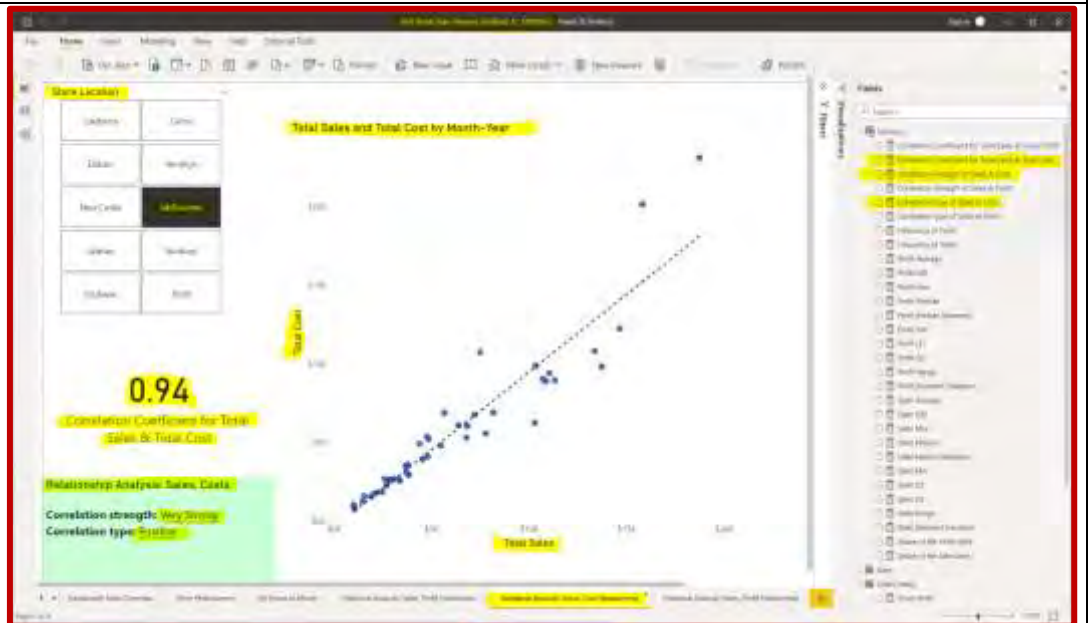


Figure 17 - Screenshot for task E1b using Power BI Desktop © Microsoft

Demonstration tasks:

Evidence (Screenshot):

E1c. Statistical Analysis: Sales, Profit Relationship

Note: The screenshot of the report page must clearly show the following details:

- 'Slider' visuals with the specific 'Year' and 'Month' selected.
- 'Card' visual indicating the coefficient of correlation value for the selected 'Year' and 'Month' according to the context.
- Summary of the 'Relationship analysis: Sales, Profit' indicating correlation strength and type.
- 'Fields' pane showing the relevant statistical measures created for the relationship analysis.
- Title bar showing the file name with your name initials and current date.

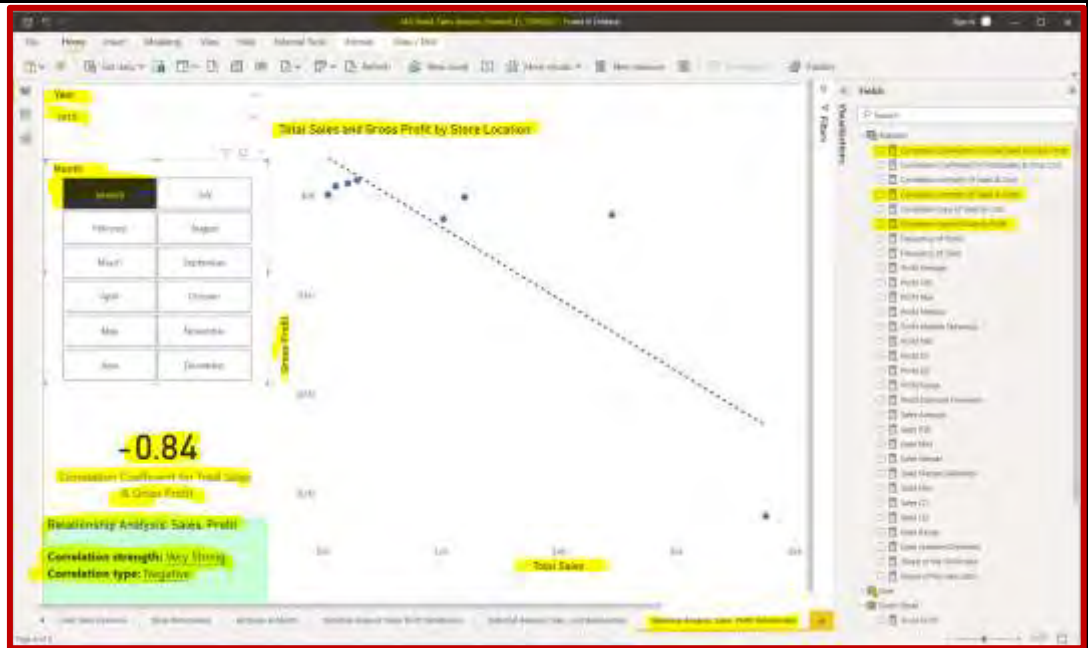


Figure 18 - Screenshot for task E1c using Power BI Desktop © Microsoft

Note: Once the statistical analysis is completed save and close your Power BI work file. You will be asked to create a finalised version of this file in Task F1.

E2. Confirm the accuracy of the Product data analysis

In this task, you will be conducting statistical analysis to confirm the accuracy of the **Product** data analysis you have previously completed in Part D2.

Task:

Conduct the statistical analysis for the **Products** dataset by doing the following tasks using the same file you have worked on in **Task D2** (AUS Retail_Products Analysis_NameInitials_ddmmyyyy.pbix).

- a. Create the necessary statistical measures and use histograms to analyse the distribution of the 'Price Difference' field. To do this, you must complete the following sub-tasks in the checklist.

Note: When creating some of the statistical measures in the checklist, the following shortened forms or abbreviations are used instead of the full names for ease of reference.

- Price Diff. = 'Price Difference'
- PC = 'Price Change'

Table 18 - Checklist to perform Task E2a

#	Sub-task details	Tick/check <input checked="" type="checkbox"/> once completed.
	Do the following in the 'Data' view of Power BI Desktop.	
1.	Create the required bins to group the 'Price Difference' data in the 'Cost History' table according to the following specification. <ul style="list-style-type: none"> • Configure the bin size of the 'Price Differences (bins)' to '5'. 	<input checked="" type="checkbox"/>

#	Sub-task details	Tick/check <input checked="" type="checkbox"/> once completed.
	<p>Notes for assessors: The procedure to do this is as follows:</p> <ul style="list-style-type: none"> Right-click on the 'Price Differences' field. Select 'New group' to configure the bin size to '5'. 	
2.	<p>Create a new empty table called 'Statistics'.</p> <p>Notes for assessors: The procedure to do this is as follows: 'Data view' > 'Table tools' > 'New Table' > Type 'Statistics =' > Press 'Enter' key.</p>	<input type="checkbox"/>
	Within the 'Statistics' table, create the following calculated measures.	
3.	<p>Create a new measure to calculate the frequency (count) of the 'Price Difference' field called:</p> <ul style="list-style-type: none"> 'Price Difference Frequency'. <p>Note: The shortened form 'Price Diff.' is not used here as the histogram created using this field should display the full name in the axis title.</p>	<input type="checkbox"/>
4.	<p>Create a new measure to calculate the arithmetic mean of the 'Price Difference' field called:</p> <ul style="list-style-type: none"> 'Price Diff. Average'. 	<input type="checkbox"/>
5.	<p>Create a new measure to calculate the median of the 'Price Difference' field called:</p> <ul style="list-style-type: none"> 'Price Diff. Median'. 	<input type="checkbox"/>
6.	<p>Create a new measure to calculate the standard deviation of the 'Price Difference' field called:</p> <ul style="list-style-type: none"> 'Price Diff. Standard Deviation'. 	<input type="checkbox"/>
7.	<p>Create a new measure to calculate the minimum value of the 'Price Difference' field called:</p> <ul style="list-style-type: none"> 'Price Diff. Min'. 	<input type="checkbox"/>
8.	<p>Create a new measure to calculate the maximum value of the 'Price Difference' field called:</p> <ul style="list-style-type: none"> 'Price Diff. Max'. 	<input type="checkbox"/>
9.	<p>Create a new measure to calculate the range of the 'Price Difference' field called:</p> <ul style="list-style-type: none"> 'Price Diff. Range'. 	<input type="checkbox"/>
10.	<p>Create a new measure to calculate the first quartile of the 'Price Difference' field called:</p> <ul style="list-style-type: none"> 'Price Diff. Q1'. 	<input type="checkbox"/>
11.	<p>Create a new measure to calculate the third quartile of the 'Price Difference' field called:</p> <ul style="list-style-type: none"> 'Price Diff. Q3'. 	<input type="checkbox"/>
12.	<p>Create a new measure to calculate the interquartile range for the 'Price Difference' field called:</p> <ul style="list-style-type: none"> 'Price Diff. IQR'. 	<input type="checkbox"/>
13.	<p>Create a new measure to calculate 'Pearson's Median Skewness' for the 'Price Difference' field called:</p> <ul style="list-style-type: none"> 'Price Diff Median Skewness'. <p>Refer to the equation provided in sub-section '7.1 Calculate skewness' in the <i>AUS Retail Data analysis policy.pdf</i> document.</p>	<input type="checkbox"/>
	Do the following in the 'Report' view of Power BI Desktop.	
14.	Create a new report page and rename it 'Statistical Analysis: PC Distribution'.	<input type="checkbox"/>
15.	Add 'Multi-row Card' visual to display the values from the following calculated measures for 'Sales' data.	<input type="checkbox"/>

#	Sub-task details	Tick/check <input checked="" type="checkbox"/> once completed.
	<ul style="list-style-type: none"> • Price Diff. Average • Price Diff. Median • Price Diff. Standard Deviation • Price Diff. Min • Price Diff. Max • Price Diff. Range • Price Diff. Q1 • Price Diff. Q3 • Price Diff. IQR • Price Diff. Median Skewness <p>Configure the following format options for this visual:</p> <ul style="list-style-type: none"> • Title: On • Title Text: Price Difference Statistics 	
16.	<p>Create calculated measures to determine the shape of the 'Price Difference' data using the recommended DAX formula.</p> <p>Name this measure as 'Shape of the Price Diff. data'</p> <p>Refer to the sub-section '7.2 Determine the shape of the data' in the <i>AUS Retail Data analysis policy.pdf</i> document to find the recommended DAX formula.</p>	
17.	<p>Create new calculated column in the 'Cost History' table using the NORM.DIST function called 'Price Diff. Distribution Curve' using the following criteria in the formula:</p> <ul style="list-style-type: none"> • Price Difference (bins) – use the measure name in the formula • Price Diff. Average - use the numerical value in the formula • Price Diff. Standard Deviation - use the numerical value in the formula <p>Note: When using the NORM.DIST calculation formula, you will need to type in the exact values obtained for the associated arithmetic mean and standard deviation values for the 'Price Difference' field as displayed in the 'Multi-row Card' visual in the report.</p>	<input type="checkbox"/>
18.	<p>Add a 'Line and clustered column chart' visual to display the distribution of the 'Price Difference' data with the following criteria.</p> <ul style="list-style-type: none"> • Column Y-axis: 'Price Difference Frequency' • Line Y-axis: Average of 'Price Diff. Distribution Curve' • X-axis: 'Price Difference [bins]' <p>Ensure that the 'Zoom Slider' is turned on to help zoom-in to further investigate the shape of the distribution.</p>	<input type="checkbox"/>
19.	<p>Generate a summary of the data presented in the previously created 'Line and clustered column chart' visual for 'Sales' data with the following settings:</p> <ul style="list-style-type: none"> • Title: On • Title Text: Statistical Analysis of Price Difference data • Effects > Background Color: #C9E1F7 	<input type="checkbox"/>
20.	<p>Customise the summary text created for 'Price Difference' data to only include information related to the statistical analysis. Such as the following.</p> <ul style="list-style-type: none"> • Shape of the data: Indicate whether it is symmetrical, positively skewed or negatively skewed. Add here a dynamic value to display the relevant information from the previously created measures 'Shape of the Price Diff. data'. • Central tendency: Indicate where the data is centered in terms of mean, median and mode? Are there any observations that are radically different from the rest which may have an effect on the arithmetic mean? • Spread of the data: How broadly the data is dispersed? Range (Max value – Min value) and IQR. Are there huge differences between the range and the IQR indicating that extreme values may be present in the data? 	<input type="checkbox"/>

#	Sub-task details	Tick/check <input checked="" type="checkbox"/> once completed.
	Note: Delete the automatically generated summary text and only include a summary of your statistical analysis.	
21.	Provide a screenshot as evidence of the 'Statistical Analysis: Sales, Profit Distribution' report, in 'Table 21> E2a' under 'Screenshot evidence:'. Read the 'Note' sections carefully to understand the details that must be visible in each screenshot.	<input type="checkbox"/>

- b. Create the necessary statistical measures and visuals to analyse the relationship between 'Total Vendor Price changes' and 'No. of distinct Vendors Contributed' in the analysis.
To do this, you must complete the following sub-tasks in the checklist.

Note: When creating some of the statistical measures in the checklist, the following abbreviations are used instead of the full variable names for ease of reference.

- TVPC = Total Vendor Price Changes
- DVC = No. of distinct Vendors Contributed

Table 19 - Checklist to perform Task E2b

#	Sub-task details	Tick/check <input checked="" type="checkbox"/> once completed.
	<p>Within the 'Statistics' table, create the calculated measures for the X and Y variables.</p> <p>Note: The variables are defined as follows.</p> <ul style="list-style-type: none"> • Variable X = 'Total Vendor Price changes' (TVP) • Variable Y = 'No. of Distinct Vendors Contributed' (DVC) <p>Refer to the section '8. Conducting statistical analysis (bivariate)' in the <i>AUS Retail Data analysis policy.pdf</i> document for the recommended DAX formulas.</p>	
1.	Create a new measure to calculate the correlation coefficient between the X and Y variables, by 'Week of Year - Year'. Name this measure as 'Correlation Coefficient for TVPC & DVC'.	<input type="checkbox"/>
2.	Create a calculated measure to determine the strength of the relationship between the X and Y variables using the recommended DAX formula. Name this measure as 'Correlation strength of TVPC & DVC'.	<input type="checkbox"/>
3.	Create a calculated measure to determine the type of relationship between X and Y variables, using the recommended DAX formula. Name this measure as 'Correlation type of TVPC & DVC'.	<input type="checkbox"/>
	Do the following in the 'Report' view of Power BI Desktop.	
4.	Create a new report page and rename it 'Statistical Analysis: TVPC, DVC Relationship'.	<input type="checkbox"/>
5.	Add a 'Scatter chart' visual to display the associated X and Y variables by 'Week of Year - Year'. Turn on the 'Trend line' in the visual.	
6.	Add a 'Card' visual to display the value for the calculated measure 'Correlation Coefficient for TVPC & DVC'.	<input type="checkbox"/>
7.	Add a 'Slicer' visual to enable customisation of the report visuals based on 'Sub-category'. Format the 'Slicer Settings' > 'Orientation' to 'Horizontal'.	<input type="checkbox"/>

#	Sub-task details	Tick/check <input checked="" type="checkbox"/> once completed.
8.	<p>Add a 'Smart Narrative' visual and format it as follows:</p> <ul style="list-style-type: none"> • Title: On • Title Text: Relationship Analysis: TVPC & DVC • Effects > Background Color: #CBFFD7 <p>Important: Customise the contents of this visual to include the following information related to the correlation between the associated X and Y variables.</p> <ul style="list-style-type: none"> • Correlation strength: Add here a dynamic value to display the relevant information from the measure 'Correlation strength of TVPC & DVC'. • Correlation type: Add here a dynamic value to display the relevant information from the measure 'Correlation type of TVPC & DVC'. 	<input type="checkbox"/>
9.	<p>Conduct statistical analysis to confirm the accuracy of the relationship between X and Y variables using this report. To do this, you must customise this report according to the following context, using the suggested method.</p> <p>Context: The previously created 'Price Changes within a Sub-category' drill-through report (in Task D2a) analysed the product 'Sub-category' that reported the highest 'No. of distinct Vendors contribution' over 2018-2021. The drill-through report for this specific product sub-category identified the type of correlation between TVPC and DVC. This correlation type needs to be confirmed by conducting statistical analysis.</p> <p>Method: In the 'Relationship Analysis: TVPC, DVC' report page, select the sub-category of interest from the 'Slider' visual according to the context. Ensure other visuals in the report ('Scatter Chart', 'Card', 'Smart Narrative') display the relevant statistical information that confirms the correlation strength and type.</p>	<input type="checkbox"/>
10.	<p>Provide a screenshot as evidence of the customised 'Relationship Analysis: TPVC, DVC' report, in 'Table 21 > E2b' under 'Screenshot evidence:'. Read the 'Note' sections carefully to understand the details that must be visible in each screenshot.</p>	<input type="checkbox"/>

- c. Create the necessary statistical measures and visuals to analyse the relationship between 'No. of distinct Vendors Contributed' and 'No. of distinct Products Impacted' in the analysis.

To do this, complete the following sub-tasks in the checklist.

Note: When creating some of the statistical measures in the checklist, the following abbreviations should be used instead of the full variable names for ease of reference.

- DVC = No. of distinct Vendors Contributed
- DPI = No. of distinct Products Impacted
- PC = Price Change

Table 20 - Checklist to perform Task E2b

#	Sub-task details	Tick/check <input checked="" type="checkbox"/> once completed.
	<p>Within the 'Statistics' table, create the calculated measures for the X and Y variables.</p> <p>Note: The variables are defined as follows.</p> <p>Variable X = 'No. of distinct Vendors Contributed' [DVC]</p> <p>Variable Y = 'No. of distinct Products Impacted' [DPI]</p> <p>Refer to the section '8. Conducting statistical analysis (bivariate)' in the <i>AUS Retail Data analysis policy.pdf</i> document for the recommended DAX formulas.</p>	
1.	<p>Create a new measure to calculate the correlation coefficient between the X and Y variables, by 'Sub-category'.</p>	<input type="checkbox"/>

#	Sub-task details	Tick/check
		☑ once completed.
	Name this measure as 'Correlation Coefficient for DVC & DPI'.	
2.	Create a calculated measure to determine the strength of the relationship between the X and Y variables using the recommended DAX formula. Name this measure as 'Correlation strength of DVC & DPI'.	<input type="checkbox"/>
3.	Create a calculated measure to determine the type of relationship between X and Y variables, using the recommended DAX formula. Name this measure as 'Correlation type of DVC & DPI'.	<input type="checkbox"/>
	Do the following in the 'Report' view of Power BI Desktop.	
4.	Create a new report page and rename it 'Statistical Analysis: DVC, DPI Relationship'.	<input type="checkbox"/>
5.	Add a 'Scatter chart' visual to display the associated X and Y variables by 'Sub-category'. Turn on the 'Trend line' in the visual.	
6.	Add a 'Card' visual to display the value for the calculated measure 'Correlation Coefficient for TVPC & DVC'.	<input type="checkbox"/>
7.	Add a 'Slicer' visual to provide the option to select the 'Year' range from 2018-2021. Format the 'Slicer Settings' > 'Selection' > 'Single Select'.	
8.	Add another 'Slicer' visual to provide the option to select the 'Week' range from 1-53. Format the 'Slicer Settings' > 'Selection' > 'Single Select'.	
9.	Add a 'Smart Narrative' visual and format it as follows: <ul style="list-style-type: none"> • Title: On • Title Text: Relationship Analysis: DVC & DPI • Effects > Background Color: #CBFFD7 <p>Important: Customise the contents of this visual to include the following information related to the correlation between the associated X and Y variables.</p> <ul style="list-style-type: none"> • Correlation strength: Add here a dynamic value to display the relevant information from the measure 'Correlation strength of DVC & DPI'. • Correlation type: Add here a dynamic value to display the relevant information from the measure 'Correlation type of DVC & DPI'. 	<input type="checkbox"/>
10.	Conduct statistical analysis to confirm the accuracy of the relationship between X and Y variables using this report. To do this, you must customise this report according to the following context, using the suggested method. <p>Context: The previously created 'Price Changes in a Week' drill-through report (in Task D2b) analysed the 'Week of Year' that reported the highest 'Total Vendor Price Changes' over 2018-2021. The drill-through report for this specific product week of year identified the type of correlation between DVC and DPI. This correlation type needs to be confirmed by conducting statistical analysis.</p> <p>Method: In the 'Relationship Analysis: DVC, DPI' report page, select the 'Year' and 'Week' of interest from the <i>Slicer</i> visuals according to the context. Ensure other visuals in the report (Scatter Chart, Card, Smart Narrative) display the relevant statistical information that confirms the correlation strength and type.</p>	<input type="checkbox"/>

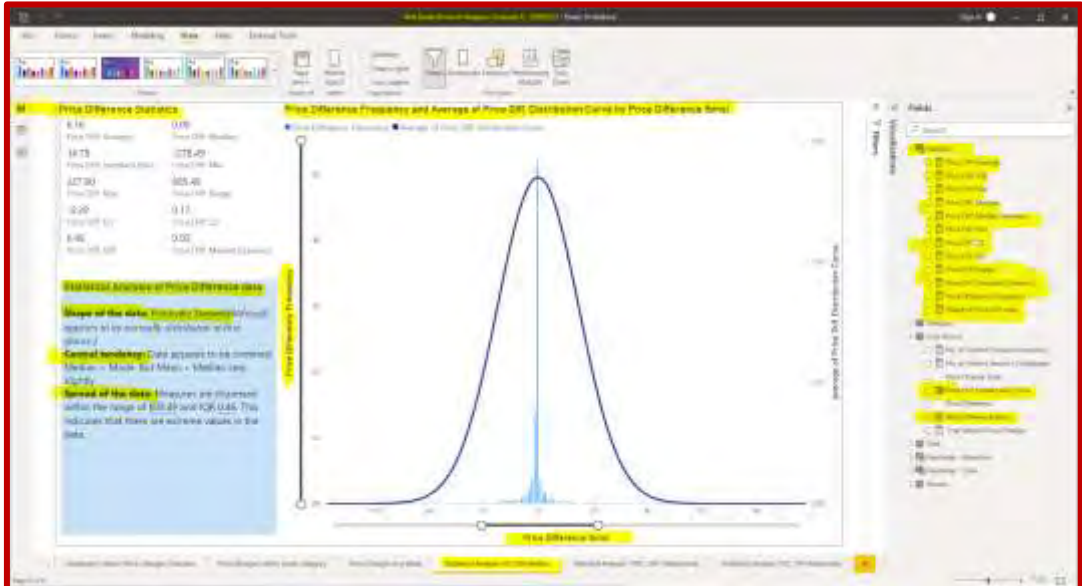
#	Sub-task details	Tick/check
		<input checked="" type="checkbox"/> once completed.
11.	Provide a screenshot as evidence of the customised 'Relationship Analysis: DVC, DPI' report, in 'Table 21 > E2c' under 'Screenshot evidence:'. Read the 'Note' sections carefully to understand the details that must be visible in each screenshot.	

Screenshot evidence:

Assessor instructions: Each report screenshot provide by the student must include a summary of the analysis in the form of a short paragraph that reports on the results and insights from the statistical analysis.

Sample screenshots are provided below.

Table 21 - Evidence of statistically analysing the non-transactional dataset, Products

Demonstration tasks:	Evidence (Screenshot):
<p>E2a. Statistical Analysis: Price Difference Distribution</p> <p>Note: The screenshot of the report page must clearly show the following details:</p> <ul style="list-style-type: none"> <input type="checkbox"/> 'Multi-row card' visuals indicating the context selected from the dashboard page. <input type="checkbox"/> 'Summary of the statistical analysis of Price Difference data' included. <input type="checkbox"/> 'Fields' pane showing all the new statistical measures created for the analysis. <input type="checkbox"/> Title bar showing the file name with your name initials and current date. 	 <p>Figure 19 - Screenshot for task E2a using Power BI Desktop © Microsoft</p>

Demonstration tasks:

Evidence (Screenshot):

E2b. Statistical Analysis: TVPC, DVC Relationship

Note: The screenshot of the report page must clearly show the following details:

- 'Card' visual indicating the coefficient of correlation value for the selected 'Store Location' according to the context.
- Summary of the 'Relationship Analysis: Sales, Cost' indicating correlation strength and type.
- 'Fields' pane showing the relevant statistical measures created for the relationship analysis.
- Title bar showing the file name with your name initials and current date.

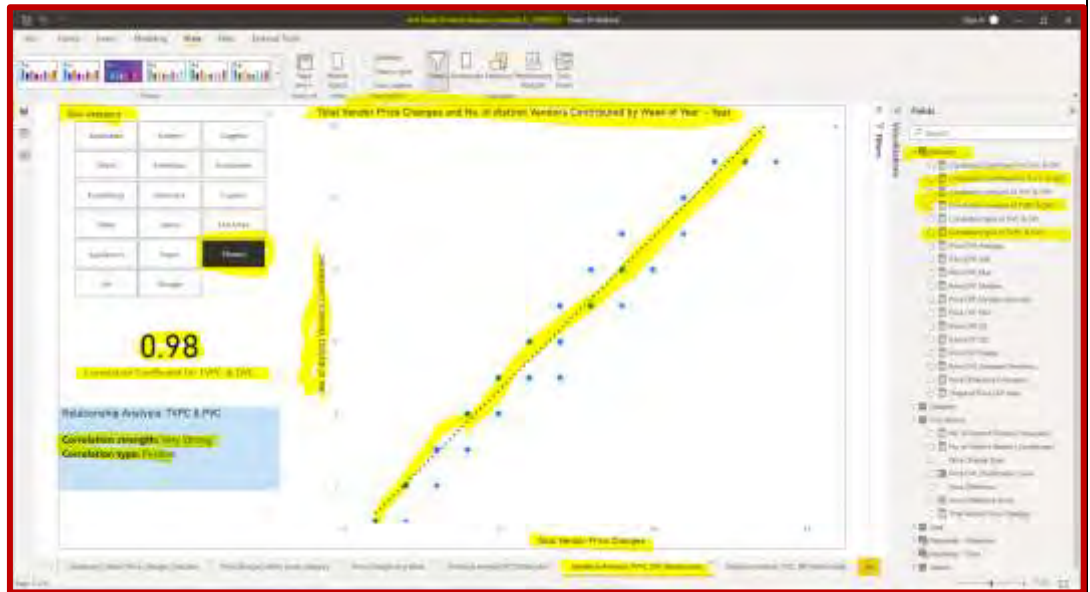


Figure 20 - Screenshot for task E2b using Power BI Desktop © Microsoft

E2c. Statistical Analysis: DVC, DPI Relationship

Note: The screenshot of the report page must clearly show the following details:

- 'Card' visual indicating the coefficient of correlation value for the selected 'Store Location' according to the context.
- Summary of the 'Relationship Analysis: Sales, Cost' indicating correlation strength and type.
- 'Fields' pane showing the relevant statistical measures created for the relationship analysis.
- Title bar showing the file name with your name initials and current date.

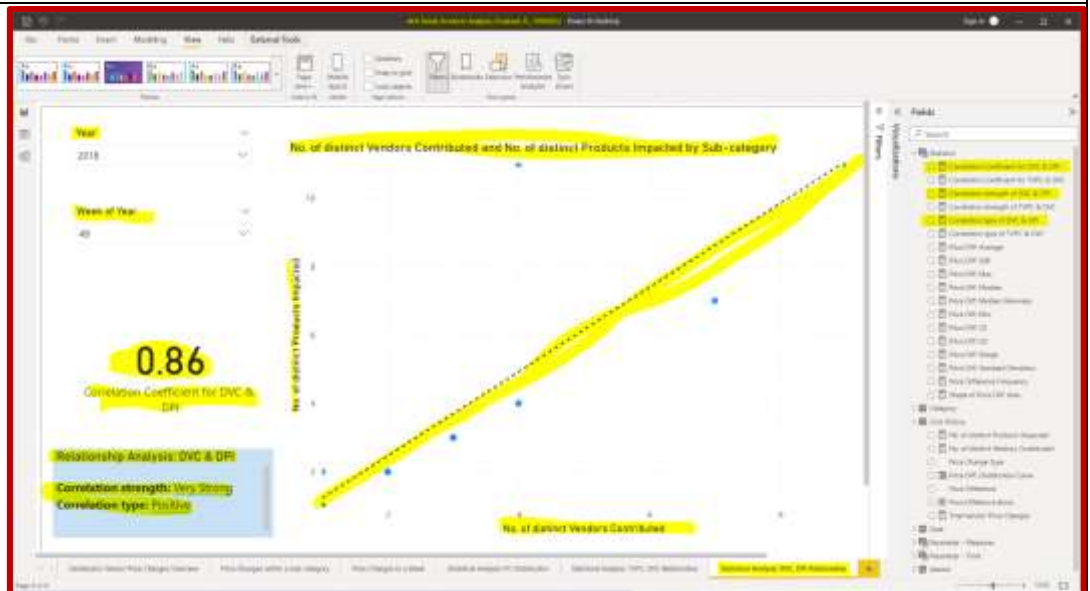


Figure 21 - Screenshot for task E2c using Power BI Desktop © Microsoft

Note: Once the statistical analysis is completed save and close your Power BI work file. You will be asked to create a finalised version of this file in Task F2.

Part F: Finalise big data 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 the required tools in the Microsoft Power BI technology platform to conduct the final analysis
- use clear, specific and industry-related terminology to represent outcomes of big data analysis in reports
- create a final version of the Power BI files that you previously worked on in Part E and rename them as follows with your name details and current date:
 - AUS Retail_Sales Analysis **Finalised**_NameInitials_ddmmyyyy (.pbix)
 - AUS Retail_Products Analysis **Finalised**_NameInitials_ddmmyyyy (.pbix)

F1. Finalise the Sales data analysis

In this task, you will be finalising the analysis of the 'Sales' dataset by isolating and removing any incorrect results.

Task:

- Isolate and remove identified incorrect results from the 'Sales' dataset. To do this, you must complete the following sub-tasks in the checklist.

Table 22 - Checklist to perform Task F1a

#	Sub-task details	Tick/check <input checked="" type="checkbox"/> once completed.
	Do the following transformations to the Sales dataset in 'Power Query Editor'.	
1.	Create a query called 'AVG_Sales' that calculates the average of the values in the 'Sales' field in the 'Order Detail' table. Notes for assessors: The procedure to do this is as follows: <ul style="list-style-type: none"> Right-click on 'Sales' column > select 'Add as Query'. Then use 'Statistical functions' to calculate the 'Average'. Rename the new query as 'AVG_Sales'. 	<input type="checkbox"/>
2.	Create a query called 'STD_Sales' that calculates the standard deviation of the values in the 'Sales' field in the 'Order Detail' table. Notes for assessors: The procedure to do this is as follows: <ul style="list-style-type: none"> Right-click on 'Sales' column > select 'Add as Query'. Then use 'Statistical functions' to calculate the 'Standard Deviation'. Rename the new query as 'STD_Sales'. 	<input type="checkbox"/>
3.	Create a custom column in the 'Order Detail' table called 'Z-score Sales' that calculates values based on the formula below. $Z - score Sales = \frac{([Sales] - AVG_Sales)}{STD_Sales}$	<input type="checkbox"/>
4.	Change the data type of the new calculated column 'Z-score Sales' to 'Decimal Number'. Ensure there are no errors.	<input type="checkbox"/>
5.	Calculate the 'Z-score Profit' column by applying steps 1 to 4 based on the 'Profit' column. Notes for assessors: The student should follow the required steps in the correct order. <ul style="list-style-type: none"> create a custom query 'AVG_Profit' to calculate the average value in step 1 create a custom query 'STD_Profit' to calculate the standard deviation in step 2 create the 'Z-score Profit' column based on the formula given in step 3. 	<input type="checkbox"/>
6.	Disable loading of the queries created to calculate the average and standard deviation for the 'Sales' and 'Profit' columns.	<input type="checkbox"/>

7. Add a conditional column to the 'Order Detail' table called 'Display Outliers' to isolate the outliers in the 'Z-score Sales' and 'Z-score Profit' fields as follows.
- If 'Z-score Sales' value is less than '-3', the output is 'YES'
 - If 'Z-score Sales' value is greater than '3', the output is 'YES'
 - If 'Z-score Profit' value is less than '-3', the output is 'YES'
 - If 'Z-score Profit' value is greater than '3', the output is 'YES'
 - Else the output is 'NO'

Once the conditional column is created, 'Close & Apply' the changes in 'Power Query Editor'.

Notes for assessors: The screenshot below shows how this task can be done in 'Power Query Editor'. The screenshot is for assessor reference only.

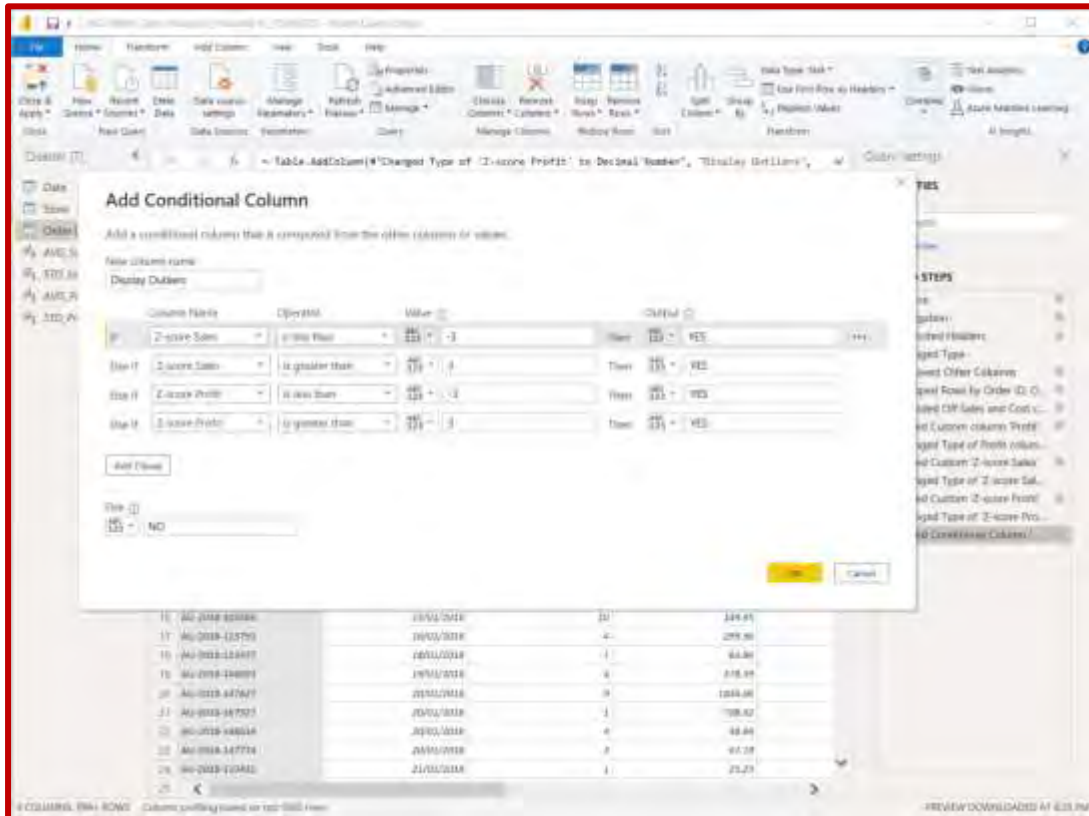


Figure 22 - Screenshot of how to set up the conditional column for Task F1a using Power BI Desktop © Microsoft

8.	Add a 'Slider' visual to the dashboard page 'Dashboard: Sales Overview' to view the options available from the 'Display Outliers' field in the 'Order Details' table. Format the 'Slider Settings' > 'Orientation' to 'Horizontal'.	<input type="checkbox"/>
9.	Filter the dashboard page to display analysis results with 'NO' outliers.	<input type="checkbox"/>
10.	Change the 'View' > 'Theme' of the final version of the Power BI work file to 'Executive'. Note: According to the theme selected the colours in the dashboard will change. This will help differentiate the final version of the dashboard easily.	<input type="checkbox"/>
11.	Provide a screenshot as evidence of the finalised 'Dashboard: Sales Overview', in 'Table 25 > F1a' under 'Screenshot evidence'. Read the 'Note' section carefully to understand the details that should be visible in the screenshot.	<input type="checkbox"/>

b. Finalise the drill-through report that analyses individual stores' performance and develop a report on the key outcomes of the analysis.

To do this, complete the sub-tasks in the checklist given below.

Table 23 - Checklist to perform Task F1b

#	Sub-task details	Tick/check ☑ once completed.
1.	<p>On the 'Store Performance' report page add the following two sentences to the 'Smart Narrative' with the title 'Summary of key outcomes, results and insights'.</p> <p>The [Variable X] and the [Variable Y] have a <Correlation Strength of X & Y>, <Correlation Type of X & Y> correlation with each other.</p> <p>The average 'Total Sales' is <Sales Average>.</p> <p>Note: The above sentences include variable names and dynamic values. Therefore, when developing the report summary you must:</p> <ul style="list-style-type: none"> • replace the text '[Variable X]' and '[Variable Y]' with the full names of the variables used in the analysis • create and add dynamic values in the sentences as indicated within the '< >' signs. • The variables specific to this report are defined as follows. <ul style="list-style-type: none"> ○ Variable X – Total Sales ○ Variable Y – Total Cost 	<input type="checkbox"/>
2.	<p>Create a revised drill-through report to analyse the monthly performance of the store location that reported the lowest 'Gross Profit Margin % over 2018-2021.</p> <p>Method: From the finalised 'Dashboard: Sales Overview' page, right-click on the specific store location of interest from a visual (according to the context of the analysis), select 'Drill through' and then select the relevant drill-through report page name.</p>	<input type="checkbox"/>
3.	<p>Provide a screenshot as evidence of the finalised 'Store Performance' report, in 'Table 25 > F1b' under 'Screenshot evidence:'. Read the 'Note' section carefully to understand the details that should be visible in the screenshot.</p>	<input type="checkbox"/>

- c. Finalise the drill-through report that analyses sales, costs, profit and margin figures of all stores in a specific month and develop a report on the key outcomes of the analysis.

To do this, complete the sub-tasks in the checklist given below.

Table 24 - Checklist to perform Task F1c

#	Sub-task details	Tick/check ☑ once completed.
1.	<p>On the 'All Stores in Month' report page add the following two sentences to the 'Smart Narrative' with the title 'Summary of key outcomes, results and insights'.</p> <p>The [Variable X] and the [Variable Y] have a <Correlation Strength of X & Y>, <Correlation Type of X & Y> correlation with each other.</p> <p>The average 'Total Sales' is <Sales Average> and the average 'Gross Profit' is <Profit Average>.</p> <p>Note: The above sentences include variable names and dynamic values. Therefore, when developing the report summary you must:</p> <ul style="list-style-type: none"> • replace the text '[Variable X]' and '[Variable Y]' with the full names of the variables used in the analysis • create and add dynamic values in the sentences as indicated within the '< >' signs. • The variables specific to this report are defined as follows. <ul style="list-style-type: none"> ○ Variable X – Total Sales ○ Variable Y – Gross Profit 	<input type="checkbox"/>
2.	<p>Create a revised drill-through report to analyse the sales, costs and profit figures for all store locations for the month-year that reported the lowest 'Gross Profit Margin %' over 2018-2021.</p>	<input type="checkbox"/>

#	Sub-task details	Tick/check <input checked="" type="checkbox"/> once completed.
	Method: From the finalised 'Dashboard: Sales Overview' page, right-click on the specific month-year of interest from a visual (according to the context of the analysis), select 'Drill through' and then select the relevant drill-through report page name.	
3.	Provide a screenshot as evidence of the finalised 'All Stores in Month' report, in 'Table 25 > F1c' under 'Screenshot evidence:'. Read the 'Note' section carefully to understand the details that should be visible in the screenshot.	<input type="checkbox"/>

Screenshot evidence:

Assessor instructions: The screenshots of the finalised dashboard page and analysis report pages for the Sales dataset provided by the student must indicate the following characteristics:

- The values displayed in the 'Card' visualisations;
 - match what is displayed in the sample screenshots
 - have a significant change when compared with the values displayed in the dashboard previously created in Task C1 and reports previously created in Task D1. This demonstrates that the student had successfully isolated and removed incorrect results from the analysis.
- The screenshots of the drill-through reports (in Table 25: F1b and F1c);
 - include a summary text in the form of a short paragraph that reports on the key outcomes, results and insights from the analysis.
 - display figures/values that are relevant to the contexts of each analysis. (The expected values/figures from the analysis are shown in the sample screenshots).

Sample screenshots are provided below.

Table 25 - Evidence of finalising the analysis of the transactional dataset, Sales.

Demonstration tasks:	Evidence (Screenshot):
<p>F1a. Dashboard: Sales Overview</p> <p>Note: The screenshot of the dashboard must clearly show the following details:</p> <ul style="list-style-type: none"> <input type="checkbox"/> 'Increase' selected in 'Price Change Type' slicer <input type="checkbox"/> 'Week of Year' selected in 'Parameter-Time' slicer. <input type="checkbox"/> 'Total Vendor Price changes' selected in 'Parameter-Measures' slicer. <input type="checkbox"/> Year range between 2018-2021 selected in 'Year' slicer. <input type="checkbox"/> Title bar showing the file name with your name initials and current date. 	<p>Figure 23 - Screenshot for task F1a using Power BI Desktop © Microsoft</p>

Demonstration tasks:

Evidence (Screenshot):

F1b. Store Performance

Note: The screenshot of the finalised drill-through report must clearly show the following details:

□ All required elements of the report page (Company Logo, Report title, chart titles, and correct visualisation types)

□ 'Card' visuals indicating the correct values after removing the incorrect results (outliers).

□ 'Summary of key outcomes, results and insights' with new information included using dynamic values.

□ Title bar showing the file name with your name initials and current date.

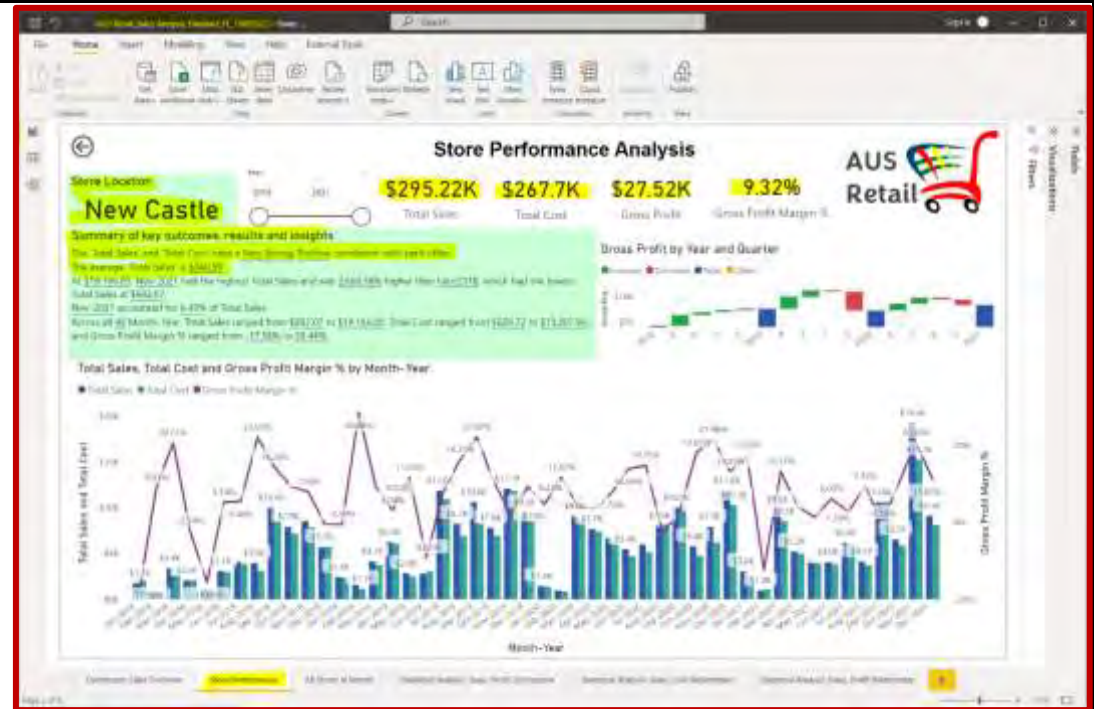


Figure 24 - Screenshot for task F1b using Power BI Desktop © Microsoft

F1c. All Stores in Month

Note: The screenshot of the finalised drill-through report must clearly show the following details:

□ All required elements of the report page (Company Logo, Report title, chart titles, and correct visualisation types)

□ 'Card' visuals indicating the correct values after removing the incorrect results.

□ 'Summary of key outcomes, results and insights' with new information included using dynamic values.

□ Title bar showing the file name with your name initials and current date.

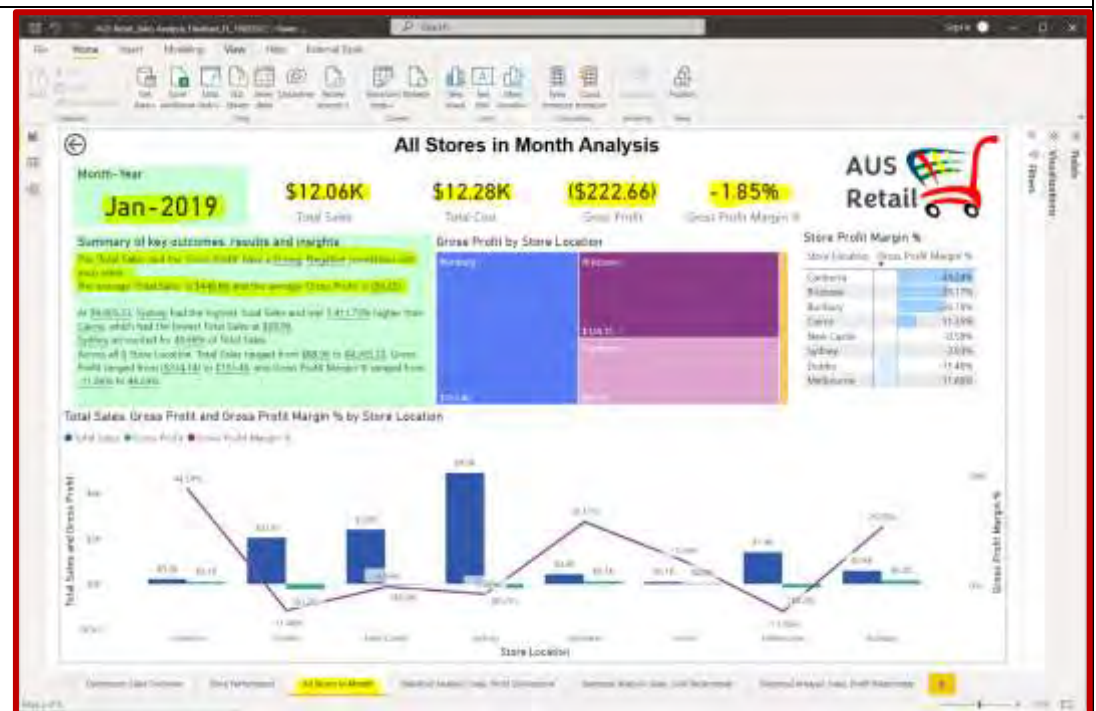


Figure 25 - Screenshot for task F1c using Power BI Desktop © Microsoft

Note: Once the analysis is completed and finalised save and close your Power BI work file. You will be asked to store a copy of this file in Task F3.

F2. Finalise the Product data analysis

In this task, you will be finalising the analysis of the 'Products' dataset by isolating and removing any incorrect results.

Task:

- a. Isolate and remove identified incorrect results from the 'Products' dataset.
To do this, you must complete the following sub-tasks in the checklist.

Table 26 - Checklist to perform Task F2a

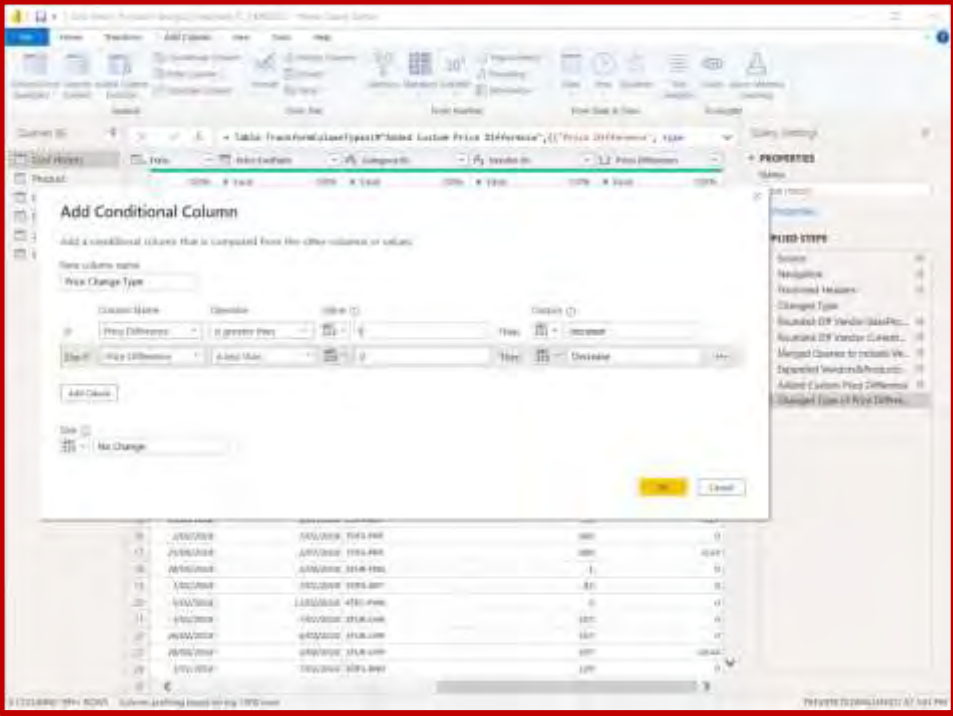
#	Sub-task details	Tick/check <input checked="" type="checkbox"/> once completed.
	Do the following transformations to the Products dataset in 'Power Query Editor'.	
1.	<p>Add a conditional column to the 'Cost History' table called 'Price Change Type' to isolate the values in the 'Price Difference' field as follows.</p> <ul style="list-style-type: none"> • If 'Price Difference' value is greater than '0', the output is 'Increase' • If 'Price Difference' values is less than '0', the output is 'Decrease' • Else the output is 'No Change' <p>Assessor instructions: The screenshot below shows how this task can be done in Power Query Editor. The screenshot is for assessor reference only.</p> 	<input type="checkbox"/>
2.	<p>Filter out the values that contain 'No Change' from the 'Price Change Type' column. So that observations that do not accurately represent a change in price are isolated and removed from the Product dataset.</p> <p>'Close & Apply' the changes in Power Query Editor.</p> <p>Notes for assessors: The screenshot below shows how this task can be done in Power Query Editor. The screenshot is for assessor reference only.</p>	<input type="checkbox"/>

Figure 26 - Screenshot of how to set up the conditional column for Task F2a(1) using Power BI Desktop © Microsoft

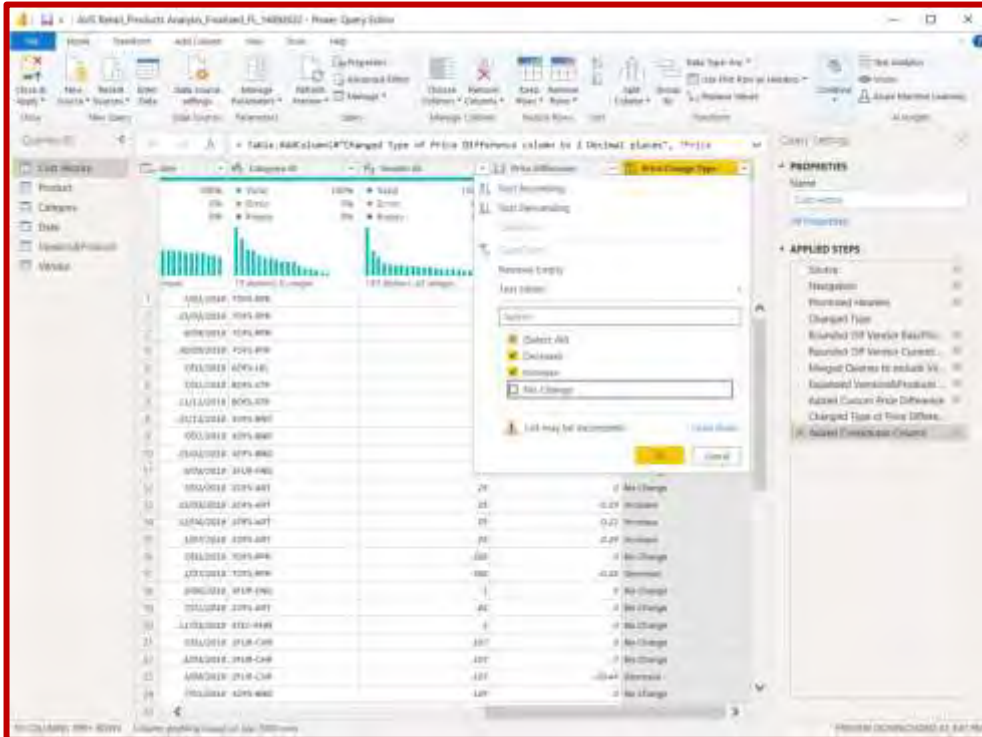


Figure 27 - Screenshot of how to filter the 'Price Change Type' column using Power BI Desktop © Microsoft

3.	Add a 'Slicer' visual to the dashboard page 'Dashboard: Vendor Price Changes Overview' to view the options available from the 'Price Change Type' field in the 'Cost History' table. Format the 'Slicer Settings' > 'Orientation' to 'Horizontal'.	<input type="checkbox"/>
4.	Filter the dashboard page to display analysis results based only on price 'Increases'.	<input type="checkbox"/>
5.	Change the 'View' > 'Theme' of the final version of the Power BI work file to 'Executive'. Note: According to the theme selected the colours in the dashboard will change. This will help differentiate the final version of the dashboard easily.	<input type="checkbox"/>
6.	Provide a screenshot as evidence of the finalised 'Dashboard: Vendor Price Changes Overview', in 'Table 29 > F2a' under 'Screenshot evidence:'. Read the 'Note' section carefully to understand the details that should be visible in the screenshot.	<input type="checkbox"/>

b. Finalise the drill-through report that analyses price changes within a sub-category and develop a report on the key outcomes of the analysis.

To do this, complete the sub-tasks in the checklist given below.

Table 27 - Checklist to perform Task F2b

#	Sub-task details	Tick/check <input checked="" type="checkbox"/> once completed.
1.	<p>On the 'Price Changes within a Sub-category' report page add the following two sentences to the 'Smart Narrative' with the title 'Summary of key outcomes, results and insights'.</p> <p>The [Variable X] and the [Variable Y] have a <Correlation Strength of X & Y>, <Correlation Type of X & Y> correlation with each other.</p> <p>The amount of the price <Price Change Type> averages <Price Diff. Average> and ranges between <Price Diff. Min> to <Price Diff. Max>.</p>	<input type="checkbox"/>

#	Sub-task details	Tick/check <input checked="" type="checkbox"/> once completed.
	<p>Note: The above sentences include variable names and dynamic values. Therefore, when developing the report summary you must,</p> <ul style="list-style-type: none"> • replace the text '[Variable X]' and '[Variable Y]' with the full names of the variables used in the analysis • create and add dynamic values in the sentences as indicated within the '< >' signs. • The variables specific to this report are defined as follows. <ul style="list-style-type: none"> ○ Variable X – Total Vendor Price Changes ○ Variable Y – No. of distinct Vendors Contributed 	
2.	<p>Create a revised drill-through report to analyse the product sub-category that has the highest DVC value.</p> <p>Method: From the finalised 'Dashboard: Vendor Price Changes Overview' page, right-click on the specific sub-category of interest from the visual (according to the context of the analysis), select 'Drill through' and then select the relevant drill-through report page name.</p>	<input type="checkbox"/>
3.	<p>Provide a screenshot as evidence of the finalised 'Price Changes with a Sub-category' report, in 'Table 29> F2b' under 'Screenshot evidence:'. Read the 'Note' section carefully to understand the details that should be visible in the screenshot.</p>	<input type="checkbox"/>

- c. Finalise the drill-through report that analyses price changes in a specific week and develop a report on the key outcomes of the analysis.

To do this, you must complete the sub-tasks in the checklist given below.

Table 28 - Checklist to perform Task F2c

#	Sub-task details	Tick/check <input checked="" type="checkbox"/> once completed.
1.	<p>On the 'Price Changes in a Week' report page add the following two sentences to the 'Smart Narrative' with the title 'Summary of key outcomes, results and insights'.</p> <p>The [Variable X] and the [Variable Y] have a <Correlation Strength of X & Y>, <Correlation Type of X & Y> correlation with each other.</p> <p>The amount of the price <Price Change Type> averages <Price Diff. Average> and ranges between <Price Diff. Min> to <Price Diff. Max>.</p> <p>Note: The above sentences include variable names and dynamic values. Therefore, when developing the report summary you must,</p> <ul style="list-style-type: none"> • replace the text '[Variable X]' and '[Variable Y]' with the full names of the variables used in the analysis • create and add dynamic values in the sentences as indicated within the '< >' signs. • The variables specific to this report are defined as follows. <ul style="list-style-type: none"> ○ Variable X – No. of distinct Vendors Contributed ○ Variable Y – No. of distinct Products Impacted 	<input type="checkbox"/>
2.	<p>Create a revised drill-through report to analyse the specific week that recorded the highest 'Total Vendor Price Changes'.</p> <p>Method: From the finalised 'Dashboard: Vendor Price Changes Overview' page, right-click on the specific week of interest from a visual (according to the context of the analysis), select 'Drill through' and then select the relevant drill-through report page name.</p>	<input type="checkbox"/>
3.	<p>Provide a screenshot as evidence of the finalised 'Price Changes in a Week' report, in 'Table 29 > F2c' under 'Screenshot evidence:'. Read the 'Note' section carefully to understand the details that should be visible in the screenshot.</p>	<input type="checkbox"/>

Note: Once the analysis is completed and finalised save and close your Power BI work file. You will be asked to store a copy of this file in Task F3.

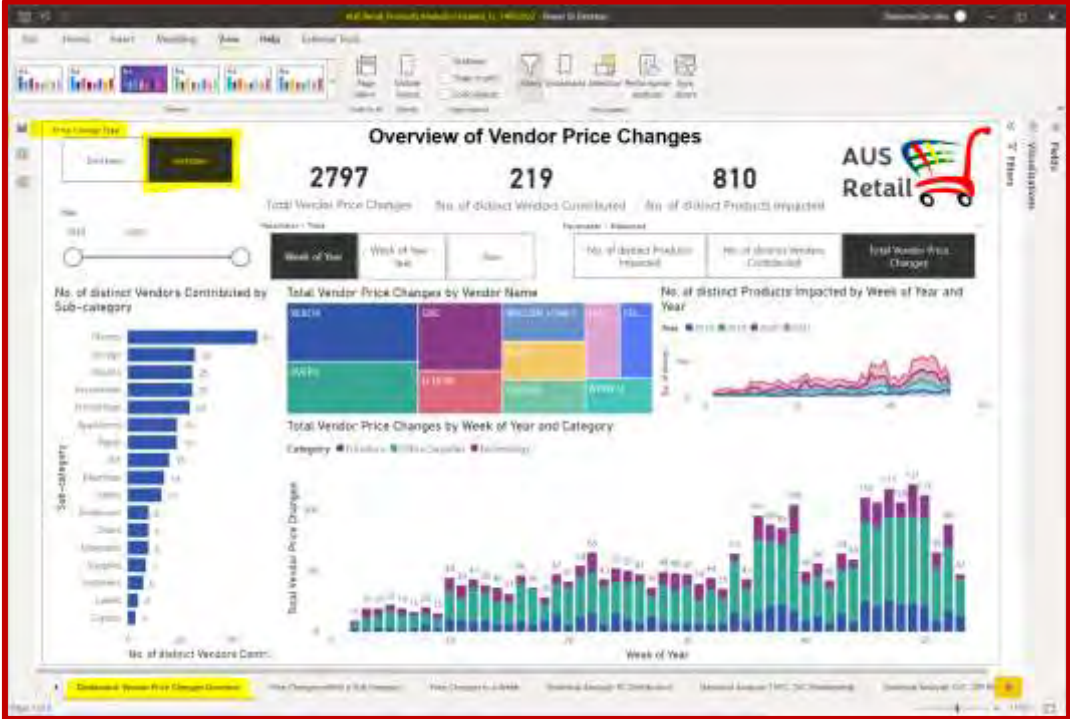
Screenshot evidence:

Assessor instructions: The screenshots of the finalised dashboard page and analysis report pages for the **Products** dataset provided by the student must indicate the following characteristics:

- The values displayed in the 'Card' visualisations;
 - matches what is displayed in the sample screenshots
 - have a significant change when compared with the values displayed in the dashboard previously created in Task C2 and reports previously created in Task D2. This demonstrates that the student had successfully isolated and removed incorrect results from the analysis.
- The screenshots of the drill-through reports (in Table 29: F2b and F2c):
 - includes a summary text in the form of a short paragraph that reports on the key outcomes, results and insights from the analysis.
 - displays figures/values that are relevant to the contexts of each analysis. (The expected values/figures from the analysis are shown in the sample screenshots).

Sample screenshots are provided below.

Table 29 - Evidence of performing Task F2

Demonstration tasks:	Evidence (Screenshot):
<p>F2a. Dashboard: Vendor Price Changes Overview</p> <p>Note: The screenshot of the dashboard must clearly show the following details:</p> <ul style="list-style-type: none"> <input type="checkbox"/> 'Increase' selected in 'Price Change Type' <i>Slicer</i> <input type="checkbox"/> 'Week of Year' selected in 'Parameter-Time' <i>Slicer</i>. <input type="checkbox"/> 'Total Vendor Price changes' selected in 'Parameter-Measures' slicer <input type="checkbox"/> Year range between 2018-2021 selected in 'Year' slicer. <input type="checkbox"/> Title bar showing the file name with your name initials and current date. 	 <p>Figure 28 - Screenshot for task F2a using Power BI Desktop © Microsoft</p>

Demonstration tasks:

Evidence (Screenshot):

F2b. Price Change within a Sub-category

Note: The screenshot of the finalised drill-through report must clearly show the following details:

- All required elements of the report page (Company Logo, Report title, chart titles, and correct visualisation types)
- 'Card' visuals indicating the correct values after removing the incorrect results (outliers).
- 'Summary of key outcomes, results and insights' with new information included using dynamic values.
- Title bar showing the file name with your name initials and current date.

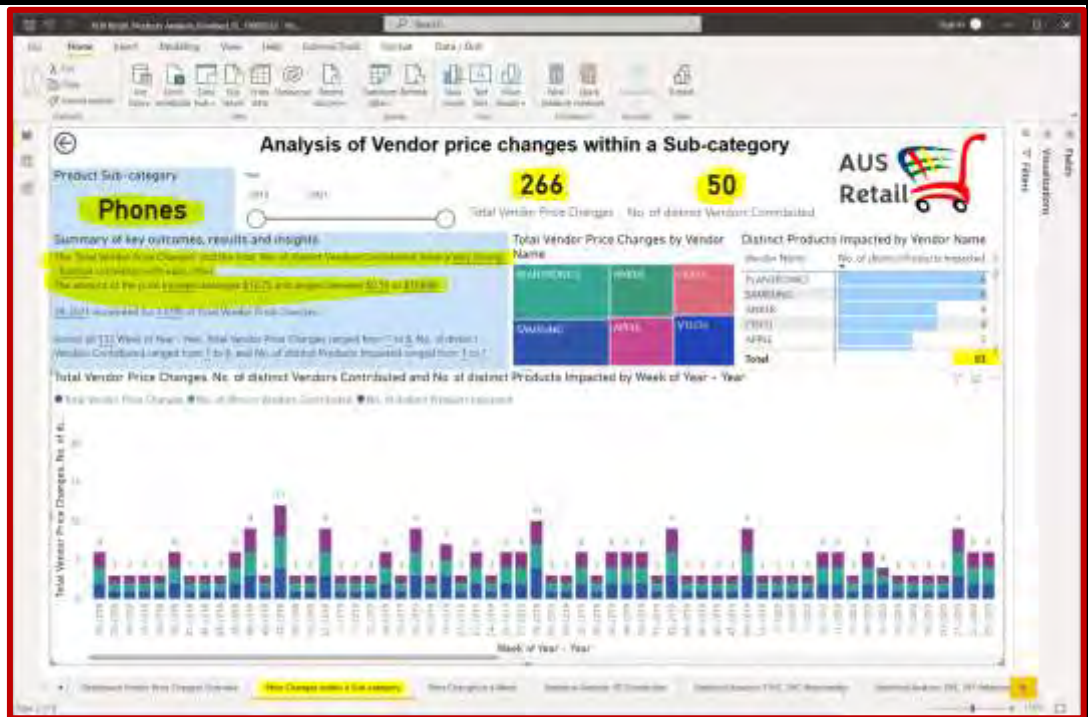


Figure 29 - Screenshot for task F2b using Power BI Desktop © Microsoft

F2c. Price Changes in a Week

Note: The screenshot of the finalised drill-through report must clearly show the following details:

- All required elements of the report page (Company Logo, Report title, chart titles, and correct visualisation types)
- 'Card' visuals indicating the correct values after removing the incorrect results (outliers).
- 'Summary of key outcomes, results and insights' with new information included using dynamic values.
- Title bar showing the file name with your name initials and current date.

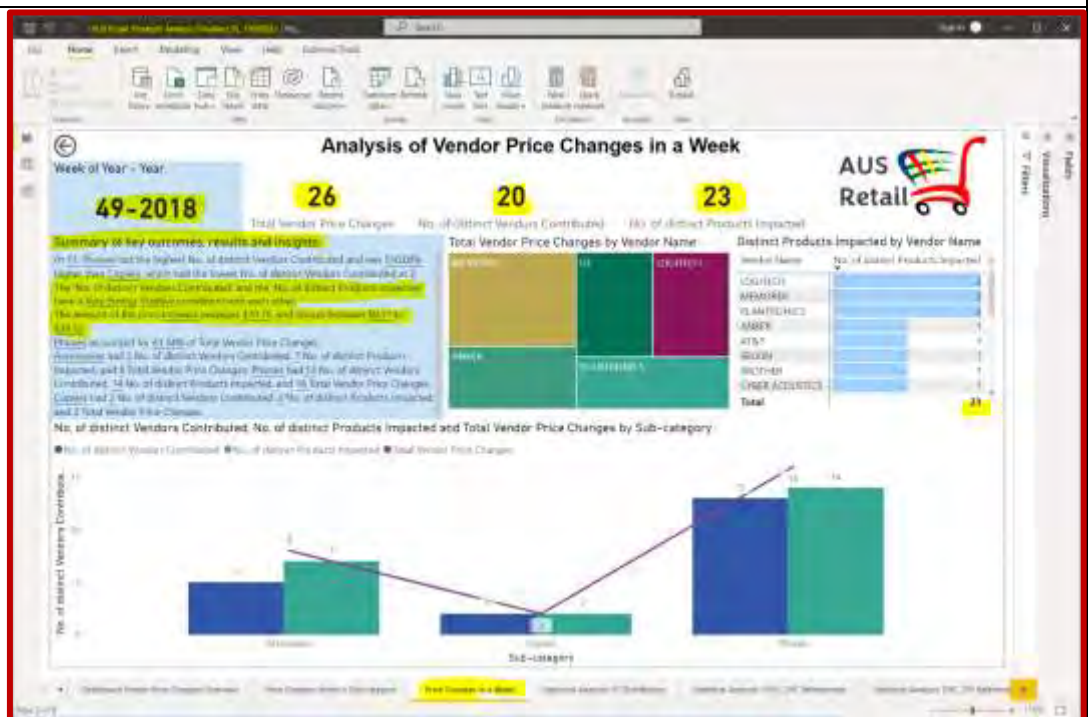


Figure 30 - Screenshot for task F2c using Power BI Desktop © Microsoft

F3. Store analytic results, associated reports and supporting evidence

In this task, you are required to store the associated supporting evidence, results and reports of the analysis according to organisational policies, procedures and legislative requirements.

As preparation for this task, create a new folder in your local computer with your name details called 'BSBXBD403_04_Firstname_Lastname'. For example, a folder created by John Smith should have the name *BSBXBD403_04_John_Smith*.

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

Scenario:

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

From: Gonzales, Mia

To: Student Lastname, Student Firstname

Attached documents: AUS Retail_Data exploration policy.pdf

Subject: Advice on storing supporting evidence, results and reports of the analysis

Hi <student name>,

Great work on completing the analysis and creating the final versions of the drill-through report pages.

Please store all the relevant supporting evidence, results and reports related to the analysis tasks for both Sales and Product datasets according to **AUS Retail's** standard procedure and legislative requirements.

Refer to the following sections of the *AUS Retail_Data analysis policy.pdf* document.

- 9 Storing analytic results and associated supporting evidence
- 9.1 Procedure for storing results of analytic activities
- 9.2 Legislative requirements

Thanks, 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.

Task:

- a. Create the folders required to store the supporting documents, analytic results and reports of the analysis according to **AUS Retail's** standard policy, procedure and legislative requirements.
- b. Store the final version of Power BI work files (the version after completing the tasks in Part F) for both datasets (Sales and Products) in the correct folder according to the company policy and procedure.

Note: The Power BI work files include both the analytic reports and results for each dataset within each file.

- c. Store a copy of the following specification documents that contain technical details related to both datasets in the correct folder according to the company policy and procedure.
- AUS Retail_Data transformation specifications.pdf
 - AUS Retail_Data categorisation specifications.pdf
- d. Provide two (2) screenshots of the folder structure you have created and evidence of storing the relevant documents within the correct folders, in 'Table 30' under 'Screenshot evidence:'
- e. Create a zipped file of the 'BSBXBD403_04_Firstname_Lastname' folder.
- Note:** You must ensure that this folder includes the completed final versions of the Power BI work files and associated supporting documents for each dataset stored within the correct folder structure before compressing (zipping) it to create the zipped file.

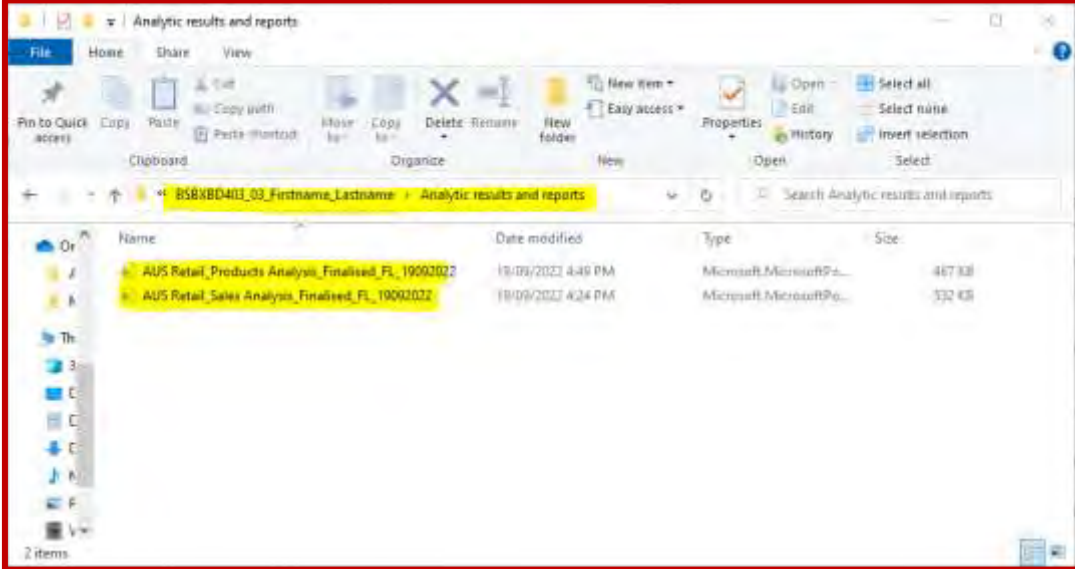
Important: You must submit the 'BSBXBD403_04_Firstname_Lastname' zipped file along with this completed assessment document for marking. Refer to assessment submission instructions for more information.

Screenshot evidence:

Assessor instructions: The folder submitted by the student 'BSBXBD403_04_Firstname_Lastname' (once downloaded and extracted by assessors) must include the 'Supporting documents' folder and 'Analytic results and reports' folder. The associated work files and documents for both datasets must be stored within these folders.

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

Table 30 - Evidence of storing analytic results, associated reports and supporting evidence

Demonstration task:	Evidence (Screenshot):
Storing analytic results and reports	 <p>Figure 31 - Screenshot of storing analytic results and reports in Microsoft Windows File Explorer © Microsoft</p>

Demonstration task: Evidence (Screenshot):

Storing associated supporting evidence

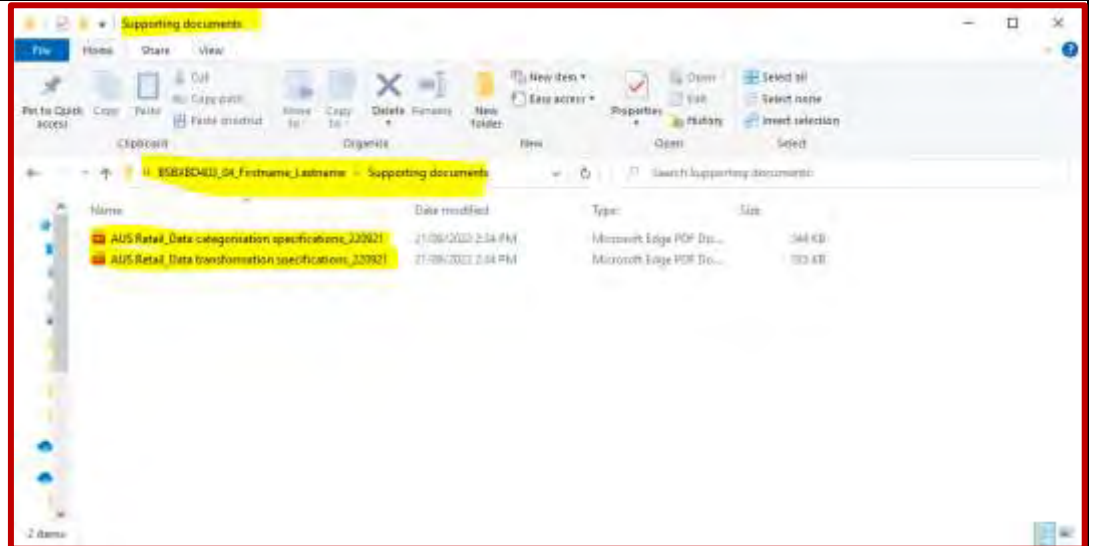


Figure 32 - Screenshot of storing associated supporting evidence in Microsoft Windows File Explorer © Microsoft

Assessment checklist:

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

Part B: Extract and transform big data in preparation for analysis		
B1	Table 3- Evidence of extracting and transforming structured transactional dataset, Sales. (Two screenshots provided)	<input type="checkbox"/>
B2	Table 6 – Evidence of extracting and transforming structured non-transactional dataset, Products. (Two screenshots provided)	<input type="checkbox"/>
B3	Table 7 – Evidence of extracting and transforming unstructured data. (Four screenshots provided)	<input type="checkbox"/>
Part C: Categorise and prepare the captured big data for analysis		
C1	Screenshot provided as evidence of categorising and preparing the transactional dataset, Sales.	<input type="checkbox"/>
C2	Screenshot provided as evidence of categorising and preparing the non-transactional dataset, Products.	<input type="checkbox"/>
Part D: Analyse trends and relationships in big data		
D1	Table 10 – Evidence of analysing the transactional dataset, Sales. (Two screenshots provided)	<input type="checkbox"/>
D2	Table 13 – Evidence of analysing the non-transactional dataset, Products. (Two screenshots provided)	<input type="checkbox"/>
Part E: Conduct statistical analysis to confirm accuracy of big data analysis		
E1	Table 17 – Evidence of statistically analysing the transactional dataset, Sales. (Three screenshots provided)	<input type="checkbox"/>
E2	Table 21 – Evidence of statistically analysing the non-transactional dataset, Products (Three screenshots provided)	<input type="checkbox"/>
Part F: Finalise big data analysis		
F1	Table 25 – Evidence of finalising the analysis of the transactional dataset, Sales. (Three screenshots provided)	<input type="checkbox"/>
F2	Table 29 – Evidence of finalising the analysis of the non-transactional dataset, Products. (Three screenshots provided)	<input type="checkbox"/>
F3	Table 30 – Evidence of storing analytic results, associated reports and supporting evidence (Two screenshots)	<input type="checkbox"/>



Congratulations you have reached the end of Assessment [4]!

© UP Education Online Pty Ltd 2022

Except as permitted by the copyright law applicable to you, you may not reproduce or communicate any of the content on this website, including files downloadable from this website, without the permission of the copyright owner.

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].