	Pós-Graduação em Ciência de Dados
	Disciplina: DATA WAREHOUSE
	Professor: Alberto Tavares da Silva
Avaliação Parte 01: Construção de Data warehouse.	
Prazo: 09 Set 2023.	
Valor: 05 (cinco) pontos.	

## 1. MODELO RELACIONAL

O grupo deverá selecionar um dataset csv ou xls de seu interesse que permita gerar um modelo dimensional com no mínimo 4 dimensões.

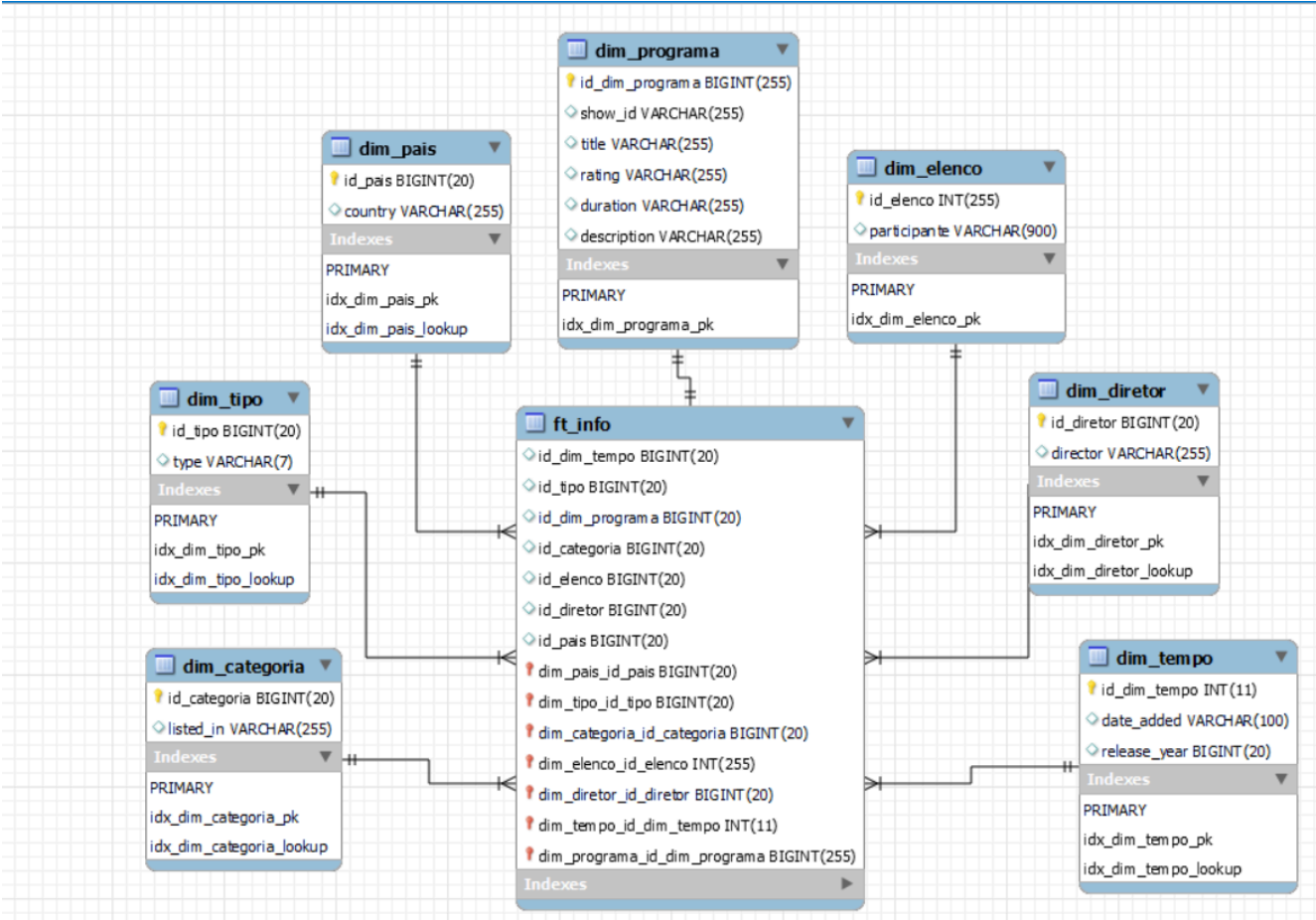
## 2. PEDIDOS

1- Utilizando o modelo acima, criar um <b>MODELO DIMENSIONAL</b> que permita uma análise de dados com operações OLAP mais efetivas. Importante: criar a dimensão tempo. <b>RESTRIÇÃO: NO MÍNIMO 4 DIMENSÕES.</b>	2 pontos
2- Gerar o arquivo de transformação ktr no PDI Pentaho.	2 pontos
3- Gerar as tabelas correspondentes com os dados no Workbench.	1 ponto

## 3. ENTREGA (UPLOAD AVA)

- Arquivo pdf com o modelo dimensional e a identificação do grupo (ver anexo).
- Arquivo Pentaho PDI com extensão “.ktr”.
- Arquivo SQL com a base de dados criada.
- Upload com os três arquivos zipados: <nome\_participantes\_grupo>.zip

Modelo Entidade Relacionamento



Fonte de dados: netflix\_titles.csv

Salvamento Automático

netflix\_titles

Pesquisar

Arquivo

Página Inicial

Inserir

Layout da Página

Fórmulas

Dados

Revisão

Exibir

Ajuda

Colar

Aptos Narrow

12

A

A

Quebrar Texto Automaticamente

Mesclar e Centralizar

Área de Transfer...

Fonte

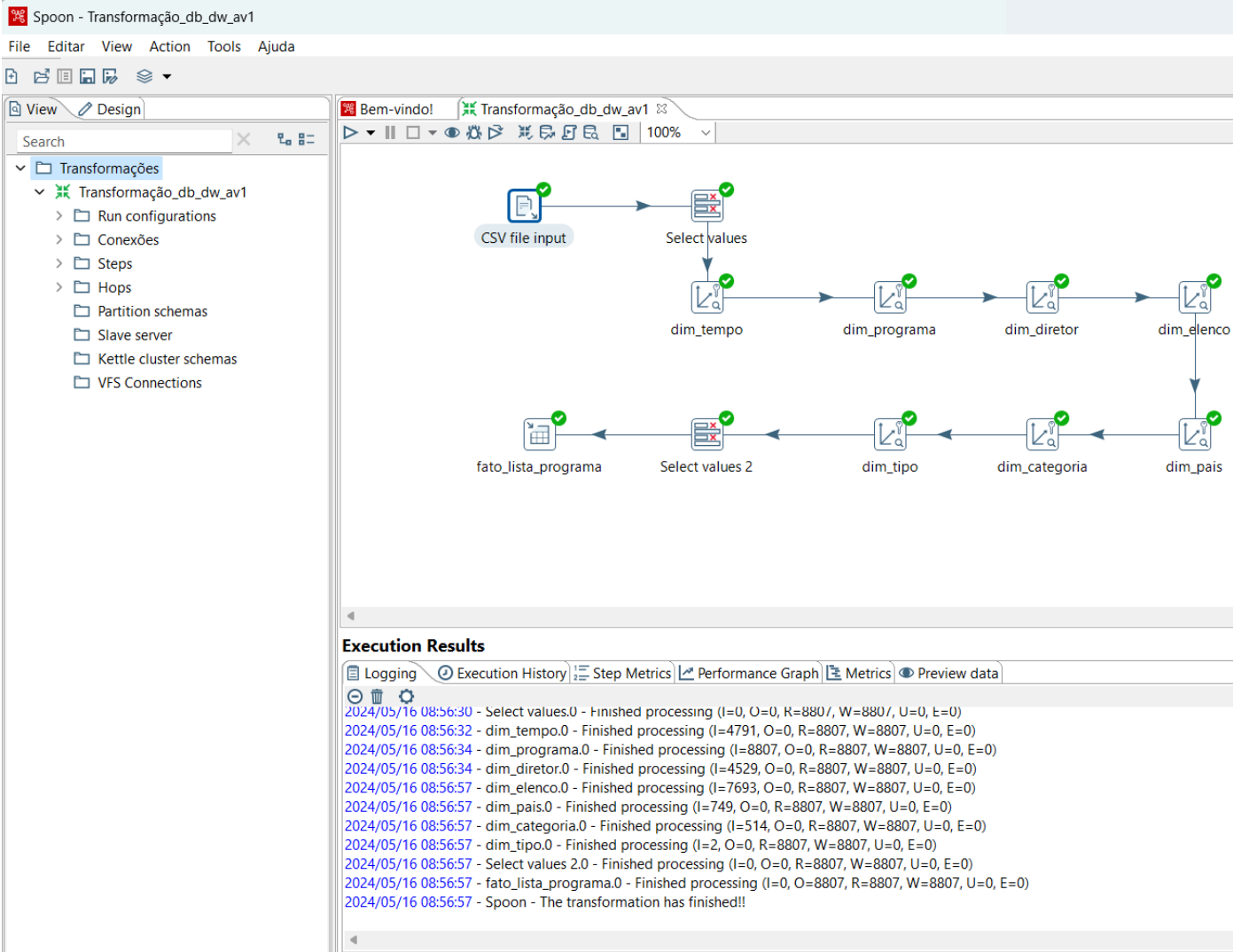
Alinhamento

L1

fx

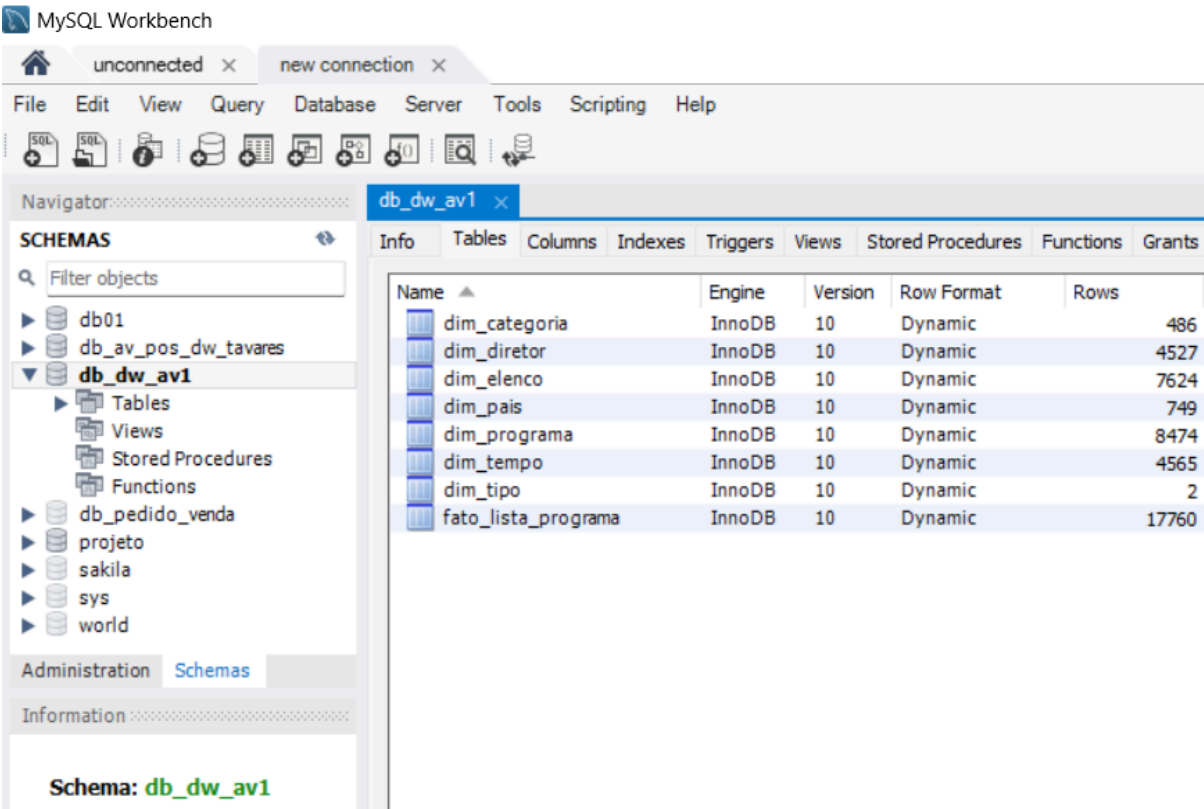
	A	B	C	D	E	F	G	H	I	J	K
1	show_id,type,title,director,cast,country,date_added,release_year,rating,duration,listed_in,description										
2	s1,	Movie,	Dick Johnson Is Dead,	Kirsten Johnson,,	United States,"	September 25, 2021",	2020,	PG-13,	90 min,	Documentaries,"	
3	s2,	TV Show,	Blood & Water,,	"Ama Qamata, Khosi Ngema, Gail Mabalane, Thabang Molaba, Dillon Windvogel, Natasha Thaha							
4	s3,	TV Show,	Ganglands,	Julien Leclercq,"	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabiha Akkari, Sofia Lesaffre, Salim Kecl						
5	s4,	TV Show,	Jailbirds New Orleans,,,"	September 24, 2021",	2021,	TV-MA,	1 Season,"	Docuseries, Reality TV",	"Feuds, flirtatior		
6	s5,	TV Show,	Kota Factory,,	"Mayur More, Jitendra Kumar, Ranjan Raj, Alam Khan, Ahsaas Channa, Revathi Pillai, Urvi Singh, Ai							

Pentaho - Extração/Transformação de dados

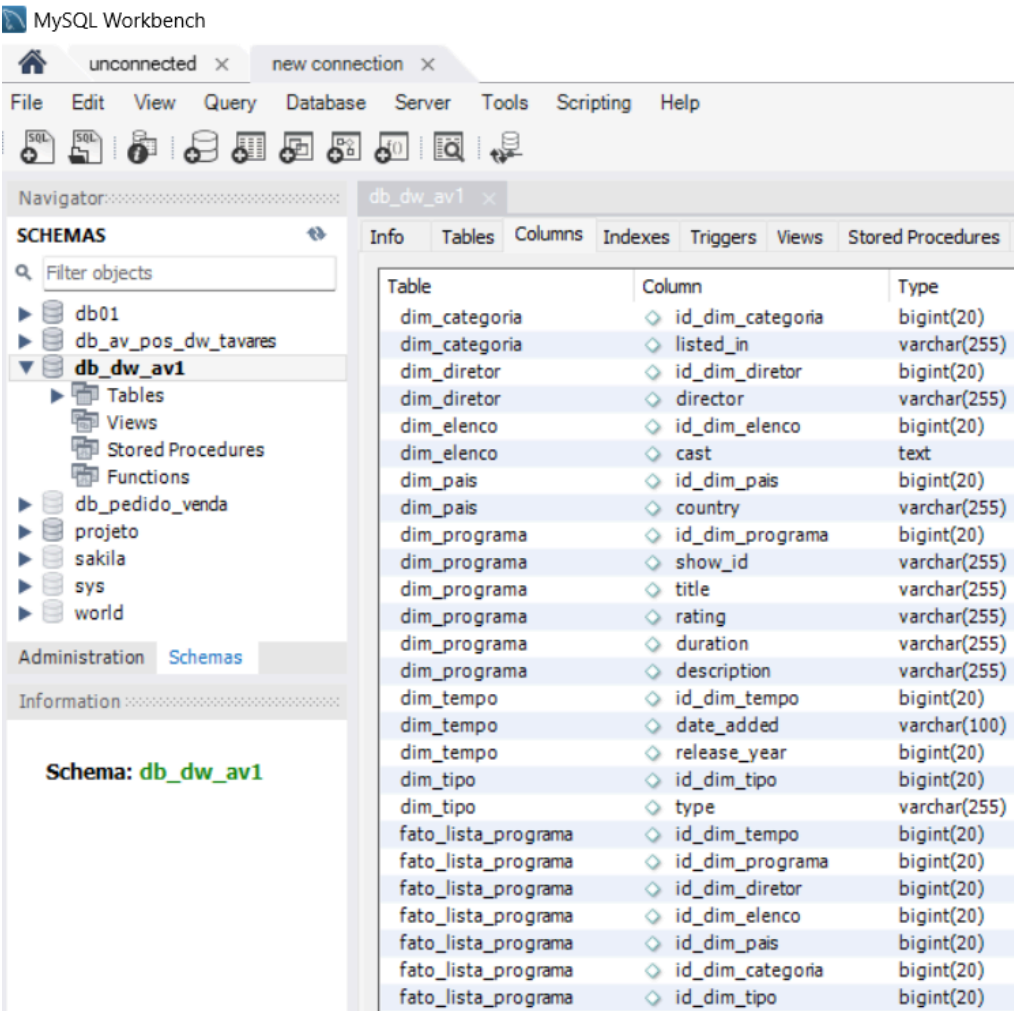


MySQL - Banco de dados

Tabelas



Atributos



## MySQL - Banco de dados - Estrutura

```
1      -- MySQL dump 10.13  Distrib 8.0.34, for Win64 (x86_64)
2      --
3      -- Host: localhost    Database: db_dw_av1
4      -- -----
5      -- Server version  5.7.23-log
6
7  •  /*!40101 SET @OLD_CHARACTER_SET_CLIENT=@@CHARACTER_SET_CLIENT */;
8  •  /*!40101 SET @OLD_CHARACTER_SET_RESULTS=@@CHARACTER_SET_RESULTS */;
9  •  /*!40101 SET @OLD_COLLATION_CONNECTION=@@COLLATION_CONNECTION */;
10 •  /*!50503 SET NAMES utf8 */;
11 •  /*!40103 SET @OLD_TIME_ZONE=@@TIME_ZONE */;
12 •  /*!40103 SET TIME_ZONE='+00:00' */;
13 •  /*!40014 SET @OLD_UNIQUE_CHECKS=@@UNIQUE_CHECKS, UNIQUE_CHECKS=0 */;
14 •  /*!40014 SET @OLD_FOREIGN_KEY_CHECKS=@@FOREIGN_KEY_CHECKS, FOREIGN_KEY_CHECKS=0 */;
15 •  /*!40101 SET @OLD_SQL_MODE=@@SQL_MODE, SQL_MODE='NO_AUTO_VALUE_ON_ZERO' */;
16 •  /*!40111 SET @OLD_SQL_NOTES=@@SQL_NOTES, SQL_NOTES=0 */;
17
18
19
20
21
22 •  DROP TABLE IF EXISTS `dim_categoria`;
23 •  /*!40101 SET @saved_cs_client      = @@character_set_client */;
24 •  /*!50503 SET character_set_client = utf8mb4 */;
25 •  CREATE TABLE `dim_categoria` (
26     `id_dim_categoria` bigint(20) NOT NULL,
27     `listed_in` varchar(255) DEFAULT NULL,
28     PRIMARY KEY (`id_dim_categoria`),
29     UNIQUE KEY `idx_dim_categoria_pk` (`id_dim_categoria`),
30     KEY `idx_dim_categoria_lookup` (`listed_in`)
31 ) ENGINE=InnoDB DEFAULT CHARSET=latin1;
32 •  /*!40101 SET character_set_client = @saved_cs_client */;
33
34  --
35  -- Dumping data for table `dim_categoria`
36  --
37
38 •  LOCK TABLES `dim_categoria` WRITE;
39 •  /*!40000 ALTER TABLE `dim_categoria` DISABLE KEYS */;
40 •  INSERT INTO `dim_categoria` VALUES (39,'Action & Adventure'),(111,'Action & Adventure,
41 •  /*!40000 ALTER TABLE `dim_categoria` ENABLE KEYS */;
42 •  UNLOCK TABLES;
```



## MySQL - Banco de dados - Estrutura

```
48 • DROP TABLE IF EXISTS `dim_diretor`;
49 • /*!40101 SET @saved_cs_client      = @@character_set_client */;
50 • /*!50503 SET character_set_client = utf8mb4 */;
51 • CREATE TABLE `dim_diretor` (
52     `id_dim_diretor` bigint(20) NOT NULL,
53     `director` varchar(255) DEFAULT NULL,
54     PRIMARY KEY (`id_dim_diretor`),
55     UNIQUE KEY `idx_dim_diretor_pk` (`id_dim_diretor`),
56     KEY `idx_dim_diretor_lookup` (`director`)
57 ) ENGINE=InnoDB DEFAULT CHARSET=latin1;
58 • /*!40101 SET character_set_client = @saved_cs_client */;
59
60 --
61 -- Dumping data for table `dim_diretor`
62 --
63
64 • LOCK TABLES `dim_diretor` WRITE;
65 • /*!40000 ALTER TABLE `dim_diretor` DISABLE KEYS */;
66 • INSERT INTO `dim_diretor` VALUES (2,NULL),(1908,'A. L. Vijay'),(1356,'A. Raajdheep'),
67 • /*!40000 ALTER TABLE `dim_diretor` ENABLE KEYS */;
68 • UNLOCK TABLES;

74 • DROP TABLE IF EXISTS `dim_elenco`;
75 • /*!40101 SET @saved_cs_client      = @@character_set_client */;
76 • /*!50503 SET character_set_client = utf8mb4 */;
77 • CREATE TABLE `dim_elenco` (
78     `id_dim_elenco` bigint(20) NOT NULL,
79     `cast` text,
80     PRIMARY KEY (`id_dim_elenco`),
81     UNIQUE KEY `idx_dim_elenco_pk` (`id_dim_elenco`)
82 ) ENGINE=InnoDB DEFAULT CHARSET=latin1;
83 • /*!40101 SET character_set_client = @saved_cs_client */;
84
85 --
86 -- Dumping data for table `dim_elenco`
87 --
88
89 • LOCK TABLES `dim_elenco` WRITE;
90 • /*!40000 ALTER TABLE `dim_elenco` DISABLE KEYS */;
91 • INSERT INTO `dim_elenco` VALUES (1,NULL),(2,'Ama Qamata, Khosi Ngema, Gail Mabalane,
92 • /*!40000 ALTER TABLE `dim_elenco` ENABLE KEYS */;
93 • UNLOCK TABLES;
```

## MySQL - Banco de dados - Estrutura

```
99 • DROP TABLE IF EXISTS `dim_pais`;
100 • /*!40101 SET @saved_cs_client      = @@character_set_client */;
101 • /*!50503 SET character_set_client = utf8mb4 */;
102 • CREATE TABLE `dim_pais` (
103     `id_dim_pais` bigint(20) NOT NULL,
104     `country` varchar(255) DEFAULT NULL,
105     PRIMARY KEY (`id_dim_pais`),
106     UNIQUE KEY `idx_dim_pais_pk` (`id_dim_pais`),
107     KEY `idx_dim_pais_lookup` (`country`)
108 ) ENGINE=InnoDB DEFAULT CHARSET=latin1;
109 • /*!40101 SET character_set_client = @saved_cs_client */;
110
111 --
112 -- Dumping data for table `dim_pais`
113 --
114
115 • LOCK TABLES `dim_pais` WRITE;
116 • /*!40000 ALTER TABLE `dim_pais` DISABLE KEYS */;
117 • INSERT INTO `dim_pais` VALUES (3,NULL),(68,', France, Algeria'),(45,', South Korea'),
118 • /*!40000 ALTER TABLE `dim_pais` ENABLE KEYS */;
119 • UNLOCK TABLES;

125 • DROP TABLE IF EXISTS `dim_programa`;
126 • /*!40101 SET @saved_cs_client      = @@character_set_client */;
127 • /*!50503 SET character_set_client = utf8mb4 */;
128 • CREATE TABLE `dim_programa` (
129     `id_dim_programa` bigint(20) NOT NULL,
130     `show_id` varchar(255) DEFAULT NULL,
131     `title` varchar(255) DEFAULT NULL,
132     `rating` varchar(255) DEFAULT NULL,
133     `duration` varchar(255) DEFAULT NULL,
134     `description` varchar(255) DEFAULT NULL,
135     PRIMARY KEY (`id_dim_programa`),
136     UNIQUE KEY `idx_dim_programa_pk` (`id_dim_programa`),
137     KEY `idx_dim_programa_lookup` (`show_id`,`title`,`rating`,`duration`,`description`)
138 ) ENGINE=InnoDB DEFAULT CHARSET=latin1;
139 • /*!40101 SET character_set_client = @saved_cs_client */;
140
141 --
142 -- Dumping data for table `dim_programa`
143 --
144
145 • LOCK TABLES `dim_programa` WRITE;
146 • /*!40000 ALTER TABLE `dim_programa` DISABLE KEYS */;
147 • INSERT INTO `dim_programa` VALUES (1,'s1','Dick Johnson Is Dead','PG-13','90 min','As h
148 • INSERT INTO `dim_programa` VALUES (5637,'s5637','Casablanacas: The Man Who Loved Women',
149 • /*!40000 ALTER TABLE `dim_programa` ENABLE KEYS */;
150 • UNLOCK TABLES;
```

## MySQL - Banco de dados - Estrutura

```
156 • DROP TABLE IF EXISTS `dim_tempo`;
157 • /*!40101 SET @saved_cs_client      = @@character_set_client */;
158 • /*!50503 SET character_set_client = utf8mb4 */;
159 • CREATE TABLE `dim_tempo` (
160     `id_dim_tempo` bigint(20) NOT NULL,
161     `date_added` varchar(100) DEFAULT NULL,
162     `release_year` bigint(20) DEFAULT NULL,
163     PRIMARY KEY (`id_dim_tempo`),
164     UNIQUE KEY `idx_dim_tempo_pk` (`id_dim_tempo`),
165     KEY `idx_dim_tempo_lookup` (`date_added`,`release_year`)
166 ) ENGINE=InnoDB DEFAULT CHARSET=latin1;
167 • /*!40101 SET character_set_client = @saved_cs_client */;
168
169 --
170 -- Dumping data for table `dim_tempo`
171 --
172
173 • LOCK TABLES `dim_tempo` WRITE;
174 • /*!40000 ALTER TABLE `dim_tempo` DISABLE KEYS */;
175 • INSERT INTO `dim_tempo` VALUES (3743,NULL,2003),(3803,NULL,2008),(3967,NULL,2010),
176 • /*!40000 ALTER TABLE `dim_tempo` ENABLE KEYS */;
177 • UNLOCK TABLES;

183 • DROP TABLE IF EXISTS `dim_tipo`;
184 • /*!40101 SET @saved_cs_client      = @@character_set_client */;
185 • /*!50503 SET character_set_client = utf8mb4 */;
186 • CREATE TABLE `dim_tipo` (
187     `id_dim_tipo` bigint(20) NOT NULL,
188     `type` varchar(255) DEFAULT NULL,
189     PRIMARY KEY (`id_dim_tipo`),
190     UNIQUE KEY `idx_dim_tipo_pk` (`id_dim_tipo`),
191     KEY `idx_dim_tipo_lookup` (`type`)
192 ) ENGINE=InnoDB DEFAULT CHARSET=latin1;
193 • /*!40101 SET character_set_client = @saved_cs_client */;
194
195 --
196 -- Dumping data for table `dim_tipo`
197 --
198
199 • LOCK TABLES `dim_tipo` WRITE;
200 • /*!40000 ALTER TABLE `dim_tipo` DISABLE KEYS */;
201 • INSERT INTO `dim_tipo` VALUES (1,'Movie'),(2,'TV Show');
202 • /*!40000 ALTER TABLE `dim_tipo` ENABLE KEYS */;
203 • UNLOCK TABLES;
```



## MySQL - Banco de dados - Estrutura

```
209 • DROP TABLE IF EXISTS `fato_lista_programa`;
210 • /*!40101 SET @saved_cs_client      = @@character_set_client */;
211 • /*!50503 SET character_set_client = utf8mb4 */;
212 • CREATE TABLE `fato_lista_programa` (
213     `id_dim_tempo` bigint(20) DEFAULT NULL,
214     `id_dim_programa` bigint(20) DEFAULT NULL,
215     `id_dim_diretor` bigint(20) DEFAULT NULL,
216     `id_dim_elenco` bigint(20) DEFAULT NULL,
217     `id_dim_pais` bigint(20) DEFAULT NULL,
218     `id_dim_categoria` bigint(20) DEFAULT NULL,
219     `id_dim_tipo` bigint(20) DEFAULT NULL
220 ) ENGINE=InnoDB DEFAULT CHARSET=latin1;
221 • /*!40101 SET character_set_client = @saved_cs_client */;
222
223 --
224 -- Dumping data for table `fato_lista_programa`
225 --
226
227 • LOCK TABLES `fato_lista_programa` WRITE;
228 • /*!40000 ALTER TABLE `fato_lista_programa` DISABLE KEYS */;
229 • INSERT INTO `fato_lista_programa` VALUES (1,1,1,1,1,1,1),(2,2,2,2,2,2,2),(2,3,3,3,3,3,2),
230 • /*!40000 ALTER TABLE `fato_lista_programa` ENABLE KEYS */;
231 • UNLOCK TABLES;

232 • /*!40103 SET TIME_ZONE=@OLD_TIME_ZONE */;
233
234 • /*!40101 SET SQL_MODE=@OLD_SQL_MODE */;
235 • /*!40014 SET FOREIGN_KEY_CHECKS=@OLD_FOREIGN_KEY_CHECKS */;
236 • /*!40014 SET UNIQUE_CHECKS=@OLD_UNIQUE_CHECKS */;
237 • /*!40101 SET CHARACTER_SET_CLIENT=@OLD_CHARACTER_SET_CLIENT */;
238 • /*!40101 SET CHARACTER_SET_RESULTS=@OLD_CHARACTER_SET_RESULTS */;
239 • /*!40101 SET COLLATION_CONNECTION=@OLD_COLLATION_CONNECTION */;
240 • /*!40111 SET SQL_NOTES=@OLD_SQL_NOTES */;
241
242 -- Dump completed on 2023-09-05 19:30:40
243
```

