

Assignment 4: Normalization

Fears Away Database

Incidents Table:

MySQL Workbench

Local instance MySQL80 x

File Edit View Query Database Server Tools Scripting Help

Navigator

SCHEMAS

Filter objects

fears_away

Tables

cops_info

incident_type

incidents

location

Views

Stored Procedures

Functions

fears_away

sakila

sys

world

Administration Schemas

Information

No object selected

1 • select * from incidents;

2

3

4

5

6

7

Result Grid

Filter Rows:

Edit:

Export/Import:

Wrap Cell Contents:

Fetch rows:

incident_number	incident_desc	incident_district	incident_year	incident_month	incident_hour	location	zip
10637224	WARRANT ARREST - BOSTON WARRANT (MUS...	D4	2021	1	18	NORTHAMPTON ST	2124
120221042	SICK/INJURED/MEDICAL - PERSON	E5	2021	4	10	WASHINGTON ST	2215
121000816	NOISY PARTY/RADIO-NO ARREST	E5	2021	1	0	BRADFELD AVENUE	2131
121005808	FRAUD - FALSE PRETENSE / SCHEME	B3	2021	1	12	CLARKWOOD ST	2131
121016052	M/V - LEAVING SCENE - PROPERTY DAMAGE	C11	2021	3	9	DORCHESTER AVE	2131
121033495	DISTURBING THE PEACE/ DISORDERLY CONDU...	C6	2021	5	15	SOUTHAMPTON ST	2125
121055112	PROPERTY - LOST/ MISSING	E13	2021	8	21	WASHINGTON ST	2215
121069865	LARCENY THEFT FROM BUILDING	C6	2021	9	11	WEST BROADWAY	2131
121090532	SICK ASSIST	E18	2021	12	6	GORDON AVE	2122
121094465	PROPERTY - LOST/ MISSING	B2	2021	12	23	WASHINGTON ST	2115
20021670	LARCENY SHOPLIFTING	D4	2021	5	17	HARRISON AVENUE	2124
201906722	TRESPASSING		2021	6	18	WESLEYAN PLACE	2131
202001105	INVESTIGATE PERSON	C6	2020	1	6	E FOURTH ST	2119
202001106	SEX OFFENSE - RAPE - OTHER	C6	2020	1	0	ALLSTATE RD	2125
202001107	AUTO THEFT	A1	2020	1	6	NEW SUDBURY ST	2127
202001108	INVESTIGATE PERSON	B2	2020	1	7	PARKER HILL AVE	2132
202001109	VERBAL DISPUTE	C6	2020	1	8	OLD COLONY AVE	2124
202001110	LARCENY ALL OTHERS	B3	2020	1	8	WILDWOOD ST	2135
202001111	LARCENY THEFT FROM MV - NON-ACCESSORY	B3	2020	1	18	BLUE HILL AVE	2119

Incidents 143 x

Incidents table contains the Boston crime incidents across different neighborhoods along with their street address.

- Primary key for this table is Incident Number

1st Normal Form

- This table has one primary key which is incident_number
- No column of this table contains multi value attributes.
- There is unique name for every column.
- Two columns of this table do not store similar information.

2nd Normal Form

- All requirements for 1st NF are meeting.
- No separate table need to be created as it has no redundant data.
- The tables are related to each other by use of foreign keys-incident_id.

3rd Normal Form

- All requirements for 2nd NF.
- There are no fields in this table which do not depend on primary key of the table.
- There are no fields dependent on the primary key or any another field.

Note: First I was confused that location and zip code has transitive dependency but even if I break it down none of them can be the primary key as 1 zip can have different streets.

Incident_Type Table:

Local instance MySQL80 x

File Edit View Query Database Server Tools Scripting Help

Navigator

SCHEMAS

Filter objects

▼ fear_away

▼ Tables

- cops_info
- incident_type
- incidents
- location

Views

Stored Procedures

Functions

▼ fears_away

▼ sakila

▼ sys

▼ world

Administration Schemas

Information

No object selected

1 • select * from incident_type;

2

3

4

5

Result Grid

Filter Rows:

Export: Wrap Cell Co

incident_id	incident_code	incident_type	locality
212091685	617	Ballistics	A ST
I192072665	301	Robbery	A ST
I192071858	614	Larceny From Motor Vehicle	A ST
I192070640	1402	Vandalism	A ST
I192067402	1402	Vandalism	A ST
I192070512	1402	Vandalism	A ST
202001831	3301	Investigate Person	ABBOT ST
I192072506	2610	Other	ABBOT ST
I192072265	3501	Missing Person Reported	ABBOT ST
I182046873	2906	Violations	ACADEMY RD
I182047264	2647	Other	ACADEMY TER
222094543	3114	Larceny From Motor Vehicle	ADAMS ST
212091736	3831	Larceny	ADAMS ST
212092216	1402	Larceny	ADAMS ST
202053871	3001	Other	ADAMS ST
202053839	2647	Counterfeiting	ADAMS ST
202053772	3005	Firearm Violations	ADAMS ST
202052851	241	Property Found	ADAMS ST
202052758	3006	Medical Assistance	ADAMS ST
202052649	2647	Investigate Property	ADAMS ST
202052431	3301	Towed	ADAMS ST
202052427	3301	Other	ADAMS ST
202052036	3114	Motor Vehicle Accident Res...	ADAMS ST
202052312	2906	Other Burglary	ADAMS ST

Incident_type contains the different types of incidents across different locations in Boston.

- Primary key for this table is Incident_Id

1st Normal Form

- This table has one primary key which is incident_number
- No column of this table contains multi value attributes.
- There is unique name for every column.
- Two columns of this table do not store similar information.

2nd Normal Form

- All requirements for 1st NF are meeting.
- No separate table need to be created as it has no redundant data.
- The tables are related to each other by use of foreign keys-incident_number, zip_code.

3rd Normal Form

- All requirements for 2nd NF.
- There are no fields in this table which do not depend on primary key of the table.
- There are no fields dependent on the primary key or any another field.

Cops_Info Table:

MySQL Workbench

Local instance MySQL80 x

File Edit View Query Database Server Tools Scripting Help

Navigator

SCHEMAS

Filter objects

▼ fear_away

▼ Tables

► cops_info

► incident_type

► incidents

► location

Views

Stored Procedures

Functions

► fears_away

► sakila

► sys

► world

Administration Schemas

Information

No object selected

1 • select * from cops_info;

2

3

4

5

Result Grid

Filter Rows:

Edit:

Export/Import:

Wrap Cell Content:

	cop_id	cop_badge	cop_name	cop_title	cop_district	zip_code	neighborhood
	7481	1975	Worthy, Jeffrey Ma...	Police Officer	C6	2119	Roxbury
	7546	4188	Ajemian, Gerald F	Police Officer	E5	2081	Walpole
	7700	1261	Mullan, John P	Police Officer	A1	1821	Billerica
	7906	61	Harris, Joseph E	Supn Bpd	D4	2186	Milton
	7995	2194	Cameron, Leverage	Police Offc/Juvenile Offc 4\$10	A1	2119	Roxbury
	8007	2100	Gonzalez, Ivan P	Police Officer	A7	2131	Roslindale
	8010	2117	Hawkins, James D	Police Officer	B2	2155	Medford
	8013	2162	Kincade, Horace N	Police Officer Hdq Dispatch	C11	2125	Dorchester
	8016	604	Lynch, Timothy E	Police Detective	E13	2171	Quincy
	8021	583	Novo, Miguel A	Police Detective	C11	2184	Braintree
	8022	221	O'Brien, Kenneth R	Police Sergeant (Det)	B2	2135	Allston/Brighton
	8023	5018	Precia, Carmel E	Police Detective	A1	2124	Dorchester
	8116	899	Craven, Roberta C	Police Officer	B2	2136	Hyde Park
	8179	1684	Sullivan, Laurence J	Police Officer	B3	2128	East Boston
	8180	98	Wilson, Charles E	Supn Bpd	D14	2129	Charlestown
	8181	846	Kennedy, Joseph M	Police Detective	D4	1906	Saugus
	8190	1015	Roberto, David P	Police Officer	D4	2067	Sharon
	8191	82	Torigian, Arthur G	Police Lieutenant (Det)	B2	1770	Sherborn
	8195	234	Woodley, William J.	Police Sergeant (Det)	D4	2301	Brockton
	8196	405	Manning, Thomas S	Police Sergeant/Comm Serv Offc	A7	2339	Hanover
	8205	1568	Butler, Michael V	Police Offc/Auto Invest 4\$10	B2	2056	Norfolk
	8208	1553	Curry, Michael E	Police Offc Acad Instr 2\$6	E13	2119	Roxbury
	8210	10	Dowd, Thomas A	Police Captain (Det)	C11	2780	Taunton
	8215	1643	Conrad, Peter A	Police Officer	E5	2128	East Boston

Cops info table contains details of officers in charge for each neighborhood in Boston.

- Primary key for this table is Cops_Id

1st Normal Form

- This table has one primary key which is incident_number
- No column of this table contains multi value attributes.
- There is unique name for every column.
- Two columns of this table do not store similar information.

2nd Normal Form

- All requirements for 1st NF are meeting.
- No separate table need to be created as it has no redundant data.
- The tables are related to each other by use of foreign keys-zip.

3rd Normal Form

- All requirements for 2nd NF.
- There is transitional dependency between neighborhood and zip code in order to remove it a new table location is created which stores the data of all the neighborhoods along with their zip codes.
- The cops_info table is altered, and it just contains zip codes.

MySQL Workbench

The screenshot shows the MySQL Workbench interface for a local instance of MySQL 8.0. The 'SCHEMAS' sidebar on the left lists databases: fear_away, sakila, sys, and world. The 'fear_away' database is expanded, showing tables: cops_info, incident_type, incidents, and location. The 'Query' tab is active, displaying the SQL query: `select * from location`. The 'Result Grid' shows the results of this query, displaying columns 'zip_code' and 'neighbourhood'.

zip_code	neighbourhood
2210	South Boston
2132	West Roxbury
2135	Allston-Brighton
2127	South Boston
2114	Financial District
2136	Hyde Park
2131	Roslindale
2122	Dorchester
2129	Charlestown
2119	Roxbury
2124	Dorchester
2128	East Boston
2130	Jamaica Plain
2125	Roxbury
2126	Hyde Park
2118	South End
2115	Back Bay
2120	Mission Hill
2134	Allston-Brighton
2116	South End
2109	Financial District
2111	Financial District

Altered table cops_info

The screenshot shows the MySQL Workbench interface. On the left, the 'SCHEMAS' pane shows a tree view of the 'fear_away' database, including tables like 'cops_info', 'incident_type', 'incidents', and 'location'. The 'cops_info' table is selected. The main query editor shows the SQL query: `select * from cops_info;`. The 'Result Grid' pane displays the query results in a table format.

cop_id	cop_badge	cop_name	cop_title	cop_district	neighborhood
7700	1261	Mullan, John P	Police Officer	A1	Billerica
7906	61	Harris, Joseph E	Supn Bpd	D4	Milton
7995	2194	Cameron, Leverage	Police Offc/Juvenile Offc 4\$10	A1	Roxbury
8007	2100	Gonzalez, Ivan P	Police Officer	A7	Roslindale
8010	2117	Hawkins, James D	Police Officer	B2	Medford
8013	2162	Kincade, Horace N	Police Officer Hdq Dispatch	C11	Dorchester
8016	604	Lynch, Timothy E	Police Detective	E13	Quincy
8021	583	Novo, Miguel A	Police Detective	C11	Braintree
8022	221	O'Brien, Kenneth R	Police Sergeant (Det)	B2	Allston/Brighton
8023	5018	Precia, Carmel E	Police Detective	A1	Dorchester
8116	899	Craven, Roberta C	Police Officer	B2	Hyde Park
8179	1684	Sullivan, Laurence J	Police Officer	B3	East Boston
8180	98	Wilson, Charles E	Supn Bpd	D14	Charlestown
8181	846	Kennedy, Joseph M	Police Detective	D4	Saugus
8190	1015	Roberto, David P	Police Officer	D4	Sharon
8191	82	Torigian, Arthur G	Police Lieutenant (Det)	B2	Sherborn
8195	234	Woodley, William J.	Police Sergeant (Det)	D4	Brockton
8196	405	Manning, Thomas S	Police Sergeant/Comm Serv Offc	A7	Hanover
8205	1568	Butler, Michael V	Police Offc/Auto Invest 4\$10	B2	Norfolk
8208	1553	Curry, Michael E	Police Offc Acad Instr 2\$6	E13	Roxbury
8210	10	Dowd, Thomas A	Police Captain (Det)	C11	Taunton
8215	1643	Gonzalez, Dino A	Police Officer	E5	East Boston

Location Table:

Location table contains all the neighborhoods tagged to a zip_code.

- Primary key for this table is zip_code.

1st Normal Form

- This table has one primary key which is incident_number
- No column of this table contains multi value attributes.
- There is unique name for every column.

- Two columns of this table do not store similar information.

2nd Normal Form

- All requirements for 1st NF are meeting.
- No separate table need to be created as it has no redundant data.
- The tables are related to each other by use of foreign keys-zip.

3rd Normal Form

- All requirements for 2nd NF.
- There are no fields in this table which do not depend on primary key of the table.
- There are no fields dependent on the primary key or any another field.

Final SQL

1. Count the total number of incidents occurred in boston in last 5 years?

```
SELECT count(incident_number)
FROM incidents
where incident_year between 2018 and 2022;
```

2. List the top 5 neighborhoods in Boston with highest crime incidents?

```
SELECT l.neighborhood, count(i.incident_number) as total_incidents
FROM location l RIGHT JOIN incidents i
ON l.zip_code = i.zip
where i.incident_year = 2022
group by neighborhood
order by count(i.incident_number) desc
limit 5;
```

3. Count the harassment incidents on Boston streets

```
SELECT l.neighborhood, t.incident_type, count(i.incident_number) as total_incidents
FROM incidents i LEFT JOIN incident_type t
ON t.incident_id = i.incident_number
LEFT JOIN location l
ON l.zip_code = i.zip
where t.incident_type = 'Harassment'
group by l.neighborhood
order by count(i.incident_number) desc;
```

4. Who are the cops in charge for Roxbury?

```
SELECT c.cop_name, c.cop_title, l.neighborhood
FROM cops_info c
left JOIN location l
ON c.zip_code = l.zip_code
where l.neighborhood = 'Roxbury'
order by c.cop_name;
```

5. What is the most recurring incident on Adams Street in a one year period.

```
SELECT count(incident_number), t.incident_type
FROM incidents i LEFT JOIN incident_type t
ON i.location = t.locality
where t.locality = 'ADAMS ST' and incident_year between 2020 and 2021
group by t.incident_type
order by count(i.incident_number) desc
limit 1;
```

6. The above use case determines the highly unsafe street in Boston.

```
SELECT count(i.incident_number) as incident_count, i.location, l.neighbourhood, l.zip_code
FROM incidents i RIGHT JOIN location l
ON i.zip = l.zip_code
where incident_year <= 2022 and incident_year >=2018
group by i.location, l.neighbourhood, l.zip_code
order by count(incident_number) desc
limit 1;
```

7. Which year marks the highest number of incidents?

```
SELECT count(i.incident_number) as incident_count, i.incident_year
FROM incidents i RIGHT JOIN location l
ON i.zip = l.zip_code
where i.zip = l.zip_code
group by i.incident_year
order by count(incident_number) desc
limit 1;
```

8. Between what hours of day most incidents happen?

```
SELECT count(i.incident_number) as incident_count, i.incident_hour, l.neighbourhood, l.street
FROM incidents i RIGHT JOIN location l
ON i.zip = l.zip_code
where i.zip = l.zip_code
group by i.incident_hour, l.neighbourhood
```

```
order by count(incident_number) desc
limit 1;
```

9. How many incidents happen after midnight night?

```
SELECT count(incident_number) from incidents
where incident_hour between 1 and 6
group by incident_hour
order by count(incident_number) desc;
```

VIEWS CREATED FOR THE USECASES

1. Count the total number of incidents occurred in boston in last 5 years?

```
VIEW: CREATE VIEW total_incidents AS count(incident_number) FROM incidents
where incident_year between 2018 and 2022;
```

2. List the top 5 neighborhoods in Boston with highest crime incidents?

```
VIEW: CREATE VIEW crime_statistics_nbr AS SELECT l.neighborhood,
count(i.incident_number) as total_incidents FROM location l RIGHT JOIN incidents i ON
l.zip_code = i.zip where i.incident_year = 2022 group by neighborhood order by
count(i.incident_number) desc limit 5;
```

3. Count the harassment incidents on Boston streets

```
VIEW: CREATE VIEW harassment_stats AS
SELECT l.neighborhood, t.incident_type, count(i.incident_number) as total_incidents
FROM incidents i LEFT JOIN incident_type t
ON t.incident_id = i.incident_number
LEFT JOIN location l
ON l.zip_code = i.zip
where t.incident_type = 'Harassment'
group by l.neighborhood
order by count(i.incident_number) desc;
```

4. Who are the cops in charge for Roxbury neighbourhood?

```
VIEW: CREATE VIEW cops_info AS
SELECT c.cop_name, c.cop_title, l.neighborhood
```

```

FROM cops_info c
left JOIN location l
ON c.zip_code = l.zip_code
where l.neighborhood = 'Roxbury'
order by c.cop_name;

```

5. What is the most recurring incident on Adams Street in a one year period.

```

VIEW: CREATE VIEW recur_incidents AS

SELECT count(incident_number), t.incident_type
FROM incidents i LEFT JOIN incident_type t
ON i.location = t.locality
where t.locality = 'ADAMS ST' and incident_year between 2020 and 2021
group by t.incident_type
order by count(i.incident_number) desc limit 1;

```

6. Determine the most unsafe street in Boston.

```

VIEW: CREATE VIEW unsafe_boston AS
SELECT count(i.incident_number) as incident_count, i.location, l.neighbourhood,
l.zip_code
FROM incidents i RIGHT JOIN location l
ON i.zip = l.zip_code
where incident_year <= 2022 and incident_year >=2018
group by i.location, l.neighbourhood, l.zip_code
order by count(incident_number) desc
limit 1;

```

7. Which year marks the highest number of incidents?

```

VIEW--> CREATE VIEW crime_stats_year AS
SELECT count(i.incident_number) as incident_count, i.incident_year
FROM incidents i RIGHT JOIN location l
ON i.zip = l.zip_code
where i.zip = l.zip_code
group by i.incident_year
order by count(incident_number) desc
limit 1;

```

8. Between what hours of day most incidents happen?

VIEW--> CREATE VIEW crime_stat_hour AS

```
SELECT count(i.incident_number) as incident_count, i.incident_hour, l.neighbourhood,  
FROM incidents i RIGHT JOIN location l  
ON i.zip = l.zip_code  
where i.zip = l.zip_code  
group by i.incident_hour, l.neighbourhood, l.neighbourhood  
order by count(incident_number) desc  
limit 1;
```

9. How many incidents happen after midnight night?

VIEW--> CREATE VIEW crime_stat_night AS
SELECT count(incident_number) from incidents
where incident_hour between 1 and 6
group by incident_hour
order by count(incident_number) desc;