

Harvard Artifacts Explorer

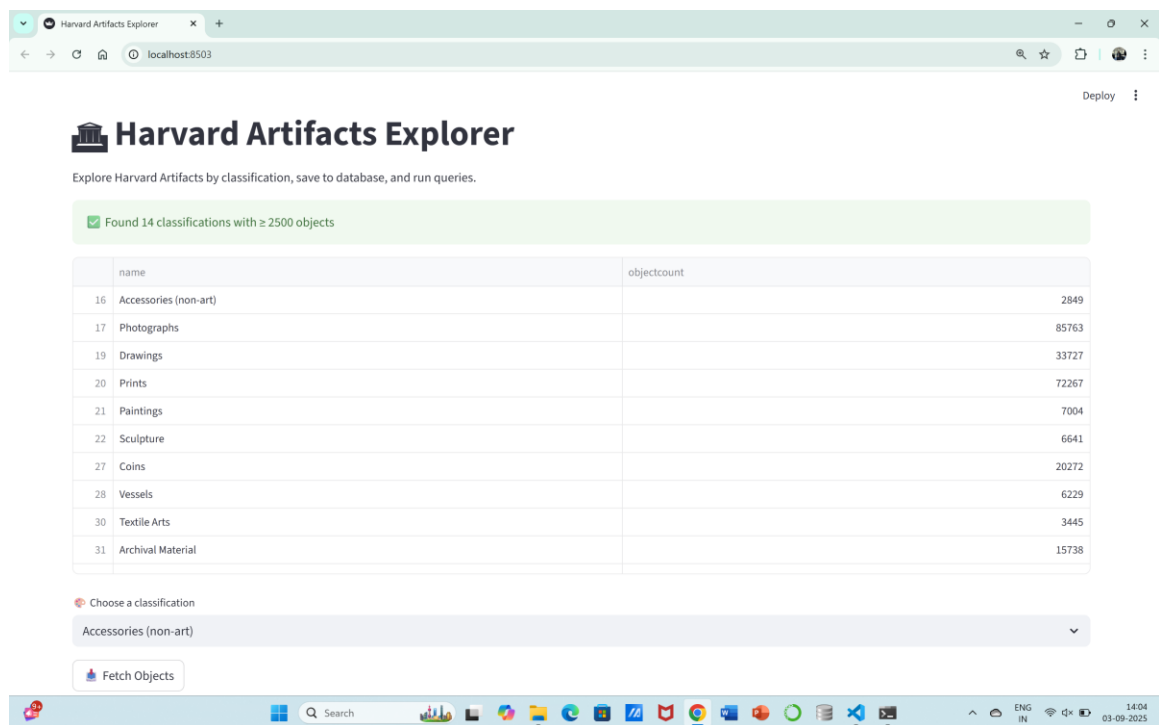
Author: Anubriya Baskaran

Project: Data Exploration and SQL Queries using Streamlit and SQLite

Date: September 2025

1. Introduction

The Harvard Artifacts Explorer project is a data-driven application built using Python, Streamlit, SQLite, and the Harvard Art Museums API. The goal of the project is to fetch, store, and explore artifact data from the Harvard Art Museums. Users can interactively choose artifact classifications, fetch corresponding objects, save the data into a database, and run SQL queries to analyze the collection.



Harvard Artifacts Explorer

Explore Harvard Artifacts by classification, save to database, and run queries.

Found 14 classifications with ≥ 2500 objects

	name	objectcount
16	Accessories (non-art)	2849
17	Photographs	85763
19	Drawings	33727
20	Prints	72267
21	Paintings	7004
22	Sculpture	6641
27	Coins	20272
28	Vessels	6229
30	Textile Arts	3445
31	Archival Material	15738

Choose a classification

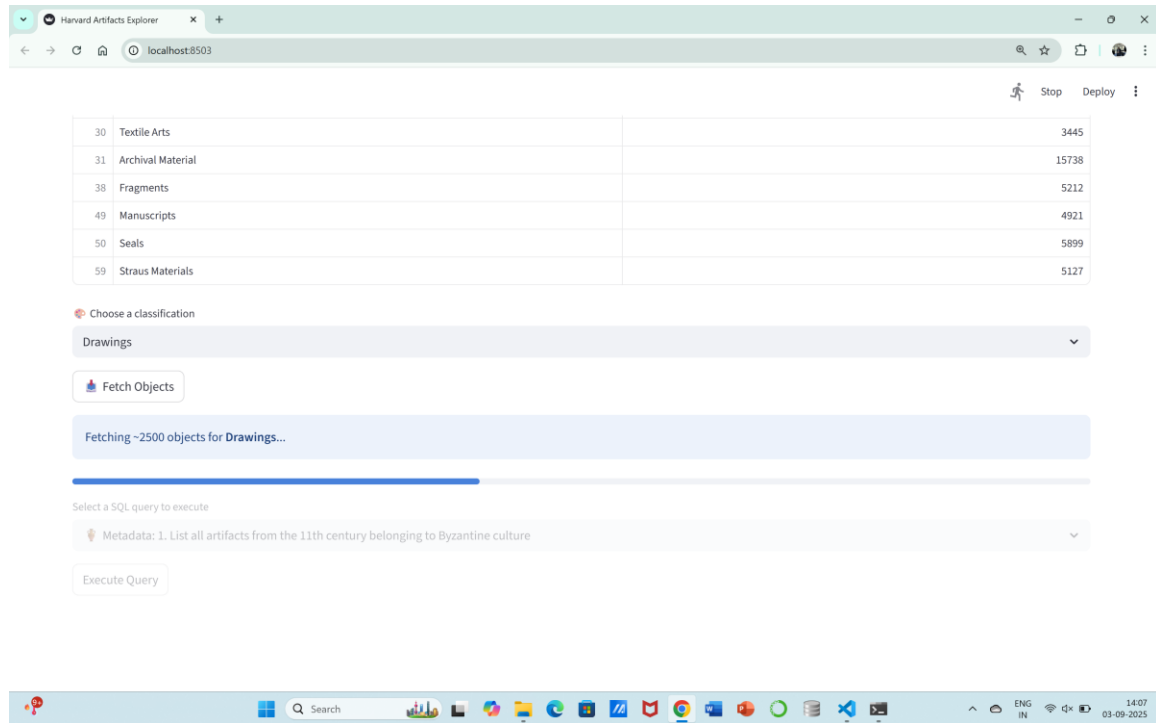
Accessories (non-art)

Fetch Objects

1.1. Classifications Greater than or Equal to 2500 objects

2. Objectives

- To integrate and fetch real-time artifact data from the Harvard Art Museums API.
- To filter and store artifact data into structured SQLite tables.
- To enable interactive exploration of metadata, media, and color information.
- To design and execute SQL queries for gaining insights into the dataset.
- To provide a simple, beginner-friendly, and interactive data exploration tool.



2.1. Fetching the desired classification

3. Tools and Technologies Used

- Programming Language: Python
- Framework: Streamlit
- Database: SQLite
- API: Harvard Art Museums API
- Libraries: pandas, requests, sqlite3, time
- Documentation Tools: MS Word, MS PowerPoint

4. System Architecture

1. Data is fetched from the Harvard Art Museums API.
2. Data is filtered to select only classifications with more than 2500 objects.
3. Selected classifications are explored by fetching artifacts (metadata, media, colors).

4. Data is stored in SQLite tables (artifact_metadata, artifact_media, artifact_colors).
5. Users interact via Streamlit interface to insert, query, and visualize data.

	copyright	contextualtextcount	creditline	accesslevel	createdate	dateoflastpageview	classificationid	division
0	None	0	Harvard Art Museums/Fogg Museum, Partial and Promised Gift of David M. Leventhal	1	2005-04-06T00:00:00-04:00	2025-08-24	21	Euro
1	None	0	Harvard Art Museums/Fogg Museum, Gift of Roberta J.M. Olson and Alexander B.V. Jo	1	2005-12-19T00:00:00-05:00	2025-07-21	21	Euro
2	None	0	Harvard Art Museums/Fogg Museum, Gift of Victoria Munroe in honor of Marjorie Coh	1	2006-02-07T00:00:00-05:00	2025-07-21	21	Euro
3	None	0	Harvard Art Museums/Busch-Reisinger Museum, Bequest of Richard B. Sisson	1	2005-09-26T00:00:00-04:00	2025-07-21	21	Euro
4	None	0	Harvard Art Museums/Busch-Reisinger Museum, Bequest of Richard B. Sisson	1	2005-09-26T00:00:00-04:00	2025-08-17	21	Mod
5	© copyright	0	Harvard Art Museums/Fogg Museum, Purchase through the generosity of an anonym	1	2005-03-10T00:00:00-05:00	2025-07-31	21	Mod
6	None	0	Harvard Art Museums/Fogg Museum, Purchase through the generosity of the Weinga	1	2003-09-05T00:00:00-04:00	2025-07-21	21	Euro
7	None	0	Harvard Art Museums/Fogg Museum, Gift of Patricia Cornwell	1	2005-12-20T00:00:00-05:00	2025-09-01	21	Euro
8	© Estate of	0	Harvard Art Museums/Fogg Museum, Stephen Lee Taller Ben Shahn Archive, Gift of D	1	2004-06-17T00:00:00-04:00	2021-06-18	21	Mod
9	None	0	Harvard Art Museums/Fogg Museum, The Marian H. Phinney Fund, Paul J. Sachs Men	1	2005-12-15T00:00:00-05:00	2025-08-14	21	Euro

4.1. Fetched data with progress bar

5. Database Design

The project uses three tables for structured storage:

- **artifact_metadata:** Stores general metadata such as title, culture, period, century, medium, department, classification, etc.
- **artifact_media:** Stores media-related data such as imagecount, mediacount, colorcount, rank, and date ranges.
- **artifact_colors:** Stores extracted colors, hues, coverage percentages, and CSS3 representations.

Harvard Artfacts Explorer

localhost:8503

Deploy

	id	title	culture	period	century	medium	dimensions
0	1423	Portrait of a Woman, after a Drawing in the Uffizi then at	French	None	19th century	Graphite on off-white wove paper	36.8 x 28 cm (14 1/2 x
1	1434	San Arcangelo	French	None	19th century	Black and white chalk on brown modern laid paper, lined to Japanese paper	34.3 x 28.2 cm (13 1/2
2	1436	Geranium Chamadryoides	British	None	19th century	Cut paper collage with watercolor on laid paper prepared with black ink	18.5 x 26.6 cm (7 5/16
3	4629	Beach Scene	German	None	20th century	Black chalk on off-white wove paper	13.5 x 21 cm (5 5/16 x
4	4630	Portrait of a Soldier	German	None	20th century	Black ink and graphite on tan wove paper	37 x 27 cm (14 9/16 x
5	4632	My Personal Atlas	American	None	21st century	Cotton paper, with inclusions, colored pencil and overlaid with watermarked abaca	26.5 x 35.7 cm (10 7/16
6	4647	Study for the Mansfield Monument	British	None	18th century	Graphite, black ink, and gray wash on off-white wove paper	20.6 x 14.8 cm (8 1/8 x
7	4654	Butterfly for a Valance at Dowdeswell's Gallery	American	None	19th century	Graphite and watercolor on beige laid paper	7.3 x 7.6 cm (2 7/8 x 3
8	4677	"Six million farms and over in this country last census."	American	None	20th century	Graphite on off-white wove paper	folded sheet: 26.6 x 21
9	4726	A View of Cronesteyn Castle, near Leiden	Dutch	None	17th century	Gray wash, black chalk and graphite on cream antique laid paper, framing lines in br	37.6 x 51.7 cm (14 13/16
10	4843	Sculpture Study	American	None	20th century	Black and/or brown ink on off-white Japanese paper	17.5 x 11.7 cm (6 7/8 x

Insert into Database

Search

ENG IN 14:09 03-09-2025

5.1. Metadata Table

Harvard Artfacts Explorer

localhost:8503

Deploy

	id	title	culture	period	century	medium	dimensions	
8	© Estate of	0	Harvard Art Museums/Fogg Museum, Stephen Lee Taller Ben Shahn Archive, Gift of D	1	2004-06-17T00:00:00-04:00	2021-06-18	21	Mod
9	None	0	Harvard Art Museums/Fogg Museum, The Marian H. Phinney Fund, Paul J. Sachs Mem	1	2005-12-15T00:00:00-05:00	2025-08-14	21	Euro

View Data Tables

Metadata Media Colors

	objectid	imagecount	mediacount	colorcount	rank	datebegin	dateend
0	1423	1	0	6	125859	1858	1859
1	1434	2	0	8	126417	0	0
2	1436	4	0	8	126461	0	0
3	4629	1	0	5	126261	1903	1913
4	4630	2	0	7	126259	1918	1918
5	4632	1	0	3	125799	2004	2004
6	4647	1	0	1	125201	0	0
7	4654	2	0	4	126426	1879	1889
8	4677	1	0	6	124399	1942	1942
9	4726	5	0	7	126418	1646	1647

Insert into Database

Search

ENG IN 14:09 03-09-2025

5.2. Media Table

Harvard Artfacts Explorer

localhost:8503

Deploy

7	ROUTE	0	Harvard Art Museums/Fogg Museum, Gift of Patricia L. Gifford	1	2003-12-17T00:00:00-05:00	2023-09-01	21	Euro
8	© Estate of	0	Harvard Art Museums/Fogg Museum, Stephen Lee Taller Ben Shahn Archive, Gift of D	1	2004-06-17T00:00:00-04:00	2021-06-18	21	Mod
9	None	0	Harvard Art Museums/Fogg Museum, The Marian H. Phinney Fund, Paul J. Sachs Men	1	2005-12-15T00:00:00-05:00	2025-08-14	21	Euro

View Data Tables

MetadataMediaColors

	objectid	color	spectrum	hue	percent	css3
0	1423	#c8c8c8	#8c5fa8	Grey	0.5631	#c0c0c0
1	1423	#c8afaf	#b55592	Red	0.3734	#c0c0c0
2	1423	#af9696	#ae5596	Brown	0.0432	#bc8f8f
3	1423	#e1e1e1	#955ba5	Grey	0.0157	#dcdcdc
4	1423	#7d7d7d	#8362aa	Grey	0.0039	#808080
5	1423	#7d6464	#8e5ea7	Brown	0.0006	#696969
6	1434	#967d4b	#7fc241	Brown	0.3704	#a0522d
7	1434	#7d644b	#b25593	Yellow	0.2711	#696969
8	1434	#af7d64	#c85783	Brown	0.2236	#cd5c5c
9	1434	#e1e1c8	#e9715f	Green	0.0814	#dcdcdc

Insert into Database

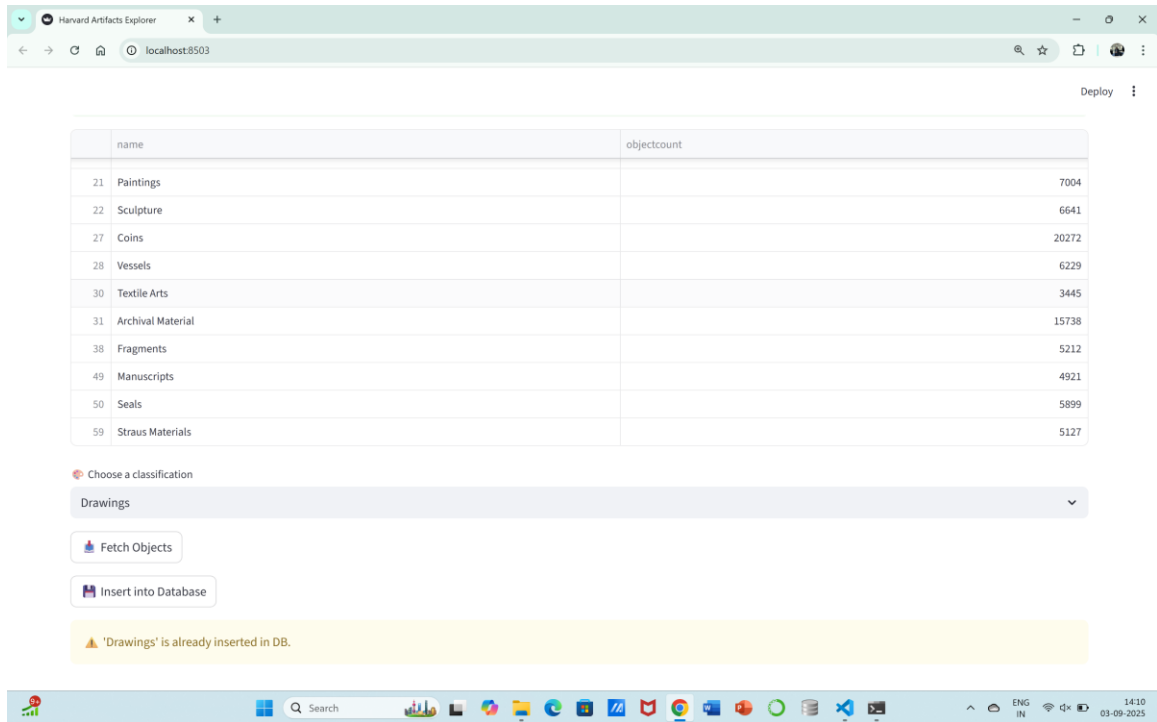
Search

ENG IN 14:09 03-09-2025

5.3. Colors Table

6. Implementation

The application is implemented in Python with Streamlit providing an interactive frontend. The workflow includes fetching data, splitting it into structured DataFrames, inserting into SQLite, and providing query execution options. A dropdown-based query explorer allows users to select one of 30 SQL queries (20 predefined + 10 DIY) and view results instantly.

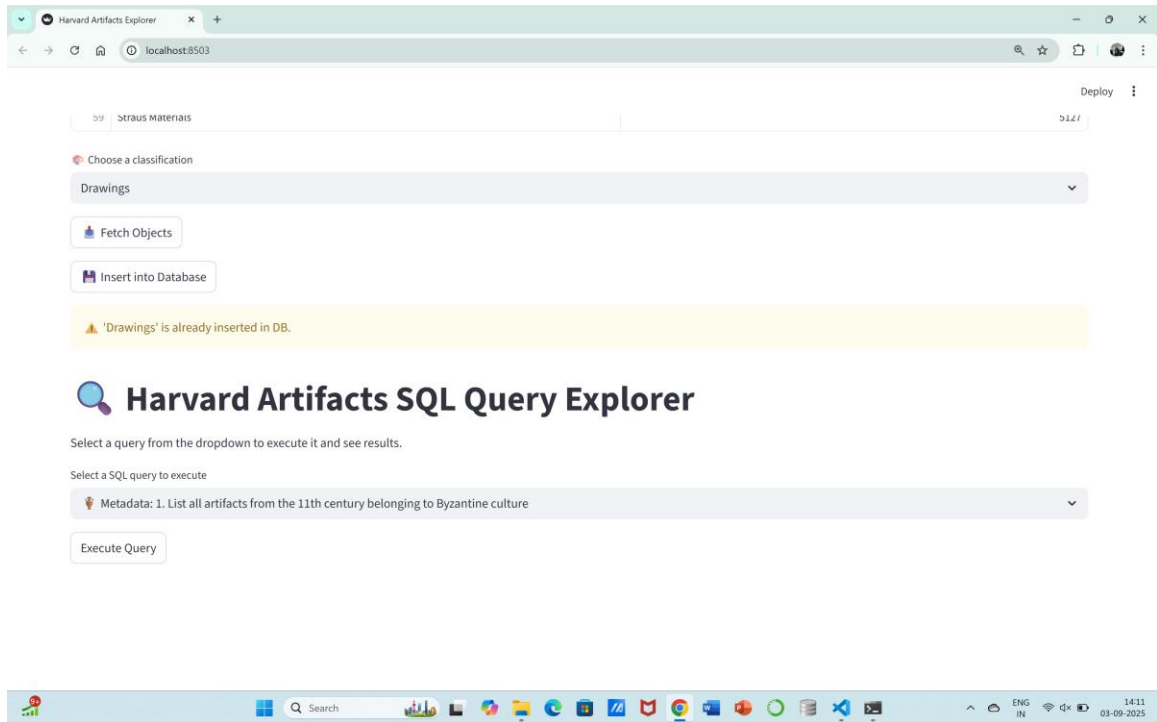


6.1. Inserting into the Database

7. Sample SQL Queries

Examples of SQL queries implemented in this project include:

- List all artifacts from the 11th century belonging to Byzantine culture.
- Show unique cultures represented in the dataset.
- Find the average rank of all artifacts.
- Show top 5 most frequently used colors.
- Count artifacts per department.
- List artifact titles and hues for Byzantine culture (join query).



7.1. SQL Query Explorer to choose desired query

8. Results and Analysis

The project successfully demonstrates how real-world datasets can be fetched, stored, and analyzed interactively.


Key findings include:

- The database contains thousands of artifacts across multiple cultures, centuries, and classifications.
- Byzantine and Egyptian artifacts are richly represented.
- Colors like Grey, Red, and Blue dominate in artifact imagery.
- Join-based queries reveal detailed relationships between metadata, media, and color usage.

Harvard Artifacts Explorer

localhost:8503

Deploy



Harvard Artifacts SQL Query Explorer

Select a query from the dropdown to execute it and see results.

Select a SQL query to execute

Metadata: 1. List all artifacts from the 11th century belonging to Byzantine culture

Execute Query

Query executed successfully. 351 rows fetched.


	id	title	culture	period	century	medium	dimensions	description
7	74536	Coin of Alexios I	Byzantine	Byzantine period, Middle	11th-12th century	Silver	3.74 g	Obv.: Bust of Christ, bearded and nimbate, wearing tunic and kolobos
8	74537	Coin of Alexios I	Byzantine	Byzantine period, Middle	11th-12th century	Silver	3.71 g	Obv.: Bust of Christ, bearded and nimbate, wearing tunic and kolobos
9	74538	Coin of Alexios I	Byzantine	Byzantine period, Middle	11th-12th century	Bronze and	3.90 g	Obv.: Bust of Christ, bearded and nimbate, wearing tunic and kolobos
10	74539	Coin of Alexios I	Byzantine	Byzantine period, Middle	11th-12th century	Silver	3.04 g	Obv.: Bust of Virgin nimbate, orans, wearing tunic and maphorion; l
11	74540	Coin of Alexios I	Byzantine	Byzantine period, Middle	11th-12th century	Silver	3.46 g	Obv.: Bust of Virgin nimbate, orans, wearing tunic and maphorion; l
12	74541	Coin of Alexios I	Byzantine	Byzantine period, Middle	11th-12th century	Bronze	6.04 g	Obv.: Cross with globule at each extremity, on two steps. Rev.: Inscr
13	74547	Coin of Constantine IX	Byzantine	Byzantine period, Middle	11th century	Gold	4.33 g	Obv.: Bust of Christ Pantocrator (Type VIIIc) facing, bearded, with cr
14	74548	Coin of Constantine X	Byzantine	Byzantine period, Middle	11th century	Bronze	8.85 g	Obv.: Christ standing on square souppedion (Type IVb) facing, bearc
15	74549	Coin of Constantine X	Byzantine	Byzantine period, Middle	11th century	Bronze	8.47 g	Obv.: Christ standing on square souppedion (Type IVb) facing, bearc
16	74550	Coin of Constantine X	Byzantine	Byzantine period, Middle	11th century	Bronze	6.27 g	Obv.: Christ standing on square souppedion (Type IVb) facing, bearc

8.1. Example query from Metadata

Harvard Artifacts Explorer

localhost:8503

Deploy



Harvard Artifacts SQL Query Explorer

Select a query from the dropdown to execute it and see results.

Select a SQL query to execute

Media: 9. Artifacts created between 1500 and 1600

Execute Query

Query executed successfully. 861 rows fetched.

	objectid	imagecount	mediacount	colorcount	rank	datebegin	dateend
0	47660	4	0	5	137985	1543	1543
1	94413	1	0	10	154685	1591	1591
2	144755	5	0	6	123399	1515	1525
3	1424	1	0	7	245147	1585	1585
4	1425	1	0	8	245149	1585	1585
5	4605	0	0	0	245137	1585	1585
6	4606	1	0	7	245143	1585	1585
7	4607	1	0	8	245145	1585	1585
8	4608	1	0	8	245146	1585	1585
9	4620	1	0	8	245148	1585	1585

8.2. Example query from Media

Harvard Artifacts Explorer

localhost:8503

Deploy

Harvard Artifacts SQL Query Explorer

Select a query from the dropdown to execute it and see results.

Select a SQL query to execute

Colors: 11. Distinct hues used

Execute Query

Query executed successfully. 10 rows fetched.

	hue
0	Orange
1	Grey
2	Brown
3	Yellow
4	Green
5	Black
6	White
7	Red
8	Blue
9	Violet

8.3. Example query from Colors

Harvard Artifacts Explorer

localhost:8503

Deploy

Harvard Artifacts SQL Query Explorer

Select a query from the dropdown to execute it and see results.

Select a SQL query to execute

Join: 18. Artifact titles, cultures, and media ranks where period is not null

Execute Query

Query executed successfully. 8836 rows fetched.


	title	culture	rank
0	Kiln Furniture (?): Ceramic Fragment Possibly used as a Kiln-Stacking Tool	Japanese	89438
1	Drawing of Samanosuke Mitsutoshi with Fox Fires, from the series One Hundred Ghost Stories of China and Japar	Japanese	125597
2	Of of Four Sketches of Birds in Trees	Japanese	91421
3	One of Three Monochrome Sketches	Japanese	91413
4	One of Five Monochrome Sketches	Japanese	91408
5	One of Five Monochrome Sketches	Japanese	91407
6	Illustration of a Deity from the Temukayama Shrine (Temuakayama Jinja)	Japanese	91411
7	One of Five Monochrome Sketches	Japanese	91406
8	One of Five Monochrome Sketches	Japanese	91409
9	One of Eight Designs for Cloisonne	Japanese	91398

8.4. Example Join query

Harvard Artifacts Explorer

localhost:8503

Deploy



Harvard Artifacts SQL Query Explorer

Select a query from the dropdown to execute it and see results.

Select a SQL query to execute

DIY 2. Artifacts from culture 'Egyptian'

Execute Query

Query executed successfully. 579 rows fetched.


	id	title	culture	period	century	medium	dimensions	description
0	167757	Painted Panel from Wooden Sarcophagus	Egyptian	Roman Imperial period	2nd century CE	Wood	24 x 40.4 x 3 cm	Wooden panel over a foot in width, covered with gesso, with a figure
1	291707	Portrait of a woman	Egyptian	Roman Imperial period	2nd-3rd century CE	Binder: An	32 x 19 x 0.5 cm	None
2	291901	Painting on cloth	Egyptian	Roman period	None	Painted lin	None	Painted fragment of a mummy shroud.
3	292324	Portrait of a man	Egyptian	Roman Imperial period, Lz	3rd century CE	Encaustic	46 x 20.2 x 2 cm	None
4	292519	Fragment of Drawing on Gesso	Egyptian	None	None	None	12.1 x 8.3 cm	Torso and legs of figure moving left with headdress.
5	292520	Fragments of painting on plaster over linen	Egyptian	None	6th-3rd century BCE	None	None	None
6	72349	Selket	Egyptian	Late Period to Roman	1st millennium BCE	Leaded br	8 x 7.1 x 3.6 cm	None
7	92353	Lid for a canopic jar in the shape of a human	Egyptian	Ptolemaic period to Roma	4th century BCE-4th	Limestone	8.5 x 11.5 cm	Ptolemaic or later. Surface of head is dirty grey. Raised circular proj
8	117323	Egitto-Phoenician Ivory Head of a Duck	Egyptian	None	None	Ivory	None	None
9	117373	Hellenistic Alabaster Head	Egyptian	Ptolemaic period	3rd-2nd century BC	Alabaster	None	None

8.5. DIY Example query

Harvard Artifacts Explorer

localhost:8503

Deploy



Harvard Artifacts SQL Query Explorer

Select a query from the dropdown to execute it and see results.

Select a SQL query to execute

DIY 9. Artifacts with 'Vase' in title

Execute Query

Query executed successfully. 267 rows fetched.

	id	title	culture	period	century	medium	dimensions	description
64	5141	Vase with Bamboo Design Handles	French	None	19th century	Falence (le	31.1 x 25.1 x 3 cm	A lead-glazed earthenware vase, ovoid with sloping neck, banded ri
65	53659	Vase	French	None	16th century	Pewter	None	None
66	92349	Miniature Copy of the Portland Vase	British, Eng	None	20th century	Stoneware	10.48 x 6.99 cm	A miniature blue jasper ware copy of the Portland vase, a Greco-Rom
67	99572	Tall, Flattened Vase with Short Neck, Bulging	Japanese	Heisei era, 1989-2019	21st century	Off-white s	H. 38.0 cm x 1	None
68	99739	Rectangular Bottle or Vase	Japanese	Heisei era, 1989-2019	21st century	Off white	H. 25.5 cm x 1	None
69	100181	Bottle or Vase with Decoration of Blossoming	Chinese	Southern Song period (11	13th-14th century	Jizhou wa	20.2 x 11.8 cm	None
70	146550	Tall Vase	American	None	20th century	Porcelain	38.9 x 17.3 cm	None
71	165421	Small Vase with Stripes	None	Modern	19th-20th century	Fritware p	9.2 cm x 8.4 cm	This small vase is decorated with vertical cobalt-blue stripes under i
72	167161	Vase with Inscriptions	Arab	Mamluk period	14th century	Brass, inla	H: 17.5 x sma	Vase with tall, round, slightly flared neck; the pear-shaped body stai
73	173314	Conical Vase with Decoration of Three Stack	Japanese	Heisei era, 1989-2019	21st century	Light gray	H. 21.0 cm x 1	None

8.6. DIY Example query

9. Conclusion

The Harvard Artifacts Explorer provides an end-to-end solution for exploring museum collections with modern data science tools. It highlights the importance of structured storage, interactive visualization, and query-based insights in cultural data analysis. This project can be extended further with advanced visualizations, deployment, and predictive analytics.

10. Future Work

- Deploying the Streamlit app on the cloud for public access.
- Enhancing visualizations using charts and graphs.
- Adding Natural Language Processing (NLP) for querying in plain English.
- Expanding dataset coverage beyond 2500-object classifications.
- Integrating AI/ML for pattern detection and predictive insights.