

Unidade Curricular: Sistemas de Informação II

## Ficha T. Prática n.º 5

### Objectivo:

O objetivo desta ficha é perceber como é elaborada a modelação dimensional de um Data Warehouse. Criar uma DW no SQL Server e realizar consultas com recurso ao Microsoft Power BI.

### Caso: Área de marketing

De forma a reforçar o entendimento acerca do processo em quatro etapas para a identificação de um modelo multidimensional adequado a uma determinada realidade organizacional, vamos considerar para esta ficha um exercício retirado do livro Business Intelligence Da Informação ao Conhecimento, 3ª Edição, FCA, ISBN: 9789727228805 de Isabel Ramos e Maribel Yasmina Santos (2017).

O exercício apresenta um exemplo ligado ao setor de Marketing, área que necessita de conhecer o **perfil dos clientes** para poder dirigir, de uma forma eficiente, as campanhas a realizar.

Ao conhecer o perfil dos clientes mais propensos para a aquisição de um dado produto, as organizações podem restringir as suas campanhas reduzindo o desperdício associado, por exemplo, ao envio de cartazes promocionais a clientes que, à partida, não se enquadram no perfil dos clientes consumidores do produto que está a ser comercializado ou na campanha que está a ser realizada.

Considere que o conjunto de dados associado a este exemplo na área do Marketing integra registos associados a clientes que foram alvo de uma campanha anterior (Exemplo tabela 1). Na sequência da campanha realizada, determinado conjunto de clientes aderiu à mesma, adquirindo o produto promovido, enquanto um outro conjunto de clientes não aderiu à referida campanha.

Ao analisar este conjunto de dados pretende-se identificar o perfil dos clientes que aderiram à campanha anterior, de forma a torná-los o alvo privilegiado de uma futura campanha.

Os atributos que integram o conjunto de dados para análise são: Identificação, Idade, Sexo, Região, Vencimento, Estado Civil, Idade, Número de Filhos, Carro, Conta a Prazo, Conta à Ordem, Crédito Habitação, data de adesão à promoção e compra na Campanha Anterior.

Este conjunto de atributos permite conhecer a idade, o sexo e o número de filhos de cada indivíduo, a região em que habita e o seu vencimento. Além desta informação, existe um conjunto de atributos de carácter booleano que nos permite saber se o cliente tem ou não carro, conta à ordem, conta a prazo e um crédito à habitação. Além desse aspeto ainda temos um atributo que permite identificar se o cliente aderiu ou não à campanha anterior.

Partindo destes pressupostos, é pretendido que o modelo permita:

- Analisar o perfil dos clientes que aderiram, ou não, à campanha anterior;
- Identificar as classes de idades e de vencimentos que estão mais propensas à adesão a uma futura campanha;

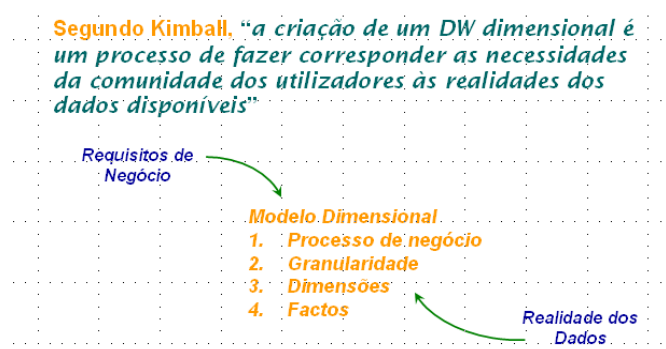
- Verificar as características dos clientes que permitem identificar o seu perfil no que diz respeito ao padrão de consumo;
- Verificar se a situação bancária dos clientes influencia os seus padrões de consumo;
- Analisar se a região onde o cliente habita permite diferenciar os diversos indivíduos em termos de adesão, ou não, à campanha.

**Tabela 1 - Exemplo dados**

identificação	Idade	Sexo	Região	Vencimento	Casado	Filhos	Carro	Conta Prazo	Conta Ordem	Crédito Habitação	Data adesão	Compra Campanha Anterior
1242434	19	M	Portugal	1200	Sim	1	Sim	sim	não	não	10/2020	Não
2142343	32	F	Espanha	1600	não	1	não	não	sim	não	11/2022	Não
6524654	44	F	Brasil	2200	não	0	Sim	não	sim	não	12/2021	Não
6546546	54	M	Portugal	1100	não	2	não	sim	não	sim	4/2022	Sim
5665466	33	F	Portugal	1900	Sim	3	Sim	sim	não	não	3/2022	Sim
7655756	32	F	Espanha	800	Sim	1	não	sim	não	sim	1/2022	Não

### Tarefas propostas:

Os alunos devem seguir processo proposto por Kimball para desenvolver o modelo dimensional para o DW deste exercício.



### Primeira Etapa - Começamos por selecionar o processo de negócio a modelar.

- O processo de negócio está associado à compra do produto, ou produtos, alvo da campanha de marketing anterior.

### Segunda Etapa - Avaliar a granularidade: qual o nível de detalhe mais fino que será disponibilizado no modelo dimensional.

- A granularidade adotada, para o carregamento da informação disponível, é ao nível do indivíduo e transação não existindo qualquer agregação prévia da informação.

### Terceira Etapa - Selecionar as dimensões que serão incluídas no modelo dimensional.

- As dimensões que serão utilizadas para analisar o processo de negócio são: Cliente, Região, Carro, Classe de Vencimento, Classe de Idade, Conta à Ordem, conta a prazo, tempo e Crédito Habitação.
- Notas as classes definidas para os atributos:
  - Idade: [18,30], ]30,42], ]42,54], ]54,62], ]62,74]
  - Vencimento: ]500-1000], ]1000,1500], ]1500,2000], ]2000,3000], ]3000-4000]
  - Tempo: mês e ano
  - Conta à Ordem, conta a prazo, carro, crédito habitação: sim,não

#### **Quarta Etapa - Selecionar os factos que serão incluídos no modelo dimensional.**

- O facto a analisar, segundo as diferentes dimensões definidas no ponto anterior, é a Adesão ao Cliente à Campanha Anterior. Este facto permitirá analisar o perfil dos clientes que aderiram, ou não aderiram, à campanha anterior. Os valores SIM e NÃO, apresentados pelo atributo campanha anterior, serão convertidos em valores 1 e 0 respetivamente, para ser possível a sua análise pelas diversas dimensões consideradas. Após essa conversão, este atributo é representado por uma flag com dois valores possíveis, sendo possível utilizar a função de agregação COUNT para obter agregados dos clientes que aderiram, ou não à campanha.

#### **1) Construa o modelo em estrela, incluindo os atributos que julgar mais significativos.**

- Percorridas as quatro etapas, deve agora definir o modelo multidimensional. Utilize as dimensões sugeridas nas alíneas anteriores.

#### **2) Elabore o código SQL para o seu modelo em estrela. Implemente o mesmo no SQL Server. Deve povoar a bd com registos que permitam responder à questão seguinte.**

- Se não conseguir, deve solicitar ao docente o código SQL para o efeito.

#### **3) Sobre o modelo em estrela, elabore pesquisas SQL que permitam:**

1. Identificar qual a região onde mais clientes aderiram à campanha anterior.
2. Relacione a Região, a classe de vencimento e o número de adesão à campanha anterior.
3. Apresentar uma listagem que reflita a relação do número de filhos com a classe de vencimento, para clientes que não tenham aderido à campanha anterior e que estejam no intervalo de idades 20-30 e 30-40.
4. Apresentar uma listagem, agrupada pela região que apresente o número de clientes que aderiram à campanha anterior e tenham um crédito de habitação.
5. Relacione a existências de contas à ordem com os intervalos de vencimentos.
6. Qual o número de aderentes à campanha anterior, para clientes que recebem mais de 1000€ mensais, que estão na região de Coimbra, detentores de Contas à Ordem e a Prazo, com menos de 40 anos e sem crédito de habitação. Este perfil de cliente poderá ser o target de uma campanha direcionada.

**Utilizando o power bi, realize os seguintes exercícios**

- 4) Realizar uma consulta que permita relacionar a classe de vencimento com a classe de idade de clientes que aderiram à última campanha.**

CLASSEIDADE	1001-1500	1501-2000	2001-2500	2501-3000	3001-3500	3501-4000	500-1000	Total
18 - 30	75	61	69	66	70	69	69	479
31 -42	61	75	70	67	63	64	59	459
43 - 54	78	77	88	75	61	75	75	529
55 - 62	82	77	62	74	92	68	73	528
63 - 74	71	66	70	77	88	70	70	512
Total	367	356	359	359	374	346	346	2507

**Figura 1 - Análise das classes de idade e de vencimento para os indivíduos que aderiram à campanha anterior**

- 5) Realizar, agora uma consulta que permita relacionar a classe de vencimento com a classe de idade de todos os clientes. Nota, para esse fim deverá acrescentar uma nova medida do tipo count do idClasseidade na tabela de factos.**

Formula para a nova medida: Medida = COUNT(CAMPANHAMARKETING[IDCLASSEIDADE])

CLASSEIDADE	1001-1500	1501-2000	2001-2500	2501-3000	3001-3500	3501-4000	500-1000	Total
18 - 30	154	149	151	138	125	139	129	985
31 -42	128	143	138	158	115	125	138	945
43 - 54	143	154	151	140	143	143	149	1023
55 - 62	142	157	136	150	163	135	155	1038
63 - 74	156	128	135	156	157	133	144	1009
Total	723	731	711	742	703	675	715	5000

**Figura 2 - Análise das classes de idade e de vencimento para os todos os indivíduos**

- 6) Nesta fase de exploração de dados, pretende-se aprofundar o conhecimento acerca dos clientes que aderiram à campanha anterior e do seu respetivo perfil. Assim, apresente uma listagem que permita relacionar o número de filhos, Sexo e Classe de vencimento de forma perceber a evolução de adesão à campanha com o número de filhos.**

FILHOS	1001-1500	1501-2000	2001-2500	2501-3000	3001-3500	3501-4000	500-1000	Total
1	53	64	53	50	55	60	61	396
F	29	37	23	24	34	22	30	199
M	24	27	30	26	21	38	31	197
2	59	57	69	60	77	67	55	444
F	35	28	33	24	42	34	32	228
M	24	29	36	36	35	33	23	216
3	55	65	63	66	56	58	57	420
F	33	35	26	37	28	31	26	216
M	22	30	37	29	28	27	31	204
4	58	61	63	62	57	62	56	419
F	32	26	33	28	24	24	30	197
M	26	35	30	34	33	38	26	222
5	75	51	45	60	56	44	45	376
F	38	28	23	30	27	25	18	189
M	37	23	22	30	29	19	27	187
6	67	58	66	61	73	55	72	452
F	33	20	32	32	41	29	36	223
M	34	38	34	29	32	26	36	229
Total	367	356	359	359	374	346	346	2507

**Figura 3 - Análise do número de filhos e das classes de vencimento para os indivíduos que aderiram à campanha anterior**

- 7) A figura anterior poderá evidenciar os resultados obtidos, de forma mais clara, se for utilizada uma representação diferente dos indicadores. Para isso apresente um gráfico, como na imagem seguinte, que permita por exemplo analisar os o número de filhos e das classes de vencimento para os indivíduos que aderiram à campanha anterior. Apresente apenas o gráfico relativo aos intervalos de 500-1000,1001 a 1500 e 1501 a 2000.

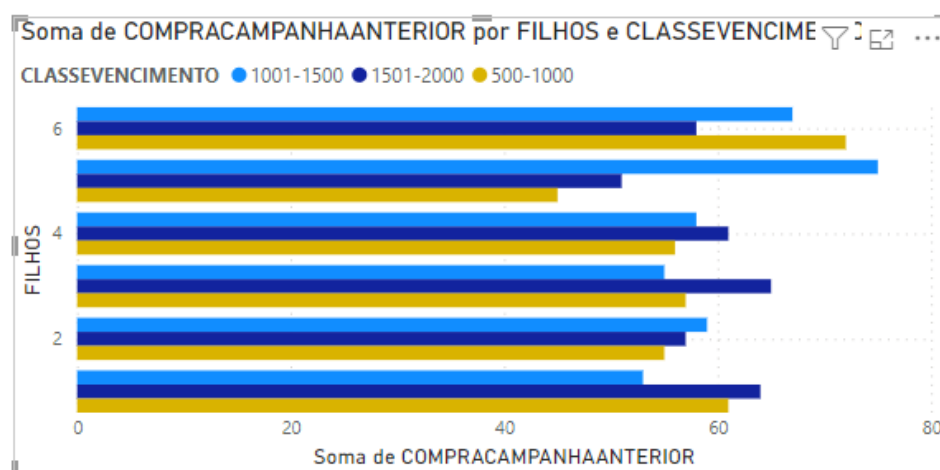


Figura 4 - Análise do número de filhos e das classes de vencimento para os indivíduos que aderiram à campanha anterior

- 8) Elabore um gráfico que permita a análise do perfil dos clientes que aderiram à campanha anterior recorrendo aos atributos classe de idade, sexo e casado (Atributo que indica se o cliente é, ou não casado).

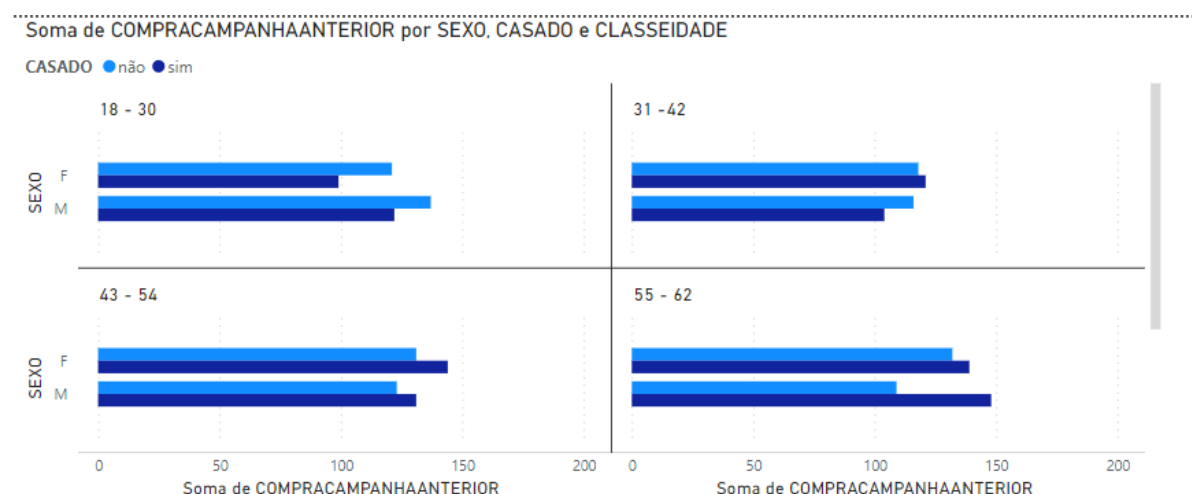


Figura 5 - Análise das classes de idade, sexo e casado para os indivíduos que aderiram à campanha anterior.

- 9) Nesta questão, pretende-se verificar a influência dos atributos que indicam se o cliente possui, ou não, conta à ordem, conta a prazo e crédito à habitação. Para isso elabore uma tabela que permita visualizar os dados dos três atributos, conforme a imagem seguinte:

TEMCONTAORDEM	não	sim	Total
☒ não	642	634	1276
não	292	308	600
sim	350	326	676
☒ sim	612	619	1231
não	309	319	628
sim	303	300	603
Total	1254	1253	2507

Figura 6 - Análise da situação bancária para os indivíduos que aderiram à campanha anterior, com base nos atributos temcontaordem,temcontaaprazo,temcreditohabitacao

- 10) Realize uma análise que permita verificar se o cliente possui, ou não, carro e ainda o número de filhos.

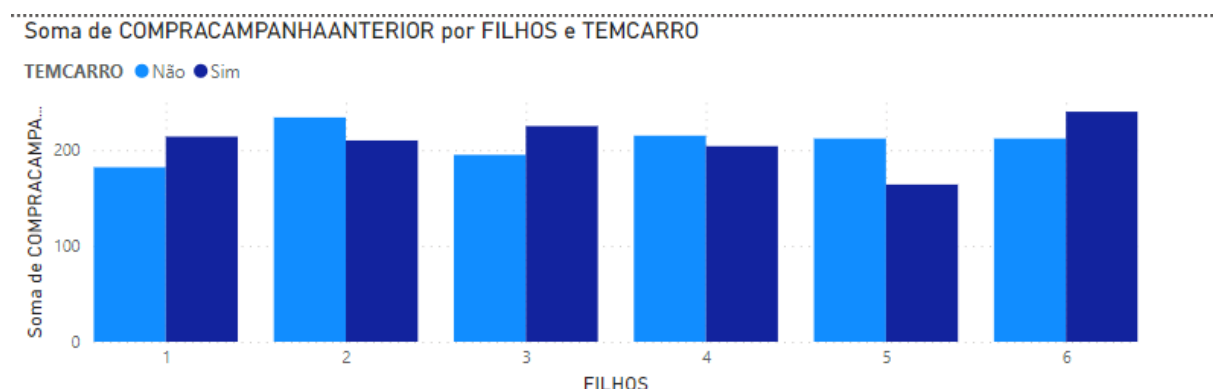


Figura 7 - Análise atendendo à disponibilidade de carro e ao número de filhos para os indivíduos que aderiram à campanha anterior

11)Elabore um gráfico, conforme o apresentado na imagem seguinte, que permita verificar a adesão à campanha anterior, atendendo à região em que os indivíduos habitam e à sua classe de vencimento.

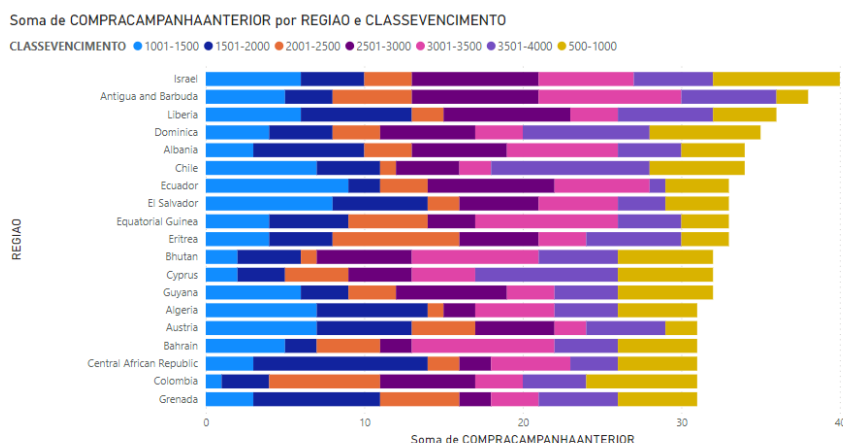


Figura 8 - Análise de região e das classes de vencimento para os indivíduos que aderiram à campanha anterior

12)Acrescente um mapa, que permita visualizar a dispersão dos dados anterior, conforme a imagem seguinte. Nota a o atributo "região" da tabela região deverá ser renomeado para "país" de forma a que o Power Bi consiga fazer a ligação.

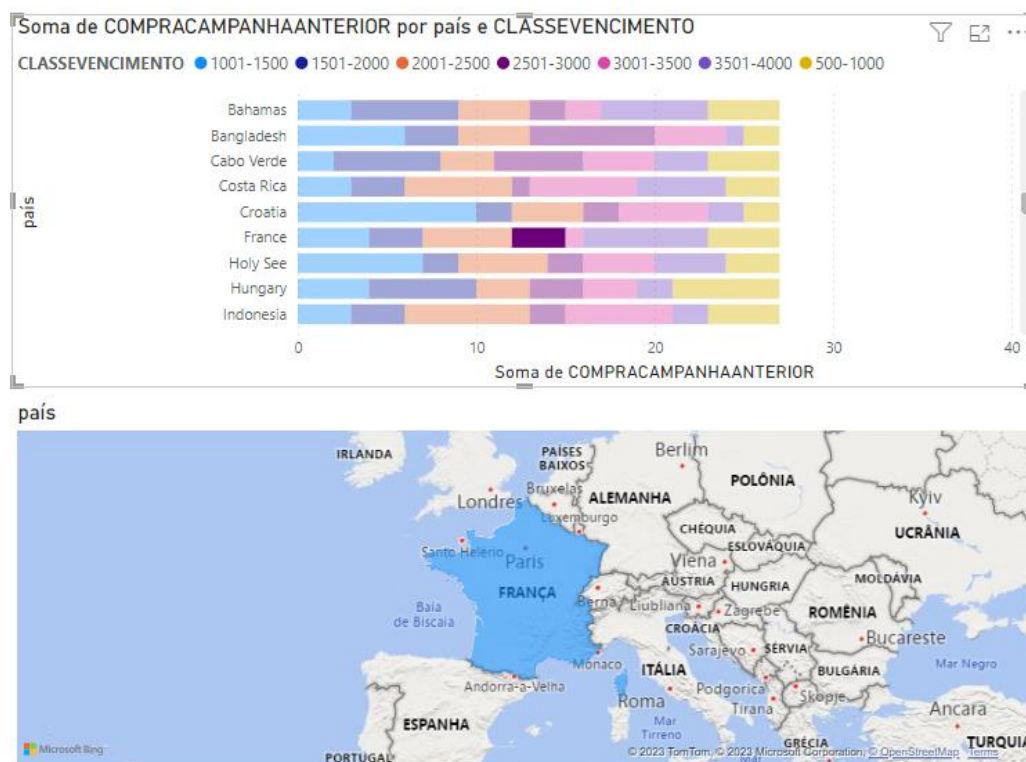
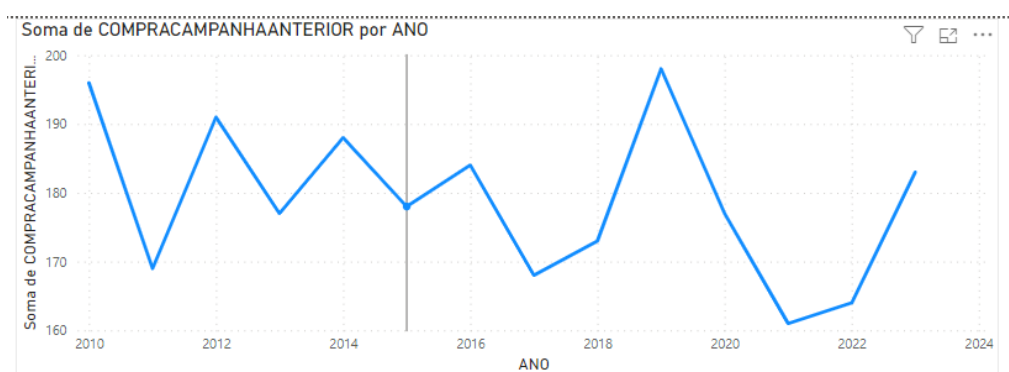


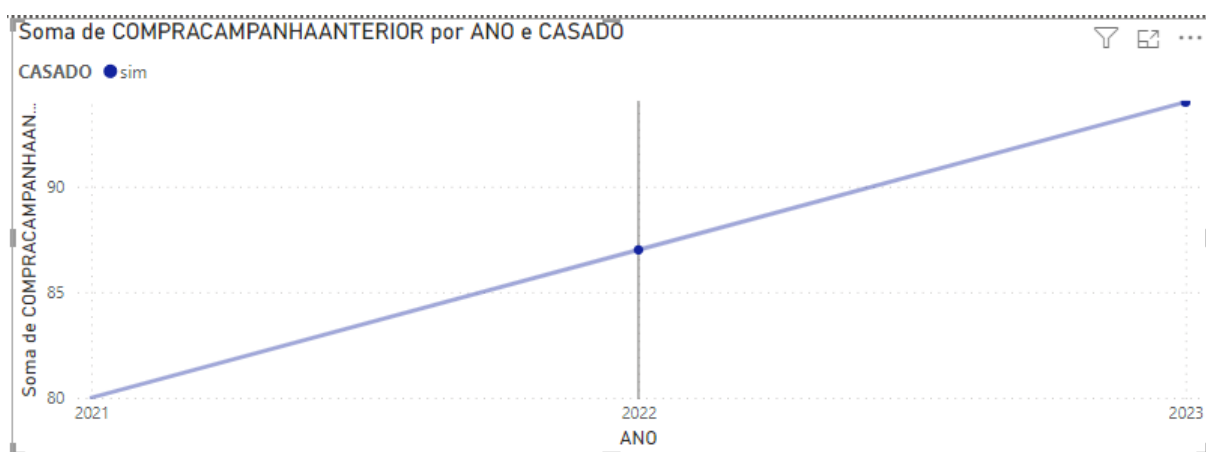
Figura 9 - Análise de região e das classes de vencimento para os indivíduos que aderiram à campanha anterior com recurso a um mapa.

**13)Elabore um gráfico que permita verificar a evolução de adesão à campanha no decorrer dos anos.**



**Figura 10 - Análise de adesão às campanhas de marketing no decorrer dos anos**

**14)Elabore um gráfico que permita verificar a evolução de adesão à campanha apenas nos anos 2021 e 2022 de clientes casados**



**Figura 5 - Análise de adesão às campanhas de marketing no decorrer dos anos 2021 e 2022 de clientes casados**

**15)Elabore um dashbord completo e interativo que permita:**

- Visualizar a dispersão dos clientes por região (num mapa) que aderiram à campanha, com mais que 3 filhos;**
- Dados que apresentem a relação entre o intervalo de vencimento e a propriedade de um carro de todos os clientes que aderiram ou não, a campanha. Considere apenas as classes de vencimento compreendidas entre 1000 e 2500;**
- O número de adesões no mês de dezembro;**
- O perfil de todos os clientes no que respeita ao intervalo de vencimento, sem crédito habitação, e número de filhos que aderiram a campanhas, no ano de 2023.**

## **Bibliografia**

Isabel Ramos e Maribel Yasmina Santos (2017),Business Intelligence Da Informação ao Conhecimento, 3ª Edição, FCA, ISBN: 9789727228805