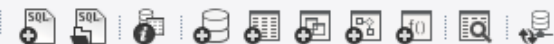


MYSQL EXPLORATORY DATA ANALYSIS



Navigator

SCHEMAS

Filter objects

- student
- sys
- world_layoffs
 - Tables
 - Views
 - Stored Procedures
 - Functions
- zomato

DATA CLEANING IN MY SQL*

EDA IN MY SQL* x



Limit to 5000 rows

```
1
2 • USE WORLD_LAYOFFS;
3
4 -- EXPLORATORY DATA ANALYSIS
5 -- Select the data
6
7 • SELECT *
8 FROM LAYOFFS_STAGING2;
9
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: IA

	company	location	industry	total_laid_off	percentage_laid_off	date	stage	country	funds_raised_millions
▶	Included Health	SF Bay Area	Healthcare	NULL	0.06	2022-07-25	Series E	United States	272
	&Open	Dublin	Marketing	9	0.09	2022-11-17	Series A	Ireland	35
	#Paid	Toronto	Marketing	19	0.17	2023-01-27	Series B	Canada	21
	100 Thieves	Los Angeles	Consumer	12	NULL	2022-07-13	Series C	United States	120
	10X Genomics	SF Bay Area	Healthcare	100	0.08	2022-08-04	Post-IPO	United States	242
	1stdibs	New York City	Retail	70	0.17	2020-04-02	Series D	United States	253
	2TM	Sao Paulo	crypto	90	0.12	2022-06-01	Unknown	Brazil	250
	2TM	Sao Paulo	crypto	100	0.15	2022-09-01	Unknown	Brazil	250
	2U	Washington D.C.	Education	NULL	0.2	2022-07-28	Post-IPO	United States	426
	54gene	Washington D.C.	Healthcare	95	0.3	2022-08-29	Series B	United States	44
	5B Solar	Sydney	Energy	NULL	0.25	2022-06-03	Series A	Australia	12
	6sense	SF Bay Area	Sales	150	0.1	2022-10-12	Series E	United States	426
	80 Acres Farms	Cincinnati	Food	NULL	0.1	2023-01-18	Unknown	United States	275

LAYOFFS_STAGING2 1 x

Administration Schemas

Information

No object selected

Object Info Session

Output

Result Grid

Form Editor

Field Types

Read Only

```

10
11
12 -- Maximum Total_laid_off and percentage_laid_off in the data:
13
14 • SELECT MAX(total_laid_off) , MAX(percentage_laid_off)
15 FROM
16 LAYOFFS_STAGING2;
17
18

```

Result Grid   Filter Rows: Export:  Wrap Cell Content: 

	MAX(total_laid_off)	MAX(percentage_laid_off)
▶	12000	1

Navigator ::::::::::::::::::::::::::::::::::

DATA CLEANING IN MY SQL*

EDA IN MY SQL*

SCHEMAS

Filter objects

▶ student

▶ sys

▼  world_layoffs

▶ Tables

Views

Stored Procedures

Functions

▶ zomato

Administration Schemas

Information

No object selected

Object Info

Session

22

```
23 -- Company_wise Total_laid_off :
```

24

```
25 • SELECT company , SUM(total_laid_off)
```

26 FROM LAYOFFS STAGING2

27 GROUP BY company

28 ORDER BY 2 DESC ;

29

30

Result Grid   Filter Rows: Export:  Wrap Cell Content: 

	company	SUM(total_laid_off)
▶	Amazon	18150
	Google	12000
	Meta	11000
	Salesforce	10090
	Microsoft	10000
	Philips	10000
	Ericsson	8500
	Uber	7585
	Dell	6650
	Booking.com	4601
	Cisco	4100
	Peloton	4084
	Byju's	4000

Result 3 ✕

Output

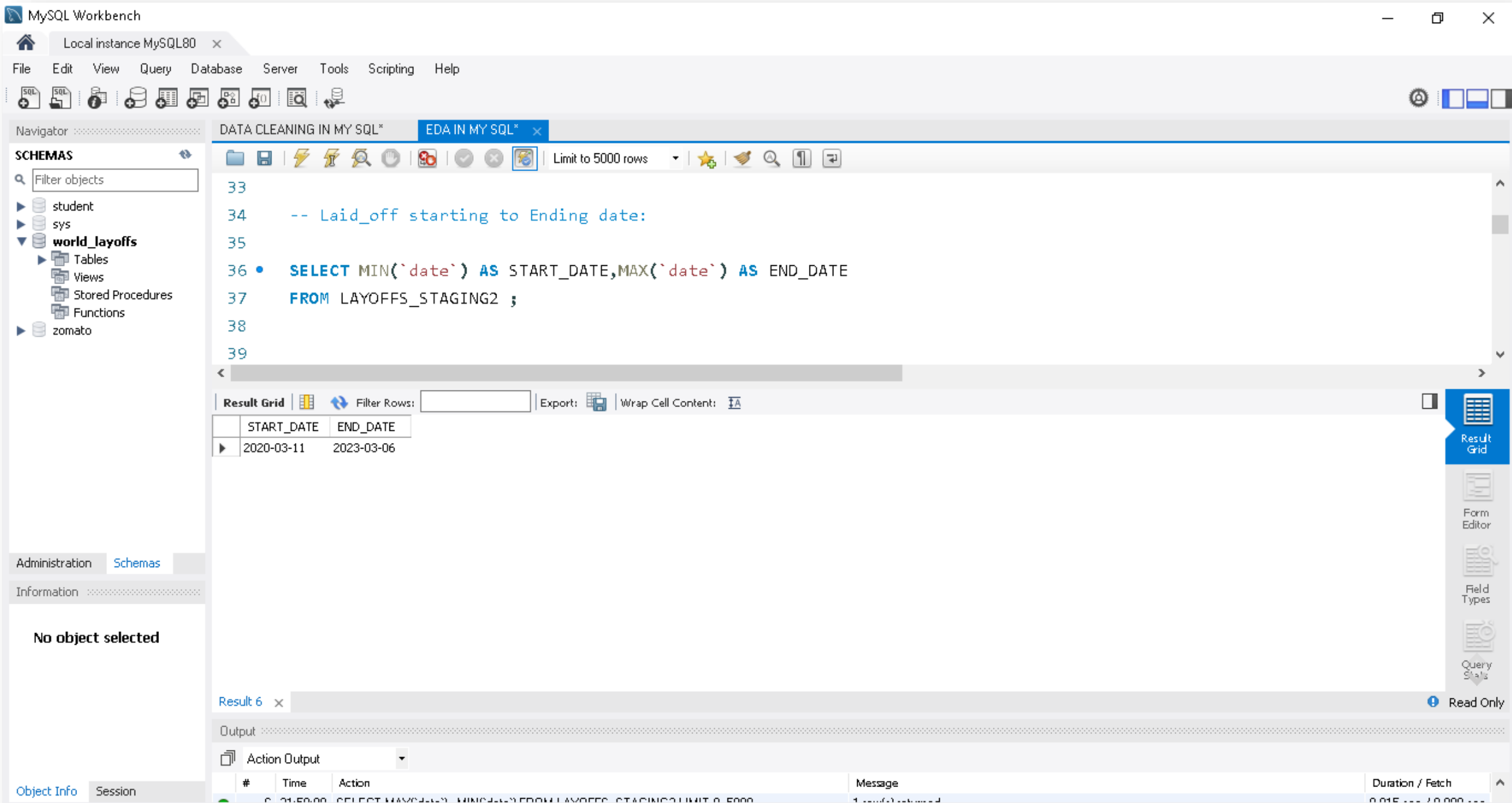


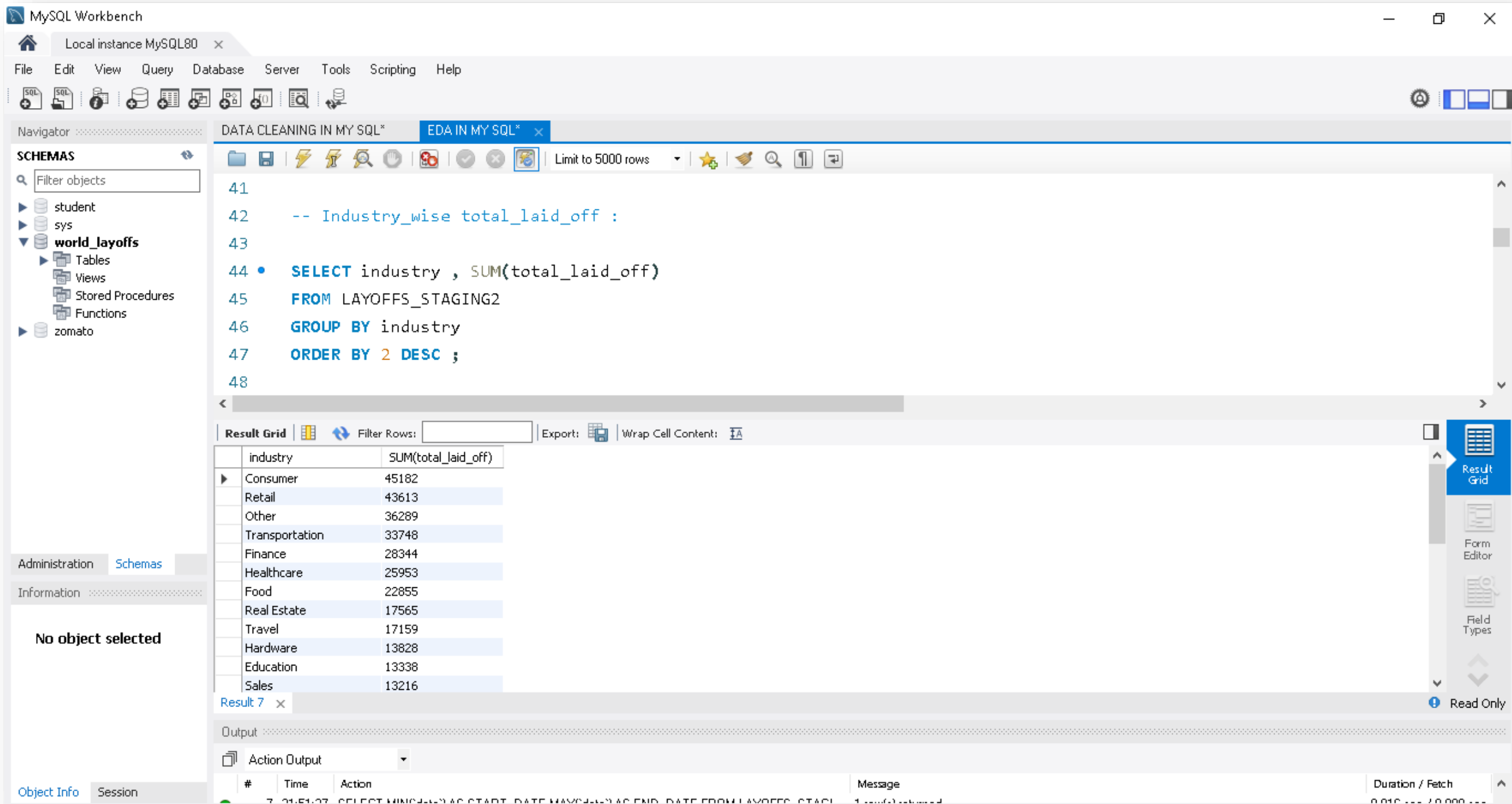
Result Grid

Form Editor

Field Types

 Read Only





MySQL Workbench

Local instance MySQL80

File Edit View Query Database Server Tools Scripting Help

Navigator

SCHEMAS

Filter objects

- student
- sys
- world_layoffs
 - Tables
 - Views
 - Stored Procedures
 - Functions
- zomato

DATA CLEANING IN MY SQL*

EDA IN MY SQL*

Limit to 5000 rows

```
66
67  -- Year_wise total_laid_off
68
69 • SELECT YEAR(`date`) , SUM(total_laid_off)
70 FROM LAYOFFS_STAGING2
71 GROUP BY YEAR(`date`)
72 ORDER BY 1 DESC ;
73
74
```

Result Grid

Filter Rows:

Export: Wrap Cell Content:

	YEAR(`date`)	SUM(total_laid_off)
▶	2023	125677
	2022	160661
	2021	15823
	2020	80998
	NULL	500

Administration Schemas

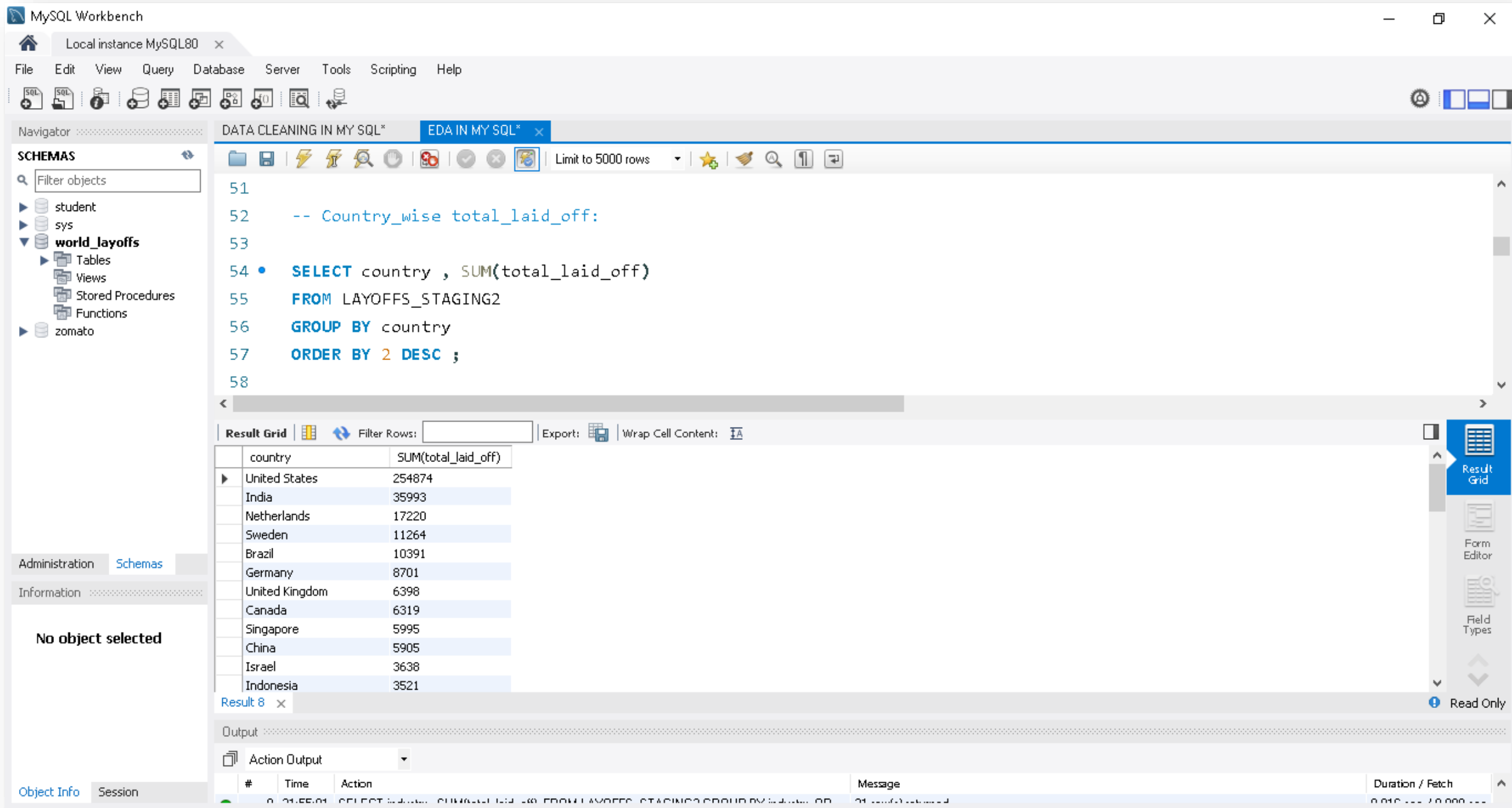
Information

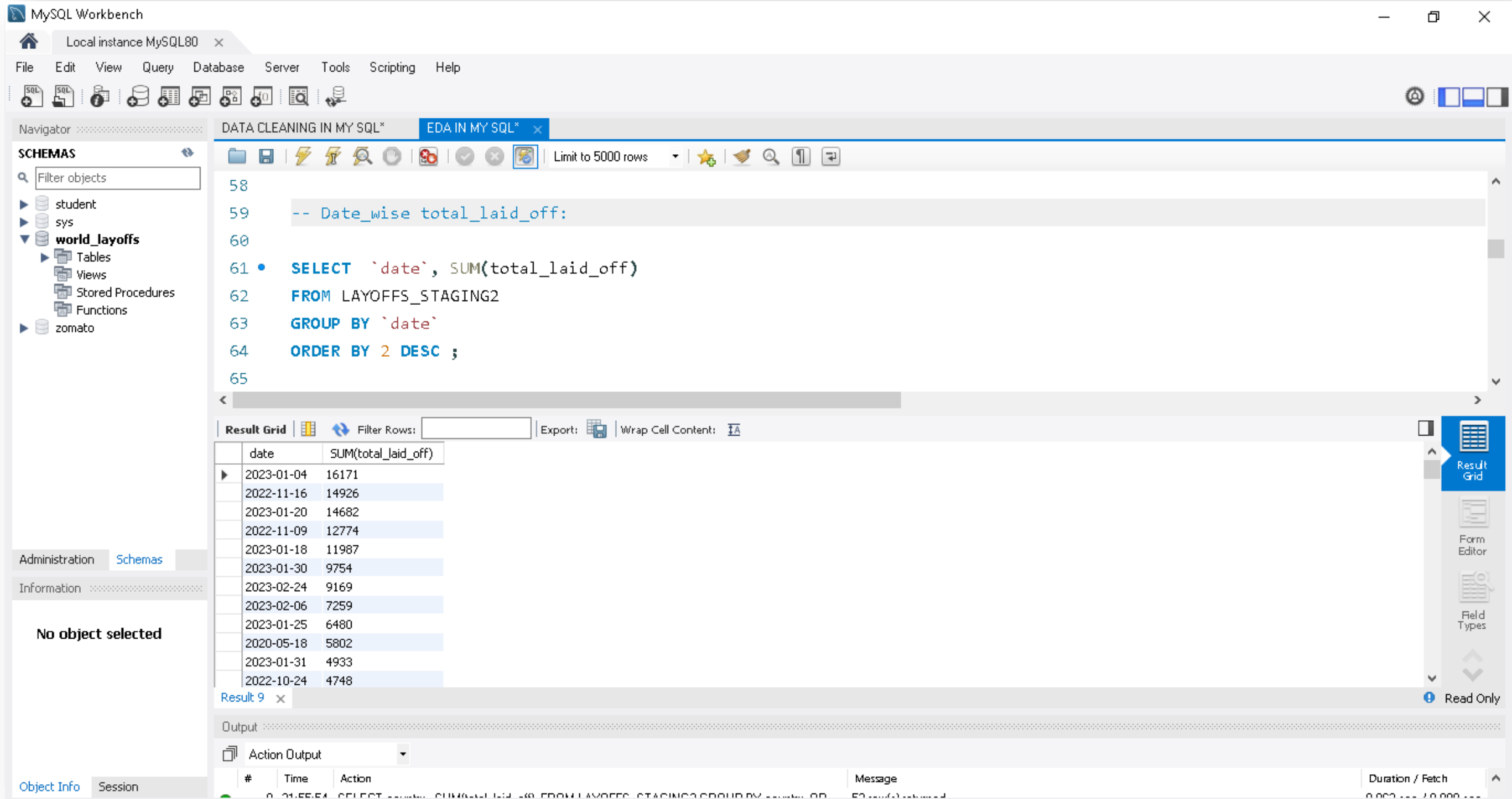
No object selected

Result 1

Output

Read Only





MySQL Workbench

Local instance MySQL80

FileEditViewQueryDatabaseServerToolsScriptingHelp

SQL

SQL

+

+

+

+

+

+

+

+

+

+

Limit to 5000 rows

★

🔍

📄

↺

Navigator

SCHEMAS

Filter objects

▶ student

▶ sys

▼ world_layoffs

▶ Tables

▶ Views

▶ Stored Procedures

▶ Functions

▶ zomato

DATA CLEANING IN MY SQL*

EDA IN MY SQL*

73

74 -- Stage_wise total_laid_offs :

75

76 • SELECT stage , SUM(total_laid_off)

77 FROM LAYOFFS_STAGING2

78 GROUP BY stage

79 ORDER BY 2 DESC ;

80

Result Grid

Filter Rows:

Export:

Wrap Cell Content:

	stage	SUM(total_laid_off)
▶	Post-IPO	204132
	Unknown	40716
	Acquired	27576
	Series C	20017
	Series D	19225
	Series B	15311
	Series E	12697
	Series F	9932
	Private Equity	7957
	Series H	7244
	Series A	5678
	Series G	3697

Result 11

Object Info

Session

Output

Action Output

#	Time	Action	Message	Duration / Fetch
11	21:57:07	SELECT YEAR(stage), SUM(total_laid_off) FROM LAYOFFS_STAGING2 GROUP BY YEAR(stage)	Execution completed	0.000 ... 0.000 ...



Navigator

SCHEMAS

Filter objects

- student
- sys
- world_layoffs
 - Tables
 - Views
 - Stored Procedures
 - Functions
- zomato

Administration Schemas

Information

No object selected

DATA CLEANING IN MY SQL*

EDA IN MY SQL* x

```
84 -- select month_wise only total_laid_offs:
85
86
87 • SELECT SUBSTRING(`date`,6,2) AS `month` , SUM(total_laid_off)
88 FROM LAYOFFS_STAGING2
89 GROUP BY `month`
90 ;
91
```

Result Grid Filter Rows: Export: Wrap Cell Content: IA

	month	SUM(total_laid_off)
▶	07	23415
	11	55758
	01	92037
	08	16891
	04	31099
	06	27455
	09	6651
	10	17878
	05	38689
	12	12381
	03	19859

Result 16 x

Output

Action Output

#	Time	Action	Message	Duration / Fetch
✓ 22	22:06:08	SELECT SUBSTRING(`date`,6,2) AS `month` , SUM(total_laid_off) FROM LAYOFFS_STAGIN...	13 row(s) returned	0.000 sec / 0.000 sec

Object Info Session

MySQL Workbench

Local instance MySQL80

FileEditViewQueryDatabaseServerToolsScriptingHelp

Navigator

SCHEMAS

Filter objects

student

sys

world_layoffs

Tables

Views

Stored Procedures

Functions

zomato

DATA CLEANING IN MY SQL*

EDA IN MY SQL*

Limit to 5000 rows

-- Using cte's for rolling the laid_off for each rows ;

WITH ROLLING_TOTAL AS

(

SELECT SUBSTRING(`date`,1,7) AS `month` , SUM(total_laid_off) AS TOTAL_OFF

FROM LAYOFFS_STAGING2

WHERE SUBSTRING(`date`,1,7) IS NOT NULL

GROUP BY `month`

ORDER BY 1 ASC

)

SELECT `month` , TOTAL_OFF ,

SUM(TOTAL_OFF) OVER(ORDER BY `month`) AS ROLLING_TOTAL

FROM ROLLING_TOTAL ;

Result Grid

Filter Rows:

Export:

Wrap Cell Content:

	month	TOTAL_OFF	ROLLING_TOTAL
▶	2020-03	9628	9628
	2020-04	26710	36338
	2020-05	25804	62142
	2020-06	7627	69769
	2020-07	7112	76881
	2020-08	1969	78850
	2020-09	609	79459
	2020-10	450	79909

Result 18

No object selected

Object Info

Session

Read Only



Navigator

SCHEMAS

Filter objects

- student
- sys
- world_layoffs
 - Tables
 - Views
 - Stored Procedures
 - Functions
- zomato

Administration Schemas

Information

No object selected

Object Info Session

DATA CLEANING IN MY SQL*

EDA IN MY SQL*

Limit to 5000 rows

```
125
126
127 -- Company and date_wise laid_offs ;
128
129 • SELECT company , `date` , SUM(total_laid_off)
130 FROM LAYOFFS_STAGING2
131 GROUP BY company , `date`
132 ;
133
134
```

Result Grid Filter Rows: Export: Wrap Cell Content: IA

	company	date	SUM(total_laid_off)
▶	Included Health	2022-07-25	NULL
	&Open	2022-11-17	9
	#Paid	2023-01-27	19
	100 Thieves	2022-07-13	12
	10X Genomics	2022-08-04	100
	1stdibs	2020-04-02	70
	2TM	2022-06-01	90
	2TM	2022-09-01	100
	2U	2022-07-28	NULL
	54gene	2022-08-29	95
	5B Solar	2022-06-03	NULL
	6sense	2022-10-12	150
	80 Acres Farm	2023-01-18	NULL

Result 20

Output

Result Grid

Form Editor

Field Types

Read Only

- student
- sys
- world_layouts**
 - Tables
 - Views
 - Stored Procedures
 - Functions
- zomato

Information

No object selected

Object Info

```

142 -- Company_wise ranking by total_lai_offs:
143
144 • WITH COMPANY_YEAR (company , years , total_laid_off) AS
145 (
146     SELECT company , YEAR(`date`), SUM(total_laid_off)
147     FROM LAYOFFS_STAGING2
148     GROUP BY company , YEAR(`date`)
149 ), COMPANY_YEAR_RANK AS
150 (
151     SELECT * ,
152     DENSE_RANK() OVER(PARTITION BY years ORDER BY total_laid_off DESC ) AS RANKING
153     FROM COMPANY_YEAR
154     WHERE years IS NOT NULL
155 )
156 SELECT *
157 FROM COMPANY_YEAR_RANK
158 WHERE RANKING <= 5 ;

```

Result Grid  Filter Rows:  Export:  Wrap Cell Content: 

	company	years	total_laid_off	RANKING
▶	Uber	2020	7525	1
	Booking.com	2020	4375	2
	Groupon	2020	2800	3
	Swiggy	2020	2250	4
	Airbnb	2020	1900	5

Result 22



Result Grid

 Read Only

MySQL Workbench

Local instance MySQL80

File Edit View Query Database Server Tools Scripting Help

Navigator

SCHEMAS

Filter objects

- student
- sys
- world_layoffs
 - Tables
 - Views
 - Stored Procedures
 - Functions
- zomato

Administration Schemas

Information

No object selected

Object Info Session

DATA CLEANING IN MY SQL* EDA IN MY SQL*

Limit to 5000 rows

```
166 -- Industry_wise ranking by total_laid_offs:
167
168 • WITH INDUSTRY_YEAR (industry , years , total_laid_off) AS
169 (
170     SELECT industry , YEAR(`date`) , SUM(total_laid_off)
171     FROM LAYOFFS_STAGING2
172     GROUP BY industry , YEAR(`date`)
173 ) , INDUSTRY_YEAR_RANK AS
174 (
175     SELECT * ,
176     DENSE_RANK() OVER(PARTITION BY YEARS ORDER BY total_laid_off DESC ) AS RANKING
177     FROM INDUSTRY_YEAR
178     WHERE YEARS IS NOT NULL
179 )
180 SELECT *
181 FROM INDUSTRY_YEAR_RANK
182 WHERE RANKING <= 5 ;
```

Result Grid

	industry	years	total_laid_off	RANKING
▶	Transportation	2020	14656	1
	Travel	2020	13983	2
	Finance	2020	8624	3
	Retail	2020	8002	4
	Food	2020	6218	5

Result 23

Read Only

MySQL Workbench

Local instance MySQL80

FileEditViewQueryDatabaseServerToolsScriptingHelp

Navigator

SCHEMAS

Filter objects

student

sys

world_layoffs

Tables

Views

Stored Procedures

Functions

zomato

Administration

Schemas

Information

No object selected

DATA CLEANING IN MY SQL*

EDA IN MY SQL*

Limit to 5000 rows

```
183  -- Stage_wise ranking by total_laid_offs
184  •  WITH STAGE_YEAR (stage , years , total_laid_off) AS
185      (
186          SELECT stage , YEAR(`date`) ,SUM(total_laid_off)
187          FROM LAYOFFS_STAGING2
188          GROUP BY stage, YEAR(`date`)
189      ) , STAGE_YEAR_RANK AS
190      (
191          SELECT * ,
192          DENSE_RANK() OVER(PARTITION BY years ORDER BY total_laid_off DESC ) AS RANKING
193          FROM STAGE_YEAR
194          WHERE YEARS IS NOT NULL
195      )
196      (
197          SELECT *
198          FROM STAGE_YEAR_RANK
199          WHERE RANKING <= 5
```

Result Grid

Filter Rows:

Export:

Wrap Cell Content:

	stage	years	total_laid_off	RANKING
▶	Post-IPO	2020	22672	1
	Acquired	2020	12169	2
	Series D	2020	8837	3
	Unknown	2020	7632	4
	Series B	2020	5609	5

Result 24

Output

Object Info

Session

Result Grid

Read On

THANK YOU