

Name : **RANJANI M**

Batch No : **DA|TN-FNB03**

Contact : +918754609051

Email ID : ranjani.mani22@gmail.com

Domain :  **Education** 

Project Title : **Educational Analytics: Student Performance Dashboard through Data Cleaning, Transformation, Modeling and Visualisation using Excel & Power BI**

Submission Date : 21-08-2025

Mentor : Mr.Kumaran

Data Source : Kaggle.com

Dataset Link :
https://docs.google.com/spreadsheets/d/1exdeUQjiGe05xGyug9TfdmGSbpXkEWX6/edit?usp=drive_link&ouid=115479253538516917723&rtpof=true&sd=true

Objective: To provide actionable insights into student academic performance, attendance, and homework completion through **data modeling, cleaning, transformation, and interactive dashboards**. This enables educators and stakeholders to identify **performance trends, strengths, and areas requiring improvement**.

The project uses **four** main tables: **Students, Attendance, Homework, and Performance**. The dataset includes a total of **1,004 students** identified by Student ID, forming the rows used for data cleaning, transformation, modeling, and visualisation.

*Below the description, screenshots attached.

Data Cleaning:

1) Students Table:

1a) **Student ID Column:** Checked for duplicates since ID should be unique.

Function Used: Remove Duplicates (Power Query Editor)

Table1 - Power Query Editor

Queries [1] Table1

Properties Advanced Editor Manage Columns Remove Columns Rows Rows Sort Data Type: Text Use First Row as Headers Merge Queries Append Queries Combine Files Manage Parameters Data source settings New Source Recent Sources Enter Data

File Home Transform Add Column View

Close & Load Refresh Preview Advanced Editor Close Manage Manage Columns Rows Rows Sort Split Column Group By Replace Values Transform Manage Parameters Data Sources New Query

Query Settings

PROPERTIES Name Table1 All Properties

APPLIED STEPS

Source Changed Type Removed Duplicates(ID Unique)

	Student_ID	Name	Name2	Gender	Date_of_Birth
1	S00002	John	Stafford	Male	11/26/2014 12:00:00
2	S00003	Chad	Harper	Male	2/2/2017 12:00:00
3	S00004	Anthony	Martin	Male	11/10/2014 12:00:00
4	S00005	Mary	Stone	f&	4/1/2016 12:00:00
5	S00006	Joe	Wells	Female	6/29/2017 12:00:00
6	S00007	Margaret	Nixon	Female	12/4/2009 12:00:00
7	S00008	James	Wheeler	Male	1/25/2011 12:00:00
8	S00009	Benjamin	Cook	male	10/4/2007 12:00:00
9	S00010	Aaron	Callahan	Male	2/18/2012 12:00:00
10	S00011	christopher	church	Male	2/23/2008 12:00:00
11	S00012	Jonathan	Solis	Male	7/2/2016 12:00:00
12	S00013	Robert	Sanders	m	4/30/2008 12:00:00
13	S00014	Melanie	Gilmore	Female	8/19/2006 12:00:00
14	S00015	Alexandria	Wallace	Female	6/21/2006 12:00:00
15	S00016	Ann	Wilkerson	Female	9/2/2007 12:00:00
16	S00017	Kelly	Delgado	Female	6/28/2015 12:00:00
17	S00018	Michael	Baker	Male	7/21/2008 12:00:00
18	S00019	Jennifer	Adams	Female	1/3/2015 12:00:00
19	S00020	Elizabeth	Lopez	Female	4/21/2016 12:00:00
20	S00021	Patricia	Atkins	Female	1/7/2019 12:00:00
21					

6 COLUMNS, 999+ ROWS Column profiling based on top 1000 rows PREVIEW DOWNLOADED AT 2:49 PM

1) Students Table:

1b) Student Names Column: Two Columns, ‘First Name’ and ‘Last Name’, are merged into one column as ‘Student Name’ and capitalise each word in the names.

Function Used:

~Merged columns (Power Query Editor) (New Column name given as Student Name).

~Capitalise each word (Power Query Editor) (First Name and Last Name first letter changed)

Table1 - Power Query Editor

File Home Transform Add Column View

Queries [1] Table1

Table1 - Table.CombineColumns(#"Removed Duplicates(ID Unique)", {"Name", "Name2"},

Properties: Name = Table1

Applied Steps: Merged Columns(Students Na...)

Student_ID	Students Name	Gender	Date_of_Birth	Grade_Level
S00002	John Stafford	Male	11/26/2014 12:00:00 AM	Grade 5
S00003	Chad Harper	Male	2/7/2017 12:00:00 AM	Grade 3
S00004	Anthony Martin	Male	11/10/2014 12:00:00 AM	Grade 5
S00005	Mary Stone	f&	4/1/2016 12:00:00 AM	Grade 3
S00006	Joe Wells	Female	6/29/2017 12:00:00 AM	Grade 2
S00007	Margaret Nixon	Female	12/4/2009 12:00:00 AM	Grade 4
S00008	James Wheeler	Male	1/25/2011 12:00:00 AM	Grade 1
S00009	Benjamin Cook	male	10/4/2007 12:00:00 AM	Grade 4
S00010	Aaron Callahan	Male	2/18/2012 12:00:00 AM	Grade 4
S00011	christopher church	Male	2/23/2008 12:00:00 AM	Grade 1
S00012	Jonathan Solis	Male	7/7/2016 12:00:00 AM	Grade 1
S00013	Robert Sanders	m	4/30/2008 12:00:00 AM	Grade 2
S00014	Melanie Gilmore	Female	8/19/2006 12:00:00 AM	Grade 4
S00015	Alexandria Wallace	Female	6/21/2006 12:00:00 AM	Grade 2
S00016	Ann Wilkerson	Female	9/2/2007 12:00:00 AM	Grade 3
S00017	Kelly Delgado	Female	6/28/2015 12:00:00 AM	Grade 2
S00018	Michael Baker	Male	7/21/2008 12:00:00 AM	Grade 1
S00019	Jennifer Adams	Female	1/3/2015 12:00:00 AM	Grade 4
S00020	Elizabeth Lopez	Female	4/21/2016 12:00:00 AM	Grade 3
S00021	Patricia Atkins	Female	1/7/2019 12:00:00 AM	Grade 5

5 COLUMNS, 999+ ROWS Column profiling based on top 1000 rows PREVIEW DOWNLOADED AT 2:51 PM

Table1 - Power Query Editor

File Home Transform Add Column View

Queries [1] Table1

Table1 - Table.TransformColumns(#"Merged Columns(Students Names)", {"[Students Name]", Text.Proper, type text})

Properties: Name = Table1

Applied Steps: Capitalized Each Word(Studen...)

Student_ID	Students Name	Gender	Date_of_Birth	Grade_Level
S00002	John Stafford	Male	11/26/2014 12:00:00 AM	Grade 5
S00003	Chad Harper	Male	2/7/2017 12:00:00 AM	Grade 3
S00004	Anthony Martin	Male	11/10/2014 12:00:00 AM	Grade 5
S00005	Mary Stone	f&	4/1/2016 12:00:00 AM	Grade 3
S00006	Joe Wells	Female	6/29/2017 12:00:00 AM	Grade 2
S00007	Margaret Nixon	Female	12/4/2009 12:00:00 AM	Grade 4
S00008	James Wheeler	Male	1/25/2011 12:00:00 AM	Grade 1
S00009	Benjamin Cook	male	10/4/2007 12:00:00 AM	Grade 4
S00010	Aaron Callahan	Male	2/18/2012 12:00:00 AM	Grade 4
S00011	Christopher Church	Male	2/23/2008 12:00:00 AM	Grade 1
S00012	Jonathan Solis	Male	7/7/2016 12:00:00 AM	Grade 1
S00013	Robert Sanders	m	4/30/2008 12:00:00 AM	Grade 2
S00014	Melanie Gilmore	Female	8/19/2006 12:00:00 AM	Grade 4
S00015	Alexandria Wallace	Female	6/21/2006 12:00:00 AM	Grade 2
S00016	Ann Wilkerson	Female	9/2/2007 12:00:00 AM	Grade 3
S00017	Kelly Delgado	Female	6/28/2015 12:00:00 AM	Grade 2
S00018	Michael Baker	Male	7/21/2008 12:00:00 AM	Grade 1
S00019	Jennifer Adams	Female	1/3/2015 12:00:00 AM	Grade 4
S00020	Elizabeth Lopez	Female	4/21/2016 12:00:00 AM	Grade 3
S00021	Patricia Atkins	Female	1/7/2019 12:00:00 AM	Grade 5

5 COLUMNS, 999+ ROWS Column profiling based on top 1000 rows PREVIEW DOWNLOADED AT 2:54 PM

1) Students Table:

1c) Student Name Column: Removed the Inconsistent titles in the names, which are placed in a few rows, such as (mr, Mrs., Ms.)

Function Used:

~Replace Values (Power Query Editor) values checked by using a filter, and each inconsistency was replaced one by one. (4 times replace value steps used for removing title)

The screenshot shows the Power Query Editor interface with the following details:

- Home Tab:** Selected.
- Queries [1]:** Shows a table named "Table1".
- Transform ribbon:** Contains various data manipulation tools like Transpose, Unpivot Columns, Detect Data Type, Fill, Move, Split Column, Format, Parse, Statistics, Standard, Scientific, Trigonometry, Date, Time, Duration, Expand, Aggregate, Extract Values, Create Data Type, and Structured Column.
- Table View:** Displays the "Table1" data with columns: Student_ID, Students Name, Gender, Date_of_Birth, and Grade_Level.
- Query Settings Panel:** Shows the properties of the query, including the name "Table1" and a list of applied steps.
- Applied Steps:** A list of four steps used to remove titles from student names:
 - Replaced Value(removing the title Mr.)
 - Replaced Value(removing the title Mrs.)
 - Replaced Value(removing the title Ms.)
 - Replaced Value(removing the title M)
- Preview:** Shows the transformed data where titles have been removed from the "Students Name" column.

1) Students Table:

1d) Gender Column: Inconsistent data such as (f, f&,m) are changed into Male and Female.

Function Used:

~Replace Values (Power Query Editor) values checked by using a filter, and each inconsistency was replaced one by one. (3 times replace value steps used for Gender Column)
 ~Capitalise each word (Power Query Editor).

The screenshot shows the Power Query Editor interface with the following details:

- Home Tab:** Selected tab.
- Transform Tab:** Active tab.
- Table1 - Power Query Editor:** Title bar.
- Queries [1]:** Shows the current query name.
- Table1:** Data grid containing student information with columns: Student_ID, Students Name, Gender, Date_of_Birth, Grade_Level.
- Applied Steps:**
 - Source
 - Changed Type
 - Removed Duplicates(ID Unique)
 - Merged Columns(Students Na...
 - Capitalized Each Word(Studen...
 - Replaced Value(removingtheti...
 - Replaced Value(removingtheti...
 - Replaced Value(removingtheti...
 - Filtered Rows(Gendercolumnc...
 - Replaced Value(F&ttoFemale)
 - Replaced Value(ItoFemale)
 - Replaced Value(mtoMale)
 - Capitalized Each Word(maletoMale)
- Properties:**
 - Name: Table1
 - All Properties
- Query Settings:**
 - Preview: DOWNLOADED AT 3:12 PM

1) Students Table:

1e) Date of Birth Column: Inconsistent date format changed.

Function Used:

~Changed Date type with locale (Power Query Editor), Type - Date, English United States

The screenshot shows the Power Query Editor interface with the title "Table1 - Power Query Editor". The ribbon has tabs for File, Home, Transform, Add Column, and View. The Home tab is selected. The Transform ribbon group includes Close & Load, Refresh, Advanced Editor, Properties, Choose Columns, Remove Columns, Keep Rows, Remove Rows, Sort, Split Column, Group By, Data Type: Date, Use First Row as Headers, Merge Queries, Append Queries, Combine Files, Manage Parameters, Data source settings, Data Sources, New Source, Recent Sources, Enter Data, and New Query. The main area displays a table with columns: Student_ID, Students Name, Gender, Date_of_Birth, and Grade_Level. The Date_of_Birth column contains various date strings in different formats. A tooltip "Changed Type with Locale(DateofBirthDate)" is visible over the Date Type button. The Query Settings pane on the right shows the properties for the query, including the name "Table1" and the applied steps, which include changing the date type and removing duplicates. The status bar at the bottom indicates "5 COLUMNS, 999+ ROWS" and "Column profiling based on top 1000 rows".

2) Homework Table:

2a) Due Date Column: Inconsistent date format changed.

Function Used:

~Changed Date type with locale (Power Query Editor), Type - Date, English United States

The screenshot shows the Power Query Editor interface with the title "Table3 - Power Query Editor". The ribbon has tabs for File, Home, Transform, Add Column, and View. The Home tab is selected. The Transform ribbon group includes Close & Load, Refresh, Advanced Editor, Properties, Choose Columns, Remove Columns, Keep Rows, Remove Rows, Sort, Split Column, Group By, Data Type: Date, Use First Row as Headers, Merge Queries, Append Queries, Combine Files, Manage Parameters, Data source settings, Data Sources, New Source, Recent Sources, Enter Data, and New Query. The main area displays a table with columns: Assignment_Name, Due_Date, Status, Grade_Feedback, and Grade. The Due_Date column contains various date strings in different formats. A tooltip "Changed Type with Locale(DueDate)" is visible over the Date Type button. The Query Settings pane on the right shows the properties for the query, including the name "Table3" and the applied steps, which include changing the date type. The status bar at the bottom indicates "7 COLUMNS, 999+ ROWS" and "Column profiling based on top 1000 rows".

2) Homework Table:

2b) Status Column: Inconsistent data such as “(âœŒ, âœ”, “âœ...) are changed into “Pending”.

Function Used:

~Find and Replace values (Excel): Copy the string and paste it into the Find, and replace it with 'pending'.

Before:

	A	B	C	D	E	F	G	H	I	J	K	L
1	Student_ID	Subject	Assignment_Name	Due_Date	Status	Grade_Feedback	Guardian_Signature					
2	S00002	English	Wrong enter chair.	12/6/2024	Pending	B-						
3	S00002	Geography	Mission medical.	3/6/2025	âœ”	A						
4	S00002	History	Level somebody by everyone.	1/13/2025	not done	F						
5	S00002	Math	Culture claim.	12/5/2024	pending	F	No					
6	S00003	Math	General participant.	1/7/2025	âœ”	B						
7	S00003	Geography	Drive staff.	12/16/2024	not done	B						
8	S00003	Math	Such lead around.	2/27/2025	âœ”	A+	No					
9	S00003	Geography	Spring.	1/14/2025	not done	C-	No					
10	S00003	Math	If message by.	12/9/2024	pending	B	No					
11	S00004	English	Movement before.	3/2/2025	âœ”	C-	No					
12	S00004	Arabic	Daughter case wish power.	10/5/2024	not done	B						
13	S00004	English	Marriage effect experience.	10/6/2024	âœ”	C-						
14	S00004	History	Then difference simple.	12/16/2024	âœ... ...	F	No					
15	S00004	Science	Cut firm.	12/26/2024	âœ... ...	B-	No					
16	S00005	Geography	Degree society word those.	12/14/2024	Done	A+						
17	S00005	Math	Station send.	1/2/2025	pending	A+						
18	S00005	Math	Whose government.	12/17/2024	Done	A+						
19	S00005	Geography	Top.	10/26/2024	not done	C-						
20	S00005	History	Table scientist.	11/24/2024	âœ”	B						
21	S00005	Geography	This young.	11/28/2024	Done	C	Yes					
22	S00005	English	Pressure best act.	1/31/2025	Done	B	No					
23	S00005	History	Rather war dream.	2/5/2025	âœ”	B	No					
24	S00006	Math	Establish ready foot.	1/30/2025	pending	B-						
25	S00006	English	Question tend local.	2/13/2025	pending	A	Yes					
26	S00006	Math	Will coach.	10/10/2024	âœ”	D	Yes					

After:

Student_ID	Subject	Assignment_Name	Due_Date	Status	Grade_Feedback	Guardian_Signature
S00002	English	Wrong enter chair.	12/6/2024	Pending	B-	
S00002	Geography	Mission medical.	3/6/2025	Pending	A	
S00002	History	Level somebody by everyone.	1/13/2025	Not Done	F	
S00002	Math	Culture claim.	12/5/2024	Pending	F	No
S00003	Math	General participant.	1/7/2025	Pending	B	
S00003	Geography	Drive staff.	12/16/2024	Not Done	B	
S00003	Math	Such lead around.	2/27/2025	Pending	A+	No
S00003	Geography	Spring.	1/14/2025	Not Done	C-	No
S00003	Math	If message by.	12/9/2024	Pending	B	No
S00004	English	Movement before.	3/2/2025	Pending	C-	No
S00004	Arabic	Daughter case wish power.	10/5/2024	Not Done	B	
S00004	English	Marriage effect experience.	10/6/2024	Pending	C-	
S00004	History	Then difference simple.	12/16/2024	Pending	F	No
S00004	Science	Cut firm.	12/26/2024	Pending	B-	No
S00005	Geography	Degree society word those.	12/14/2024	Done	A+	
S00005	Math	Station send.	1/2/2025	Pending	A+	
S00005	Math	Whose government.	12/17/2024	Done	A+	
S00005	Geography	Top.	10/26/2024	Not Done	C-	
S00005	History	Table scientist.	11/24/2024	Pending	B	
S00005	Geography	This young.	11/28/2024	Done	C	Yes
S00005	English	Pressure best act.	1/31/2025	Done	B	No
S00005	History	Rather war dream.	2/5/2025	Pending	B	No
S00006	Math	Establish ready foot.	1/30/2025	Pending	B-	
S00006	English	Question tend local.	2/13/2025	Pending	A	Yes
S00006	Math	Will coach.	10/10/2024	Pending	D	Yes

2) Homework Table:

2c) **Guardian Signature Column:** Contained 1688 blank entries. To handle this missing data, a mode imputation approach was applied. A frequency analysis revealed that the most common value in this column was “No” = 1727, followed by “Yes” = 1664. Accordingly, all blank values were replaced with “No”, ensuring consistency and preserving the categorical distribution of the data.

Formula Used:

~COUNTBLANK(G1:G5080)
 ~COUNTIF(G1:G5080, "No")
 ~COUNTIF(G2:G5080, "Yes")

~IF(ISBLANK(G2), "No",[@[Guardian_Signature]])

Student_ID	Subject	Assignment_Name	Due_Date	Status	Grade_Feedback	Guardian_Signature	Guardian_Sig
S00002	English	Wrong enter chair.	12/6/2024	Pending	B-	No	
S00002	Geography	Mission medical.	3/6/2025	Pending	A	No	
S00002	History	Level somebody by everyone.	1/13/2025	Not Done	F	No	
S00002	Math	Culture claim.	12/5/2024	Pending	F	No	
S00003	Math	General participant.	1/7/2025	Pending	B	No	
S00003	Geography	Drive staff.	12/16/2024	Not Done	B	No	
S00003	Math	Such lead around.	2/27/2025	Pending	A+	No	
S00003	Geography	Spring.	1/14/2025	Not Done	C-	No	
S00003	Math	If message by.	12/9/2024	Pending	B	No	
S00004	English	Movement before.	3/2/2025	Pending	C-	No	
S00004	Arabic	Daughter case wish power.	10/5/2024	Not Done	B	No	
S00004	English	Marriage effect experience.	10/6/2024	Pending	C-	No	
S00004	History	Then difference simple.	12/16/2024	Pending	F	No	
S00004	Science	Cut firm.	12/26/2024	Pending	B-	No	
S00005	Geography	Degree society word those.	12/14/2024	Done	A+	No	
S00005	Math	Station send.	1/2/2025	Pending	A+	No	
S00005	Math	Whose government.	12/17/2024	Done	A+	No	
S00005	Geography	Top.	10/26/2024	Not Done	C-	No	
S00005	History	Table scientist.	11/24/2024	Pending	B	No	
S00005	Geography	This young.	11/28/2024	Done	C	Yes	Yes
S00005	English	Pressure best act.	1/31/2025	Done	B	No	No
S00005	History	Rather war dream.	2/5/2025	Pending	B	No	No
S00006	Math	Establish ready foot.	1/30/2025	Pending	B-	No	
S00006	English	Question tend local.	2/13/2025	Pending	A	Yes	Yes
S00006	Math	Will coach.	10/10/2024	Pending	D	Yes	Yes

Subject	Assignment_Name	Due_Date	Status	Grade_Feedback	Guardian_Signature	Guardian_Sig
English	Wrong enter chair.	12/6/2024	Pending	B-	No	
Geography	Mission medical.	3/6/2025	Pending	A	No	
History	Level somebody by everyone.	1/13/2025	Not Done	F	No	
Math	Culture claim.	12/5/2024	Pending	F	No	
Math	General participant.	1/7/2025	Pending	B	No	
Geography	Drive staff.	12/16/2024	Not Done	B	No	
Math	Such lead around.	2/27/2025	Pending	A+	No	
Geography	Spring.	1/14/2025	Not Done	C-	No	
Math	If message by.	12/9/2024	Pending	B	No	
English	Movement before.	3/2/2025	Pending	C-	No	
Arabic	Daughter case wish power.	10/5/2024	Not Done	B	No	
English	Marriage effect experience.	10/6/2024	Pending	C-	No	
History	Then difference simple.	12/16/2024	Pending	F	No	
Science	Cut firm.	12/26/2024	Pending	B-	No	
Geography	Degree society word those.	12/14/2024	Done	A+	No	
Math	Station send.	1/2/2025	Pending	A+	No	
Math	Whose government.	12/17/2024	Done	A+	No	
History	Table scientist.	10/26/2024	Not Done	C-	No	
Geography	This young.	11/24/2024	Pending	B	No	
English	Pressure best act.	11/28/2024	Done	C	Yes	Yes
History	Rather war dream.	1/31/2025	Done	B	No	No
Math	Establish ready foot.	2/5/2025	Pending	B	No	No
English	Question tend local.	2/13/2025	Pending	A	Yes	Yes
Math	Will coach.	10/10/2024	Pending	D	Yes	Yes

3) Attendance Table:

3a) **Attendance Status Column:** Inconsistent Data, such as (absnt, presnt, late) mix of uppercase and lowercase in first letters are changed.

Function Used:

~Replace Value (Power Query Editor) (5 times replace value steps used).

Before:

Student_ID	Date	Subject	Attendance_Status
S00002	7/6/2024	Arabic	Present
S00002	10/17/2024	Geography	Present
S00002	4/17/2024	Geography	Absent
S00002	10/8/2024	Math	Present
S00002	12/24/2024	Arabic	Late
S00002	9/12/2024	Arabic	left early
S00002	10/11/2024	Math	Absent
S00002	8/2/2024	Geography	Absent
S00002	8/5/2024	English	absnt
S00002	9/21/2024	Science	PRESENT
S00002	2/21/2025	Geography	late
S00002	9/18/2024	Science	Present
S00002	4/7/2024	Science	PRESENT
S00002	2/17/2025	Arabic	Absent
S00002	10/2/2024	Science	Present
S00002	5/6/2024	Geography	absnt
S00002	5/3/2024	Arabic	absnt
S00002	9/3/2024	Math	absnt
S00002	10/6/2024	Arabic	left early
S00002	9/3/2024	History	Present
S00002	9/7/2024	Arabic	Present
S00002	6/29/2024	Math	Present
S00002	3/5/2025	English	Late
S00002	3/31/2024	English	PRESENT
S00002	4/3/2024	Geography	absnt

After:

The screenshot shows the Power Query Editor interface with the following details:

- File**, **Home**, **Transform**, **Add Column**, **View** tabs are visible.
- Transform ribbon** with various data manipulation tools like Transpose, Detect Data Type, Replace Values, Unpivot Columns, Pivot Column, etc.
- Query Settings pane** on the right shows the **Properties** section with the name set to "Table7" and the **Applied Steps** section listing the five replace value steps used to correct the attendance status.
- Table7 preview** shows the corrected data:

Student_ID	Date	Subject	Attendance_Status
S00002	7/6/2024	Arabic	Present
S00002	10/17/2024	Geography	Present
S00002	4/17/2024	Geography	Absent
S00002	10/8/2024	Math	Present
S00002	12/24/2024	Arabic	Late
S00002	9/12/2024	Arabic	Left Early
S00002	10/11/2024	Math	Absent
S00002	8/2/2024	Geography	Absent
S00002	8/5/2024	English	Absent
S00002	9/21/2024	Science	Present
S00002	2/21/2025	Geography	Late
S00002	9/18/2024	Science	Present
S00002	4/7/2024	Science	Present
S00002	2/17/2025	Arabic	Absent
S00002	10/2/2024	Science	Present
S00002	5/6/2024	Geography	Absent
S00002	5/3/2024	Arabic	Absent
S00002	9/3/2024	Math	Absent
S00002	10/6/2024	Arabic	Left Early
S00002	9/3/2024	History	Present
S00002	9/7/2024	Arabic	Present
- Bottom status bar: 4 COLUMNS, 999+ ROWS, Column profiling based on top 1000 rows, PREVIEW DOWNLOADED AT 4:47 PM

4) Performance Table:

4a) **Subject Column:** Identified duplicate entries where the same Student_ID appeared multiple times for the same subject. Removed these duplicates to maintain one unique record per Student-Subject, ensuring accurate subject-level and grade-level analysis.

*By removing duplicates, each Student-Subject pair is unique, making analysis reliable.

*Removed duplicate subject records per student to ensure data accuracy.

Function Used:

~Remove Duplicates (Power Query Editor)

The screenshot shows the Power Query Editor interface with a table containing 21 rows of data. The columns are labeled: Student_ID, Subject, Exam_Score, L2_Homework_Completion..., and L2_HomeworkCompletion. The data includes various subjects like History, Arabic, English, and Geography, along with their corresponding scores and completion percentages. The Power Query ribbon is visible at the top, and the Query Settings pane on the right shows the step 'RemovedDuplicatesCombinationofStudent-IDandSubjectColumn'.

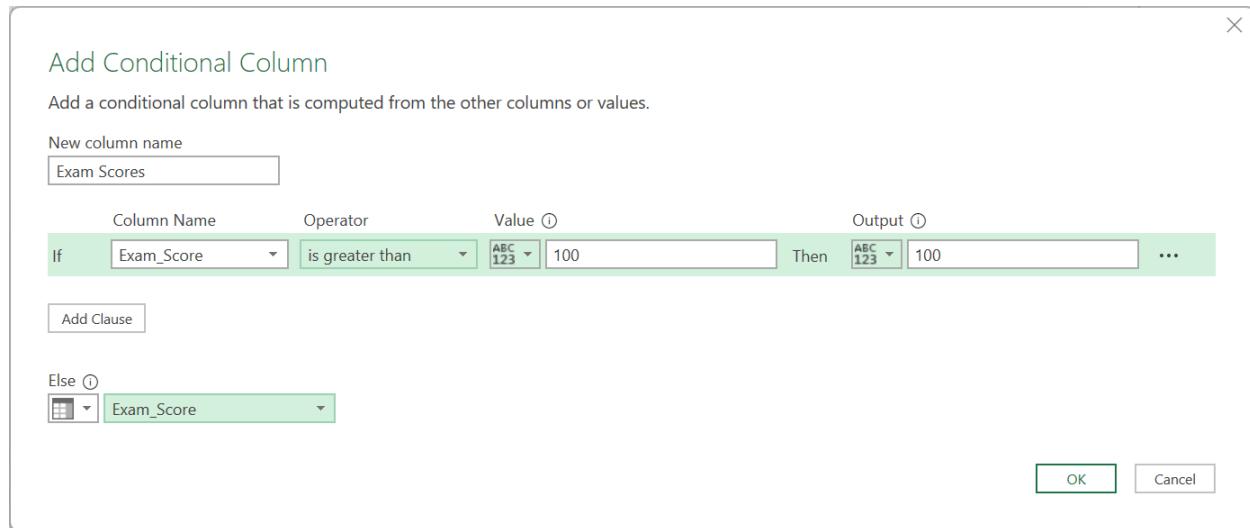
	Student_ID	Subject	Exam_Score	L2_Homework_Completion...	L2_HomeworkCompletion
1	S00002	History	98	90	
2	S00002	Arabic	69	-5	
3	S00003	History	63	-5	
4	S00004	History	59	90	
5	S00004	English	106	-5	
6	S00005	Math	75	-5	
7	S00005	English	78	90	
8	S00005	History	78	0.8	
9	S00005	Arabic	71	-5	
10	S00006	English	79	-5	
11	S00006	History	50	1	
12	S00007	Geography	99	1	
13	S00007	History	51	1	
14	S00007	Math	66	95	
15	S00008	Math	103	0.8	
16	S00008	Arabic	90	95	
17	S00009	Geography	89	-5	
18	S00010	Math	44	90	
19	S00010	Arabic	81	90	
20	S00010	Geography	110	90	
21					

4) Performance Table:

4b) Exam Score Column: Certain exam scores were identified as outliers with values exceeding 100. An outlier was applied to restrict all scores to a maximum of 100, ensuring consistency and validation of the dataset.

Function Used:

~Add Conditional Column(Power Query) - New Column Name: Exam Scores - Operator: is greater than - Value:100 - Output: 100 - Else: Exam Score Column Value.



9 (2) - Power Query Editor

Transform Add Column View

Conditional Column Index Column Duplicate Column

Format From Text From Number From Date & Time

Statistics Standard Scientific Rounding Information

Date Time Duration

Query Settings

Properties Name: Table9 (2)

All Properties

Applied Steps

Source Changed Type RemovedDuplicatesCombinati... Filtered Rows1 AddedConditionalColumnExa...

Index	Exam_Score	1.2_Homework_Completion...	1.2_HomeworkCompletion	Exam Score
1	98	90	0.9	98
2	69	-5	0	69
3	63	-5	0	63
4	59	90	0.9	59
5	106	-5	0	100
6	75	-5	0	75
7	78	90	0.9	78
8	78	0.8	0.8	78
9	71	-5	0	71
10	79	-5	0	79
11	50	1	1	50
12	phy	99	1	99
13	51	1	1	51
14	66	95	0.95	66
15		103	0.8	100
16		95	0.95	90
17	phy	89	-5	89
18		44	90	44
19		81	90	81
20	phy	110	90	100
21				

NWS Column profiling based on top 1000 rows

PREVIEW DOWNLOADED AT 11:31 PM

4c) Homework Completion Column: Contained inconsistent entries, including negative values and percentages exceeding 100. These were cleaned by converting all values to valid percentages between 0% and 100%. After that, format the column as a percentage.

Formula Used:

$\sim\text{MAX}(0,\text{MIN}(\text{IF}(D2>1,D2/100,D2),1))$.

Column D = Homework_Completion %

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
	Student_ID	Subject	Exam_Score	Completion	Completion													
1	S00002	History	98	90	90%													
2	S00002	Arabic	69	-5	0%													
3	S00003	History	63	-5	0%													
5	S00004	History	59	90	90%													
6	S00004	English	106	-5	0%													
7	S00005	Math	75	-5	0%													
8	S00005	English	78	90	90%													
9	S00005	History	78	80%	80%													
10	S00005	Math	46	95	95%													
11	S00005	Arabic	71	-5	0%													
12	S00005	Math	75	90	90%													
13	S00006	English	79	-5	0%													
14	S00006	History	50	100%	100%													
15	S00007	Geography	99	100%	100%													
16	S00007	History	51	100%	100%													
17	S00007	History	81	-5	0%													
18	S00007	Math	66	95	95%													
19	S00008	Math	103	80%	80%													
20	S00008	Arabic	90	95	95%													
21	S00009	Geography	89	-5	0%													
22	S00009	Geography	109	-5	0%													
23	S00010	Math	44	90	90%													
24	S00010	Math	109	90	90%													
25	S00010	Arabic	81	90	90%													

Data Transformation:

A) Calculated Columns:

- 1) **Pass/Fail Status (Performance Table)**: Created a Calculated Column in the performance table to identify whether a student passed or failed based on exam scores.
*Passing mark set at above 40; handles all exam scores after outlier cleaning.

Syntax Used:

$\sim \text{Pass/Fail Status} = \text{IF}(\text{Table}'[\text{ExamScore}] > 40, \text{"Pass"}, \text{"Fail"})$

The screenshot shows the Microsoft Power BI Data Editor interface. The top ribbon has 'Table tools' selected. The 'Column tools' tab is active, showing settings for a new column named 'Pass/FailStatus'. The formula bar contains the DAX code: `1 Pass/FailStatus = IF(Performance[Exam Scores]>40,"Pass","Fail")`. The main data grid displays student performance data with columns: Student_ID, Subject, HomeworkCompletion, Exam Scores, and Pass/FailStatus. The 'Data' pane on the right shows the structure of the 'Performance' table, including the newly created 'Pass/FailStatus' column. The status bar at the bottom indicates there are 2,352 rows in the table.

Student_ID	Subject	HomeworkCompletion	Exam Scores	Pass/FailStatus
S00002	Arabic	0	69	Pass
S00002	History	0.9	98	Pass
S00003	History	0	63	Pass
S00004	English	0	100	Pass
S00004	History	0.9	59	Pass
S00005	Math	0	75	Pass
S00005	English	0.9	78	Pass
S00005	Arabic	0	71	Pass
S00005	History	0.8	78	Pass
S00006	History	1	50	Pass
S00006	English	0	79	Pass
S00007	Geography	1	99	Pass
S00007	History	1	51	Pass
S00007	Math	0.95	66	Pass
S00008	Math	0.8	100	Pass
S00008	Arabic	0.95	90	Pass
S00009	Geography	0	89	Pass
S00010	Geography	0.9	100	Pass
S00010	History	1	53	Pass
S00010	Math	0.9	44	Pass
S00010	Arabic	0.9	81	Pass

2) Homework Status (Performance Table): Created a Calculated Column in the performance table to categorise students (“Completed”/ “Not Completed”) based on homework completion.

Syntax Used:

$\sim \text{Homework Status} = \text{IF}(\text{Performance}[\text{HomeworkCompletion}] >= 0.75, \text{"Completed"}, \text{"Not Completed"})$

The screenshot shows the Microsoft Power BI Data Editor interface. The top navigation bar includes File, Home, Help, Table tools, and Column tools. The Column tools tab is selected, showing settings for a new column named "HomeworkStatus" with a data type of Text. The formula bar displays the DAX code: `1 HomeworkStatus = IF(Performance[HomeworkCompletion]>=0.75,"Completed","Not Completed")`. The main data grid shows student performance data with columns: Student_ID, Subject, HomeworkCompletion, Exam Scores, Pass/FailStatus, and HomeworkStatus. The HomeworkStatus column contains values like "Not Completed" and "Completed". The right side of the screen features a Data pane with a tree view of the data model, including tables for Attendance, Homework, Performance (with sub-tables for Exam Scores and HomeworkCompletion), HomeworkStatus, Pass/FailStatus, Student_ID, Subject, and Students. The bottom status bar indicates "Table: Performance (2,352 rows) Column: HomeworkStatus (2 distinct values)".

Student_ID	Subject	HomeworkCompletion	Exam Scores	Pass/FailStatus	HomeworkStatus
S00002	Arabic	0%	69	Pass	Not Completed
S00002	History	90%	98	Pass	Completed
S00003	History	0%	63	Pass	Not Completed
S00004	English	0%	100	Pass	Not Completed
S00004	History	90%	59	Pass	Completed
S00005	Math	0%	75	Pass	Not Completed
S00005	English	90%	78	Pass	Completed
S00005	Arabic	0%	71	Pass	Not Completed
S00005	History	80%	78	Pass	Completed
S00006	History	100%	50	Pass	Completed
S00006	English	0%	79	Pass	Not Completed
S00007	Geography	100%	99	Pass	Completed
S00007	History	100%	51	Pass	Completed
S00007	Math	95%	66	Pass	Completed
S00008	Math	80%	100	Pass	Completed
S00008	Arabic	95%	90	Pass	Completed
S00009	Geography	0%	89	Pass	Not Completed
S00010	Geography	90%	100	Pass	Completed
S00010	History	100%	53	Pass	Completed
S00010	Math	90%	44	Pass	Completed
S00010	Arabic	90%	81	Pass	Completed

3) Rank-Subject Wise (Performance Table): Created a Calculated Column in the performance table to determine student rank per subject. Ranking is subject-wise and considers all cleaned exam scores.

Syntax Used:

$\sim \text{ExamRank} =$
 $\text{Rankx}(\text{CALCULATETABLE}(\text{Performance}, \text{ALLEXCEPT}(\text{Performance}, \text{Performance}[Subject])),$
 $\text{Performance}[\text{Exam Scores}], , \text{DESC}, \text{Dense})$

The screenshot shows the Microsoft Power BI Data Editor interface. A calculated column named "Exam Rank" is being defined. The formula is:

```
1 Exam Rank = Rankx(CALCULATETABLE(Performance, ALLEXCEPT(Performance, Performance[Subject])), Performance[Exam Scores], , DESC, Dense)
```

The Data pane on the right shows the structure of the "Performance" table, including columns for Subject, HomeworkCompletion, Exam Scores, Pass/FailStatus, HomeworkStatus, and the newly created Exam Rank column.

Student_ID	Subject	HomeworkCompletion	Exam Scores	Pass/FailStatus	HomeworkStatus	Exam Rank
S00002	Arabic	0%	69	Pass	Not Completed	32
S00002	History	90%	98	Pass	Completed	3
S00003	History	0%	63	Pass	Not Completed	38
S00004	English	0%	100	Pass	Not Completed	1
S00004	History	90%	59	Pass	Completed	42
S00005	Math	0%	75	Pass	Not Completed	26
S00005	English	90%	78	Pass	Completed	23
S00005	Arabic	0%	71	Pass	Not Completed	30
S00005	History	80%	78	Pass	Completed	23
S00006	History	100%	50	Pass	Completed	51
S00006	English	0%	79	Pass	Not Completed	22
S00007	Geography	100%	99	Pass	Completed	2
S00007	History	100%	51	Pass	Completed	50
S00007	Math	95%	66	Pass	Completed	35
S00008	Math	80%	100	Pass	Completed	1
S00008	Arabic	95%	90	Pass	Completed	11
S00009	Geography	0%	89	Pass	Not Completed	12
S00010	Geography	90%	100	Pass	Completed	1
S00010	History	100%	53	Pass	Completed	48
S00010	Math	90%	44	Pass	Completed	56
S00010	Arabic	90%	81	Pass	Completed	20

4) Performance Category (Performance Table): Created a Calculated Column in the performance table to group students into performance categories like “Excellent”, “Average”, “Good” and “Poor”. Categories are based on cleaned average scores; thresholds are defined for analysis clarity.

Syntax Used:

$\sim\text{Performance Category} = \text{Switch}(\text{True}(), \text{Performance}[Exam Scores] >= 90, \text{"Excellent"}, \text{Performance}[Exam Scores] >= 75, \text{"Good"}, \text{Performance}[Exam Scores] >= 50, \text{"Average"}, \text{"Poor"})$

The screenshot shows the Power BI desktop application interface. The main area displays a table named "Performance" with 2,352 rows. The table has columns: Student_ID, Subject, HomeworkCompletion, Exam Scores, Pass/FailStatus, HomeworkStatus, Exam Rank, and Performance Category. The "Performance Category" column contains the formula: $\text{Performance Category} = \text{Switch}(\text{True}(), \text{Performance}[Exam Scores] >= 90, \text{"Excellent"}, \text{Performance}[Exam Scores] >= 75, \text{"Good"}, \text{Performance}[Exam Scores] >= 50, \text{"Average"}, \text{"Poor"})$. The data model pane on the right shows the relationships between tables: Subject, Homework, Assignment_Name, Due_Date, Grade_Feedback, Guardian_Sign, Status, Student_ID, and Subject. The Power BI ribbon at the top is visible, showing "Table tools" and "Column tools". The "Column tools" tab is selected, showing options for Name, Data type, Format, Summarization, Data category, Sort by column, Data groups, Manage relationships, and New column.

Student_ID	Subject	HomeworkCompletion	Exam Scores	Pass/FailStatus	HomeworkStatus	Exam Rank	Performance Category
S00002	Arabic	0%	69	Pass	Not Completed	32	Average
S00002	History	90%	98	Pass	Completed	3	Excellent
S00003	History	0%	63	Pass	Not Completed	38	Average
S00004	English	0%	100	Pass	Not Completed	1	Excellent
S00004	History	90%	59	Pass	Completed	42	Average
S00005	Math	0%	75	Pass	Not Completed	26	Good
S00005	English	90%	78	Pass	Completed	23	Good
S00005	Arabic	0%	71	Pass	Not Completed	30	Average
S00005	History	80%	78	Pass	Completed	23	Good
S00006	History	100%	50	Pass	Completed	51	Average
S00006	English	0%	79	Pass	Not Completed	22	Good
S00007	Geography	100%	99	Pass	Completed	2	Excellent
S00007	History	100%	51	Pass	Completed	50	Average
S00007	Math	95%	66	Pass	Completed	35	Average
S00008	Math	80%	100	Pass	Completed	1	Excellent
S00008	Arabic	95%	90	Pass	Completed	11	Excellent
S00009	Geography	0%	89	Pass	Not Completed	12	Good
S00010	Geography	90%	100	Pass	Completed	1	Excellent
S00010	History	100%	53	Pass	Completed	48	Average
S00010	Math	90%	44	Pass	Completed	56	Poor

5) Attendance Value (Attendance Table): Created the AttendanceFlag column, which assigns numeric values to each student's attendance status: Present = 1, and all other statuses (Absent, Excused, Late, Left Early) = 0. This column provides a consistent numeric basis for calculating overall attendance percentages and KPI measure.

Syntax Used:

$\sim\text{Attendance Flag} = \text{IF}(\text{Attendance}[\text{Attendance_Status}] = \text{"Absent"}, 0, 1)$

The screenshot shows a Microsoft Excel spreadsheet titled 'StudentsPerformance_Dataset'. The 'Attendance Flag' column is selected, and its formula is displayed as `= IF(Attendance[Attendance_Status] = "Absent", 0, 1)`. The table contains 30,192 rows of student attendance data. The columns are labeled: Student_ID, Date, Subject, Attendance_Status, and Attendance_Flag. The Attendance_Status column lists various dates and subjects, while the Attendance_Flag column consistently shows the value 0 across all rows.

Student_ID	Date	Subject	Attendance_Status	Attendance Flag
S00002	Monday, August 5, 2024	English	Absent	0
S00003	Thursday, May 9, 2024	English	Absent	0
S00003	Tuesday, January 21, 2025	English	Absent	0
S00004	Tuesday, January 14, 2025	English	Absent	0
S00004	Tuesday, December 24, 2024	English	Absent	0
S00004	Tuesday, February 4, 2025	English	Absent	0
S00005	Saturday, April 27, 2024	English	Absent	0
S00005	Thursday, May 16, 2024	English	Absent	0
S00005	Thursday, September 12, 2024	English	Absent	0
S00006	Sunday, October 27, 2024	English	Absent	0
S00006	Wednesday, June 19, 2024	English	Absent	0
S00007	Monday, October 14, 2024	English	Absent	0
S00008	Tuesday, March 12, 2024	English	Absent	0
S00008	Friday, March 22, 2024	English	Absent	0
S00009	Monday, October 7, 2024	English	Absent	0
S00009	Tuesday, September 24, 2024	English	Absent	0
S00011	Saturday, September 7, 2024	English	Absent	0
S00011	Tuesday, August 20, 2024	English	Absent	0
S00013	Monday, April 29, 2024	English	Absent	0
S00014	Sunday, May 26, 2024	English	Absent	0
S00014	Sunday, June 16, 2024	English	Absent	0

- 6) **Assignment Value (Homework Table):** Created a Calculated Column in the homework table to convert assignment completion into a numeric scale. Supports Homework Completion % measure; ensures numeric consistency.

Syntax Used:

~Assignment Value =

SWITCH(Homework[Grade_Feedback], "A+", 5.5, "A", 5, "A-", 4.5, "B", 4, "B-", 3.5, "C", 3, "C-", 2.5, "D", 2, "D-", 1.5, "E", 1, "E-", 0.5, "F", 0, BLANK())

The screenshot shows the Microsoft Power BI Data Editor interface. The top navigation bar includes 'File', 'Home', 'Help', 'Table tools', and 'Column tools'. The 'Column tools' tab is selected, showing settings for the 'Assignment Value' column: 'Name' (Assignment Value), 'Format' (Whole number), 'Summarization' (Sum), 'Data type' (Decimal number), 'Data category' (Uncategorized), and various sorting and grouping options. Below the ribbon is a table view with columns: Student_ID, Subject, Assignment_Name, Due Date, Status, Grade_Feedback, Guardian Sign, and Assignment Value. The 'Assignment Value' column contains numerical values corresponding to the grade feedback. A tooltip for the first row shows the formula: `1 Assignment Value = SWITCH(Homework[Grade_Feedback], "A+", 5.5, "A", 5, "A-", 4.5, "B", 4, "B-", 3.5, "C", 3, "C-", 2.5, "D", 2, "D-", 1.5, "E", 1, "E-", 0.5, "F", 0, BLANK())`. To the right of the table is a 'Data' pane displaying relationships between tables like 'Attendance', 'Date', 'Homework', and 'Assignment Value'. At the bottom, it says 'Table: Homework (5,079 rows) Column: Assignment Value (8 distinct values)' and 'Update available (click to download)'.

7) **Guardian Sign Value (Homework Table):** Created a Calculated Column in the homework table to convert the guardian signature into a numeric value. Numeric 0/1 values enable calculation of the Guardian Signature % measure.

Syntax Used:

$\sim \text{Guardian Sign Value} = \text{IF}(\text{Homework}[\text{Guardian Sign}] = "Yes", 1, 0)$

The screenshot shows the Microsoft Power BI Data Editor interface. The top navigation bar includes 'File', 'Home', 'Help', 'Table tools', and 'Column tools'. The 'Column tools' tab is selected, showing settings for the current column: Name is 'Guardian Sign Value', Data type is 'Whole number', Format is '\$ %', and Summarization is 'Sum'. The 'Guardian Sign Value' column is highlighted in the table, and its formula is displayed as `if(Homework[Guardian Sign] = "Yes", 1, 0)`. The table contains 5,079 rows of assignment data. To the right of the table is a 'Data' pane showing various tables and columns available in the model, such as Attendance, Date, Homework, Assignment Value, Due Date, Grade Feedback, Guardian Sign, Status, Student ID, and Subject. The bottom status bar indicates 'Table: Homework (5,079 rows) Column: Guardian Sign Value (2 distinct values)' and 'Update available (click to download)'.

Assignment_Name	Due_Date	Status	Grade_Feedback	Guardian Sign	Assignment Value	Guardian Sign Value
Personal environmental fight its.	Sunday, December 22, 2024	Pending	B-	No	4	0
Bag.	Tuesday, September 24, 2024	Pending	B	No	4	0
Time year.	Wednesday, December 18, 2024	Pending	C-	No	3	0
Operation agree popular.	Tuesday, October 22, 2024	Pending	F	No	0	0
Two choose.	Thursday, December 12, 2024	Pending	C	No	3	0
Defense teach.	Friday, December 20, 2024	Pending	A	No	5	0
Per.	Sunday, December 8, 2024	Pending	A	No	5	0
Yet sometimes.	Tuesday, December 17, 2024	Pending	B-	No	4	0
Hold week car.	Monday, November 4, 2024	Pending	B-	No	4	0
Water wrong picture two.	Saturday, November 9, 2024	Pending	B	No	4	0
Hot budget.	Saturday, December 14, 2024	Pending	B	No	4	0
Of dream technology.	Wednesday, November 27, 2024	Pending	C	No	3	0
Move leg.	Sunday, November 3, 2024	Pending	C-	No	3	0
Father physical heavy.	Sunday, September 29, 2024	Pending	C-	No	3	0
See activity.	Thursday, November 14, 2024	Pending	C-	No	3	0
Carry above foreign natural.	Thursday, February 20, 2025	Pending	C	No	3	0
Rest serious thought.	Wednesday, January 8, 2025	Pending	C-	No	3	0
Save face pretty.	Sunday, November 10, 2024	Pending	C-	No	3	0
Medical ahead join.	Saturday, October 26, 2024	Pending	A	No	5	0
Local sell benefit.	Saturday, January 18, 2025	Pending	D	No	2	0

Calculated Measures:

- 1) **Students Table (TotalStudents):** Created a measure in the student table to count the total number of students in the dataset. Provides a single value displayed in a Card visual.

Syntax Used:

$$\sim \text{Total Students} = \text{DISTINCTCOUNT}(\text{Students}[Student_ID])$$

The screenshot shows the Power BI Desktop interface. In the top ribbon, the 'Measure tools' tab is selected. A new measure named 'Total Students' has been created, defined as `DISTINCTCOUNT(Students[Student_ID])`. The measure is set to a whole number format. In the main workspace, a card visual displays the value '1004' for 'Total Students'. The 'Data' pane on the right shows the 'Students' table with fields like Student_ID, Date_of_Birth, Gender, Grade_Level, and Students_Name. The 'Fields' section of the 'Data' pane has a checkmark next to 'Total Students'.

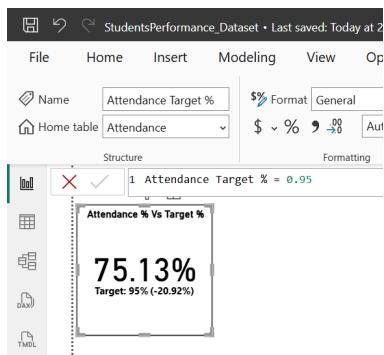
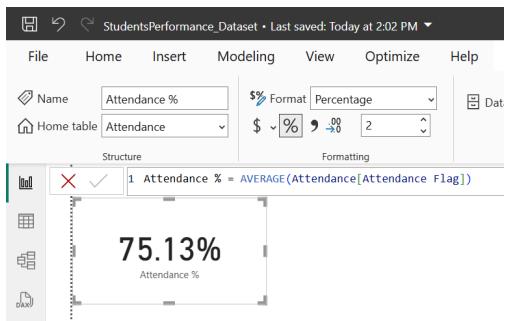
2) Attendance Table (Attendance %): The measure Attendance % computes the average of Attendance Flag across all students, giving the overall Attendance %. The Attendance % KPI measure calculates overall student attendance using the AttendanceFlag column (1 = Present, 0 = Absent), ignoring all filters to ensure the KPI. The target attendance is set at 95%.

Syntax Used:

$\sim\text{Attendance \%} = \text{Average}(\text{Attendance}[\text{Attendance Flag}])$

$\sim\text{Attendance \% KPI} = \text{Attendance \% KPI} = \text{CALCULATE}(\text{AVERAGE}(\text{Attendance}[\text{Attendance Flag}]), \text{ALL}(\text{Attendance}), \text{ALL}(\text{Students}), \text{ALL}(\text{Performance}), \text{ALL}(\text{Homework}))$

Attendance Target % = 0.95



3) Performance Table (Average Exam Score): Created a measure in the performance table to calculate the mean exam score after outlier handling. To use for visualising per subject in a matrix or chart.

Syntax Used:

$\sim \text{Average Exam Score} = \text{Average}(\text{Performance}[Exam Scores])$

The screenshot shows the Power BI interface with the title bar "StudentsPerformance_Dataset • Last saved: Today at 1:43 PM". The ribbon menu includes File, Home, Insert, Modeling, View, Optimize, Help, and Format. In the Modeling tab, a new measure is being defined. The formula bar shows the DAX code: `Average Exam Score = Average(Performance[Exam Scores])`. The visual area displays a single value "74" with the label "Average Exam Score". The bottom right corner shows a note: "Attendance % Overall and".

4) Performance Table (Homework completion %): Created a measure in the performance table to calculate the average homework completion.

Syntax Used:

$\sim \text{Homework Completion \%} =$
 $\text{AVERAGEX}(\text{Performance}, \text{IF}(\text{Performance}[\text{HomeworkCompletion}] < 0, 0, \text{Performance}[\text{HomeworkCompletion}]))$

The screenshot shows the Power BI interface with the title bar "StudentsPerformance_Dataset • Last saved: Today at 1:43 PM". The ribbon menu includes File, Home, Insert, Modeling, View, Optimize, Help, Format, Data / Drill, and Table tools. In the Table tools tab, a new measure is being defined. The formula bar shows the DAX code: `Homework Completion \% = AVERAGEX(Performance, IF(Performance[HomeworkCompletion] < 0, 0, Performance[HomeworkCompletion]))`. The visual area displays a single value "72.52%" with the label "Homework Completion \%". The bottom right corner shows a note: "Filters on this v is (All)".

- 5) Homework Table (Guardian Signature %): Created a measure in the homework table to calculate the percentage of students whose guardian has signed. Uses numeric 0/1 Guardian Sign Value; displayed in Card visual.

Syntax Used:

$\sim \text{Guardian Signature \%} = \text{AVERAGE}(\text{Homework}[\text{Guardian Sign Value}])$

The screenshot shows the Power BI Model view. In the top navigation bar, 'StudentsPerformance_Dataset' is selected, and the status bar indicates 'Last saved: Today at 1:43 PM'. The ribbon tabs include File, Home, Insert, Modeling, View, Optimize, Help, and Format. Under the Insert tab, 'Card' is selected. The main area displays a card visual with the text '32.76%' and 'Guardian Signature %'. Above the card, the formula 'Guardian Signature \% = AVERAGE(Homework[Guardian Sign Value])' is shown in the formula bar. The 'Formatting' section of the ribbon is open, showing 'General' selected under 'Format'.

- 6) Performance Table (Pass and Fail %): Created a measure in the performance table to calculate the percentage of pass and fail displayed in Card visual.

Syntax Used:

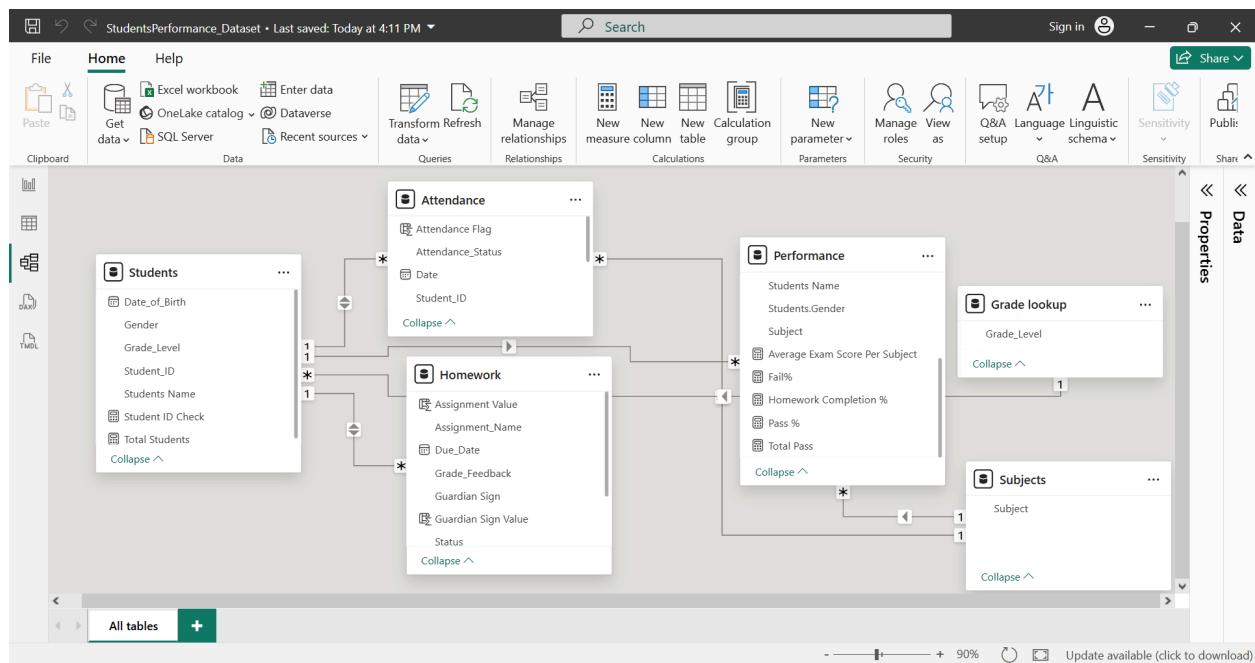
$\sim \text{Pass \%} = \text{Divide}(\text{CALCULATE}(\text{COUNTROWS}(\text{Performance}), \text{FILTER}(\text{Performance}, \text{Performance}[\text{Pass/FailStatus}] = "Pass")), \text{COUNTROWS}(\text{Performance}), 0)$

$\sim \text{Fail \%} = \text{Divide}(\text{CALCULATE}(\text{COUNTROWS}(\text{Performance}), \text{FILTER}(\text{Performance}, \text{Performance}[\text{Pass/FailStatus}] = "Fail")), \text{COUNTROWS}(\text{Performance}), 0)$

The screenshot shows the Power BI Model view. The ribbon tabs include File, Home, Insert, Modeling, View, Optimize, Help, Format, Data / Drill, Table tools, and Measure to. Under the Insert tab, 'Card' is selected. The main area displays two cards: one labeled 'Pass %' with the value '98.26%' and another labeled 'Fail %' with the value '1.74%'. Above the cards, the formula 'Pass \% = Divide(CALCULATE(COUNTROWS(Performance), FILTER(Performance, Performance[Pass/FailStatus] = "Pass")), COUNTROWS(Performance), 0)' is shown in the formula bar. The 'Formatting' section of the ribbon is open, showing 'Percentage' selected under 'Format'.

Data Modeling:

The data model connects fact tables (**Students**, **Attendance**, **Performance**, and **Homework**) with lookup tables (**Grade** and **Subject**) to enable effective filtering and slicer functionality. **One-to-many relationships were established between the lookup tables and fact tables**, allowing, for example, one Grade to have many Students and one Subject to relate to multiple records in Performance and Attendance. Relationships are configured with single-direction filtering to avoid ambiguity, and the Grade and Subject lookup tables act as a bridge for interactive slicers. Data cleaning steps, such as removing blanks and duplicates, were performed before modeling. This setup ensures consistent interactive filtering across all charts and KPI on the dashboard, supporting efficient analysis of student metrics.

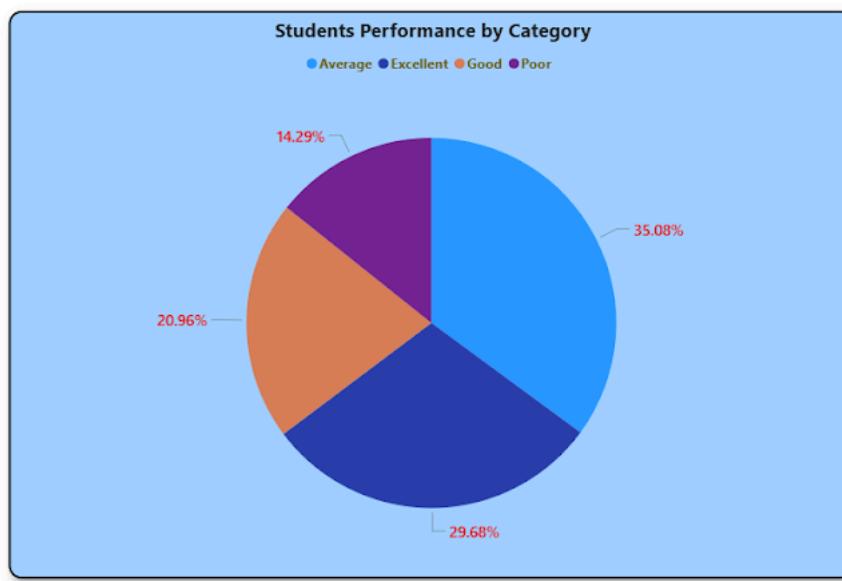


Data Visualisation:

The Chart and dashboard visualise student performance across subjects, attendance, homework completion and performance, enabling quick insights and interactive analysis.

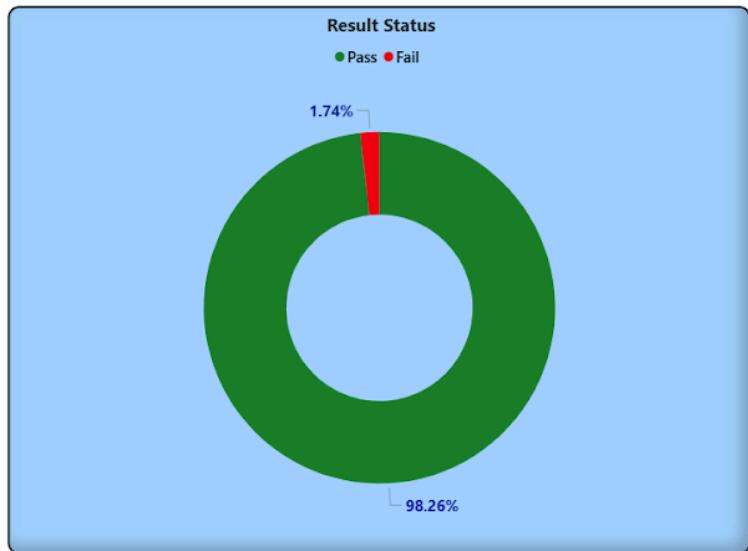
1) Student Performance by Category (Pie Chart)

- **What it shows:** The share of students in each performance category (Excellent, Good, Average, Poor).
- **Why it matters:** Gives a quick view of how student performance is distributed across different levels.
- **Key insight:** Helps identify the largest and smallest performance groups, guiding teachers where to focus support.



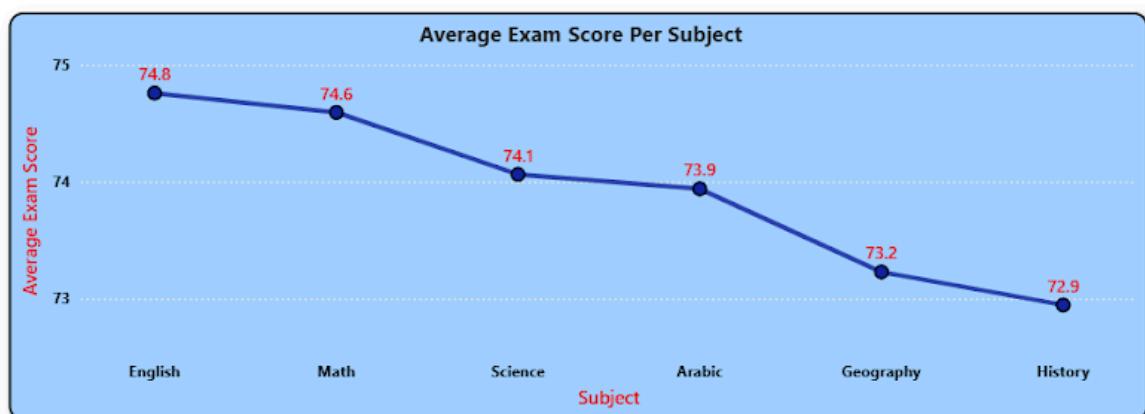
2) Pass/Fail Status (Doughnut Chart)

- **What it shows:** The proportion of students who have passed versus those who have failed.
- **Why it matters:** Provides a quick visual of overall academic success and areas of concern.
- **Key insight:** Helps identify the percentage of students meeting minimum performance standards and those needing additional support.



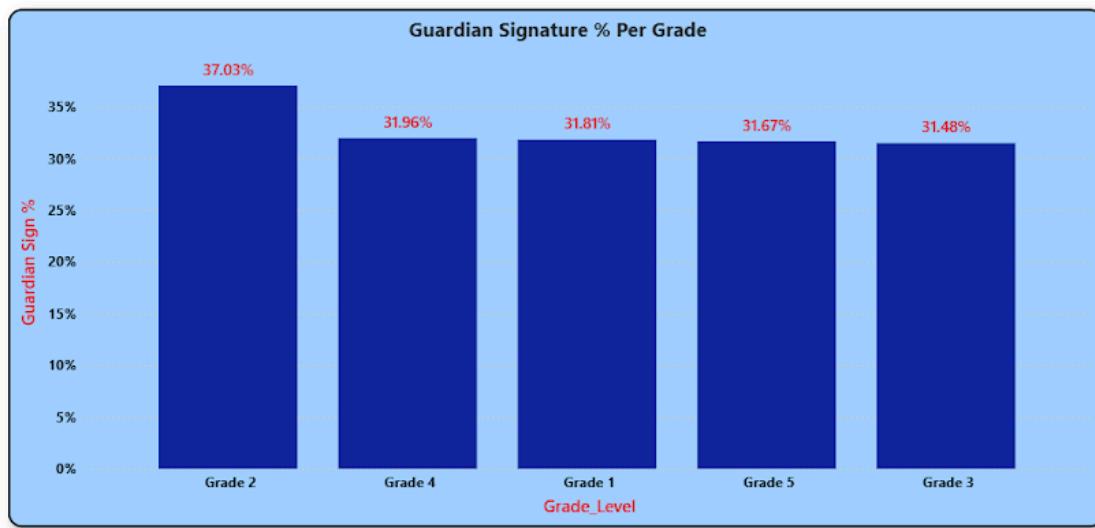
3) Average Exam Score per Subject (Line Chart)

- **X-axis:** Subjects
- **Y-axis:** Average Exam Score
- **What it shows:** The trend of students' average exam scores across each subject.
- **Why it matters:** Helps identify subject-wise performance strengths and weaknesses.
- **Key insight:** Subjects with lower average scores indicate areas needing additional attention or support.



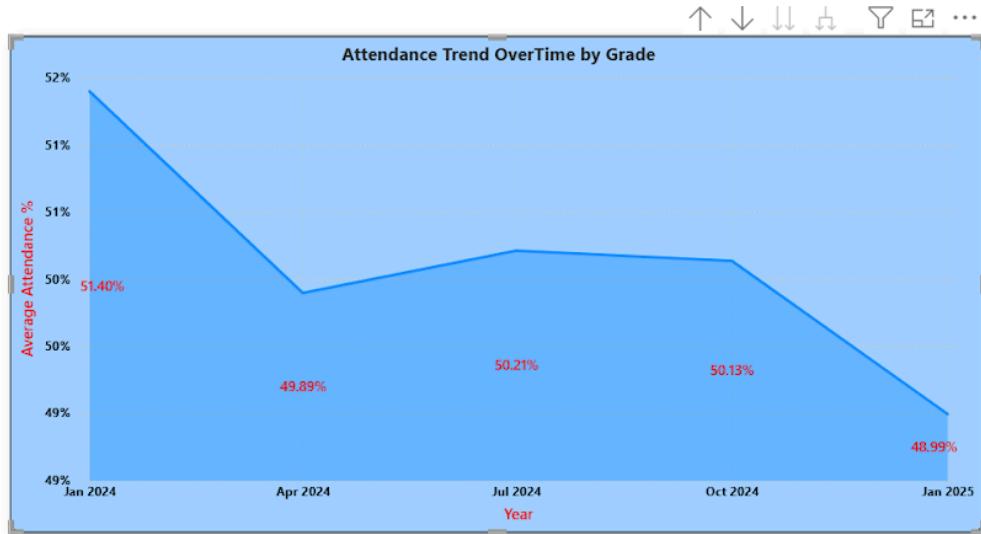
4) Guardian Signature by Class/Section (Column Chart)

- **X-axis:** Class/Section
- **Y-axis:** Guardian Signature Percentage
- **What it shows:** The percentage of guardians who signed homework for each grade.
- **Why it matters:** Reflects the level of parental involvement in academics.
- **Key insight:** Classes with lower signature percentages may require more parent-teacher communication.



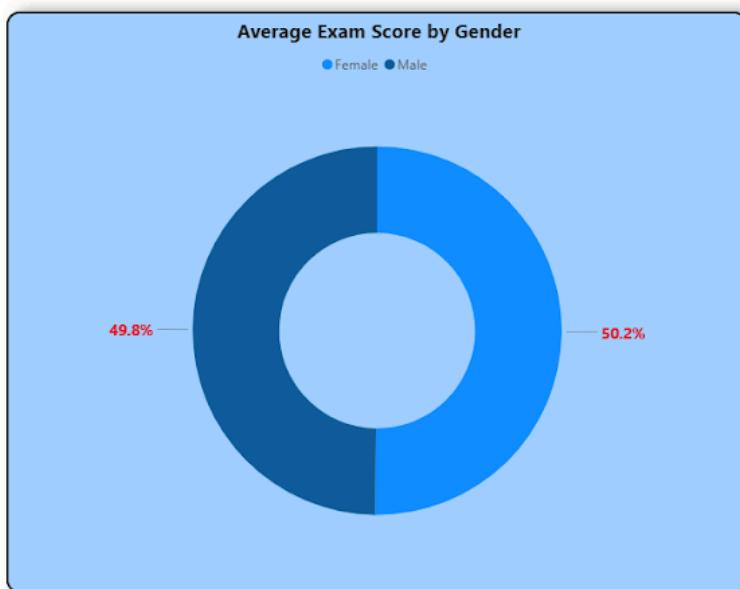
5) Attendance Trend over Time (Area Chart)

- **X-axis:** Time (Dates or Months)
- **Y-axis:** Attendance Percentage
- **What it shows:** The variation of student attendance over a specific time period, filled for better visual emphasis.
- **Why it matters:** Makes it easy to observe attendance patterns, peaks, and dips across time.
- **Key insight:** Identifies time periods with low attendance, allowing for proactive measures.



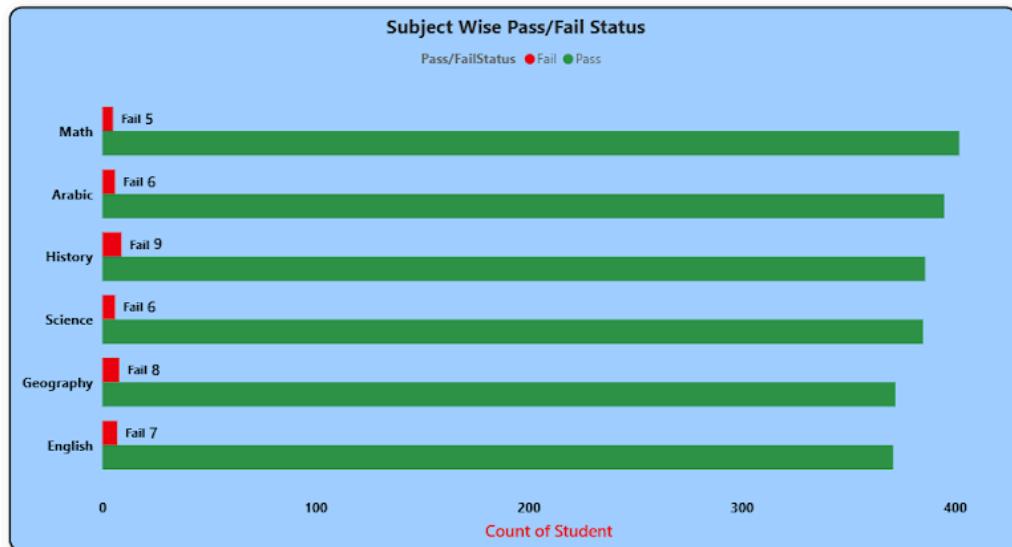
6) Gender Distribution (Doughnut Chart)

- What it shows:** The proportion of male and female students in the dataset.
- Why it matters:** Provides an overview of gender balance among students.
- Key insight:** Helps ensure equal representation and identify any gender-based participation gaps.



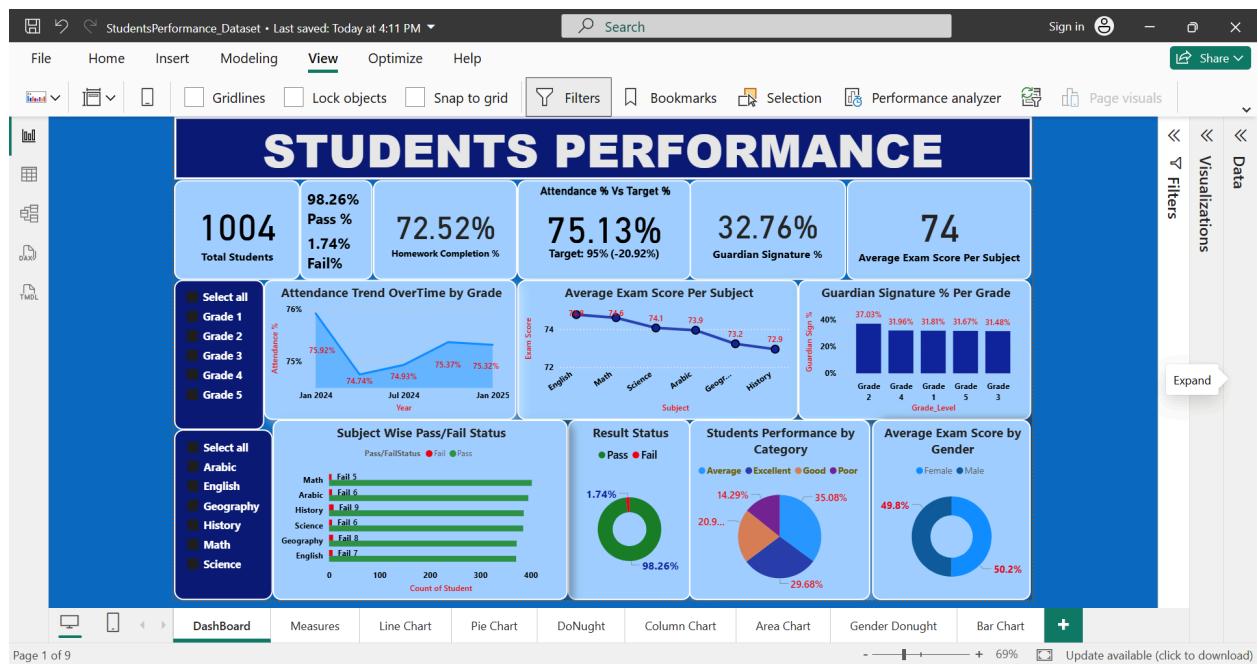
7) Subject-wise Pass/Fail Status (Bar Chart)

- **X-axis:** Subjects
- **Y-axis:** Number of Students
- **What it shows:** The count of students who passed and failed in each subject, displayed side by side.
- **Why it matters:** Highlights which subjects have higher failure rates and need additional attention.
- **Key insight:** Subjects with more failures indicate areas requiring focused academic support.



Dashboard (Students Performance):

Students Performance dashboard provides a clear view of student performance using KPIs, cards, and interactive charts. It includes visualisations for Pass/Fail distribution, Performance Categories, Average Exam Score by Subject, Guardian Signature, Attendance Trend, and Gender Distribution. Grade and Subject slicers allow users to filter and analyse the data dynamically across all visuals.



Conclusion:

The **dashboard** provides an interactive view of **student performance** for **1004 students**. The **Pass/Fail chart** shows the proportion of students who passed and failed, with an overall **pass percentage of 98.26%**. **Grade and Subject slicers** enable dynamic filtering, helping users explore trends in **attendance, exam scores, homework, and guardian engagement**. This project demonstrates how **cleaned and modelled data** can be effectively visualised for clear academic insights.