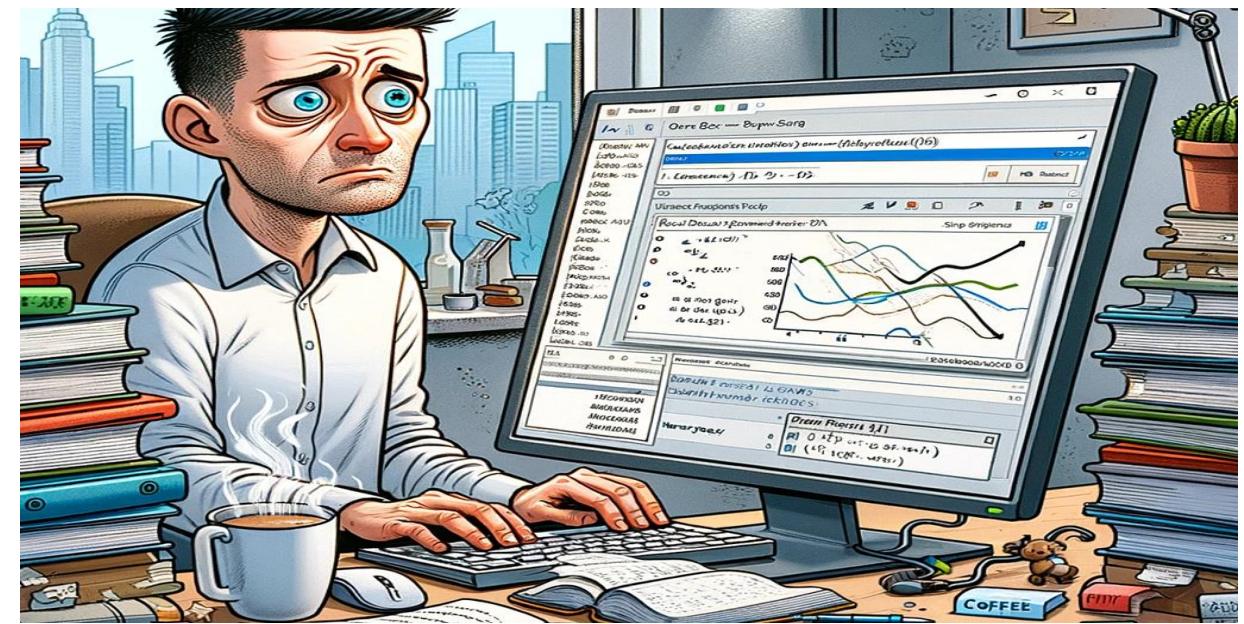


Rapid Power BI Development with ChatGPT

Pedro Reis





About the speaker

- Microsoft Data Platform MVP / MCT
- Expert Leader Data & AI
- Leader of Fabric PBI Portugal Group
- www.fabricbi.net (training / consulting)



pedroreis@reispmonmicrosoft.com



<https://www.linkedin.com/in/pedrogoncalvesreis/>



www.fabricbi.net/



Agenda

Today's 20m Session

- 1) DAX w/ SVG and CSS
- 2) Data Modeling & Analysis
- 3) Report Design Assessment
- 4) Troubleshooting
- 5) Knowledge Assessment and Improvement

Tomorrow's 40m Deep Dive

- 1) Generate a Data Model
- 2) Generate Artificial Data
- 3) Generate Power Query code
- 4) Generate DAX Measures
- 5) Enhance Report Theme
- 6) Publication & Administration w/ Cmdlets
- 7) Key Takeaways / The Future



1) DAX with SVG and CSS

DAX with SVG

DAX with CSS on HTML visual





Generate DAX Measure with SVG

AG

You

Create a DAX Measure to return an SVG image with a button-like appearance and displays the text "Not Ok" on it. The SVG is defined to have a rectangular red button with rounded corners, and the text is centered on the button in white

ChatGPT

To create a DAX measure that returns an SVG image of a button with the text "Not Ok," you'll need to use a combination of DAX and SVG markup. The SVG will define the shape and style of the button, while DAX will be used to generate and return the SVG code as a measure in Power BI.

Here's a step-by-step guide:

- 1. Create a New Measure:** In your Power BI report, create a new measure. You can do this by right-clicking on your data table and selecting "New measure".
- 2. Write the DAX for SVG Code:** The DAX measure will return a string that is the SVG code for the button. Here is an example measure:

```
dax
Copy code

SVG Button =
"<svg width='100' height='50' xmlns='http://www.w3.org/2000/svg'>" &
"  <rect width='100%' height='100%' style='fill:red;stroke-width:3;stroke:rgb(0,0,0);rx:15;ry:15' />" &
"  <text x='50%' y='50%' dominant-baseline='middle' text-anchor='middle' style='fill:white;font-size:20px'>Not Ok</text>" &
"</svg>"
```

The screenshot shows the Power BI desktop application. On the left, a DAX editor window is open with the following code:

```
1 SVG Button =
2 "<svg width='100' height='50' xmlns='http://www.w3.org/2000/svg'>" &
3 "  <rect width='100%' height='100%' style='fill:red;stroke-width:3;stroke:rgb(0,0,0);rx:15;ry:15' />" &
4 "  <text x='50%' y='50%' dominant-baseline='middle' text-anchor='middle' style='fill:white;font-size:20px'>Not Ok</text>" &
5 "</svg>"
```

On the right, the 'Visualizations' pane is visible, showing various chart and matrix icons. Below it, the 'Data' pane lists tables: 'Date', 'Geography', 'Manufacturer', 'Product', 'Sales', and 'Table'. Under 'Table', there are two items: 'Number' and 'SVG Button', with 'SVG Button' being the selected item.

Generate DAX Measure with SVG

AG

You

Here is an example of a DAX Measure which displays an SVG image with text. Based on this I will ask further adjustments.

SVG Button with Text =

```
VAR TextToDisplay = "Not Ok"  
RETURN "data:image/svg+xml;utf8," &  
    "<svg xmlns='http://www.w3.org/2000/svg' viewBox='0 0 160 100'>  
        <rect x='10' y='25' width='140' height='50' rx='15' fill='red' />  
        <text x='80' y='50' font-size='24' fill='white' text-anchor='middle' alignment-  
            baseline='middle'>" & TextToDisplay & "</text>  
    </svg>"
```

AG

You

Adjust this measure to display "OK" in Green if the value of column Table[Number] >= 5 or "Not OK" in Red if the value is < 5

ChatGPT

To adjust the DAX measure to display "OK" in green if the value of the `Table[Number]` column is greater than or equal to 5, or "Not OK" in red if the value is less than 5, you can use an `IF` statement within your measure. Here's how you can modify the measure:

dax

Copy code

```
SVG Button with Text =  
VAR NumberValue = MAX(Table[Number]) -- Assuming you want to evaluate  
VAR TextToDisplay = IF(NumberValue >= 5, "OK", "Not OK")  
VAR Color = IF(NumberValue >= 5, "green", "red")  
RETURN "data:image/svg+xml;utf8," &  
    "<svg xmlns='http://www.w3.org/2000/svg' viewBox='0 0 160 100'>  
        <rect x='10' y='25' width='140' height='50' rx='15' fill=''" & Color &  
        "' />  
        <text x='80' y='50' font-size='24' fill='white' text-anchor='middle' alignment-baseline='middle'>" & TextToDisplay & "</text>  
    </svg>"
```

AG

You

It didn't work because it needed to be max('Table'[Number]).
Why didn't you place ' in the Table name?

ChatGPT

You are absolutely correct, and I apologize for the oversight. In DAX, when referencing a table name or a column name that might contain special characters or spaces, or when it's best practice to avoid ambiguity, it is necessary to enclose the table and column names in single quotes.

Here's the corrected measure with the appropriate use of single quotes:

```
dax  
X ✓  
1 SVG Button =  
2 VAR NumberValue = MAX('Table'[Number]) -- Use single quotes around table and  
column names if needed  
3 VAR TextToDisplay = IF(NumberValue >= 5, "OK", "Not OK")  
4 VAR Color = IF(NumberValue >= 5, "green", "red")  
5 RETURN "data:image/svg+xml;utf8," &  
    "<svg xmlns='http://www.w3.org/2000/svg' viewBox='0 0 160 100'>  
        <rect x='10' y='25' width='140' height='50' rx='15' fill=''" & Color &  
        "' />  
        <text x='80' y='50' font-size='24' fill='white' text-anchor='middle' alignment-baseline='middle'>" & TextToDisplay & "</text>  
    </svg>"
```

Visualizations >> Data
Build visual
Search
> Date
> Geography
> Manufacturer
> Product
> Sales
Table
Number
SVG Button
Py
Date
Geography
Manufacturer
Product
Sales
Table
Number
SVG Button
Columns
Number
SVG Button
Drill through

Bonus 5.5 Dash Pokemon - Final - Power BI Desktop

Jefferson Alves

Arquivo Página Inicial Inserir Modelagem Exibição Ajuda Ferramentas externas

Recortar Colar Copiar Pincel de formatação

Obter dados Pasta de trabalho do Excel Hub de SQL Server dados Fontes recentes

Inserir Dataverse Novo visual Transformar dados Atualizar dados

Caixa de texto Mais visuais

Nova medida rápida Cálculos Confidencialidade

Publicar Compartilhar

Área de Transferência Dados Consultas Inserir Cálculos Confidencialidade Compartilhar

PokéMoy

Omanyte

#138

Como alguns omanitas conseguem escapar após serem restaurados ou são soltos na natureza por pessoas, essa espécie está se tornando um problema.

Altura (M): 0,4 Peso (KG): 7,5 IMC: 46,9

Evolução

Omanyte #138 Omastar #139

Poderes

	Ataque	Defesa	HP	Velocidade
Omanyte	40	100	35	35

Dados

Filtros

Pesquisar

Medidas

- 1. Principais
- 2. Detalhes
- 3. Batalhas
 - Imagen animada

Aux Pokemon

dPokemon

fHabilidades

Capa Detalhes Detalhes v2 Batalhas Batalhas v2 Aula

Página 6 de 6

113%

Generate DAX CSS Measure

AG

You

Provide a DAX measure for Power BI that embeds an image from a column and scales it to twice its size using CSS within a DAX measure referencing
SELECTEDVALUE(Jogadores[Foto])

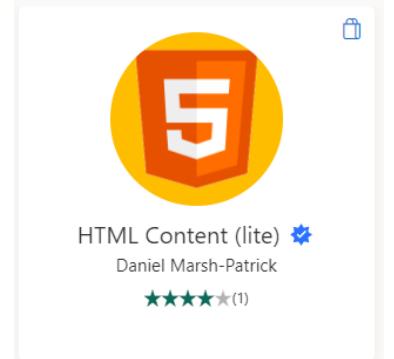
ChatGPT

Certainly! Here's a DAX measure for Power BI that embeds an image from the `Jogadores[Foto]` column and scales it to twice its original size using CSS. This measure references the `SELECTEDVALUE` function to obtain the image URL:

```
dax
DoubleSizeImage =
VAR Foto = SELECTEDVALUE(Jogadores[Foto])
RETURN
IF(
    ISBLANK(Foto),
    BLANK(),
    "<img src='" & Foto & "' style='transform: scale(2); object-fit: co
)
```



Nome
■ Cristiano Ronaldo
□ Neymar



HTML Custom Visual

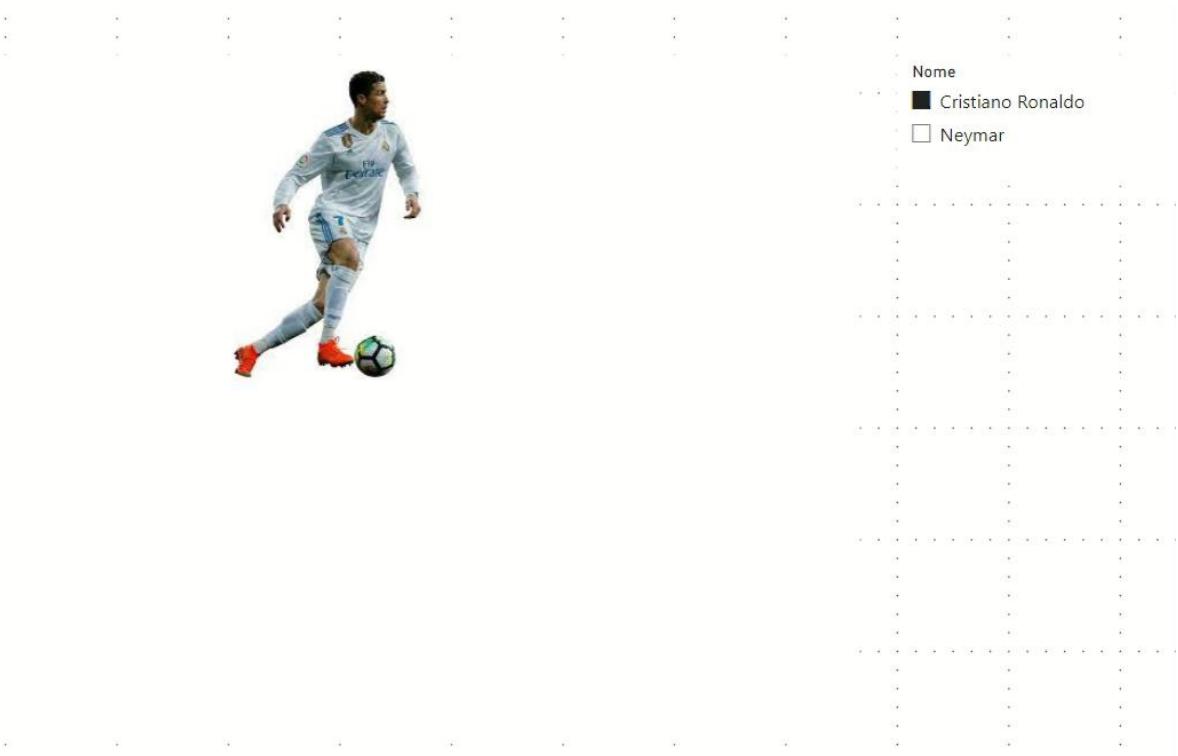
Generate DAX CSS Measure

AG You

- adjust the animation to the following:
 - Move it right 50% of its size
 - Scale it to 2x in 1 second
- Move it left to the original position
- Scale it to the original size

AG Yo

adjust the code so that if Nome = "Cristiano Ronaldo" keep the defined trajectory, and if it's "Neymar", the image will do zig-zag's



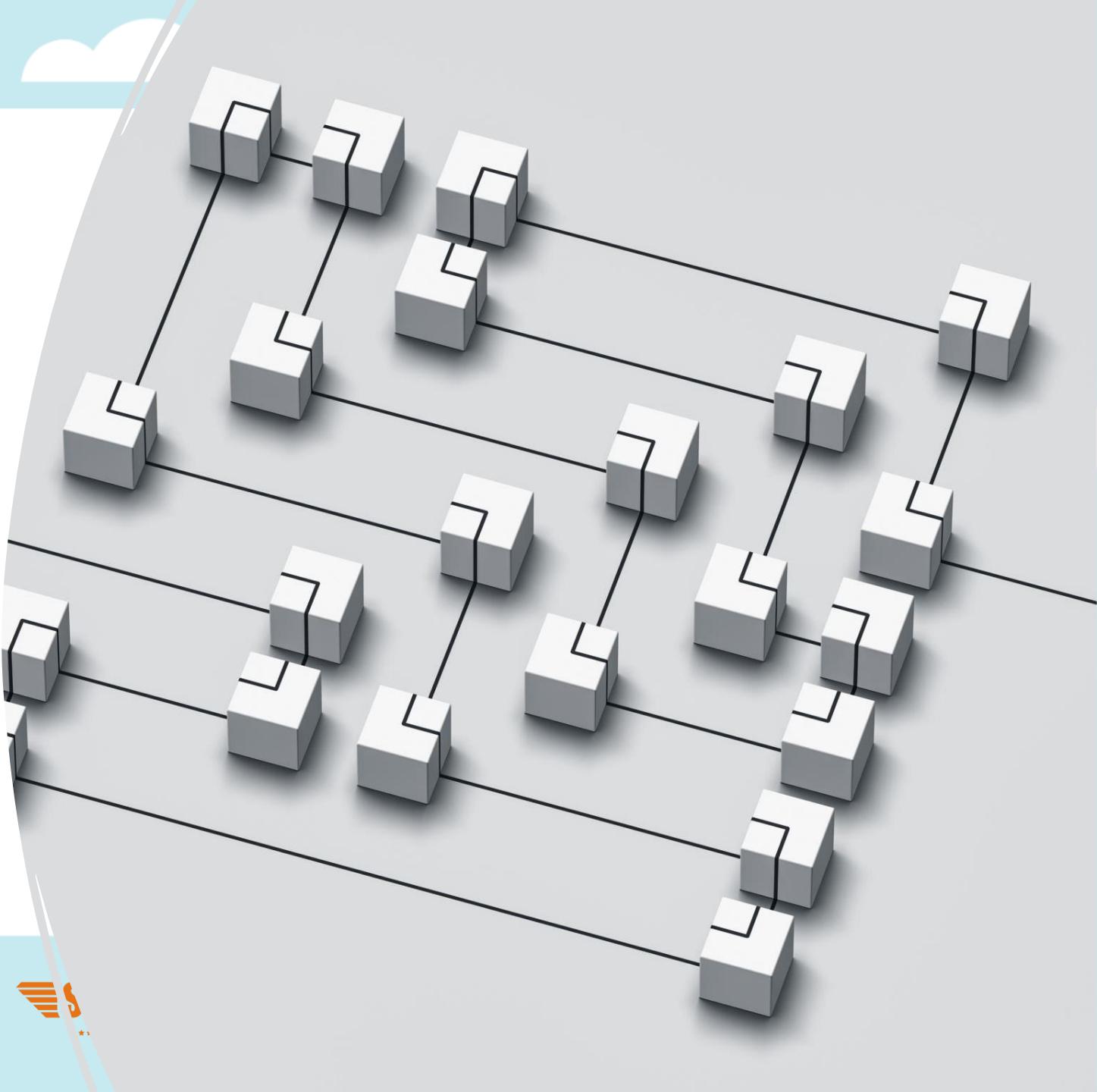


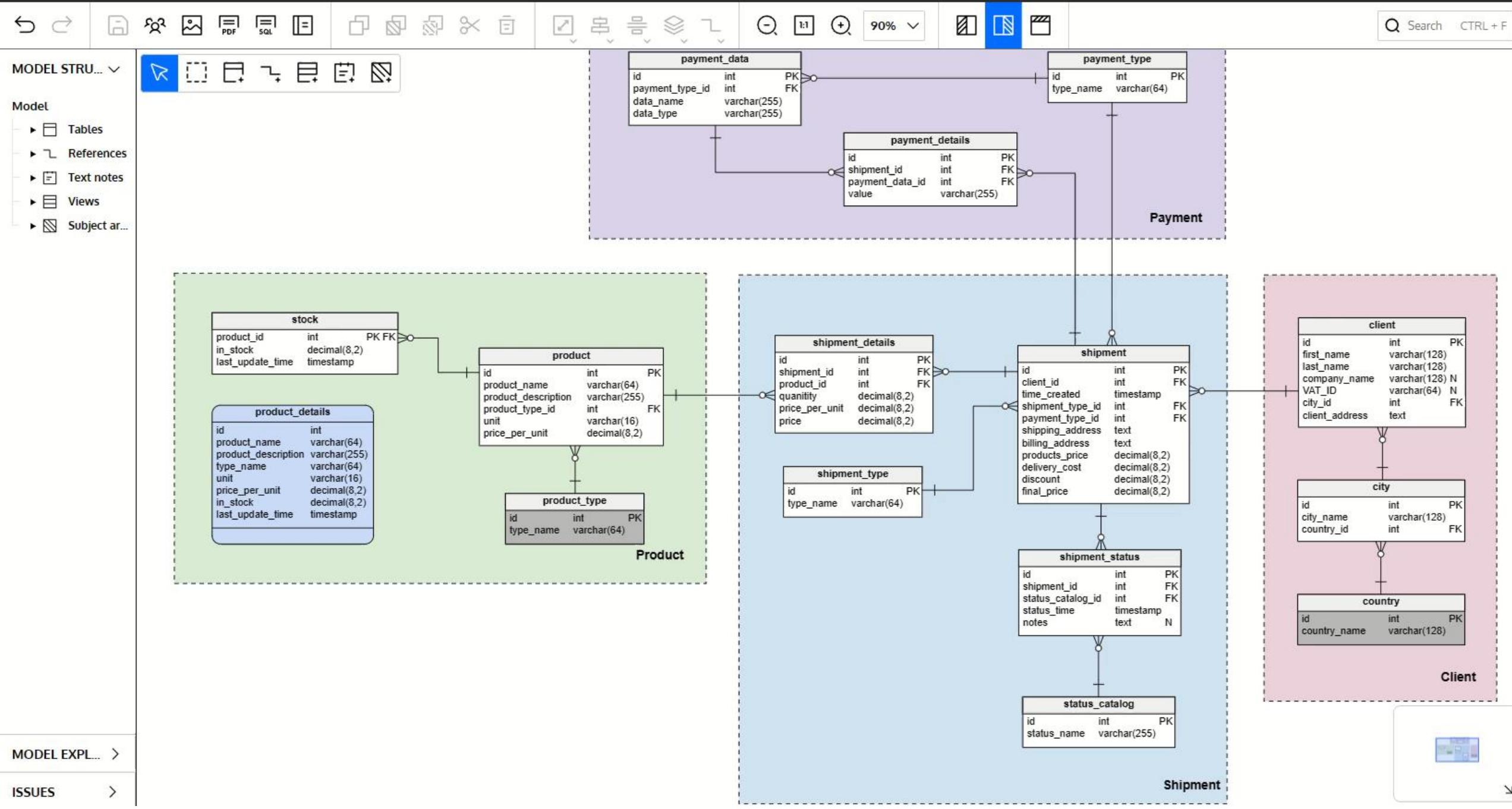
2) Data Modeling & Analysis

Analyze a Data Model

Analyze Data

Generate a Data Model

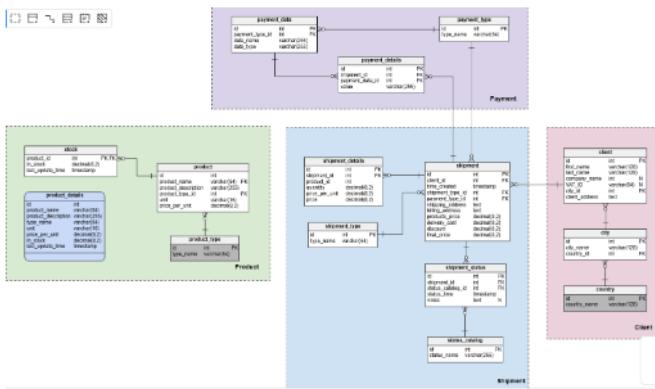




Data Model Analysis

AG

You



Analyze the following Data Model. Focus on the product and client tables. Assess the model and identify any potential issues with relationships, keys, data types or table designs.



ChatGPT

Based on the provided data model diagram, here are the potential issues identified within the `product` and `client` tables:

Product Table Issues:

- Primary Key (PK):** The `product` table has its PK set on the `id` field, which is standard and appropriate. There do not appear to be issues with the primary key for this table.
- Foreign Keys (FK):** There is a `product_type_id` field in the `product` table which is presumably a FK to the `product_type` table. This is standard practice and there seems to be no issue here as long as the relationship is properly enforced in the database.

- Redundancy:** There is a `product_details` table with an `id` that seems to reference back to the `product` table. This table also includes `product_name`, `product_description`, and `price_per_unit`, which are already present in the `product` table. This is redundant and could lead to inconsistency unless it serves a specific purpose such as tracking historical changes.

- Data Types:** The data types for the `product` table seem appropriate for the fields defined.

Client Table Issues:

- Data Types:** The `company_name` is listed with a data type of `int`, which is not appropriate for storing textual data. This should likely be a `varchar` or similar string data type.
- VAT ID:** The `VAT_ID` field is of type `varchar(64)`, which may be larger than necessary depending on the format of VAT IDs used. However, this is not a major issue unless there are strict storage constraints.
- Foreign Key Relationships:** There is a `city_id` FK in the `client` table that should link to the `city` table. This is the standard design, but care must be taken to ensure that the FK constraint is properly defined in the actual SQL schema.
- Client Address:** The `client_address` field is of type `text`, which is suitable for long text fields. The choice between `text` and `varchar` would depend on the expected length of the addresses to be stored.

Overall, the primary concern in this data model is the redundancy in the `product_details` table and the incorrect data type for the `company_name` in the `client` table. The rest of the model seems standard, assuming that the FK relationships are correctly implemented in the database.



Data Model Analysis

AG You



Model_Sample_create.sql

File

Analyze the following SQL Script which generated a Data Model. Focus on the product and client tables. Assess the model and identify any potential issues with relationships, keys, data types or table designs.

ChatGPT

Based on the SQL script segments provided for the `product` and `client` tables, here are the potential issues identified:

Product Table Issues:

- The `product_pk` PRIMARY KEY constraint is incorrectly set on the `product_name` column. This is not a good practice because product names can change and are not guaranteed to be unique. It is better to set the PRIMARY KEY on the `id` column, which is already set up to auto-increment.
- The `product_type_id` column is present, indicating a relationship to a `product_type` table, but the script does not include a `FOREIGN KEY` constraint to enforce this relationship. Without this, there is no referential integrity to ensure that `product_type_id` values in the `product` table correspond to valid entries in the `product_type` table.

Client Table Issues:

- The `company_name` field in the `client` table is of type `int`, which is incorrect for storing a company's name. This should be a string data type such as `varchar`.
- There is no visible `FOREIGN KEY` constraint linking `city_id` in the `client` table to the `city` table. While the column is present, a foreign key relationship should be defined to maintain referential integrity.
- The `client_address` field is of type `text`, which is suitable for storing long-form text. If the expected length of addresses is not very large, a `varchar` with a specified length might be more appropriate and efficient.

To rectify these issues, the primary key for the `product` table should be changed to the `id` column, and appropriate `FOREIGN KEY` constraints should be added to both tables to enforce the relationships with other entities in the database. Additionally, the data type for `company_name` should be corrected to a string data type. [→]

Data Quality

A	B	C	D	E	F	
1	Zip	City	State	Region	District	Country
2	22654	Star Tanne VA		East	District #0 USA	
3	22655	Stephens C VA		East	District #0 USA	
4	22656	Stephenso VA		East	District #0 USA	
5	22657	Strasburg, VA		East	District #0 USA	
6	22660	Toms Broc VA		East	District #0 USA	
7	22663	White Posi VA		East	District #0 USA	
8	22664	Woodstoc VA		East	District #0 USA	
9	22701	Culpeper, VA		East	District #0 USA	
10	22709	Aroda, VA, VA		East	District #0 USA	
11	22711	Banco, VA, VA		East	District #0 USA	
12	22712	Bealeton, VA		East	District #0 USA	
13	22713	Boston, VA VA		East	District #0 USA	
14		Brandy Sta VA		East	District #0 USA	
15		Brightwoo VA		East	District #0 USA	
16	22716	Castleton, VA		East	District #0 USA	

A	B	C	D	E	F	G	H	
1	ProductID	Product	Category	Segment	Manufact	Manufact	Unit Cost	Unit Price
2	392	Maximus F Rural	Productiv	7	VanArsdel	37,27106	51,05625	
3	393	Maximus F Rural	Productiv	7	VanArsdel	37,27106	51,05625	
4	394	Maximus F Rural	Select	7	VanArsdel	119,7618	164,0573	
5	396	Maximus L Accessory	Accessory	7	VanArsdel	66,28309	90,79875	
6	397	Maximus L Accessory	Accessory	7	VanArsdel	109,2224	149,6198	
7	398	Maximus L Accessory	Accessory	7	VanArsdel	114,9712	157,4948	
8	400	Maximus L Accessory	Accessory	7	VanArsdel	61,68409	84,49875	
9	402	Maximus L Accessory	Accessory	7	VanArsdel	xxx	102,3698	
10	403	Maximus L Accessory	Accessory	7	VanArsdel	83,11543	113,8568	
11	404	Maximus L Accessory	Accessory	7	VanArsdel	136,0499	186,3698	
12	405	Maximus L Accessory		7	VanArsdel	139,8824	191,6198	
13	406	Maximus L Accessory	Accessory	7	VanArsdel	139,8824	191,6198	
14	407	Maximus L Accessory	Accessory	7	VanArsdel	116,8874	160,1198	
15	408	Maximus L Urban	Moderatio	7	VanArsdel	99,64117	136,4948	
16	409	Maximus L Urban	Moderatio	7	VanArsdel	70,55633	96,6525	
17	410	Maximus L Urban	Moderatio	7	VanArsdel	xxx	147,4988	
18	412	Maximus L Urban	Moderatio	7	VanArsdel	118,8037	162,7448	

AG You



VanArsdel_Actuals - Cópia.xlsx
Spreadsheet

Analyze the Excel file to identify any data quality issues such as duplicates, outliers, or invalid entries. Focus on the Product and Geo Sheets

ChatGPT

- ! Error analyzing ▾
- ! Error analyzing ▾

It seems there was an issue reading the Excel file. This can sometimes be caused by various factors, such as the file being in a format not recognized by the reading function or the file being corrupted.

To proceed, I will attempt a different method to read the Excel file and check if we can access the 'Product' and 'Geo' sheets. Let's try that now. [→]

- ! Error analyzing ▾

It appears that there's a persistent issue with accessing the contents of the Excel file. The exceptions indicate that the file may not be readable using the standard methods available in this environment. This could be due to a variety of reasons such as file corruption, an unsupported file format, or restrictions within the current environment.

Data Quality



You



Geo.csv
Spreadsheet



Product.csv
Spreadsheet

Analyze the Product and Geo CSV Files file to identify any data quality issues such as duplicates, outliers, or invalid entries.



ChatGPT



I'm experiencing continued difficulties in accessing and analyzing the provided CSV files for the 'Product' and 'Geo' datasets within this environment. It seems there is an issue preventing the files from being read correctly, which might be related to the environment's limitations or the files themselves.

Data Quality



ChatGPT

The 'Product' and 'Geo' sheets from the Excel file have been successfully loaded. The 'Product' sheet contains 212 rows and 8 columns, and the 'Geo' sheet contains 39,948 rows and 6 columns.

Next, I will identify any data quality issues such as duplicates, outliers, or invalid entries within these two sheets. Let's start with the 'Product' sheet and then proceed to the 'Geo' sheet. [\[–\]](#)

The analysis of the 'Product' and 'Geo' sheets for data quality issues has revealed the following:

Product Sheet:

- **Outliers:** There are 10 outliers in the 'Unit Price' column. These are numerical entries that fall far outside the interquartile range and might be errors or simply extreme values.
- **Invalid Entries:** There is 1 invalid entry in the 'Segment' column, which could be an empty string or a NaN value.

Geo Sheet:

- **Invalid Entries:** There are 2 invalid entries in the 'Zip' column, which again could be empty strings or NaN values.

No duplicates were found in either of the sheets, which is positive in terms of data uniqueness.

Next steps should involve:

- Further investigation of the outliers in the 'Unit Price' column of the 'Product' sheet to determine if they are indeed errors or valid data points.
- Examination of the invalid entries in the 'Segment' column of the 'Product' sheet and the 'Zip' column of the 'Geo' sheet to correct or fill in the missing data. [\[–\]](#)



3) Report Design Assessment

Analyze report design

Rate report design

Provide recommendations



CURRENT CHALLENGE

Maven Coffee Challenge

Analyze real survey results to craft a data-driven strategy for opening a new coffee shop

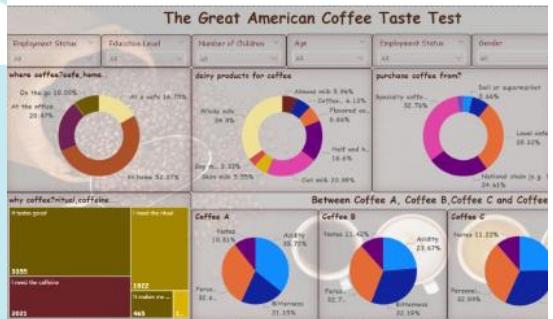
Feb 22, 2024 Start date | **Mar 18, 2024** Deadline | **6 days** Days remaining | **3 Annual Memberships + 1 Winner T-Shirt** Grand prize

Submit Your Project

Nidhi Gupta • 2º
Data Engineer || Azure Certified || Databricks Associate Cer...
2 d •

Maven Analytics Maven Analytics Group
Here is my dashboard for the Maven Coffee Challenge.
Great American Coffee Taste Test

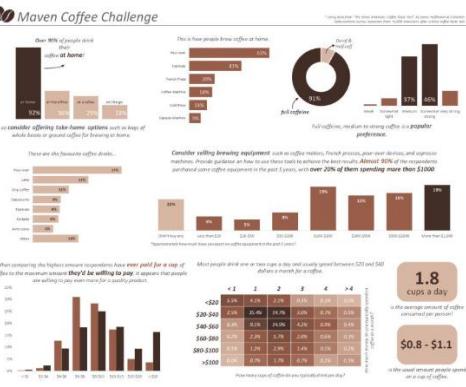
[Visualizar tradução](#)



Rade Marinkovic • 3º e +
13 h •

Just wrapped up my analysis of the [Maven Analytics #mavencoffeechallenge!](#) 😊 Kept it simple yet insightful. Such a fun dataset to dive into. 🎉 #DataAnalytics #CoffeeData

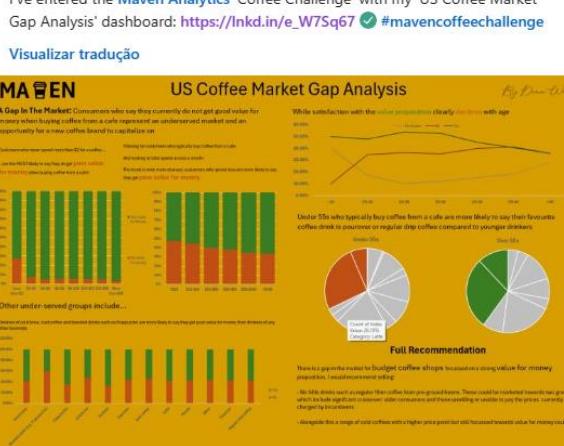
[Visualizar tradução](#)



dean walsh • 3º e +
2 d •

Exploring the possibility of a career in data analytics

[Visualizar tradução](#)



Great American Coffee Taste Test

Survey responses from ~4,000 Americans after a blind coffee taste test.

[Download](#)

CSV
File type | **Single table**
Data structure | **4043** # of records | **111** # of fields

Szimonetta Farkas • 2º
Data Analyst | Business Analyst | Data Scientist | Open to Jo...
17 h •

Dear LinkedIn Family,

I am delighted to announce that I have participated in the ...ver

[Visualizar tradução](#)



in [#mavencoffeechallenge](#)

AG

You



GACTT_RESULTS_ANONYMIZED_v2.csv
Spreadsheet



QUESTION_KEY.csv
Spreadsheet

Act as a Juri for a Power BI Analytics and Data Visualization Contest.

The rules for the contest are here: <https://mavenanalytics.io/challenges/maven-coffee-challenge/30>

I will ask you to evaluate a set of submissions on a 1 to 10 rating.

Participants will submit a 1 page screenshot of their solution.

- 1) Analyze the attached contest dataset (data and key for each question's answer)
- 2) Prepare a framework to classify each of the submissions

Return in table format the proposed framework to classify the solutions.

I will then submit screenshots of each contestant submission.



ChatGPT

To evaluate the submissions effectively, we'll need a comprehensive framework that considers several critical aspects of analytics and data visualization. Given that we are evaluating submissions for a Power BI Analytics and Data Visualization Contest, our framework will focus on criteria such as data accuracy, clarity, insightfulness, design and aesthetics, and overall effectiveness in communication.

Let's start by analyzing the datasets provided to understand what kind of data the contestants are working with and what key insights or answers they are expected to derive. This will help in shaping the criteria further.

I'll first take a look at the content of both the contest dataset and the question key to better inform our evaluation framework.

Based on the overview of the datasets:

- **Contest Dataset:** It contains anonymized submissions with a wide range of questions related to coffee consumption habits, including age groups, cups of coffee consumed per day, preferred locations for drinking coffee, and more. The dataset is quite comprehensive, with 111 columns, indicating a rich set of data points for analysis.
- **Question Key:** This outlines the order and type of questions asked in the survey, including whether they're single select, multiple selection, or text-based. It also details the survey sections and answer choices, providing context for what types of insights the contestants might derive.

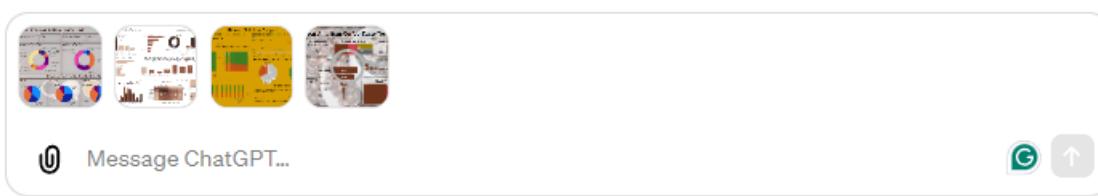
Given this information, here is a proposed framework to classify each of the submissions on a scale of 1 to 10. The framework will consist of five main criteria:

- Data Accuracy (20%):** Assessing how accurately the submission reflects the dataset's information. This includes correct representation of data points and attention to the details provided in the question key.
- Clarity (20%):** The ease with which viewers can understand the visualizations. This involves clear labeling, appropriate use of charts, and avoiding clutter.
- Insightfulness (20%):** The ability to go beyond basic visualization to extract meaningful insights that are not immediately obvious. This includes identifying trends, patterns, or correlations.
- Design and Aesthetics (20%):** The visual appeal of the submission, including the use of colors, spacing, and overall design principles that enhance readability and viewer engagement.
- Communication Effectiveness (20%):** How well the submission communicates its findings to the audience. This includes storytelling elements, logical flow, and the use of annotations or explanations to guide the viewer.

Each criterion will be rated on a scale of 1 to 10, and the final score will be a weighted average of these five scores. Here's a table format for the evaluation framework:

Criterion	Weight	Description
Data Accuracy	20%	Accuracy in reflecting dataset information and details from the question key.
Clarity	20%	Clear and understandable visualizations with appropriate chart selection and labeling.
Insightfulness	20%	Ability to derive and present meaningful insights beyond basic visualization.
Design and Aesthetics	20%	Visual appeal, including color use, spacing, and overall design principles.
Communication Effectiveness	20%	Effectiveness in storytelling, logical presentation flow, and use of annotations/explanations.

With this framework, we can evaluate the submissions objectively, focusing on key aspects that contribute to high-quality analytics and data visualization. You can now submit screenshots of each contestant's submission for evaluation. [↓](#)



The screenshot shows a messaging interface with a light blue header. At the top left are three small circular icons representing different data types. Below them is a text input field with the placeholder "Message ChatGPT...". To the right of the input field are two small circular buttons with "G" and "U" respectively.

ChatGPT

Let's evaluate each submission using the established framework. I'll score each criterion on a scale from 1 to 10 and then provide the final weighted score for each submission. Here are the evaluations:

Submission 1:

Criterion	Score (1-10)	Comments
Data Accuracy	7	Data appears to be represented accurately, with corresponding percentages reflecting the dataset.
Clarity	6	While the donut charts are common in data representation, the presence of many on one page without clear segmentation can be confusing. The treemap chart's color coding is not immediately intuitive.
Insightfulness	6	There is an attempt to segment data by various demographic factors, which can provide insights into consumer behavior. However, insights are not explicitly stated, leaving interpretation solely to the viewer.
Design and Aesthetics	5	The design is somewhat cluttered with different chart types and colors that do not seem to have a consistent theme.
Communication Effectiveness	5	The visualizations provide information, but there is no clear narrative or takeaway communicated to the viewer.
Final Score	5.8	Weighted Average

Submission 1:

Criterion	Score (1-10)	Comments
Data Accuracy	7	Data appears to be represented accurately, with corresponding percentages reflecting the dataset.
Clarity	6	While the donut charts are common in data representation, the presence of many on one page without clear segmentation can be confusing. The treemap chart's color coding is not immediately intuitive.
Insightfulness	6	There is an attempt to segment data by various demographic factors, which can provide insights into consumer behavior. However, insights are not explicitly stated, leaving interpretation solely to the viewer.
Design and Aesthetics	5	The design is somewhat cluttered with different chart types and colors that do not seem to have a consistent theme.
Communication Effectiveness	5	The visualizations provide information, but there is no clear narrative or takeaway communicated to the viewer.
Final Score	5.8	Weighted Average

Submission 2:

Criterion	Score (1-10)	Comments
Data Accuracy	9	The data is represented in a clean and straightforward manner, corresponding well with the dataset.
Clarity	9	The layout is clear, with a good flow that makes it easy to follow. Annotations and labels are helpful in guiding the viewer through the data.
Insightfulness	8	The submission does a good job of identifying and communicating key insights, such as the popularity of home brewing and spending habits.
Design and Aesthetics	8	The design is clean and professional, with a consistent color scheme and well-organized layout.
Communication Effectiveness	9	The submission tells a clear story with actionable insights and recommendations, making it highly effective in communication.
Final Score	8.6	Weighted Average

Submission 3:

Criterion	Score (1-10)	Comments
Data Accuracy	8	The data visualizations seem accurate and well-corresponded to the dataset.
Clarity	7	The visuals are mostly clear, but some of the charts could benefit from better labeling or legends to clarify the data points.
Insightfulness	7	The analysis identifies market gaps and under-served customer segments, showing a good level of insight.
Design and Aesthetics	7	The color choices are consistent, and the layout is generally clean, though it might seem a bit busy with multiple bar charts.
Communication Effectiveness	8	The narrative is mostly clear with practical recommendations provided, though the flow could be improved to guide the viewer more effectively.
Final Score	7.4	Weighted Average

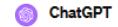
Submission 4:

Criterion	Score (1-10)	Comments
Data Accuracy	7	The submission shows data accurately through various visualizations. However, the readability of some charts is impacted by the choice of colors and the size of text.
Clarity	6	The layout is somewhat disjointed, and the visual flow between sections is not as smooth. The visualizations are standard but require effort to correlate due to spacing.
Insightfulness	6	Provides basic data representation with some insights into preferences but lacks deeper analytical conclusions.
Design and Aesthetics	6	The design lacks a cohesive theme and the coffee bean background may distract from the data.
Communication Effectiveness	5	The communication of insights is not as strong as it could be. There is data presented, but less in the way of narrative or conclusions.
Final Score	6.0	Weighted Average



You

Suggest improvements for submission 1



ChatGPT

Submission 1 can be improved in several key areas to better align with the principles of effective data visualization and communication:

1. Simplify and Focus:

- Reduce the number of donut charts or consider consolidating some data points into a single, multi-faceted visualization.
- Highlight key data by focusing on fewer, more impactful metrics.

2. Improve Clarity:

- Include clear and concise titles for each chart to immediately inform the viewer of the subject matter.
- Adjust the color palette to be more harmonious and less visually jarring, using colors to differentiate data points without overwhelming the viewer.
- Simplify the treemap chart and use labels or tooltips to explain what each rectangle represents.

3. Enhance Insightfulness:

- Explicitly state the key insights next to each visualization. This could be a simple sentence summarizing the finding or a bullet point highlighting the takeaway.
- Use data to tell a story by guiding the viewer through a logical flow from problem to solution or question to answer.

4. Refine Design and Aesthetics:

- Adopt a consistent design theme throughout the dashboard to create a visual identity that can be easily recognized.
- Use whitespace effectively to give each chart room to breathe, reducing clutter and focusing attention on the data.
- Ensure that text is legible, with sufficient contrast against the background and an appropriate size for viewing.

5. Improve Communication Effectiveness:

- Structure the presentation to lead with the most important information or insights, followed by supporting data.
- Incorporate interactive elements if possible, allowing users to explore the data themselves and gain a personal understanding of the insights.
- Provide a concise summary or conclusion section that recaps the findings and suggests potential actions or implications.



You

Be more specific:

- suggest alternative visuals for each donut.
- suggest which metrics from this dataset are relevant to display
- suggest how to format the clear and concise titles in terms of font and font size and in which element to display it in power bi
- suggest an existing theme for this project
- make other specific suggestions

ChatGPT

Absolutely, let's delve into specifics for enhancing Submission 1:

Alternatives to Donut Charts:

1. "**Where coffee? cafe, home...**": Replace the donut chart with a stacked bar chart or a column chart. This allows for easier comparison of categories and can accommodate more data points without becoming overwhelming.
2. "**Dairy products for coffee**" and "**Sugar or sweetener?**": Consider using a bar chart, which can facilitate quicker comparisons between categories. For added clarity, sort the bars by size to immediately see the most and least popular options.
3. "**Between Coffee A, Coffee B, Coffee C, and Coffee D Reviews**": Instead of four pie charts, use a multi-series bar chart comparing each coffee type across the different attributes (notes, acidity, bitterness). This could improve the ability to compare across coffee types.

Metrics to Display:

The metrics chosen should reflect the key insights you wish to communicate. Consider the following:

- **Demographic Breakdown:** Show the age range, gender, or employment status of respondents to segment the coffee preferences.
- **Consumption Habits:** Visualize the number of cups consumed daily and correlate it with preferred coffee types or spending habits.
- **Purchase Locations:** Compare home brewing to café purchases, possibly hinting at market opportunities for home brewing products.

Formatting Titles in Power BI:

For titles:

- **Font:** Choose a sans-serif font like Segoe UI or Arial for readability.
- **Font Size:** Titles should stand out, so use at least a 14pt font for titles and 12pt for subtitles.
- **Display:** Place titles within the header of each visual in Power BI, ensuring they're aligned (left alignment is standard) and have enough space from the top edge of the visual.

Theme Suggestion:

Power BI has a variety of built-in themes, but for a coffee-centric dataset, consider a custom theme with a warm and inviting color palette. For instance:

- **Primary Color:** A coffee brown (#6f4e37) to evoke the subject matter.
- **Accent Colors:** Cream (#ffffdd0), Dark Roast (#403b33), and Latte Foam (#f3e9e5) for diversity and readability.
- **Background:** A light cream or white background (#ffffdd0 or #ffffff) to ensure contrast and readability.

Other Specific Suggestions:

- **Interactive Elements:** Use slicers for interactive filtering by demographics or preferences.
- **Legibility:** Ensure all text is clearly legible against its background; avoid placing text over busy images.
- **Annotation:** Use text boxes to annotate key findings or to add brief commentary on what the viewer should note.
- **Consistency:** Maintain consistent margins and alignment throughout the dashboard for a tidy appearance.
- **Tooltips:** Customize tooltips to provide additional context or data not immediately visible in the primary visual.
- **Card Visuals:** Use card visuals to display key numbers or metrics at a glance, such as "Average Cups of Coffee Per Day."
- **Iconography:** Integrate coffee-related icons to establish a thematic connection and visual interest.

Remember, the goal is to tell a story with the data, so ensure that each visual contributes meaningfully to that narrative.



4) **Troublesh ooting**

Troubleshooting Problem

Understanding errors

Performance Optimization



Forums

Join the discussion

Options ▾

Get Help with Power BI

Type	Title	Posts	New
Desktop	Data shaping, modeling, and report building in the Power BI Desktop app. Latest Topic - MoM% Showing the Infinity Value	1175020	1174547
Service	Power BI Premium, Dataflows, dashboards, reports, sharing, and everything else you do at app.powerbi.com. Latest Topic - Updating App on Service	211708	211307
Report Server	Report viewing, installation, configuration, and everything else in Power BI Report Server (on-premises). Latest Topic - Google Sheets Connector doesnt refresh when publis...	29253	29229
Power Query	Power Query is the Data Connectivity and Preparation experience across multiple Microsoft products, including Power BI, Excel, Microsoft Dataverse and more. Latest Topic - Date.fromtext passing too early of a date to datav...	116863	116809

AG You

Act as a Power BI Expert, in charge of community support on the Forum.

I will provide you with requests and you will provide answers for the questions, and the links to the relevant official documentation.

ChatGPT

Of course! I'm ready to assist with any Power BI questions you have. Please go ahead with your first request.

AG You

MoM% Showing the Infinity Value

22m ago

Hi BI Community Team,

I have a sales order table and looking for MTD, PMTD & MoM Growth % with measure below that I explore and apply.

MTD-Sales = TOTALMTD(SUM('Sales Order'[Amt]), 'CALENDAR'[Date])

PMTD-Sales = TOTALMTD(SUM('Sales Order'[Amt]), DATEADD('CALENDAR'[Date], -1, MONTH))

MoM-Sales = ((MTD-Sales) - (PMTD-Sales))/(PMTD-Sales)

But MoM-Sales, it's not showing correctly if we choose the slicer in Jan'24 as the screenshot below.

3,894,766.51	(Blank) 0.00
MTD-Sales	PMTD-Sales
Infinity	0%
MoM-Sales	

Any suggestion/advise?

ChatGPT

It seems that the MoM-Sales measure is resulting in an Infinity value, which typically occurs when you divide a number by zero. In this case, the problem arises when the `[PMTD-Sales]` measure returns a value of 0, which is then used as the denominator in your MoM-Sales calculation.

Here's a revised version of your MoM-Sales measure that includes a condition to check if `[PMTD-Sales]` is zero to avoid dividing by zero:

```
dax
Copy code

MoM-Sales =
VAR CurrentMonthSales = [MTD-Sales]
VAR PreviousMonthSales = [PMTD-Sales]
RETURN
IF(
    PreviousMonthSales = 0,
    BLANK(),
    (CurrentMonthSales - PreviousMonthSales) / PreviousMonthSales
)
```

With this measure, if the `[PMTD-Sales]` is zero, it will return BLANK instead of Infinity.

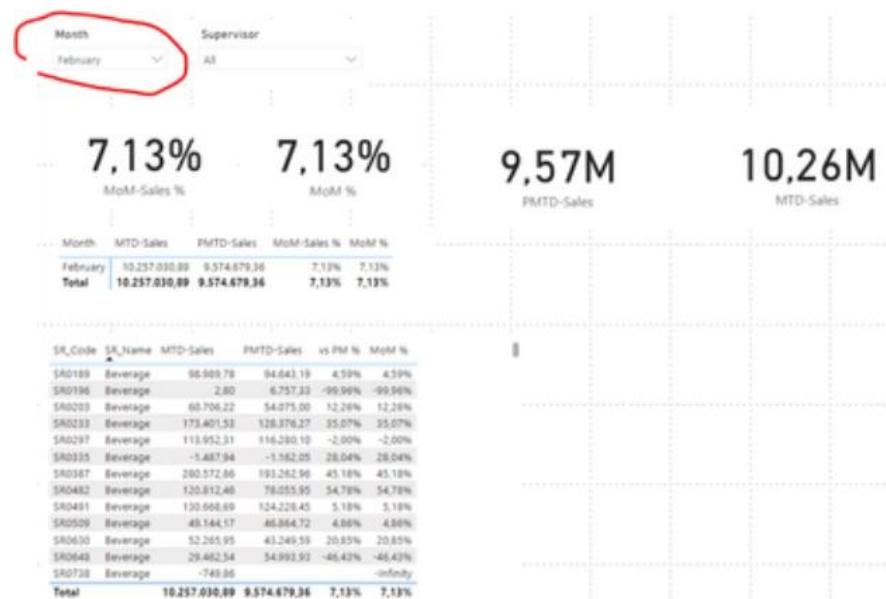
To understand more about how to avoid division by zero and other DAX formula pitfalls, you can refer to the official documentation on handling errors in DAX and using conditional statements:

- Handling errors in DAX formulas: [DAX error reference](#)
- Using IF function in DAX: [IF function](#)

The IF function in DAX is crucial for handling these types of conditions to ensure your reports are accurate and free from misleading calculat. ↓

You have everything right on this file you shared! It's just that the period you selected (January) doesn't have a previous month, hence the calculation for the previous month doesn't work, neither the variations as you don't have a previous month value.

Try another month and you will see it works just fine.



Count Number of Days in a Range of Days over Multiple Rows

32m ago

I want to count the number of days an object is in use. I want to count both the departure and return dates in the calculation. However, in some cases the object is returned on the same day it departs again.

I don't want to double count the days the object is in use in these situations. Is there a formula I'm able to use with DAX?

Below is a table with an example. I don't want to count March 15th twice

Departure Date	Return Date	Days in Use	Desired Result
March 11	March 12	2	2
March 13	March 15	3	3
March 15	March 18	4	3
Total		9	8

ChatGPT

In DAX, you can calculate the total number of unique days an object is in use across multiple rows without double-counting overlapping days with the following steps:

1. Create a calculated column to create a list of dates for each row.
2. Create a distinct list of all dates from the calculated column in step 1.
3. Count the number of dates in the distinct list.

Here's a sample formula for each step:

Step 1: Create a calculated column to enumerate each date between 'Departure Date' and 'Return Date'.

```
dax
Copy code

Datelist =
CALCULATETABLE (
    VALUES ( 'Calendar'[Date] ),
    'Calendar'[Date] >= EARLIER ( 'Table'[Departure Date] ),
    'Calendar'[Date] <= EARLIER ( 'Table'[Return Date] )
)
```

Step 2: Create a new table with distinct dates.

```
dax
Copy code

DistinctDates = DISTINCT ( UNION ( SELECTCOLUMNS ( 'Table', "Date", 'Table'[Datelist] ) ) )
```

Step 3: Create a measure to count the unique dates.

```
dax
Copy code

UniqueDaysInUse = COUNTROWS ( 'DistinctDates' )
```

Remember that 'SELECTCOLUMNS' and 'UNION' are used to combine and flatten the list of dates into a single column, and 'DISTINCT' is used to ensure that each date is only counted once.

You'll need a Calendar table that contains all possible dates that could be in your data. This is a common practice in Power BI to facilitate time intelligence calculations.

AG You

Can we use Visual calculations instead?

<https://learn.microsoft.com/en-us/power-bi/transform-model/desktop-visual-calculations-overview>

ChatGPT

Certainly, visual-level calculations in Power BI can be used to perform aggregations that are only displayed in a specific visual without the need to create new columns or measures that affect the entire model. However, the visual-level calculations are somewhat limited compared to DAX measures and won't allow the level of data manipulation required to count unique days across multiple rows with overlapping periods.

To accomplish the task of counting the number of unique days an object is in use without double-counting the overlapping days, you would typically need to create a measure using DAX, as visual-level calculations don't provide a direct way to create a distinct count of dates that span across multiple rows while considering overlaps.

That said, if you were to attempt to use visual-level calculations, you could perhaps try creating calculated columns that enumerate the days as described in the previous messages, then using a visual that has a distinct count or count (distinct) aggregation available. But this approach is still not directly available through visual-level calculations alone.

The measure I described earlier, using DAX, is the recommended approach in Power BI to calculate unique days across rows considering overlaps. Once you have created the measure, you can easily use it in any visual across your report.

If you have a specific visual in mind that you would like to use, I could help you integrate the measure into that visual for the desired outcome.

AG You

Departure	Return	Sum of DateDaysInUse
11-03-2024	12-03-2024	2
13-03-2024	15-03-2024	3
15-03-2024	18-03-2024	4
Total		9

based on visual calculations and on the following data in a power bi table visual, suggest a formula for a column which displays the intended result

ChatGPT

For visual calculations in a Power BI table visual, you would typically use Quick Measures or apply filters directly in the visual. Unfortunately, these don't support the complex type of calculation needed to account for overlapping dates directly in the visual. The calculation of non-overlapping days would generally require DAX measures that operate across the entire table.

However, as per your request to provide a suggestion within the context of a Power BI table visual, you might consider using a measure with DAX that counts the number of unique days, considering overlaps. Here's a DAX formula that can be used as a measure:

```
dax
Copy code

Distinct Days In Use =
VAR UniqueDates =
    SUMMARIZE (
        GENERATESERIES ( MIN ( 'Table'[Departure] ), MAX ( 'Table'[Return] ), 1 ),
        "Date", [Value]
    )
VAR DistinctCount =
    CALCULATE (
        DISTINCTCOUNT ( 'UniqueDates'[Date] ),
        FILTER (
            UniqueDates,
            NOT (
                [Date] IN VALUES ( 'Table'[Return] )
                && [Date] IN VALUES ( 'Table'[Departure] )
            )
        )
    )
RETURN
    DistinctCount
```

Count Number of Days in a Range of Days over Multiple Rows • Last saved: Today at 5:30 PM

File Home Insert Modeling View Optimize Help External tools Format Data / Drill

Visual table Data point table Show Interactions Apply drill down filters to Entire page Drill actions Drill through Groups Find anomalies AI

Departure Return Sum of DateDaysInUse Calculation

11-03-2024	12-03-2024	2	2,00
13-03-2024	15-03-2024	3	3,00
15-03-2024	18-03-2024	4	3,00
Total		9	8,00

X ✓ fx 1 Calculation = [Sum of DateDaysInUse] - if([Departure] = PREVIOUS([Return]),1,0)

Departure Return Sum of DateDaysInUse Calculation

11-03-2024	12-03-2024	2	2,00
13-03-2024	15-03-2024	3	3,00
15-03-2024	18-03-2024	4	3,00
Total		9	8,00

Visualizations Data

Build visual

Filters

Rentals

- ✓ Σ DateDaysInUse
- ✓ Departure
- ✓ Return

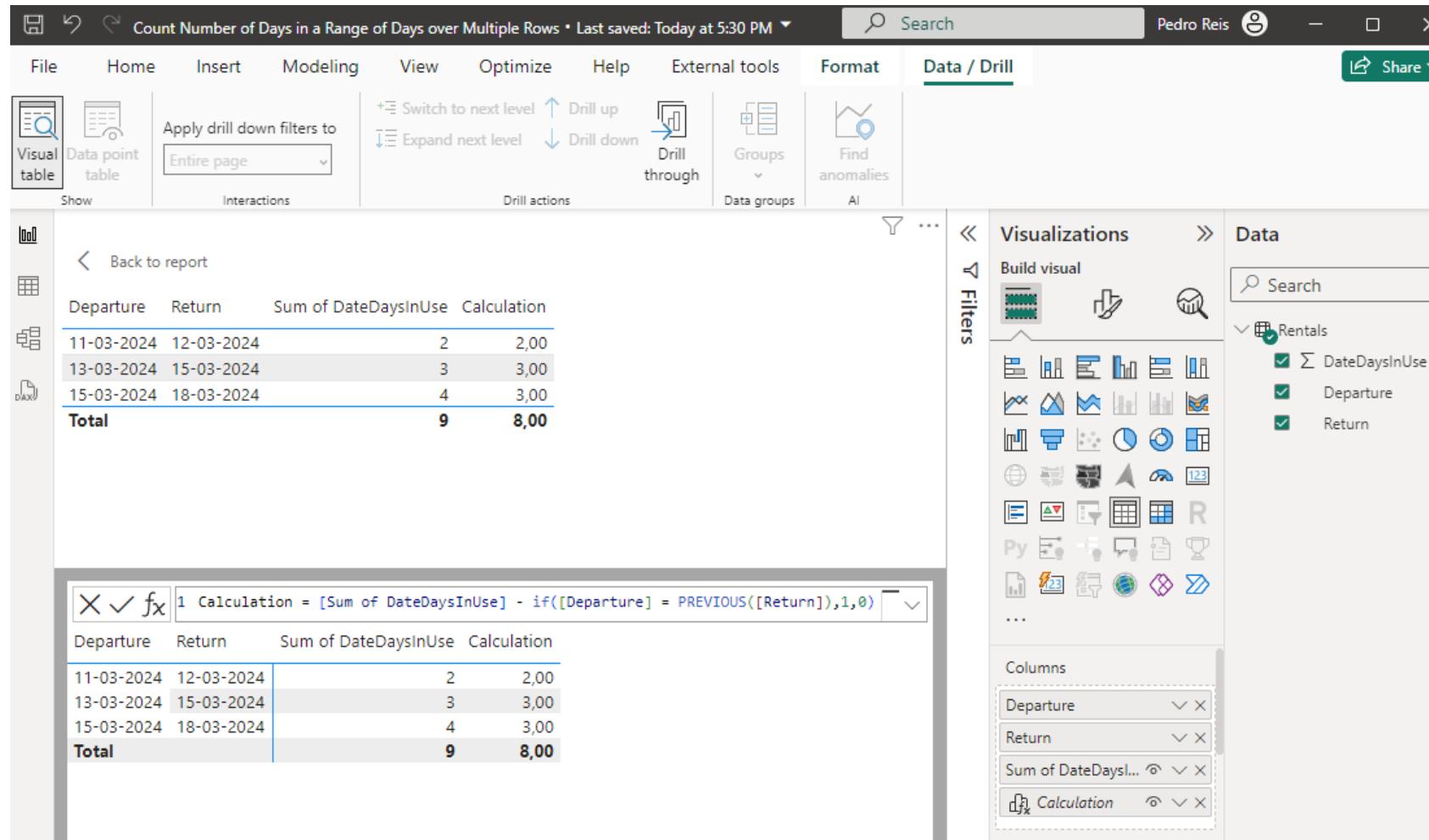
Columns

Departure

Return

Sum of DateDaysInUse

Calculation



Dataset refresh failure for Impala database

55m ago

I have been getting an error trying to refresh data from Impala on daily basis with underlying error message of ERROR [08S01]

```
{"error":{"code":"DM_GWPipeline_Gateway_MashupDataAccessError","pbi.error": {"code":"DM_GWPipeline_Gateway_MashupDataAccessError","parameters":{}, "details": [{"code":"DM_ErrorDetailNameCode_UnderlyingErrorCode","detail":{"type":1,"value":"-2147467259}}, {"code":"DM_ErrorDetailNameCode_UnderlyingErrorMessage","detail":{"type":1,"value":"<ccon>ODBC: ERROR [HY000] [Cloudera][ImpalaODBC] (110) Error while executing a query in Impala: [HY000] : Runtime Error: Admission for query exceeded timeout 900000ms in pool root.svcbigateway. Queued reason: Not enough aggregate memory available in pool root.svcbigateway with max mem resources 1024.00 GB (configured statically). Needed 232.00 GB but only 68.00 GB was available.\n</ccon>"}}, {"code":"DM_ErrorDetailNameCode_UnderlyingHResult","detail":{"type":1,"value":"-2147467259}}, {"code":"Microsoft.Data.Mashup.ValueError.DataSourceKind","detail":{"type":1,"value":"Odbc"}}, {"code":"Microsoft.Data.Mashup.ValueError.DataSourcePath","detail":{"type":1,"value":"<ccon>dsn=Impala PROD</ccon>"}}, {"code":"Microsoft.Data.Mashup.ValueError.OdbcErrors","detail":{"type":1,"value":"#table({\"SQLState\", \"NativeError\", \"Message\"}, {})"}, {"code":"Microsoft.Data.Mashup.ValueError.Reason","detail": {"type":1,"value":"DataSource.Error"}}, {"code":"Microsoft.Data.Mashup.ValueError.ExceptionCulprit","detail": {"type":1,"value":1}}]} Table: RSA Bill Payments.
```

Is there anyone received the same error and how do you fix it?

ChatGPT

The error message you've encountered indicates that the Impala ODBC connection is timing out because the query is requesting more memory than is available in the pool. Here's a breakdown of the key parts of the error message and what they signify:

- Error [HY000] [Cloudera][ImpalaODBC] (110):** This is the error from the ODBC driver indicating an issue with the query execution.
- Runtime Error: Admission for query exceeded timeout 900000ms in pool root.svcbigateway:** The query has a timeout setting (900000ms or 15 minutes), after which it will fail if not executed.
- Not enough aggregate memory available in pool root.svcbigateway with max mem resources 1024.00 GB (configured statically). Needed 232.00 GB but only 68.00 GB was available:** The memory pool allocated to the query execution service (root.svcbigateway) doesn't have enough memory to perform the operation. It's configured to have a maximum of 1024 GB, but at the time of the query, it required 232 GB, and only 68 GB was available.

To resolve this issue, here are several steps you can take:

- Optimize the Query:** Ensure that the query you're running is optimized. Avoid pulling unnecessary data. If possible, filter the data before it's returned by the query.
- Increase Memory Allocation:** If possible, increase the memory allocation for the Impala pool (root.svcbigateway) to ensure there's enough memory available for the query.
- Adjust Timeout Settings:** Increase the timeout setting for the ODBC connection if the query is expected to take longer than the current timeout.
- Check Cluster Load:** If you're running this query when the cluster is under heavy load, try scheduling the refresh during off-peak hours.
- Break Down the Query:** If the query is too large, break it down into smaller queries to reduce the memory footprint.
- Contact Support:** If you've tried the above and still encounter issues, consider contacting your Impala cluster administrator or support for assistance with resource allocation.

Additionally, you should check the Power E ↴ Impala documentation for best practices on managing data refreshes and handling large queries:



5) Knowledge

Knowledge Assessment

Knowledge Improvement



Knowledge Assessment



AG

You

I'm learning about Power BI Model types. Create a 3 question assessment to test my knowledge



ChatGPT

Certainly! Here's a 3-question assessment to test your knowledge on Power BI model types. These questions cover fundamental concepts and distinctions between different types of models in Power BI.

Question 1: Data Connectivity Modes

Which of the following statements best describes the difference between DirectQuery and Import modes in Power BI?

- A. DirectQuery mode stores a copy of the data in the Power BI report, while Import mode keeps the data in the source system and queries it directly.
- B. DirectQuery mode keeps the data in the source system and queries it directly, while Import mode stores a copy of the data in the Power BI report.
- C. Both DirectQuery and Import modes store data in the Power BI report, but DirectQuery allows for real-time updates while Import does not.
- D. Both DirectQuery and Import modes keep the data in the source system, but DirectQuery allows for more complex transformations within Power BI.



Question 2: Composite Models

As of my last update in April 2023, which statement correctly describes a capability of Composite Models in Power BI?

- A. Composite Models do not allow combining DirectQuery and Import data sources within the same report.
- B. Composite Models allow the use of DirectQuery for some data sources and Import mode for others within the same report, but measures cannot be mixed between the two.
- C. Composite Models enable the combination of DirectQuery and Import data sources within the same report, including the ability to create relationships and measures across these different data sources.
- D. Composite Models are exclusively available in Power BI Premium and cannot be used in Power BI Desktop or Pro versions.

Question 3: Aggregation Tables

What is the primary purpose of using an Aggregation table in a Power BI model?

- A. To duplicate the data from the source tables for backup purposes.
- B. To reduce the model size by only importing summarized data from the source system.
- C. To improve report performance by providing a summarized layer of data that can satisfy queries without needing to access detailed data, thus reducing the load on the source system and speeding up query response times.
- D. Aggregation tables are used to increase data accuracy by cross-verifying data points between different tables.

Please provide your answers, and I'll let you know how you did!

Knowledge Assessment



AG You
B, C, C

Question 1: Data Connectivity Modes

Your answer: **B**. DirectQuery mode keeps the data in the source system and queries it directly, while Import mode stores a copy of the data in the Power BI report.

- **Correct!** This is the right distinction between DirectQuery and Import modes. Import mode brings data into Power BI, storing it within the report, allowing for faster queries but requiring refreshes to update the data. DirectQuery does not store data in the report; instead, it queries the source system live, which ensures data is always up-to-date but might be slower depending on the source system's performance.

Question 2: Composite Models

Your answer: **C**. Composite Models enable the combination of DirectQuery and Import data sources within the same report, including the ability to create relationships and measures across these different data sources.

- **Correct!** This describes the capability of Composite Models accurately. They allow for a mix of DirectQuery and Import modes, leveraging the strengths of both. This feature enables Power BI developers to optimize their models for performance while maintaining the flexibility of accessing real-time data as needed.

Question 3: Aggregation Tables

Your answer: **C**. To improve report performance by providing a summarized layer of data that can satisfy queries without needing to access detailed data, thus reducing the load on the source system and speeding up query response times.

- **Correct!** Aggregation tables are indeed used to enhance performance in Power BI models. By summarizing detailed data, aggregation tables can significantly reduce the time it takes to retrieve and process data for visualizations, especially in scenarios where users are interacting with large volumes of data. This approach not only speeds up report loading times but also reduces the load on the source system, which can be crucial for maintaining system responsiveness and efficiency.

You answered all questions correctly, demonstrating a good understanding of Power BI model types and their functionalities. Great work! If you have any more questions or need further clarification on Power BI or related topics, feel free to ask.

Knowledge Improvement



AG You

Create 3 open-ended questions to test the knowledge of Sharing Content on Power BI Service.
Return this assessment in Word Format.

I will return the file or files from multiple respondents and ask you to rate answers on a scale of 1 to 10

ChatGPT

The assessment document with open-ended questions on Sharing Content on Power BI Service is ready. You can download it using the link below:

[Download Power BI Content Sharing Assessment](#)

Please share this document with the respondents. Once you receive their answers, you can return here to have the responses rated on a scale of 1 to 10. [→]

Power BI Service Content Sharing Knowledge Assessment

Please answer the following open-ended questions to assess your knowledge on sharing content within the Power BI Service. Provide detailed responses based on your experience and understanding.

1. Q1. Explain the process and necessary permissions for sharing a Power BI report with a colleague who does not have a Power BI account.
2. Q2. Describe how to use Power BI Apps to distribute content to a large audience within your organization. Include details on how apps are created, published, and accessed by users.
3. Q3. Discuss the differences between sharing content through Dashboard sharing, App sharing, and Workspace sharing in Power BI Service. Include considerations for when each method is most appropriately used.

Tomorrow another session

Today's 20 Session

- 1) DAX w/ SVG and CSS
- 2) Data Modeling & Analysis
- 3) Report Design Assessment
- 4) Troubleshooting
- 5) Knowledge Assessment and Improvement

Tomorrow's 40m Deep Dive

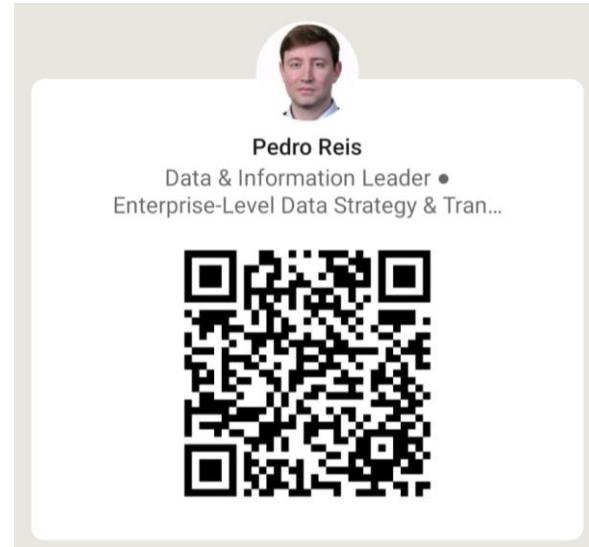
- 1) Generate a Data Model
- 2) Generate Artificial Data
- 3) Generate Power Query code
- 4) Generate DAX Measures
- 5) Enhance Report Theme
- 6) Publication & Administration w/ Cmdlets
- 7) Key Takeaways / The Future

Feedback



<https://sqlb.it/?12520>

Let's connect!



✉️ pedroreis@reispm.onmicrosoft.com