

Trabalho Final - Aluno: Vitor Duarte Bezerra de Oliveira

1. Ferramenta para o ETL e dashboard = PowerBI.
2. Tema escolhido = Simulação do setor de análise de desempenho em uma equipe de futebol (Equipe escolhida = Real Madrid da Espanha)
3. Perguntas Analíticas:

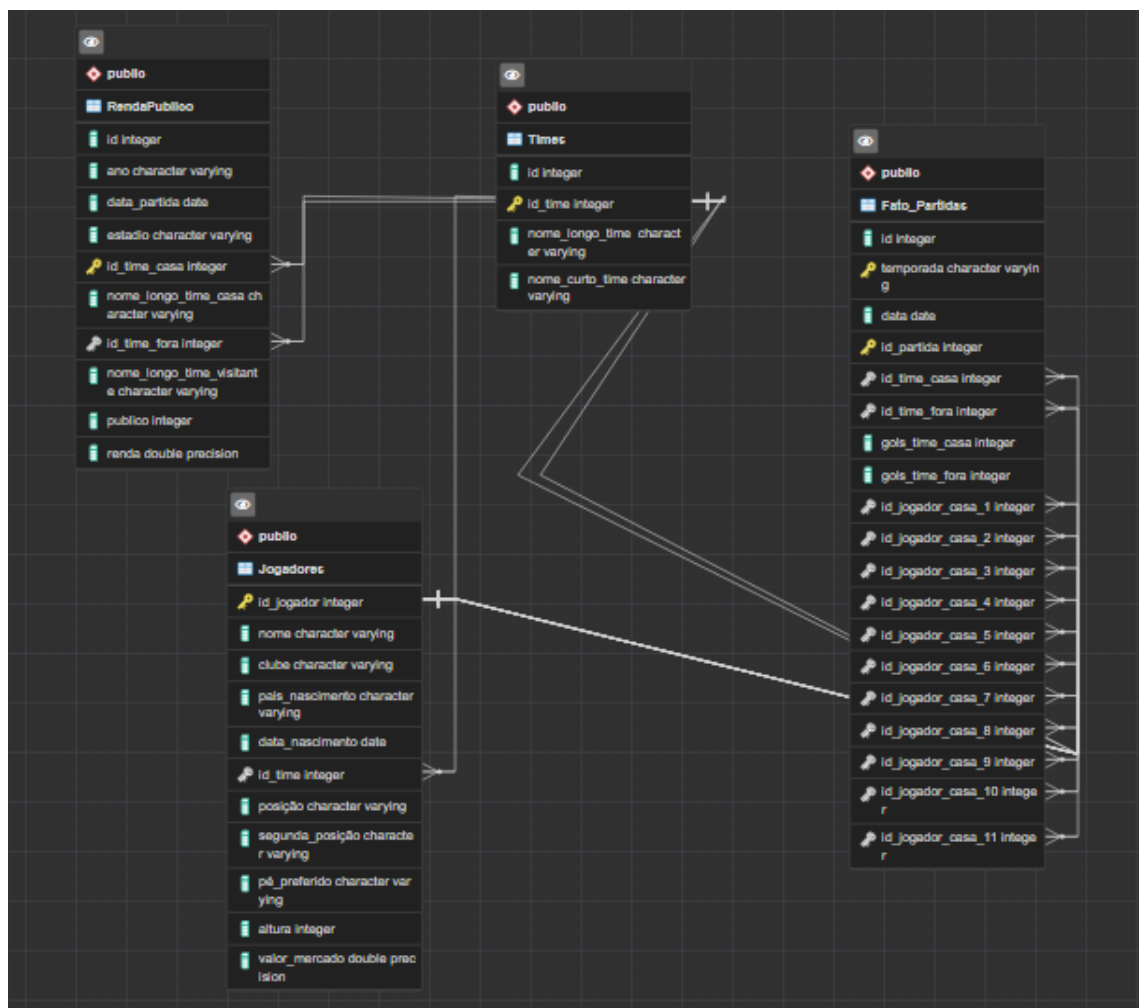
Quantos gols foram marcados fora de casa e dentro de casa na temporada?

Maior público necessariamente vai significar mais renda?

Jogadores mais velhos estão com um valor de mercado maior que nossos jogadores mais jovens?

Qual a distribuição no mapa do local em que nossos jogadores nasceram e seu valor de mercado?

4. Modelo Estrela:

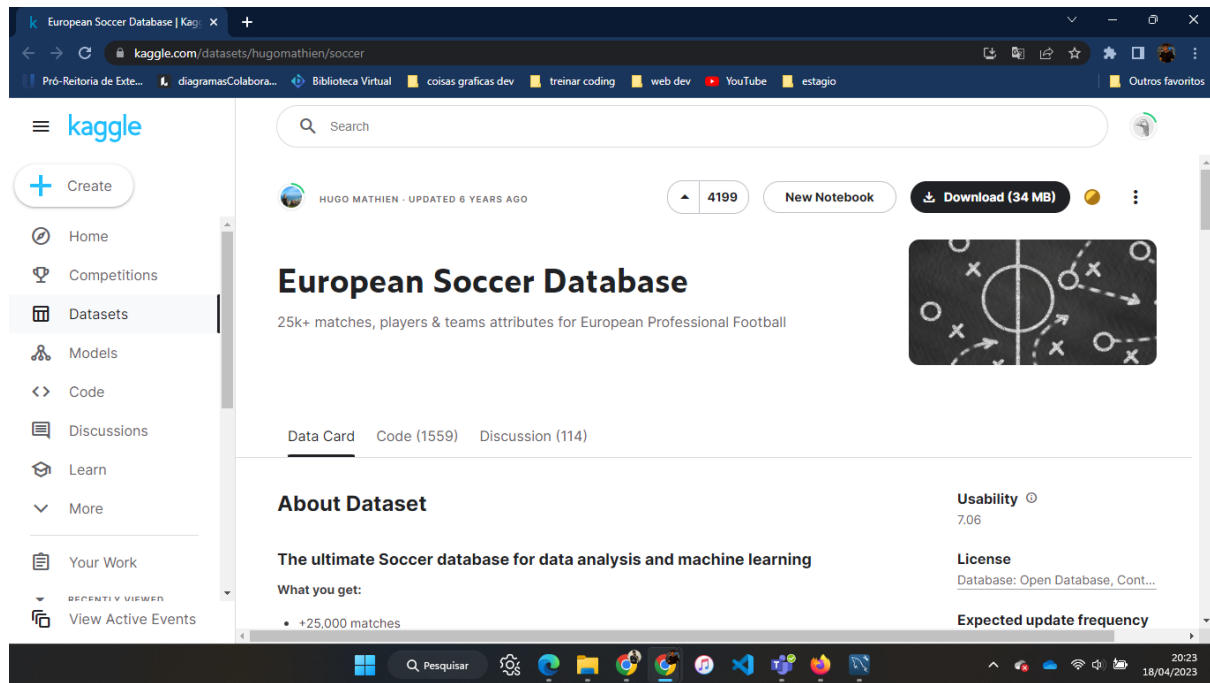


5. Obtenção das fontes de dados:

- Arquivo Sqlite (tabela relacional):

Tabela contendo todos os dados das partidas (tabela fato)

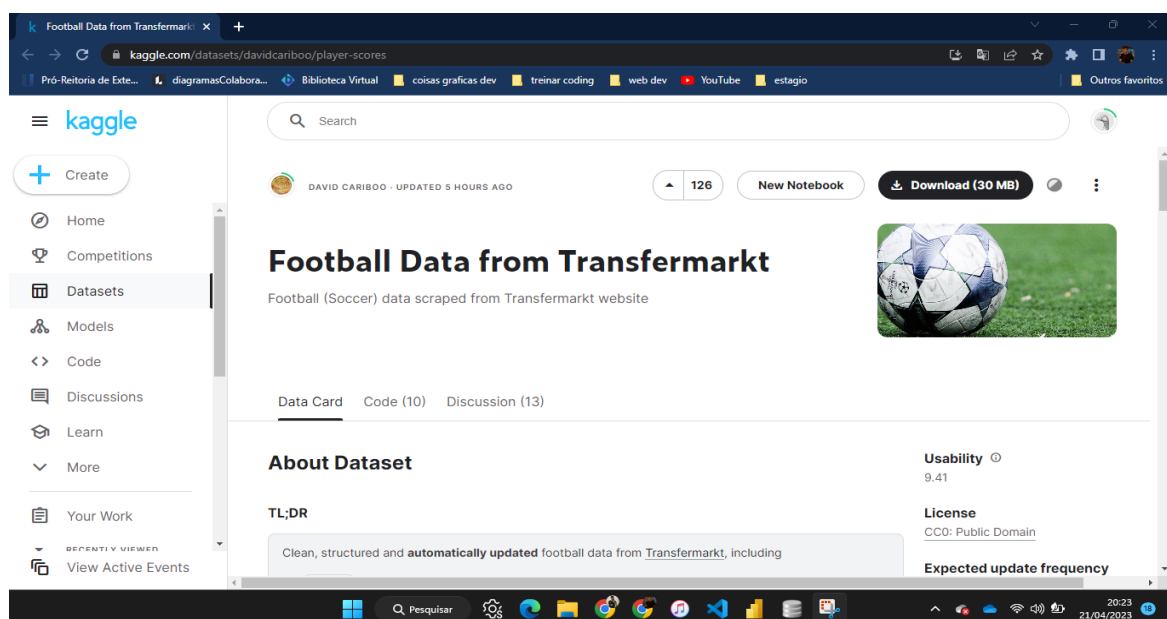
<https://www.kaggle.com/datasets/hugomathien/soccer>



- Arquivo Csv:

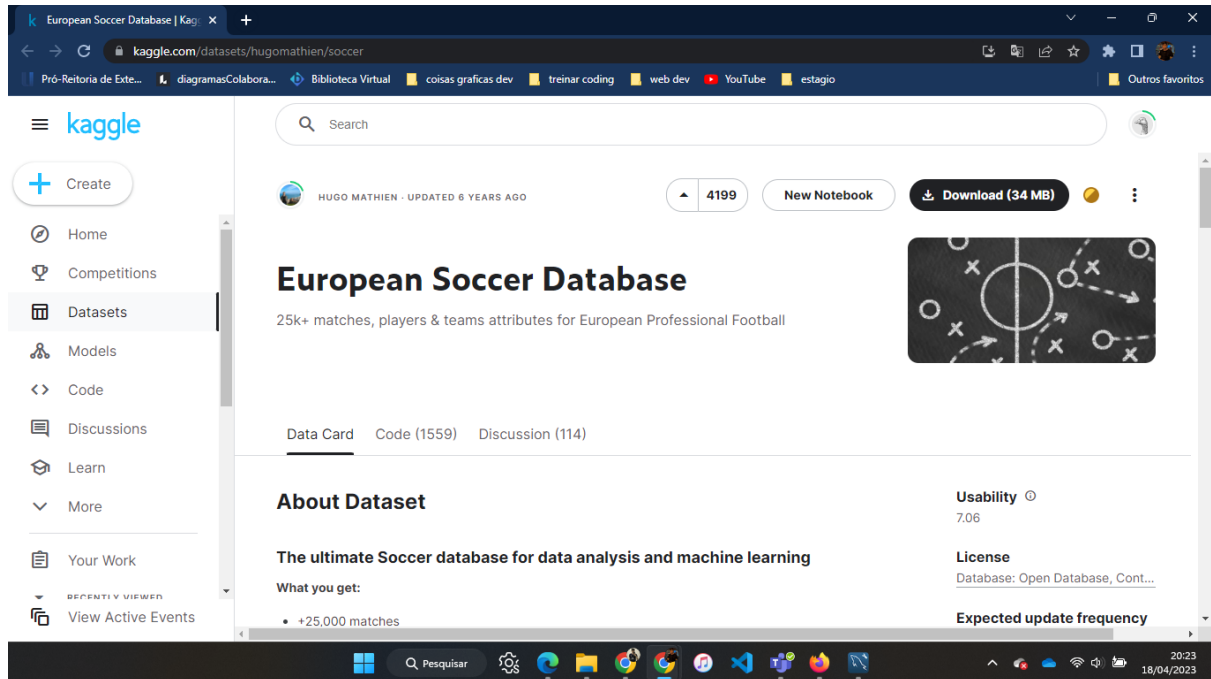
(<https://www.kaggle.com/datasets/davidcariboo/player-scores>)

quando baixa vem mais de um csv, escolhi somente o de jogadores (contendo dados de todos os jogadores) ⇒ montar a nossa equipe



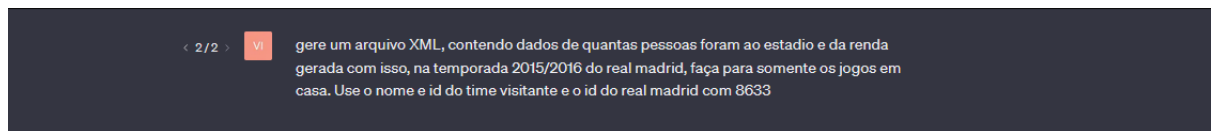
- Arquivo Json:

Foi pegado da mesma Database de onde tirei a tabela relacional, porém, é outra tabela diferente contendo informações sobre todos os times. Essa tabela foi transformada para Json (vai ser útil para saber com quem o real madrid jogou)

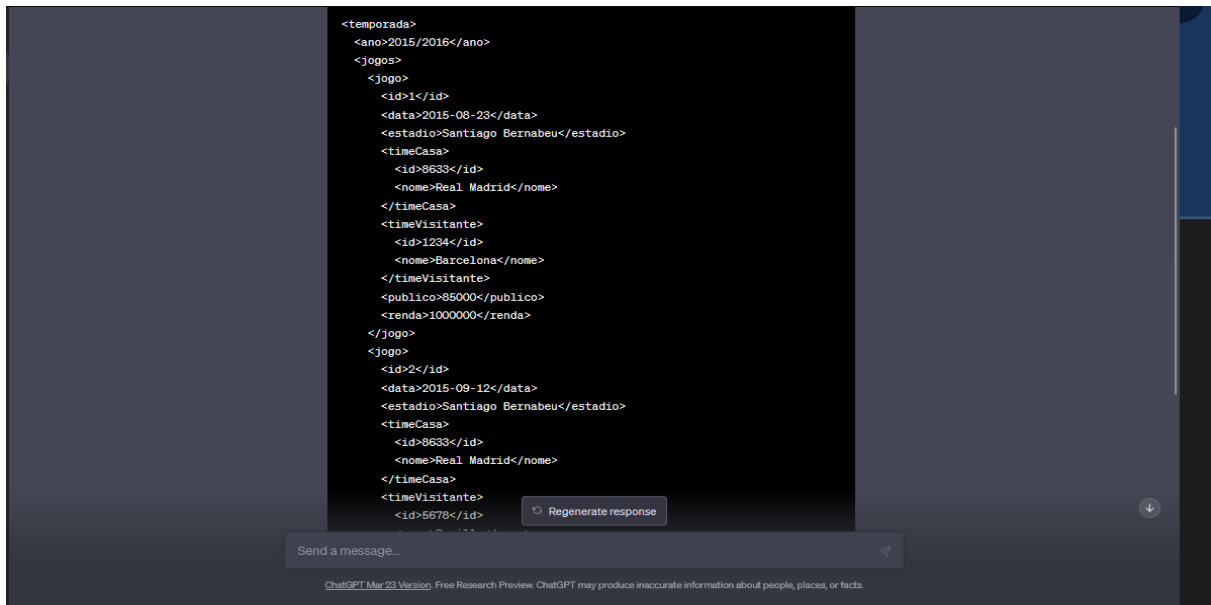


- arquivo XML:

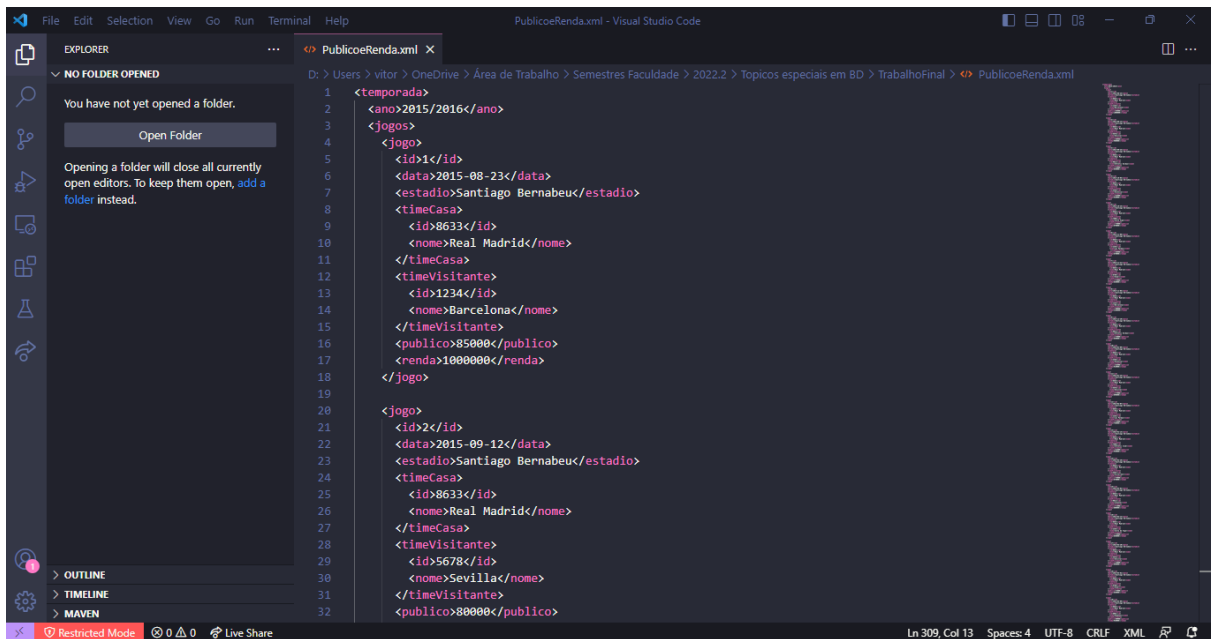
Gerado pelo chat gpt contendo informações sobre público no estádio e renda nos jogos em casa (gerou para 19 jogos, mesma quantidade de jogos em casa da tabela fato Partidas)



arquivo XML gerado, foi necessário alterar o id do time visitante para o que temos no Json de time



(foi salvo em um arquivo .xml, para ser importado ao power BI), arquivo final com 310 linhas



6. Partindo para o processo de ETL:

- Tabela sql das partidas ⇒ assim que ela veio = tabela fato do modelo estrela

The screenshot shows the D8 Browser for SQLite interface. The main window displays a table named 'Match' with the following columns: id, country_id, league_id, season, stage, date, match_api_id, home_team_api_id, away_team_api_id, and h. The table contains 17 rows of data, representing matches from the 2008/2009 season. The interface includes a menu bar (Arquivo, Editar, Exibir, Ferramentas, Ajuda) and a toolbar with various database operations. On the right, there is a panel for editing the selected cell, showing the 'Remoto' (Remote) connection details and the 'Identidade' (Identity) dropdown menu.

importada para o powerBI

The screenshot shows the Power Query Editor interface. The main window displays the 'Match' table, which has been loaded from a database. The table has 17 rows and 10 columns. The interface includes a menu bar (Arquivo, Página Inicial, Transformar, Adicionar Coluna, Exibição, Ferramentas, Ajuda) and a toolbar with various data transformation options. On the right, there is a panel for configuring the query, showing the 'Propriedades' (Properties) and 'Etapas Aplicadas' (Applied Steps) sections. The 'Etapas Aplicadas' section shows the 'Tipo Alterado' (Changed Type) step, indicating that the data has been transformed into a specific format.

Tabela enorme com mais de 30.000 linhas contendo partidas desde 2008 até 2016 ⇒ primeiramente selecionei apenas as linhas com as partidas mais atuais (deixei só 2015 e 2016)

85 COLUNAS, 199+ LINHAS Criação de perfil de coluna com base nas primeiras 1000 linhas

VISUALIZAÇÃO BAIXADA À(S) 22:26

após selecionei apenas o país que nos interessa (país do Real madrid = Espanha = country_id = 21518)

já diminuimos para 380 partidas (380 linhas na tabela) ⇒ ficamos apenas com os jogos acontecidos na espanha,

essas colunas vieram bugadas da tabelas (foram removidas)

30 COLUNAS, 380 LINHAS Criação de perfil de coluna com base nas primeiras 1000 linhas

VISUALIZAÇÃO BAIXADA À(S) 22:44

1. Colunas desnecessárias e bugadas foram removidas
 - 1.1. id dos 11 jogadores adversários
 - 1.2. gols
 - 1.3. chutes dados e recebidos
 - 1.4. faltas
 - 1.5. cartões
 - 1.6. cruzamentos
 - 1.7. escanteios
 - 1.8. stage
 - 1.9. posse de bola
 - 1.10. id_pais
 - 1.11. id_liga

Nome das colunas foram traduzidas

TrabalhoFinal - Editor do Power Query

Arquivo | Página Inicial | Transformar | Adicionar Coluna | Exibição | Ferramentas | Ajuda

Fechar e Aplicar | Nova Fonte | Fontes Recentes | Inserir Dados | Configurações da fonte de dados | Gerenciar Parâmetros | Atualizar Visualização | Propriedades | Escolher Colunas | Remover Colunas | Reduzir Linhas | Dividir Coluna | Agrupar por | Tipo de Dados: Número Inteiro | Usar a Primeira Linha como Cabeçalho | Combinar | Análise de Texto | Pesquisa Visual | Azure Machine Learning Insights da IA

Consultas [4] | = Table.TransformColumnTypes(#"Colunas Removidas2",{{"date", type date}})

	id	temporada	date	id_partida	id_time_casa	id_time_n
1	24178	2015/2016	23/08/2015	2030082	8581	
2	24179	2015/2016	22/08/2015	2030083	9906	
3	24180	2015/2016	21/08/2015	2030084	9864	
4	24181	2015/2016	23/08/2015	2030085	8315	
5	24182	2015/2016	24/08/2015	2030086	7878	
6	24183	2015/2016	22/08/2015	2030087	8558	
7	24184	2015/2016	23/08/2015	2030088	8603	
8	24185	2015/2016	23/08/2015	2030089	9869	
9	24186	2015/2016	22/08/2015	2030090	9783	
10	24187	2015/2016	22/08/2015	2030091	8370	
11	24188	2015/2016	01/11/2015	2030172	8372	
12	24189	2015/2016	31/10/2015	2030173	8305	
13	24190	2015/2016	31/10/2015	2030174	10205	
14	24191	2015/2016	31/10/2015	2030175	8633	
15	24192	2015/2016	31/10/2015	2030176	8560	
16	24193	2015/2016	31/10/2015	2030177	10267	
17	24194	2015/2016	30/10/2015	2030178	9783	
18	24195	2015/2016	01/11/2015	2030179	9869	
19	24196	2015/2016	01/11/2015	2030180	8603	
20	24197	2015/2016	01/11/2015	2030181	8558	
21	24198	2015/2016	07/11/2015	2030182	8372	
22	24199	2015/2016	08/11/2015	2030183	8634	
23	24200	2015/2016	08/11/2015	2030184	8302	

19 COLUNAS, 380 LINHAS Criação de perfil de coluna com base nas primeiras 1000 linhas

Config. Consulta

PROPRIEDADES

Nome

Partidas

Todas as Propriedades

ETAPAS APLICADAS

Fonte

Cabeçalhos Promovidos

Tipo Alterado

Linhas Filtradas

Colunas Renomeadas

Linhas Filtradas1

Colunas Renomeadas1

Colunas Removidas

Colunas Renomeadas2

Colunas Removidas1

Colunas Renomeadas3

Colunas Removidas2

Tipo Alterado1

13:36 23/04/2023

Dessa tabela gerei duas outras, para visualizar melhor as partidas foram e dentro de casa

tabela apenas para jogos dentro de casa:

TrabalhoFinal - Editor do Power Query

Arquivo | Página Inicial | Transformar | Adicionar Coluna | Exibição | Ferramentas | Ajuda

Fechar e Aplicar | Nova Fonte | Fontes Recentes | Inserir Dados | Configurações da fonte de dados | Gerenciar Parâmetros | Atualizar Visualização | Editor Avançado | Escolher Colunas | Remover Colunas | Reduzir Linhas | Dividir Coluna | Agrupar por | Tipo de Dados: Número Inteiro | Usar a Primeira Linha como Cabeçalho | Combinar | Análise de Texto | Pesquisa Visual | Azure Machine Learning | Insights da IA

Consultas [6]

Jogadores | Partidas | Times | PublicoeRenda | PartidasEmCasa | PartidasForaDeCasa

Table.SelectRows(*"Tipo Alterado1", each ([id_time_casa] = 8633))

	id	temporada	date	id_partida	id_time_casa	id_time_f
1	24191	2015/2016	31/10/2015	2030175	8633	
2	24210	2015/2016	21/11/2015	2030194	8633	
3	24229	2015/2016	05/12/2015	2030214	8633	
4	24248	2015/2016	20/12/2015	2030233	8633	
5	24258	2015/2016	30/12/2015	2030243	8633	
6	24279	2015/2016	09/01/2016	2030264	8633	
7	24295	2015/2016	29/08/2016	2030099	8633	
8	24305	2015/2016	17/01/2016	2030355	8633	
9	24324	2015/2016	31/01/2016	2030374	8633	
10	24343	2015/2016	13/02/2016	2030393	8633	
11	24362	2015/2016	27/02/2016	2030412	8633	
12	24381	2015/2016	05/03/2016	2030431	8633	
13	24410	2015/2016	20/03/2016	2030450	8633	
14	24429	2015/2016	09/04/2016	2030469	8633	
15	24448	2015/2016	20/04/2016	2030488	8633	
16	24479	2015/2016	08/05/2016	2030519	8633	
17	24504	2015/2016	19/09/2015	2030118	8633	
18	24523	2015/2016	26/09/2015	2030137	8633	
19	24542	2015/2016	17/10/2015	2030156	8633	

19 COLUNAS, 19 LINHAS Criação de perfil de coluna com base nas primeiras 1000 linhas

Config. Consulta

PROPRIEDADES

Nome

PartidasEmCasa

Todas as Propriedades

ETAPAS APLICADAS

Fonte

Cabeçalhos Promovidos

Tipo Alterado

Linhas Filtradas

Colunas Renomeadas

Linhas Filtradas1

Colunas Renomeadas1

Colunas Removidas

Colunas Renomeadas2

Colunas Renomeadas1

Colunas Renomeadas3

Colunas Removidas2

Tipo Alterado1

Linhas Filtradas2

17:41 23/04/2023

tabela para jogos fora de casa

TrabalhoFinal - Editor do Power Query

Arquivo | Página Inicial | Transformar | Adicionar Coluna | Exibição | Ferramentas | Ajuda

Fechar e Aplicar | Nova Fonte | Fontes Recentes | Inserir Dados | Configurações da fonte de dados | Gerenciar Parâmetros | Atualizar Visualização | Editor Avançado | Escolher Colunas | Remover Colunas | Reduzir Linhas | Dividir Coluna | Agrupar por | Tipo de Dados: Número Inteiro | Usar a Primeira Linha como Cabeçalho | Combinar | Análise de Texto | Pesquisa Visual | Azure Machine Learning | Insights da IA

Consultas [6]

Jogadores | Partidas | Times | PublicoeRenda | PartidasEmCasa | PartidasForaDeCasa

Table.SelectRows(*"Tipo Alterado1", each ([id_time_fora] = 8633))

	id	temporada	date	id_partida	id_time_casa	id_time_fora
1	24185	2015/2016	23/08/2015	2030089	9869	8633
2	24200	2015/2016	08/11/2015	2030184	8302	8633
3	24219	2015/2016	29/11/2015	2030204	8372	8633
4	24238	2015/2016	13/12/2015	2030223	10205	8633
5	24269	2015/2016	03/01/2016	2030254	10267	8633
6	24315	2015/2016	24/01/2016	2030365	8603	8633
7	24334	2015/2016	07/02/2016	2030384	7878	8633
8	24353	2015/2016	21/02/2016	2030403	9864	8633
9	24372	2015/2016	02/03/2016	2030422	8581	8633
10	24391	2015/2016	13/03/2016	2030441	8306	8633
11	24404	2015/2016	12/09/2015	2030108	8558	8633
12	24420	2015/2016	02/04/2016	2030460	8634	8633
13	24439	2015/2016	16/04/2016	2030479	8305	8633
14	24458	2015/2016	23/04/2016	2030498	8370	8633
15	24468	2015/2016	30/04/2016	2030508	8560	8633
16	24489	2015/2016	14/05/2016	2030529	9783	8633
17	24513	2015/2016	23/09/2015	2030127	8315	8633
18	24532	2015/2016	04/10/2015	2030146	9906	8633
19	24551	2015/2016	24/10/2015	2030165	9910	8633

19 COLUNAS, 19 LINHAS Criação de perfil de coluna com base nas primeiras 1000 linhas

Config. Consulta

PROPRIEDADES

Nome

PartidasForaDeCasa

Todas as Propriedades

ETAPAS APLICADAS

Fonte

Cabeçalhos Promovidos

Tipo Alterado

Linhas Filtradas

Colunas Renomeadas

Linhas Filtradas1

Colunas Renomeadas1

Colunas Removidas

Colunas Renomeadas2

Colunas Renomeadas1

Colunas Renomeadas3

Colunas Removidas2

Tipo Alterado1

Linhas Filtradas2

17:42 23/04/2023

podemos ver que a visualização dos jogos fora e dentro de casa fica bem mais fácil

- Csv somente com dados dos jogadores - tabela jogadores

Table.TransformColumnTypes(#"Cabeçalhos Promovidos",{{"player_id", Int64.Type}, {"name", type text}, {"current_club_id", Int64.Type}, {"current_club_name", type text}, {"country_of_citizenship", type text}, {"count", Int64.Type}})

player_id	name	current_club_id	current_club_name	country_of_citizenship	count
134354	Tan Raeymaekers	499	Koc Lokeren	Belgium	Belg
99946	Mohamed Camara	1095	Es Troyes Ac	Guinea	Guin
76948	Pablo Olivera	979	Moreirense Fc	Uruguay	Urugu
108372	Aliosman Aydin	38	Fortuna Dusseldorf	Turkey	Germ
78820	Jaime Alfonso Ruiz	354	Kv Mechelen	Colombia	Colom
167061	Leighton McIntosh	511	Dundee Fc	Scotland	Scoti
188076	Kevin Buchan	511	Dundee Fc	Scotland	Scoti
12840	Steven Milne	511	Dundee Fc	Scotland	Scoti
126278	Giannis Simosis	5219	Aok Kerkyra	Greece	Gree
139735	Giannis Gasiros	5219	Aok Kerkyra	Greece	Gree
39049	Kei Kamara	1123	Norwich City	Sierra Leone	Sier
46154	Chris Martin	1123	Norwich City	Scotland	Engli
104725	Steve Morrison	1123	Norwich City	Wales	Engli
30471	James Vaughan	1123	Norwich City	England	Engli
3666	Jason Roberts	1032	Fc Reading	Grenada	Engli
192366	Gozie Ugwu	1032	Fc Reading	England	Engli
12628	Noel Hunt	1032	Fc Reading	Ireland	Ireli
69653	Garath McCleary	1032	Fc Reading	Jamaica	Engli
67369	Nick Blackman	1032	Fc Reading	Barbados	Engli
46106	Adam Le Fondre	1032	Fc Reading	England	Engli
14047	Jobi McAnuff	1032	Fc Reading	Jamaica	Engli
16716	Danny Guthrie	1032	Fc Reading	England	Engli
42284	Tav Tabb	1032	Fc Reading	Ireland	Engli

- Colunas desnecessárias e bugadas foram removidas
 - url da imagem e da logo do time
 - primeiro nome
 - ultimo nome
 - apelido do jogador
 - código do país
 - nome do agente
 - data de expiração de contrato
 - país de cidadania
 - cidade de nascimento
 - última temporada
 - valor de mercado mais alto

Após todas as remoções foi realizado um filtro para termos apenas os jogadores do Real Madrid

Microsoft Power Query Editor - Editor do Power Query

Arquivo | Página Inicial | Transformar | Adicionar Coluna | Exibição | Ferramentas | Ajuda

Fechar e Aplicar | Nova Fonte | Fontes Recentes | Inserir Dados | Configurações da fonte de dados | Gerenciar Parâmetros | Atualizar Visualização | Gerenciar Consultas | Escolher Colunas | Remover Colunas | Reduzir Linhas | Dividir Coluna | Agrupar por | Substituir Valores | Combinar | Análise de Texto | Pesquisa Visual | Azure Machine Learning | Insights da IA

Consultas [4] | = Table.Sort(#\"Linhas Filtradas\",{{\"position\", Order.Ascending}})

	123 player_id	123 name	123 current_club_id	123 current_club_name	123 country_of_birth	123 date_of_b
1	537762	Sergio Arribas	418	Real Madrid	Spain	
2	948275	Álvaro Rodríguez	418	Real Madrid	Spain	
3	371998	Vinicius Junior	418	Real Madrid	Brazil	
4	458466	Franchu	418	Real Madrid	Argentina	
5	50202	Eden Hazard	418	Real Madrid	Belgium	
6	18922	Karim Benzema	418	Real Madrid	France	
7	652332	Óscar Aranda	418	Real Madrid	Spain	
8	603940	Marvin Park	418	Real Madrid	Spain	
9	296622	Marco Asensio	418	Real Madrid	Spain	
10	39381	Gareth Bale	418	Real Madrid	Wales	
11	85288	Iaco	418	Real Madrid	Spain	
12	412363	Rodrygo	418	Real Madrid	Brazil	
13	710004	Peter González	418	Real Madrid	Spain	
14	225020	Mariano Diaz	418	Real Madrid	Spain	
15	291417	Ferland Mendy	418	Real Madrid	France	
16	59016	David Alaba	418	Real Madrid	Austria	
17	58884	Nacho Fernández	418	Real Madrid	Spain	
18	212863	Jaime Sánchez	418	Real Madrid	Spain	
19	86202	Antonio Rüdiger	418	Real Madrid	Germany	
20	129483	Derik Osede	418	Real Madrid	Spain	
21	251896	Jesús Vallejo	418	Real Madrid	Spain	
22	629511	Pablo Ramón	418	Real Madrid	Spain	
23	44801	Marcelo	418	Real Madrid	Brazil	

15 COLUNAS, 46 LINHAS | Criação de perfil de coluna com base nas primeiras 1000 linhas

Config. Consulta

PROPRIEDADES

Nome

Jogadores

Todas as Propriedades

ETAPAS APLICADAS

Fonte

Cabeçalhos Promovidos

Tipo Alterado

Colunas Removidas

Linhas Filtradas

Linhas Classificadas

Colunas Reordenadas

Colunas Removidas1

Colunas Renomeadas

Colunas Removidas2

Valor Substituído

17:25 23/04/2023

nome das colunas foram traduzidos e id do clube foi alterado para bater como o id do real madrid na tabela fato partidas (foi de 418 para 8633)

Microsoft Power Query Editor - Editor do Power Query

Arquivo | Página Inicial | Transformar | Adicionar Coluna | Exibição | Ferramentas | Ajuda

Fechar e Aplicar | Nova Fonte | Fontes Recentes | Inserir Dados | Configurações da fonte de dados | Gerenciar Parâmetros | Atualizar Visualização | Gerenciar Consultas | Escolher Colunas | Remover Colunas | Reduzir Linhas | Dividir Coluna | Agrupar por | Substituir Valores | Combinar | Análise de Texto | Pesquisa Visual | Azure Machine Learning | Insights da IA

Consultas [4] | = Table.ReorderColumns(#\"Valor Substituído\",{\"id_jogador\", \"id_clube\", \"clube\", \"país de nascimento\", \"data de n\"})

	123 id_jogador	123 id_clube	123 clube	123 nome	123 país de nascimento	123 data de n
1	537762	8633	Real Madrid	Sergio Arribas	Spain	
2	948275	8633	Real Madrid	Álvaro Rodríguez	Spain	
3	371998	8633	Real Madrid	Vinicius Junior	Brazil	
4	458466	8633	Real Madrid	Franchu	Argentina	
5	50202	8633	Real Madrid	Eden Hazard	Belgium	
6	18922	8633	Real Madrid	Karim Benzema	France	
7	652332	8633	Real Madrid	Óscar Aranda	Spain	
8	603940	8633	Real Madrid	Marvin Park	Spain	
9	296622	8633	Real Madrid	Marco Asensio	Spain	
10	39381	8633	Real Madrid	Gareth Bale	Wales	
11	85288	8633	Real Madrid	Iaco	Spain	
12	412363	8633	Real Madrid	Rodrygo	Brazil	
13	710004	8633	Real Madrid	Peter González	Spain	
14	225020	8633	Real Madrid	Mariano Diaz	Spain	
15	291417	8633	Real Madrid	Ferland Mendy	France	
16	59016	8633	Real Madrid	David Alaba	Austria	
17	58884	8633	Real Madrid	Nacho Fernández	Spain	
18	212863	8633	Real Madrid	Jaime Sánchez	Spain	
19	86202	8633	Real Madrid	Antonio Rüdiger	Germany	
20	129483	8633	Real Madrid	Derik Osede	Spain	
21	251896	8633	Real Madrid	Jesús Vallejo	Spain	
22	629511	8633	Real Madrid	Pablo Ramón	Spain	
23	44801	8633	Real Madrid	Marcelo	Brazil	

11 COLUNAS, 46 LINHAS | Criação de perfil de coluna com base nas primeiras 1000 linhas

Config. Consulta

PROPRIEDADES

Nome

Jogadores

Todas as Propriedades

ETAPAS APLICADAS

Fonte

Cabeçalhos Promovidos

Tipo Alterado

Colunas Removidas

Linhas Filtradas

Linhas Classificadas

Colunas Reordenadas

Colunas Removidas1

Colunas Renomeadas

Colunas Removidas2

Valor Substituído

Colunas Reordenadas1

17:25 23/04/2023

dados do csv foram importados e salvos no PowerBI

- Arquivo Json

(como já explicado anteriormente, foi pegado uma tabela diferente, tabela com informações dos times e transformada para o arquivo Json)

Consultas [3]

players
Match
Team

Fechar e Aplicar
Nova Fonte
Fontes Recentes
Inserir Dados
Configurações da fonte de dados
Gerenciar Parâmetros
Atualizar Visualização
Propriedades Editor Avançado
Escolher Colunas
Remover Colunas
Reduzir Linhas
Dividir Coluna
Agrupar por
Tipo de Dados: Número Inteiro
Usar a Primeira Linha como Cabeçalho
Substituir Valores
Transformar

Config. Consulta

PROPRIEDADES
Nome
Team
Todas as Propriedades

ETAPAS APLICADAS
Fonte
Convertido para Tabela
Column1 Expandido
Tipo Alterado

id	team_api_id	team_fifa_api_id	team_long_name	team_short_name
1	1	9987	KRC Genk	GEN
2	2	9993	Beerschot AC	BAC
3	3	10000	SV Zulte-Waregem	ZUL
4	4	9994	Sporting Lokeren	LOK
5	5	9984	KSV Cerole Brugge	CEB
6	6	8635	RSC Anderlecht	AND
7	7	9991	KAA Gent	GEN
8	8	9998	RAEC Mons	MON
9	9	7947	FCV Dender EH	DEN
10	10	9985	Standard de Liège	STL
11	11	8203	KV Mechelen	MEC
12	12	8342	Club Brugge KV	CLB
13	13	9999	KSV Roeselare	ROS
14	14	8571	KV Kortrijk	KOR
15	15	4049	Tubize	TUB
16	16	111560	Royal Excel Mouscron	MOU
17	17	10001	KVC Westerlo	WES
18	18	9986	Sporting Charleroi	CHA
19	614	9997	Sint-Truidense VV	STT
20	1034	9989	Lierse SK	LIE
21	1042	6351	KAS Eupen	EUP
22	1513	1773	Oud-Heverlee Leuven	O-H
23	2004	8475	Waasland-Beveren	WAA

5 COLUNAS, 299 LINHAS Criação de perfil de coluna com base nas primeiras 1000 linhas

VISUALIZAÇÃO BAIXADA À(S) 1335

foi removida a coluna do team_fifa_api_id

Consultas [3]

players
Match
Team

Fechar e Aplicar
Nova Fonte
Fontes Recentes
Inserir Dados
Configurações da fonte de dados
Gerenciar Parâmetros
Atualizar Visualização
Propriedades Editor Avançado
Escolher Colunas
Remover Colunas
Reduzir Linhas
Dividir Coluna
Agrupar por
Tipo de Dados: Número Inteiro
Usar a Primeira Linha como Cabeçalho
Substituir Valores
Transformar

Config. Consulta

PROPRIEDADES
Nome
Team
Todas as Propriedades

ETAPAS APLICADAS
Fonte
Convertido para Tabela
Column1 Expandido
Tipo Alterado
Colunas Removidas
Linhas Classificadas

id	team_api_id	team_long_name	team_short_name
1	16848	1. FC Kaiserslautern	KAI
2	15624	1. FC Köln	FCK
3	16239	1. FC Nürnberg	NUR
4	16243	1. FSV Mainz 05	MAI
5	11817	AC Ajaccio	AJA
6	11074	AC Arles-Avignon	ARL
7	49116	AC Bellinzona	BEL
8	26560	ADO Den Haag	HAA
9	9537	AJ Auxerre	AUX
10	9547	AS Monaco	MON
11	9549	AS Nancy-Lorraine	NAN
12	9556	AS Saint-Étienne	ETI
13	26563	AZ	ALK
14	39393	Aberdeen	ABE
15	35292	Académica de Coimbra	ACA
16	26552	Ajax	AJA
17	35291	Amadora	AMA
18	14868	Angers SCO	ANG
19	31455	Arka Gdynia	ARK
20	3459	Arsenal	ARS
21	3465	Aston Villa	AVL
22	20513	Atlanta	ATA
23	43051	Athletic Club de Bil...	BIL

4 COLUNAS, 299 LINHAS Criação de perfil de coluna com base nas primeiras 1000 linhas

VISUALIZAÇÃO BAIXADA À(S) 1337

nome das colunas foram traduzidos:

4 COLUNAS, 299 LINHAS Criação de perfil de coluna com base nas primeiras 1000 linhas

json foi importado e salvo no PowerBI

- Arquivo XML (arquivo contendo renda e público nos 19 jogos em casa do real madrid)

10 COLUNAS, 19 LINHAS Criação de perfil de coluna com base nas primeiras 1000 linhas

colunas foram renomeadas:

Microsoft Power Query Editor - Editor do Power Query

Arquivo | Página Inicial | Transformar | Adicionar Coluna | Exibição | Ferramentas | Ajuda

Fechar e Aplicar | Nova Fonte | Fontes Recentes | Inserir Dados | Configurações da fonte de dados | Gerenciar Parâmetros | Atualizar Visualização | Editor Avançado | Escolher Colunas | Remover Colunas | Reduzir Linhas | Dividir Coluna | Agrupar por | Tipo de Dados: Qualquer | Usar a Primeira Linha como Cabeçalho | Substituir Valores | Combinar | Análise de Texto | Pesquisa Visual | Azure Machine Learning | Insights da IA

Consultas [4] | Jogadores | Partidas | Times | PublicoeRenda

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19

ano	id	data_partida	estadio	id_time_casa	nome_jogador
2015/2016	1	2015-08-23	Santiago Bernabeu	8633	Real Madri
2015/2016	2	2015-09-12	Santiago Bernabeu	8633	Real Madri
2015/2016	3	2015-09-23	Santiago Bernabeu	8633	Real Madri
2015/2016	4	2015-10-04	Santiago Bernabeu	8633	Real Madri
2015/2016	5	2015-10-25	Santiago Bernabeu	8633	Real Madri
2015/2016	6	2015-11-08	Santiago Bernabeu	8633	Real Madri
2015/2016	7	2015-11-21	Santiago Bernabeu	8633	Real Madri
2015/2016	8	2015-12-05	Santiago Bernabeu	8633	Real Madri
2015/2016	9	2015-12-20	Santiago Bernabeu	8633	Real Madri
2015/2016	10	2016-01-09	Santiago Bernabeu	8633	Real Madri
2015/2016	11	2016-01-24	Santiago Bernabeu	8633	Real Madri
2015/2016	12	2016-02-07	Santiago Bernabeu	8633	Real Madri
2015/2016	13	2016-02-14	Santiago Bernabeu	8633	Real Madri
2015/2016	14	2016-02-28	Santiago Bernabeu	8633	Real Madri
2015/2016	15	2016-03-02	Santiago Bernabeu	8633	Real Madri
2015/2016	16	2016-03-12	Santiago Bernabeu	8633	Real Madri
2015/2016	17	2016-03-20	Santiago Bernabeu	8633	Real Madri
2015/2016	18	2016-04-02	Santiago Bernabeu	8633	Real Madri
2015/2016	19	2016-04-09	Santiago Bernabeu	8633	Real Madri

10 COLUNAS, 19 LINHAS Criação de perfil de coluna com base nas primeiras 1000 linhas

Config. Consulta

PROPRIEDADES

Nome

PublicoeRenda

Todas as Propriedades

ETAPAS APLICADAS

Fonte

Tipo Alterado

jogos.jogo Expandido

jogos.jogo Expandido

jogos.jogo.timeCasa Expandido

jogos.jogo.timeVisitante Expa...

Colunas Renomeadas

11:58 23/04/2023

valores de renda foram filtrados para euro

Microsoft Power Query Editor - Editor do Power Query

Arquivo | Página Inicial | Transformar | Adicionar Coluna | Exibição | Ferramentas | Ajuda

Agrupar por | Usar a Primeira Linha como Cabeçalho | Contar Linhas | Transpor | Inverter Linhas | Detectar Tipo de Dados | Renomear | Tipo de Dados: Número decimal fixo | Dividir Coluna | Formato | Extrair | Misturar Colunas | Estatísticas Padrão | Científico | Trigonometria | Arredondamento | Informações | Data | Hora | Duração | Coluna de Data e Hora | R | Py | Executar script R | Executar script Python | Scripts

Consultas [4] | Jogadores | Partidas | Times | PublicoeRenda

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19

nome_longo_time_casa	id_time_fora	nome_longo_time_visitante	publico	renda
Real Madrid	8634	Barcelona	85000	1.000.000,00
Real Madrid	8302	Sevilla	80000	900.000,00
Real Madrid	8315	Athletic Bilbao	75000	850.000,00
Real Madrid	10267	Valencia	82000	920.000,00
Real Madrid	9910	Celta Vigo	78000	870.000,00
Real Madrid	8306	Las Palmas	68000	670.000,00
Real Madrid	8372	Eibar	74000	830.000,00
Real Madrid	8634	Barcelona	90000	1.100.000,00
Real Madrid	8315	Athletic Bilbao	76000	850.000,00
Real Madrid	10267	Valencia	82000	910.000,00
Real Madrid	8302	Sevilla	77000	890.000,00
Real Madrid	8306	Las Palmas	73000	750.000,00
Real Madrid	8315	Athletic Club	81000	940.000,00
Real Madrid	10205	Villarreal	78000	870.000,00
Real Madrid	8581	Levante	71000	800.000,00
Real Madrid	9910	Celta de Vigo	75000	850.000,00
Real Madrid	8302	Sevilla	78000	890.000,00
Real Madrid	8372	Eibar	74000	850.000,00
Real Madrid	8558	Espanyol	80000	930.000,00

10 COLUNAS, 19 LINHAS Criação de perfil de coluna com base nas primeiras 1000 linhas

Config. Consulta

PROPRIEDADES

Nome

PublicoeRenda

Todas as Propriedades

ETAPAS APLICADAS

Fonte

Tipo Alterado

jogos.jogo Expandido

jogos.jogo Expandido

jogos.jogo.timeCasa Expandido

jogos.jogo.timeVisitante Expa...

Colunas Renomeadas

Tipo Alterado1

Tipo Alterado com Localidade

Tipo Alterado2

Tipo Alterado com Localidade1

Tipo Alterado3

12:03 23/04/2023

xml importado para powerBI

10. Dashboard que responda às perguntas:

(disponível para acesso no site: <https://trabalho-engenharia-de-dados.vercel.app/>)

Relembrando as perguntas:

Quantos gols foram marcados fora de casa e dentro de casa na temporada?

Maior público necessariamente vai significar maior renda?

Jogadores mais velhos estão com um valor de mercado maior que nossos jogadores mais jovens?

Qual a distribuição no mapa do local em que nossos jogadores nasceram e seu valor de mercado?

