

File Edit View Query Database Server Tools Scripting Help

Navigator: cleaning paintings - Schema

Filter objects

cleaning paintings

- Tables
 - artist
 - artist_copy
- Views
- Stored Procedures
- Functions
- sys

Query 1 x cleaning paintings - Schema

Limit to 1000 rows

```
1 -- Data Cleaning : Removing Duplicates , Standardizing Data, Null or Blanks, Dropping Columns
2 • SELECT *
3 FROM artist;
4 -- Copying of Table ( Data copy will be the one we work with not raw data , Good Practice)
5 • CREATE TABLE artist_copy
6 LIKE artist;
7
8 • INSERT INTO artist_copy
9 SELECT *
10 FROM artist;
11
12 • SELECT *
13 FROM artist_copy;
```

Result Grid

	artist_id	full_name	first_name	middle_names	last_name	nationality	style	birth	death
▶	500	Pierre-Auguste Renoir	Pierre	Auguste	Renoir	French	Impressionist	1841	1919
	501	Alexandre Cabanel	Alexandre		Cabanel	French	Classicist	1823	1889
	502	James Ensor	James		Ensor	Belgian	Expressionist	1860	1949
	503	Maximilien Luce	Maximilien		Luce	French	Pointillist	1858	1941
	504	August Macke	August		Macke	German	Expressionist	1887	1914
	505	Thomas Gainsborough	Thomas		Gainsborough	English	Rococo	1727	1788
	506	Julien DuprÃ©	Julien		DuprÃ©	French	Naturalist	1851	1910
	507	George Catlin	George		Catlin	American	American West	1796	1872

LH

- Standardizing Data
- Data Cleaning

- Checking or Removing Duplicates
- Dropping Columns
- Changing Column title

```
15 -- Dropping First_Name , Middle_Names , Last_Name Columns And Renaming Full_Name Column
16 • ALTER TABLE artist_copy
17 DROP COLUMN first_name,
18 DROP COLUMN middle_names,
19 DROP COLUMN last_name;
20 -- Changing full_name column to full_middle_last_name
21 • ALTER TABLE artist_copy
22 CHANGE full_name full_middle_last_name text;
```

Result Grid

	artist_id	full_middle_last_name	nationality	style	birth	death
▶	500	Pierre-Auguste Renoir	French	Impressionist	1841	1919
	501	Alexandre Cabanel	French	Classicist	1823	1889
	502	James Ensor	Belgian	Expressionist	1860	1949
	503	Maximilien Luce	French	Pointillist	1858	1941
	504	August Macke	German	Expressionist	1887	1914
	505	Thomas Gainsborough	English	Rococo	1727	1788
	506	Julien DuprÃ©	French	Naturalist	1851	1910
	507	George Catlin	American	American West	1796	1872
	508	Federico Zandomenighi	Italian	Expressionist	1841	1917
	509	Cornelis Springer	Dutch	Landscape Art	1817	1891
	510	Hans Memling	Dutch	Northern Ren...	1435	1494
	511	Piet Mondrian	Dutch	Modern Art	1872	1944
	512	Lemuel Francis Abbott	English	Portraitist	1760	1803

LH

- Standardizing Data
- Data Cleaning

- Checking or Removing Duplicates

LH

LH

```
32 • SELECT *
33 FROM artist_copy
34 WHERE artist_id IS NULL
35 OR full_middle_last_name IS NULL
36 OR nationality IS NULL
37 OR style IS NULL
38 OR birth IS NULL
39 OR death IS NULL;
```

• CHECKING FOR NULL VALUES

Administration Schemas

Information

Table: artist_copy

Columns:
artist_id int
full_name text
first_name text
middle_names text
last_name text
nationality text
style text
birth int
death int

artist_copy 7

Output

Action Output

Time Action

Type here to search

• TRIMMING COLUMNS

LH

Administration Schemas

Information

Table: artist_copy

Columns:
artist_id int
full_name text
first_name text
middle_names text
last_name text
nationality text
style text
birth int
death int

artist_copy 9

Output

Action Output

Time Action

Type here to search

41 -- Standardizing Data : Trimming Columns
42 • SELECT full_middle_last_name, TRIM(full_middle_last_name),
43 nationality, TRIM(nationality),
44 style, TRIM(style);
45
46 • UPDATE artist_copy
47 SET full_middle_last_name = TRIM(full_middle_last_name),
48 nationality = TRIM(nationality),
49 style = TRIM(style);
50

artist_id	full_middle_last_name	nationality	style	birth	death
506	Julien DuprÃ©	French	Naturalist	1851	1910
507	George Catlin	American	American West	1796	1872
508	Federico Zandomeneghi	Italian	Expressionist	1841	1917
509	Cornelis Springer	Dutch	Landscape Art	1817	1891
510	Hans Memling	Dutch	Northern Ren...	1435	1494
511	Piet Mondrian	Dutch	Modern Art	1872	1944
512	Lemuel Francis Abbott	English	Portraitist	1760	1803
513	Mary Cassatt	American	Impressionist	1844	1926
514	Heinrich Campendonk	German	Expressionist	1889	1957
515	Thomas Cole	American	Hudson River ...	1801	1848
516	Thomas Luny	English	Marine Art	1759	1837
517	EugÃ¨ne-Louis Boudin	French	Impressionist	1824	1898
518	Hendrick Dubbels	Dutch	Marine Art	1621	1676
519	Frank Duveneck	Frank Duveneck	st	1848	1919
520	Asher Brown Durand	American	Hudson River ...	1796	1886
521	George Henry Laporte	English	Realist	1799	1873
522	Peter Monamy	English	Marine Art	1681	1749

Object Info Session

Message

LH

Data Cleaning

```
99  -- Museum : Data Cleaning
100 -- Standardizing Data
101 -- Correctly Sorting Postal Codes From City
102 -- Correctly Sorting City From State
103 • SELECT *
104 FROM museum_copy;
105 -- Correctly Sorting State Column Texts That Are Suppose To be In City Column
106 • UPDATE museum_copy
107 SET city = state
108 WHERE city = '' AND state REGEXP '[A-Za-z]';
109 -- Correctly Sorting Numbers That Were In City Column To Postal Column
110 • UPDATE museum_copy
111 SET postal = city
112 WHERE city REGEXP '.*[0-9].*';
113 -- Putting NULL Values In Postal Column
114 • UPDATE museum_copy
115 SET postal = NULL
116 WHERE TRIM(postal) = '';
117 -- Dropping State Column
118 • ALTER TABLE museum_copy
119 DROP COLUMN state;
```

LH

	museum_id	name	address	city	postal	country	phone	url
▶	30	The Museum of Modern Art	11 W 53rd St	New York	10019	USA	+1 212 708-9400	https://www.moma.org/
	31	Pushkin State Museum of Fine Arts	12 Ulitsa Volkhonka	Moscow	119019	Russia	+7 495 697-95-78	https://pushkinmuseum.art/
	32	National Gallery of Victoria	180 St Kilda Rd	Melbourne	3004	Australia	+61 (0)3 8620 2222	https://www.ngv.vic.gov.au/
	33	SÃ£o Paulo Museum of Art	Av. Paulista, 1578 - Bela Vista	SÃ£o Paulo	01310-200	Brazil	+55 11 3149-5959	https://masp.org.br/
	34	The State Hermitage Museum	Palace Square	Sankt-Peterburg	2	Russia	7 812 710-90-79	https://www.hermitagemuseum.org/wps
	35	The Metropolitan Museum of Art	1000 5th Ave	New York	10028	USA	+1 212 535-7710	https://www.metmuseum.org/
	36	Museum Folkwang	Museumsplatz 1	Essen	45128	Germany	49 201 8845000	https://www.museum-folkwang.de/en
	37	Museum of Grenoble	5 Pl. de Lavalette	Grenoble	38000	France	33 4 76 63 44 44	https://www.museedegrenoble.fr/1986-1
	38	MusÃ©e des Beaux-Arts de Quimper	40 Pl. Saint-Corentin	Quimper	29000	France	33 2 98 95 45 20	https://www.mbaq.fr/en/home-3.html
	39	Nelson-Atkins Museum of Art	4525 Oak St	Kansas City	64111	USA	+1 816 751-1278	https://www.nelson-atkins.org/
	40	MusÃ©e du Louvre	Rue de Rivoli	Paris	75001	France	33 1 40 20 50 50	https://www.louvre.fr/en
	41	National Maritime Museum	Romney Rd	London	NULL	United K...	44 20 8858 4422	https://www.rmg.co.uk/national-maritime
	42	Museum of Fine Arts Boston	465 Huntington Ave	Boston	2115	USA	617-267-9300	https://www.mfa.org/
	43	Rijksmuseum	Museumstraat 1	Amsterdam	1071 XX	Netherl...	31 20 674 7000	https://www.rijksmuseum.nl/en/visit

- Correctly Sorting State Column Texts That Are Suppose To be In City Column
- Correctly Sorting Numbers That Were In City Column To Postal Column
- Putting Null Values In Postal Column
- Dropping State Column

	museum_id	name	address	city	state	postal	country	phone	url
▶	30	The Museum of Modern Art	11 W 53rd St	New York	NY	10019	USA	+1 212 708-9400	https://www.moma.org/
	31	Pushkin State Museum of Fine Arts	12 Ulitsa Volkhonka	Moscow		119019	Russia	+7 495 697-95-78	https://pushkinmuseum.art/
	32	National Gallery of Victoria	180 St Kilda Rd	Melbourne	Victoria	3004	Australia	+61 (0)3 8620 2222	https://www.ngv.vic.gov.au/
	33	SÃ£o Paulo Museum of Art	Av. Paulista, 1578 - Bela Vista	SÃ£o Paulo		01310-200	Brazil	+55 11 3149-5959	https://masp.org.br/
	34	The State Hermitage Museum	Palace Square	2	Sankt-Peterburg	190000	Russia	7 812 710-90-79	https://www.hermitagemuseum.org/wps/portal...
	35	The Metropolitan Museum of Art	1000 5th Ave	New York	NY	10028	USA	+1 212 535-7710	https://www.metmuseum.org/
	36	Museum Folkwang	Museumsplatz 1	45128	Essen		Germany	49 201 8845000	https://www.museum-folkwang.de/en
	37	Museum of Grenoble	5 Pl. de Lavalette	38000	Grenoble		France	33 4 76 63 44 44	https://www.museedegrenoble.fr/1986-the-mu...
	38	MusÃ©e des Beaux-Arts de Quimper	40 Pl. Saint-Corentin	29000	Quimper		France	33 2 98 95 45 20	https://www.mbaq.fr/en/home-3.html
	39	Nelson-Atkins Museum of Art	4525 Oak St	Kansas City	MO	64111	USA	+1 816 751-1278	https://www.nelson-atkins.org/
	40	MusÃ©e du Louvre	Rue de Rivoli	75001	Paris		France	33 1 40 20 50 50	https://www.louvre.fr/en
	41	National Maritime Museum	Romney Rd	London	SE10 9NF		United K...	44 20 8858 4422	https://www.rmg.co.uk/national-maritime-museum
	42	Museum of Fine Arts Boston	465 Huntington Ave	Boston	MA	2115	USA	617-267-9300	https://www.mfa.org/
	43	Rijksmuseum	Museumstraat 1	Amsterdam		1071 XX	Netherl...	31 20 674 7000	https://www.rijksmuseum.nl/en/visit
	44	Israel Museum	Derech Ruppin 11	Jerusalem			Israel	972 2-670-8811	https://www.imj.org.il/en

museum_copy 39 x

Output :

Action Output

#	Time	Action	Message
---	------	--------	---------

- Removing The Numbers In the Column that contain strings

Views
Stored Procedures
Functions
sys

```
158
159 -- REMOVING THE NULL VALUES
160 • SELECT *
161 FROM subject_copy
162 WHERE subject_copy IS NULL;
163
164 • DELETE
165 FROM subject_copy
166 WHERE subject_copy IS NULL;
167
168 • SELECT *
169 FROM subject_copy;
```

Administration Schemas

Information

Table: **subject_copy**

Columns:
work_id int
subject text

Result Grid

work_id	subject
160228	Still-Life
160236	Still-Life
160244	Still-Life
160252	Still-Life
160260	Still-Life
160268	Still-Life
125752	Horses
23448	Horses
125763	Horses
24532	Marine Art/Maritime
124470	Marine Art/Maritime
124479	Marine Art/Maritime
124488	Marine Art/Maritime
124497	Marine Art/Maritime
124506	Marine Art/Maritime

subject_copy 46 x

Output

Action Output

Time Action

Object Info Session

```
170
171 -- REMOVING THE NUMBERS IN THE COLUMN THAT CONTAIN STRING'
172 • UPDATE subject_copy
173 SET subject_copy = REGEXP_REPLACE(subject, '[0-9]', '');
```

Administration Schemas

Information

Table: **subject_copy**

Columns:
work_id int
subject text

Result Grid

work_id	subject
160228	Still-Life
160236	Still-Life
160244	Still-Life
160252	Still-Life
160260	Still-Life
160268	Still-Life
125752	Horses
23448	Horses
125763	Horses
24532	Marine Art/Maritime
124470	Marine Art/Maritime
124479	Marine Art/Maritime
124488	Marine Art/Maritime
124497	Marine Art/Maritime
124506	Marine Art/Maritime

subject_copy 47 x

Output

Action Output

Time Action

Object Info Session

- Removing NULL Values