



Relatório do Trabalho 3 de Data Warehouse

“Desenvolvimento de um Cubo com SSAS”

Realizado por:

Catarina Rodrigues nº201501626;

Eduardo Palma nº201900054;

Nuno Melo nº201700465;

Ricardo Santos nº201700524;

Docente: Norberto Albino

Unidade Curricular: Data Warehouse

Curso: Licenciatura em Bioinformática, 2º ano

Ano Letivo: 2020/2021

Instituição: Escola Superior de Tecnologia do Barreiro

Índice

1.	Introdução	1
1.1	O que é o SSAS (SQL Server Analysis Services)?	1
1.2	Cubo no SSAS	1
2.	Processo Utilizado	2
2.1	Data Source	2
2.2	Data Source View.....	3
2.3	Cubo.....	6
2.4	Calculated Measure	8
2.5	Named Calculation	9
2.6	Hierarquias em Dimensões	10
2.7	Análise dos Dados em Excel.....	11
3.	Conclusão	15
4.	Bibliografia	16

Índice de Figuras

Figura 1 – Modelo do Cubo SSAS	2
Figura 2 – Escolher a base de dados	3
Figura 3 – Escolher a base de dados	3
Figura 4 – Escolha das tabelas	4
Figura 5 – Escolha de tabelas	5
Figura 6 – Visualização da Data Source View	5
Figura 7 – Escolha de tabelas	6
Figura 8 – Escolha das measures	7
Figura 9 – Escolha das dimensões	7
Figura 10 – Visualização do Cubo	8
Figura 11 – Calculated Measure do Total_Venda	9
Figura 12 – Named Calculation do NomeCompleto	9
Figura 13 – Named Calculation do Semestre	10
Figura 14 – Hierarquia do DimProduct	10
Figura 15 – Relações na Hierarquia DimProduct	10
Figura 16 – Hierarquia da DimDate	11
Figura 17 – Relações na Hierarquia DimDate	11
Figura 18 – Ficheiro de Excel vazio	11
Figura 19 – Adicionar Sales Amount	12
Figura 20 – Adicionar Order Quantity	12
Figura 21 – Adicionar Calculation Measures	12
Figura 22 – Adicionar Hierarquia do DimData	13
Figura 23 – Adicionar o nome da categoria, subcategoria e produto	13
Figura 24 – Análise dos Dados	14
Figura 25 - Análise dos Dados	14

1. Introdução

Com este trabalho, pretende-se desenvolver uma folha de Excel para consultar e analisar as vendas efetuadas por Internet e por Revendedores, em quantidade e valor, organizada por produto e categoria, análise efetuada para os anos existentes na base de dados.

Para conseguirmos perceber melhor a teoria que dá corpo a este trabalho vamos começar por escrever uma breve definição de cada ponto que mostra ser relevante para a construção deste projeto.

1.1 O que é o SSAS (SQL Server Analysis Services)?

De um modo geral, o Analysis Services é um conjunto de serviços para a gestão num data warehouse.

Analysis Services é um mecanismo de dados analíticos usado no suporte à decisão e análise de negócios. Fornece recursos de modelo de dados semânticos de nível empresarial para business intelligence (BI), análise de dados e aplicativos de relatório, como Power BI, Excel, Reporting Services e outras ferramentas de visualização de dados.

O SSAS é uma ferramenta de processamento analítico online (OLAP), e de mineração de dados no Microsoft SQL Server. É usado como uma ferramenta pelas organizações para analisar e compreender informações possivelmente distribuídas em várias bases de dados, ou em tabelas/arquivos diferentes.

O Analysis Services inclui um grupo de recursos de OLAP e mineração de dados e vem em dois tipos - Multidimensional e Tabular.

1.2 Cubo no SSAS

Um CUBO no SSAS é um método para armazenamento de dados num formato multidimensional ou hipercubo, geralmente para fins de relatórios. O processamento analítico online (os cubos OLAP) são uma funcionalidade no Service Manager que utiliza a infraestrutura de armazém de dados existente para fornecer capacidades de inteligência empresarial de autosserviço aos utilizadores finais. Um cubo OLAP é uma estrutura de dados que elimina as

limitações de bases de dados relacionais, fornecendo uma análise rápida de dados. Os cubos podem apresentar e somar grandes quantidades de dados enquanto também fornecem aos utilizadores o acesso pesquisável a quaisquer pontos de dados.



Figura 1 – Modelo do Cubo SSAS

2. Processo Utilizado

2.1 Data Source

Aqui, começou-se por abrir a base de dados que vamos utilizar no projeto, sendo essa base de dados a AdventureWorksDW2017, descarregada através do site da Microsoft, disponível no servidor.

Para começar, selecionou-se a opção “New Data Source” para adicionar a nossa source, ou seja, a base de dados que vamos usar.

De seguida, escolheu-se a base de dados.

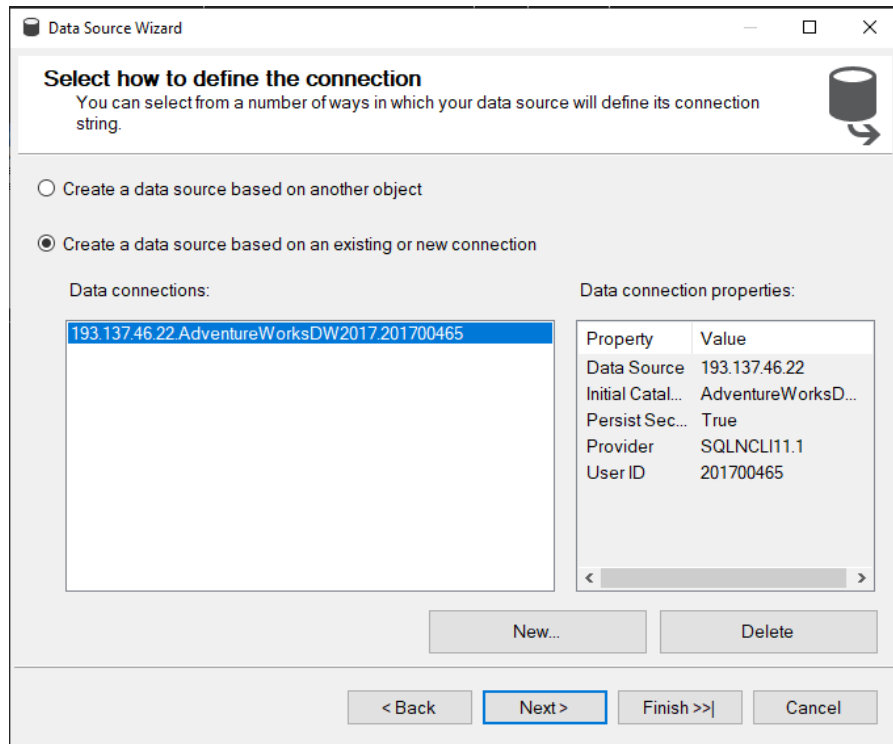


Figura 2 – Escolher a base de dados

2.2 Data Source View

Depois da base de dados escolhida, é necessário ver os dados de forma organizada. Para isso, começou-se por criar uma nova “Data Source View”.

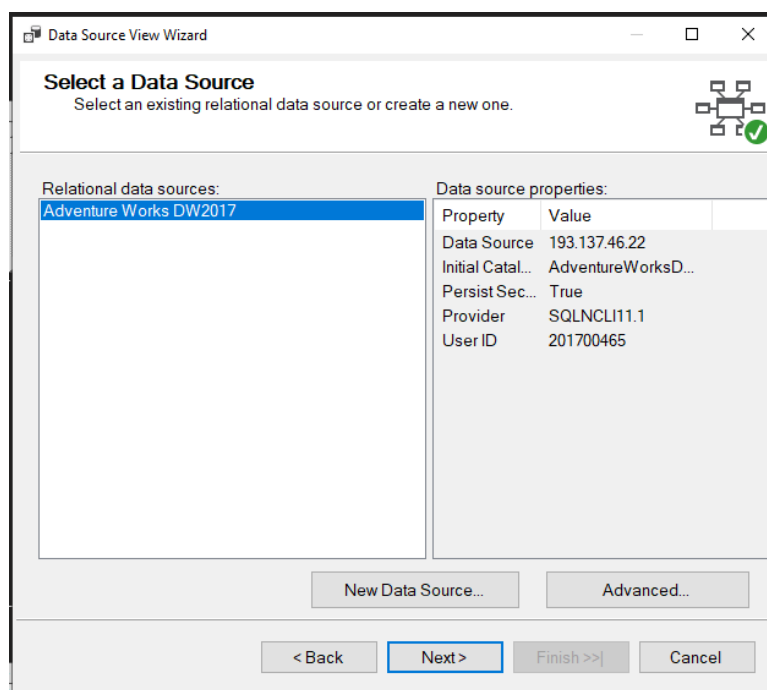


Figura 3 – Escolher a base de dados

De seguida, escolheu-se a base de dados que pretendemos utilizar e ver os dados. Depois de escolhida, é necessário escolher as tabelas que vão ser usadas para a análise de dados.

Para isso, começou-se por escolher a tabela “Fact Internet Sales”. Depois, ao selecionar “Add Related Tables”, tal como o nome indica, vai adicionar as tabelas que estão relacionadas com a tabela inicialmente escolhida (“Fact Internet Sales”). Também, como queremos observar os dados dos Revendedores, é necessário adicionar a tabela “Fact Reseller Sales”. Para que os dados sejam mais específicos, adicionamos por fim as tabelas que indicam as categorias do produto (“DimProductCategory”) e subcategorias do produto (“DimProductSubCategory”).

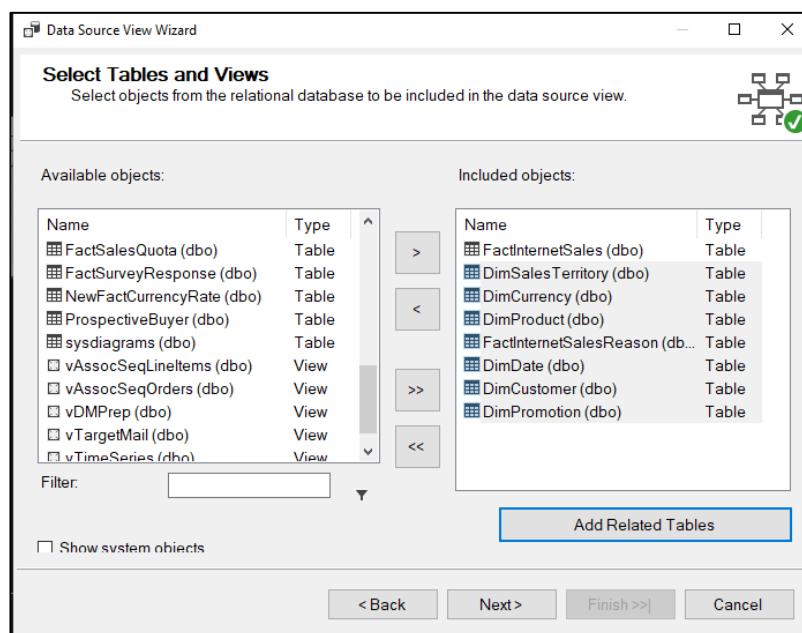


Figura 4 – Escolha das tabelas

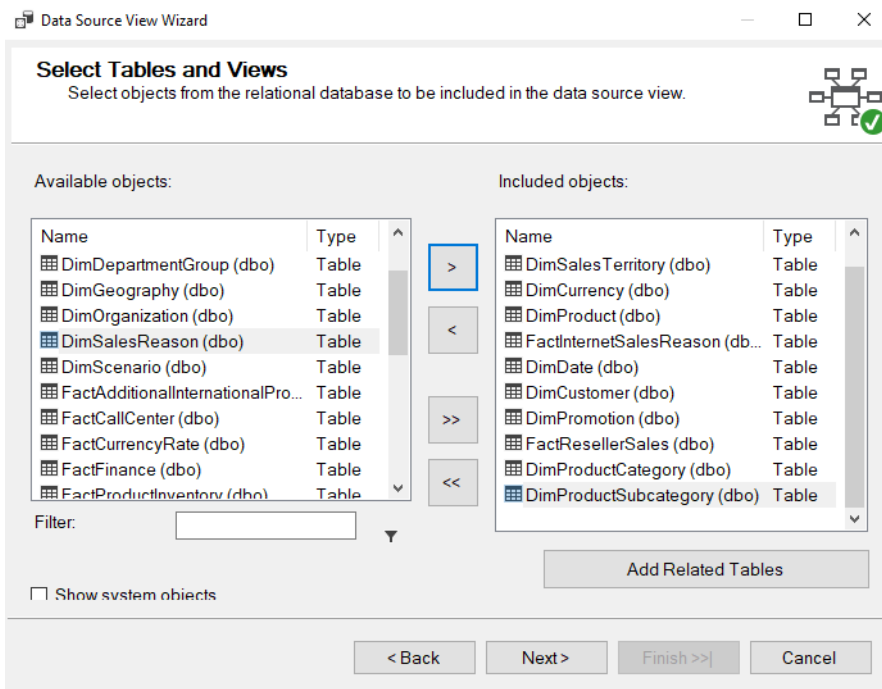


Figura 5 – Escolha de tabelas

Por fim, obtemos uma visualização das várias tabelas, com as respetivas variáveis, de uma forma organizada.

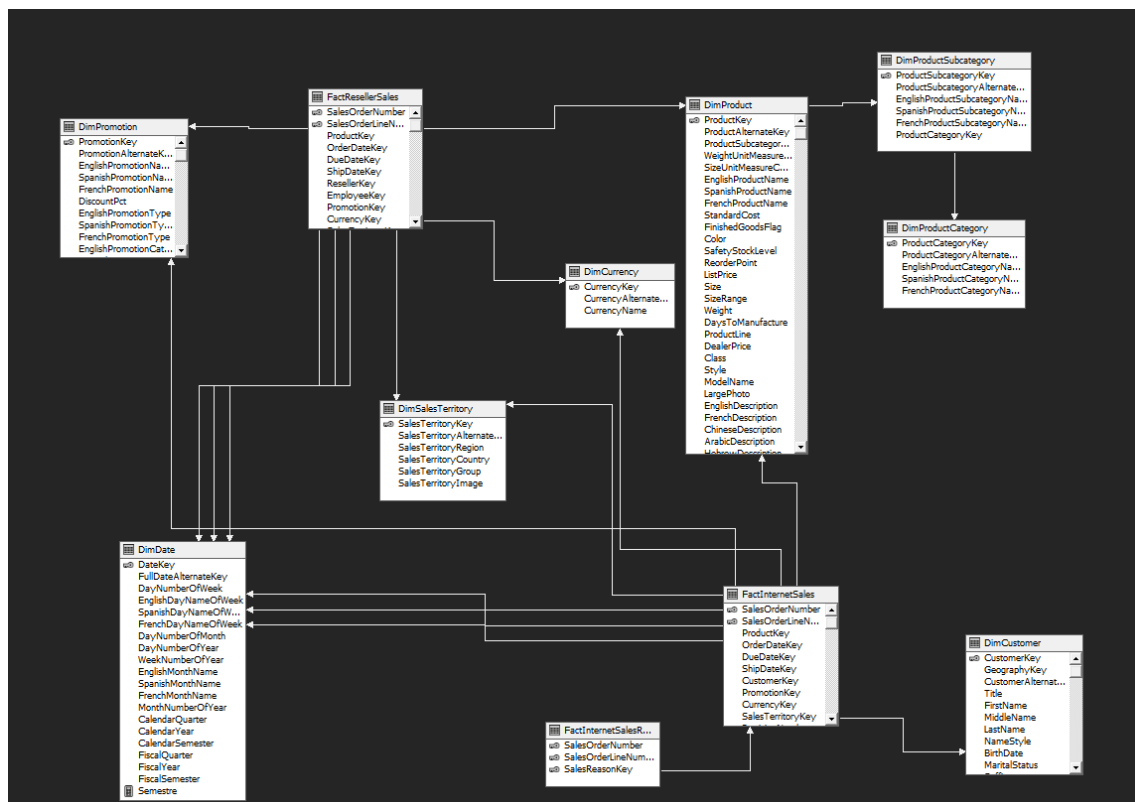


Figura 6 – Visualização da Data Source View

2.3 Cubo

Depois da “view” criada, começou-se por criar o cubo. Para isso, seleccionou-se a opção “New Cube”. Em seguida, é-nos dada várias opções para criar o cubo. Seleccionou-se a opção de usar as tabelas já existentes.

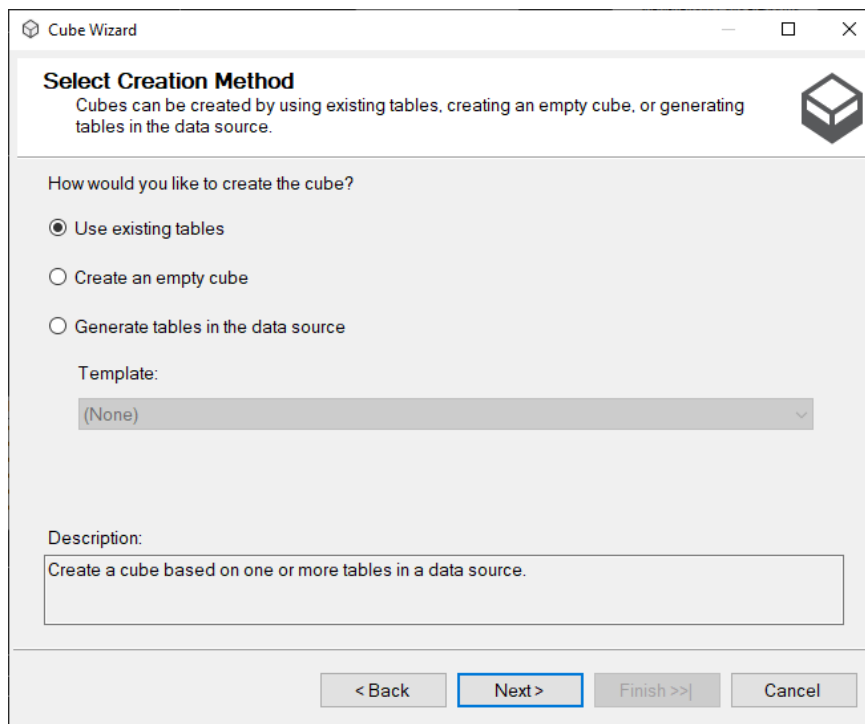


Figura 7 – Escolha de tabelas

De seguida, adicionou-se as tabelas “measure”, tabelas essas que contém medidas, sendo essas medidas quaisquer quantidades/valores numéricos que representam uma métrica alinhada aos objetivos de uma organização. Essas “measures” encontram-se nas tabelas de factos. Depois disso, é necessário escolher as tabelas que pretende-se analisar os dados.

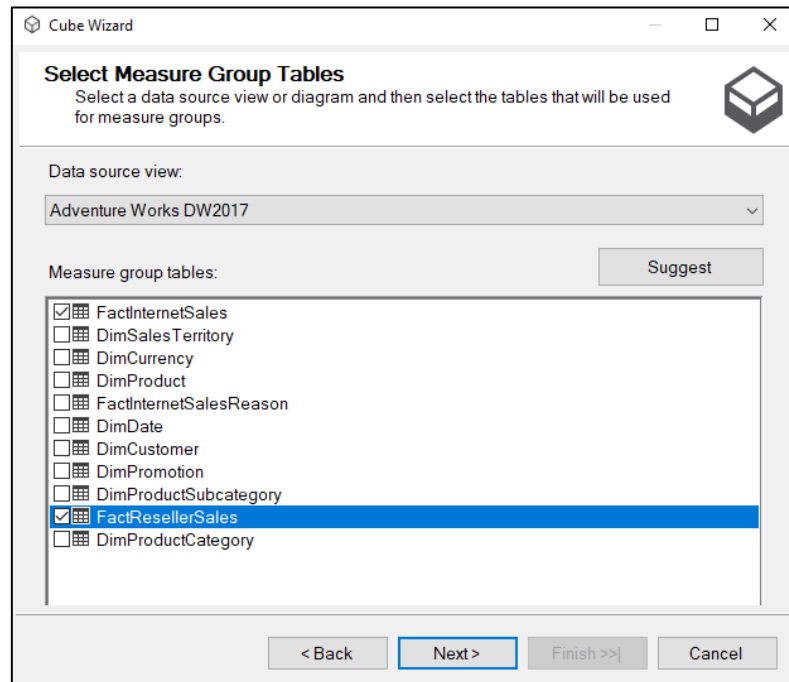


Figura 8 – Escolha das medidas

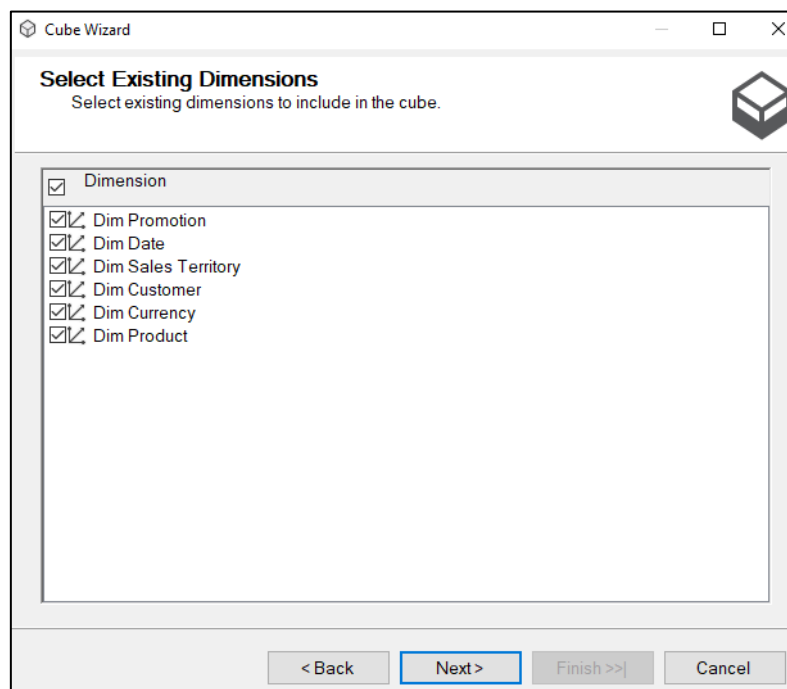


Figura 9 – Escolha das dimensões

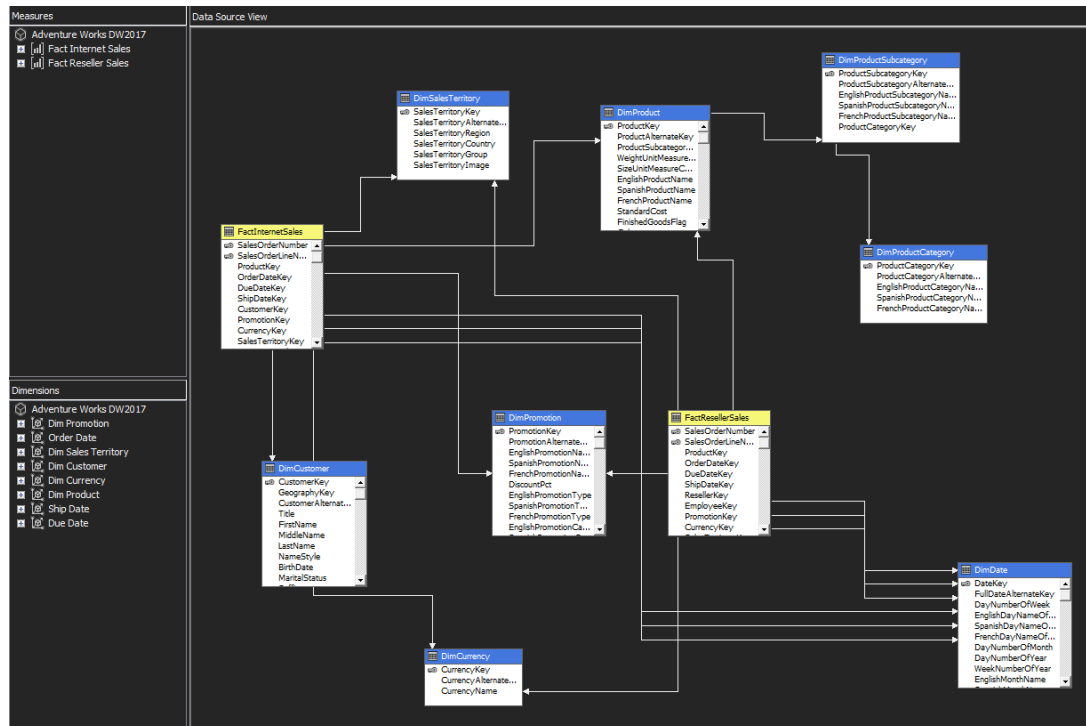


Figura 10 – Visualização do Cubo

Aqui observa-se uma visualização organizada do cubo. Consegue-se identificar uma dimensão *Snowflake*, sendo essa a dimensão DimProduct pois as tabelas DimProductCategory e a DimProductSubCategory dependem da tabela DimProduct.

2.4 Calculated Measure

Criou-se uma “Calculated Measure”, sendo ela a medida Total_Venda (que mostra o Total de Vendas entre a quantidade vendida na internet (Sales Amount da FactInternet Sales) e a quantidade vendida pelos revendedores (Sales Amount da FactResellerSales)), apresentada na figura 11.

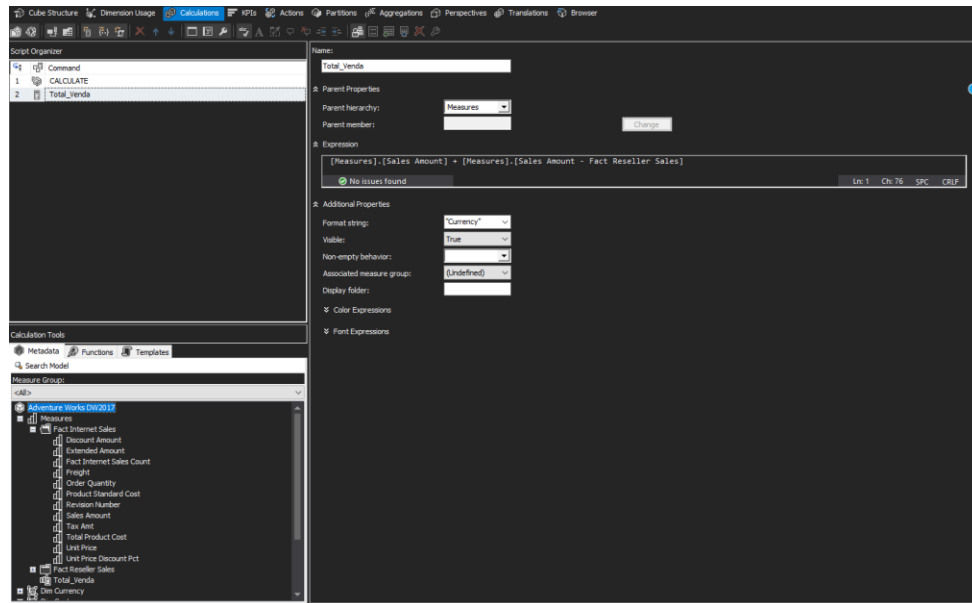


Figura 11 – Calculated Measure do Total_Venda

2.5 Named Calculation

Começou-se por criar uma “Named Calculation” na tabela DimCustomer, tendo como nome “NomeCompleto”. O objetivo desta “Named Calculation” é juntar o primeiro (FirstName), o nome do meio (MiddleName) e último nome (LastName) do cliente.

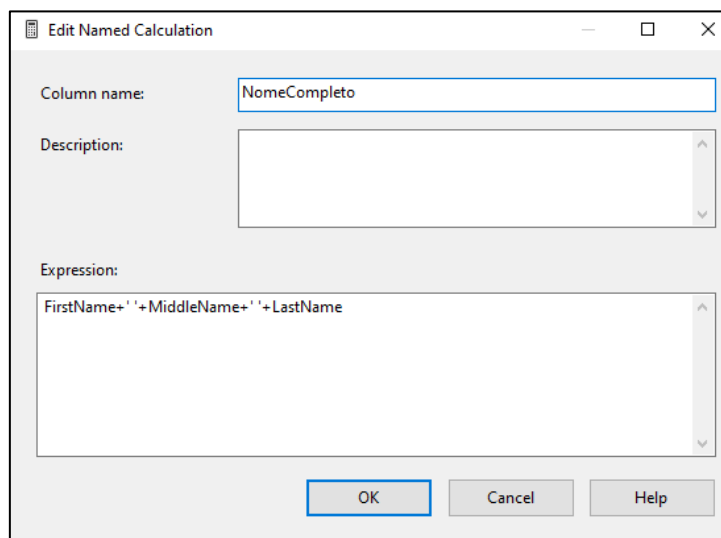


Figura 12 – Named Calculation do NomeCompleto

Depois da “Named Calculation” NomeCompleto criada, criou-se uma nova na tabela DimData, com o nome Semestre.

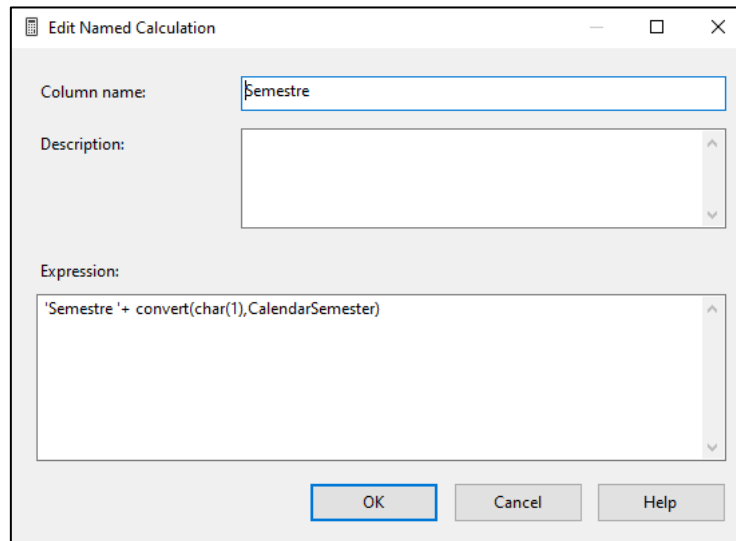


Figura 13 – Named Calculation do Semestre

2.6 Hierarquias em Dimensões

Primeiramente, começou-se por criar uma hierarquia na dimensão DimProduct.

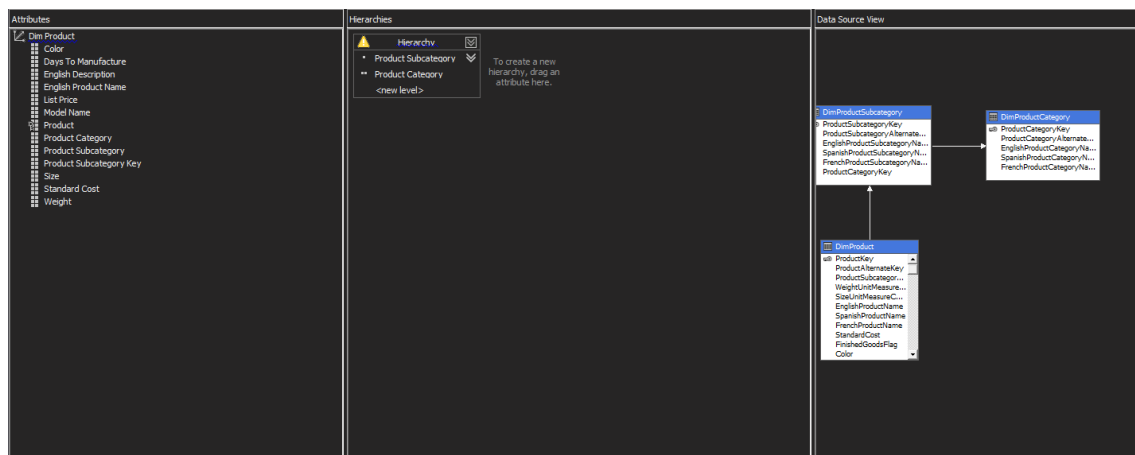


Figura 14 – Hierarquia do DimProduct

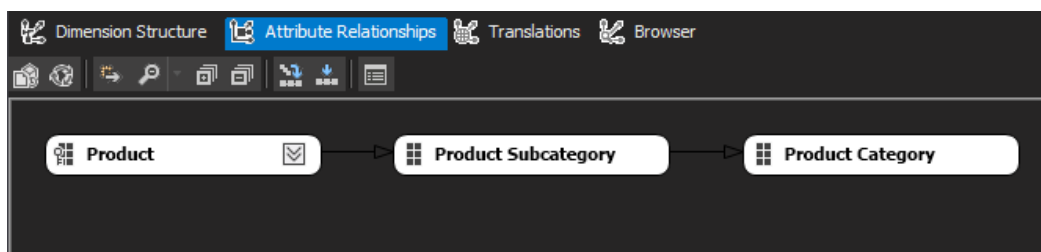


Figura 15 – Relações na Hierarquia DimProduct

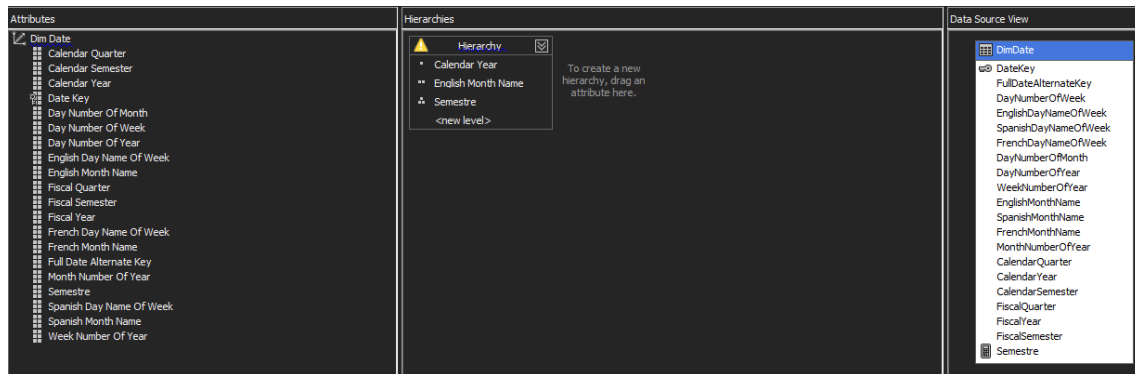


Figura 16 – Hierarquia da DimDate

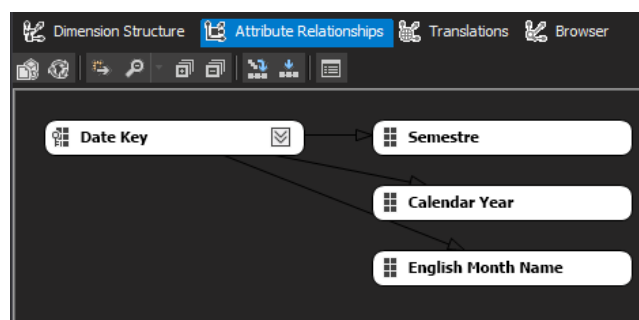


Figura 17 – Relações na Hierarquia DimDate

2.7 Análise dos Dados em Excel

Para que seja possível analisar os dados no Excel, é necessário ir ao “browser” no cubo, e procurar o ícone do Excel, e com isso, é aberto um ficheiro Excel, em que esse ficheiro tem ligação com a base de dados e o SSAS.

Ao clicar, tal como referido anteriormente, é aberto um ficheiro Excel vazio.

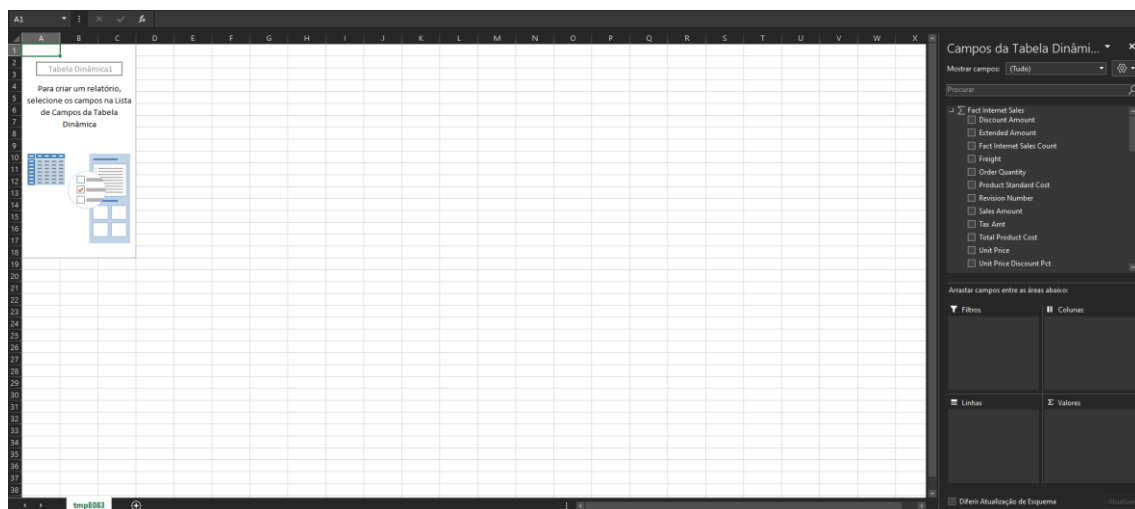


Figura 18 – Ficheiro de Excel vazio

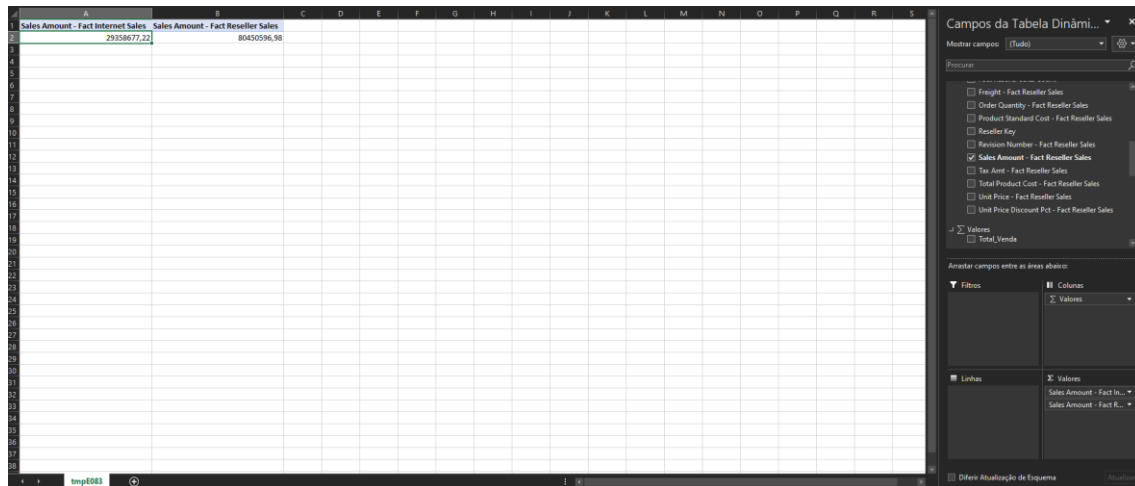


Figura 19 – Adicionar Sales Amount

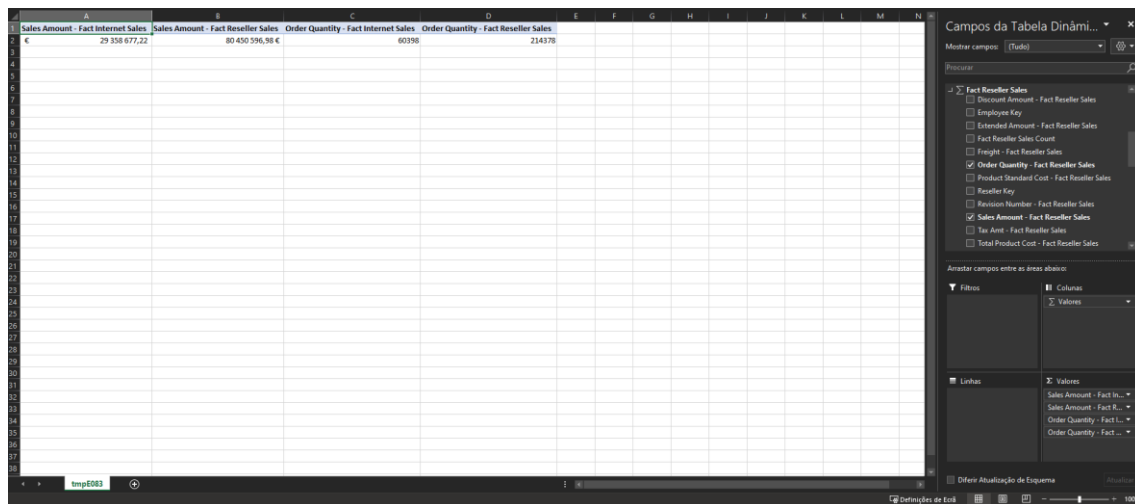


Figura 20 – Adicionar Order Quantity

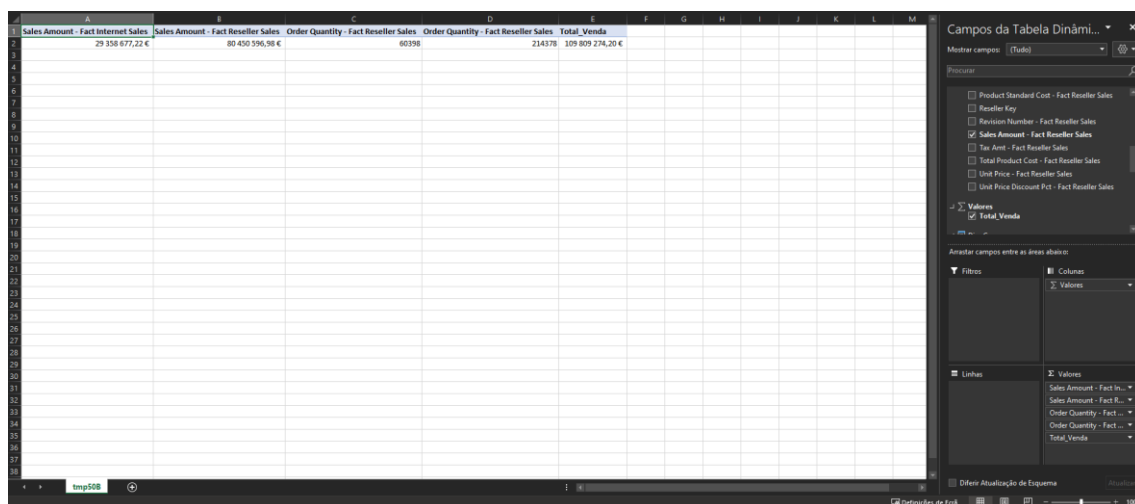


Figura 21 – Adicionar Calculation Measures

Depois de adicionados todos os valores, começamos por adicionar as “Rows”, começando por adicionar a hierarquia do DimDate, através das datas do OrderDate.

Rótulos de Linha	Sales Amount - Fact Internet Sales	Sales Amount - Fact Reseller Sales	Order Quantity - Fact Reseller Sales	Order Quantity - Fact Reseller Sales	Total_Venda
# 2010	43 421,04 €	489 328,56 €	14	830	532 749,62 €
# 2011	7 075 325,93 €	18 192 802,71 €	2216	28572	25 268 128,64 €
# 2012	5 842 485,20 €	28 193 631,53 €	3397	81328	34 036 116,73 €
# 2013	16 351 350,34 €	33 574 834,18 €	32801	103658	49 926 184,52 €
# 2014	45 694,72 €	1970	1970		45 694,72 €
Total Geral	29 358 677,22 €	80 450 596,98 €	60398	214378	109 809 278,20 €

Figura 22 – Adicionar Hierarquia do DimDate

Por fim, adicionamos as variáveis que contém o nome da categoria do produto, depois o nome da subcategoria do produto, e por fim o nome do produto.

Rótulos de Linha	Sales Amount - Fact Internet Sales	Sales Amount - Fact Reseller Sales	Order Quantity - Fact Reseller Sales	Order Quantity - Fact Reseller Sales	Total_Venda
# 2010					
Bikes					
Mountain Bikes					
Mountain-100 Black, 38		44 549,87 €	22	44 549,87 €	
Mountain-100 Black, 42		32 399,90 €	16	32 399,90 €	
Mountain-100 Black, 44		46 374,86 €	23	46 374,86 €	
Mountain-100 Black, 48	3 374,99 €	40 499,88 €	1	20	43 874,87 €
Mountain-100 Silver, 38	3 399,99 €	20 399,94 €	1	10	23 799,93 €
Mountain-100 Silver, 42		10 199,97 €	5	10 199,97 €	
Mountain-100 Silver, 44	10 199,97 €	28 559,92 €	3	14	38 759,89 €
Mountain-100 Silver, 48		14 279,96 €	7	14 279,96 €	
Road Bikes					
Road-150 Red, 44	3 578,27 €		1		3 578,27 €
Road-150 Red, 48	7 156,54 €		2		7 156,54 €
Road-150 Red, 52	3 578,27 €		1		3 578,27 €
Road-150 Red, 56	3 578,27 €	30 057,47 €	14	33 635,74 €	
Road-150 Red, 62	7 156,54 €	6 440,89 €	2	3	13 597,43 €
Road-450 Red, 44		12 247,12 €	14	12 247,12 €	
Road-450 Red, 52		40 240,52 €	46	40 240,52 €	
Road-450 Red, 58		23 619,44 €	27	23 619,44 €	
Road-450 Red, 60		9 622,73 €	11	9 622,73 €	
Road-450 Red, 64		5 452,97 €	13	5 452,97 €	
Road-450 Black, 52		12 164,31 €	29	12 164,31 €	
Road-450 Black, 58		12 583,77 €	30	12 583,77 €	
Road-450 Black, 60		4 614,05 €	11	4 614,05 €	
Road-450 Black, 62	699,10 €	699,10 €	1		699,10 €
Road-450 Red, 44		18 456,19 €	44	18 456,19 €	
Road-450 Red, 48		8 389,18 €	20	8 389,18 €	
Road-450 Red, 52	699,10 €	5 872,42 €	14	6 571,52 €	
Road-450 Red, 60		18 036,79 €	43	18 036,79 €	
Road-450 Red, 62		7 969,72 €	19	7 969,72 €	
Components		31 525,96 €	114	31 525,96 €	
Clothing					
Caps					
AWC Logo Cap		207,46 €	40	207,46 €	
Jerseys					
Long-Sleeve Logo Jersey, L		1 413,18 €	49	1 413,18 €	

Figura 23 – Adicionar o nome da categoria, subcategoria e produto

Terminando as escolhas das rows e valores, obtemos uma folha de Excel que podemos analisar, de forma geral, na figura 24, ou específica na figura 25, dos dados contidos na base de dados.

	A	B	C	D	E	F	G
1	Rótulos de Linha	Sales Amount - Fact Internet Sales	Sales Amount - Fact Reseller Sales	Order Quantity - Fact Reseller Sales	Order Quantity - Fact Reseller Sales	Total_Venda	
2	2010						
3	Bikes	43 421,04 €	453 231,80 €	14	455	496 652,84 €	
4	Components		31 525,96 €		114	31 525,96 €	
5	Clothing		2 875,15 €		167	2 875,15 €	
6	Accessories		1 695,67 €		84	1 695,67 €	
7	2011						
8	Bikes	7 075 525,93 €	16 367 229,91 €	2216	14984	23 442 755,83 €	
9	Components		1 643 351,88 €		4754	1 643 351,88 €	
10	Clothing		136 624,14 €		6372	136 624,14 €	
11	Accessories		45 596,79 €		2462	45 596,79 €	
12	2012						
13	Bikes	5 839 695,33 €	22 539 228,70 €	3269	27951	28 378 924,02 €	
14	Components		4 749 805,04 €		19778	4 749 805,04 €	
15	Clothing	642,79 €	759 490,30 €	22	25622	760 133,09 €	
16	Accessories	2 147,08 €	145 107,49 €	106	7977	147 254,57 €	
17	2013						
18	Bikes	15 359 502,36 €	26 942 691,15 €	9706	31625	42 302 193,51 €	
19	Components		5 374 393,77 €		24381	5 374 393,77 €	
20	Clothing	323 806,45 €	878 851,24 €	8686	32336	1 202 657,69 €	
21	Accessories	668 241,53 €	378 897,98 €	34409	15316	1 047 139,51 €	
22	2014						
23	Clothing	15 323,37 €		393		15 323,37 €	
24	Accessories	30 371,35 €		1577		30 371,35 €	
25	Total Geral	29 358 677,22 €	80 450 596,98 €	60398	214378	109 809 274,20 €	
26							
27							
28							
29							

Figura 24 – Análise dos Dados

	A	B	C	D	E	F	G	H
1	Rótulos de Linha	Sales Amount - Fact Internet Sales	Sales Amount - Fact Reseller Sales	Order Quantity - Fact Reseller Sales	Order Quantity - Fact Reseller Sales	Total_Venda		
2	2010							
3	Bikes	43 421,04 €	453 231,80 €	14	455	496 652,84 €		
4	Components							
5	Mountain Frames							
6	HL Mountain Frame - Black, 38		2 858,82 €		4	2 858,82 €		
7	HL Mountain Frame - Black, 42		714,70 €		1	714,70 €		
8	HL Mountain Frame - Black, 48		809,76 €		1	809,76 €		
9	HL Mountain Frame - Silver, 38		1 445,19 €		2	1 445,19 €		
10	HL Mountain Frame - Silver, 46		2 167,78 €		3	2 167,78 €		
11	HL Mountain Frame - Silver, 48		1 637,40 €		2	1 637,40 €		
12	Road Frames							
13	LL Road Frame - Black, 52		3 393,04 €		19	3 393,04 €		
14	LL Road Frame - Black, 58		1 428,65 €		8	1 428,65 €		
15	LL Road Frame - Red, 44		2 759,07 €		15	2 759,07 €		
16	LL Road Frame - Red, 48		1 655,44 €		9	1 655,44 €		
17	LL Road Frame - Red, 60		2 943,01 €		16	2 943,01 €		
18	LL Road Frame - Red, 62		2 575,13 €		14	2 575,13 €		
19	ML Road Frame - Red, 48		5 710,37 €		16	5 710,37 €		
20	ML Road Frame - Red, 52		1 427,59 €		4	1 427,59 €		
21	Clothing							
22	Caps							
23	AWC Logo Cap		207,46 €		40	207,46 €		
24	Jerseys							
25	Long-Sleeve Logo Jersey, L		1 413,18 €		49	1 413,18 €		
26	Long-Sleeve Logo Jersey, M		461,45 €		16	461,45 €		
27	Long-Sleeve Logo Jersey, XL		547,97 €		19	547,97 €		
28	Socks							
29	Mountain Bike Socks, L		28,50 €		5	28,50 €		
30	Mountain Bike Socks, M		216,60 €		38	216,60 €		
31	Accessories							
32	Helmets							
33	Sport-100 Helmet, Black		545,04 €		27	545,04 €		
34	Sport-100 Helmet, Blue		666,15 €		33	666,15 €		
35	Sport-100 Helmet, Red		484,48 €		24	484,48 €		
36	2011							
37	Bikes	7 075 525,93 €	16 367 229,91 €	2216	14984	23 442 755,83 €		
38	Components							

Figura 25 - Análise dos Dados

3. Conclusão

Com este trabalho conseguiu-se desenvolver capacidades para a construção de um cubo em SSAS que tem como objetivo de armazenar dados num formato multidimensional. Um cubo OLAP é uma estrutura de dados que elimina as limitações de bases de dados relacionais, fornecendo uma análise rápida de dados. Os cubos podem apresentar e somar grandes quantidades de dados e fornecer aos utilizadores uma forma de pesquisa de dados.

No entanto existem ainda algumas desvantagens quando nos referimos ao cubo, por exemplo uma delas é que todos os cubos afetados do OLAP ficam não processados e terão de ser processados para que seja possível haver condições para serem efetuadas consultas.

Existem outras opções melhores que o cubo, mas a construção ajudou-nos a ter uma melhor perceção do que é um cubo e que este apresenta vantagens como ter acesso à informação mais rapidamente para análise, a sua capacidade de armazenarem grandes quantidades de dados e permitirem aos utilizadores efetuarem uma pesquisa a quaisquer pontos de dados, desta forma os dados podem ser agregados, segmentados e repartidos conforme necessários para lidar com a grande variedade de perguntas relevantes ao utilizador.

Neste projeto as dificuldades foram inferiores às do projeto anterior, ainda assim apresentou alguns problemas que conseguiram-se resolver com sucesso.

4. Bibliografia

<https://www.devmedia.com.br/criando-um-cubo-no-analysis-services/25302>

Consultado a 1 de fevereiro de 2021

<https://docs.microsoft.com/pt-pt/system-center/scsm/olap-cubes-overview?view=sc-sm-2019>

Consultado a 1 de fevereiro de 2021

<https://www.sqlshack.com/build-cube-scratch-using-sql-server-analysis-services-ssas/>

Consultado a 1 de fevereiro de 2021

<https://docs.microsoft.com/en-us/analysis-services/analysis-services-overview?view=asallproducts-allversions>

Consultado a 1 de fevereiro de 2021