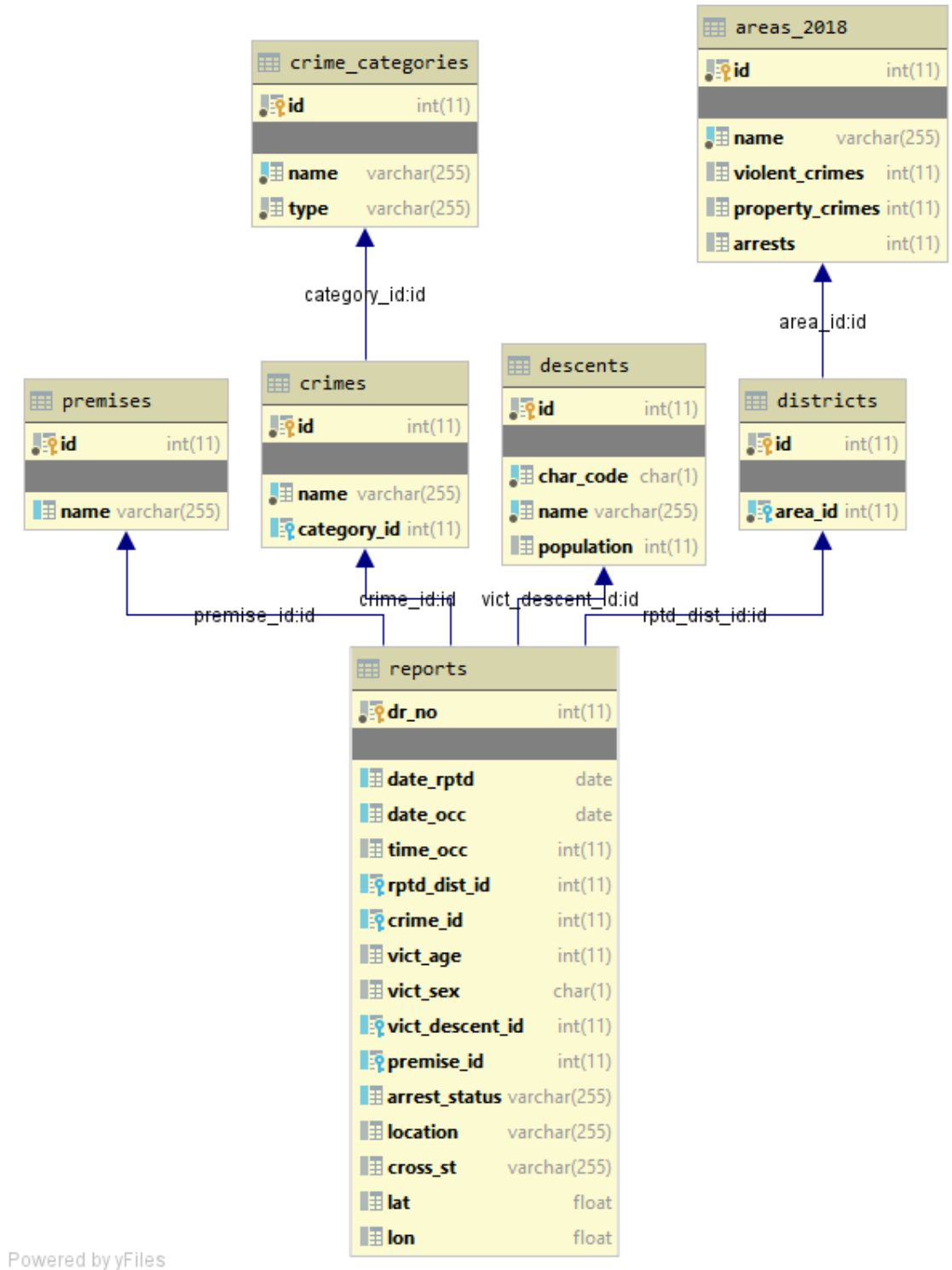
DECISION OUTLINE

- Crime_categories table
    - I wanted more information about the type and category of each crime that is mentioned in the ***reports*** table. I researched Uniform Crime Reports (UCR) to what the crime categories are and what type of crime they are.
- Crimes table
    - I separated the crime names from the main reports table (and made their codes into a foreign key) into the ***crimes*** table to serve as a connector table to the ***crime_categories*** table, which stores information about what the general category and type of each crime is.
    - Crime names should also be separated since their definitions/name could be updated in the future.
- Areas_2018 table
    - I wanted to get more information about the geographic/patrol areas of the crimes reported in the ***reports*** table. The ***areas_2018*** table stores information about the number of crimes and arrests that occurred in 2018, which is the year prior to the 2019 reports in the ***reports*** table.
- Districts table
    - Areas are broken down into districts. I created the districts table to create a connection between the ***reports*** table and the ***areas_2018*** table.
- Descents table
    - I wanted to include demographic information so I separated victim descent information from the ***reports*** table into the ***descents*** table.
- Premises table
    - Some premise names are long and wordy and their definitions could be updated in the future so I separated the premise names from the ***reports*** table into the ***premises*** table.
- Reports table
    - I made the dr_no column the primary key, since it is a unique number used to identify each crime report.
    - The date_rptd and date_occ columns are indexed because several queries may need to be made based on date from the reports table, which is quite large.
    - crime_id, rpt_dist_id, vict_descent_id, and premise_id are foreign keys to connect to the associated tables.
    - I did not create a separate table to store arrest status names since the only update that could be made to the status of a crime report would need to be made to the specific crime report regardless of whether or not there is a separate crime status names table. I do not think an instance would arise where the name of the crime status would need to change.

ERD DIAGRAM


areas_2018
id int(11)
name varchar(255)
violent_crimes int(11)
property_crimes int(11)
arrests int(11)
crime_categories
id int(11)
name varchar(255)
type varchar(255)
category_id:id
area_id:id
premises
id int(11)
name varchar(255)
crimes
id int(11)
name varchar(255)
category_id int(11)
descents
id int(11)
char_code char(1)
name varchar(255)
population int(11)
districts
id int(11)
area_id int(11)
premise_id:id
crime_id:id
vict_descent_id:id
rptd_dist_id:id
reports
dr_no int(11)
date_rptd date
date_occ date
time_occ int(11)
rptd_dist_id int(11)
crime_id int(11)
vict_age int(11)
vict_sex char(1)
vict_descent_id int(11)
premise_id int(11)
arrest_status varchar(255)
location varchar(255)
cross_st varchar(255)
lat float
lon float
Powered by yFiles

DOCUMENTATION

reports table

Physical Element Name	Logical Element Name	Format Type	Format Length	Element Definition	Examples of Valid Values
dr_no (PK)	Division of Records Number	Int	11	Official file number made up of a 2-digit year, area ID, and 5 digits	001307355
date_rptd	Date reported	Date	-	The date the crime report was made	2019-01-01
date_occ	Date occurred	Date	-	The date the reported crime occurred	2019-01-01
time_occ	Time occurred	Int	11	The 24-hour military time the crime occurred	1300
rptd_dist_id (FK)	Reported district ID	Int	11	A four-digit code that represents a sub-area within a Geographic Area. All crime records reference the "RD" that it occurred in for statistical comparisons.	1385
crime_id (FK)	Crime ID	Int	11	The crime code used to identify the name of the crime committed	900
vict_age	Victim age	Int	11	The age of the victim	0,38
vict_sex	Victim sex	Char	1	F – Female, M – Male, X - Unknown	F, M, X
vict_descent_id (FK)	Victim descent ID	Int	11	The ID used to identify the descent of the victim	0,1,2
premise_id (FK)	Premise ID	Int	11	The ID used to identify the type of premise the crime occurred on	0,1,2
arrest_status	Arrest Status	Varchar	255	The current arrest status of the crime	Adult Arrest, Invest Cont
location	Location	Varchar	255	Street address of crime incident rounded to the nearest hundred block to maintain anonymity.	300 E GAGE AV

cross_st	Cross street	Varchar	255	Cross Street of rounded Address	MANCHESTER AV
lat	Latitude	Float	-	Latitude of crime location	33.9825
lon	Longitude	Float	-	Longitude of crime location	-118.2695

crimes table

Physical Element Name	Logical Element Name	Format Type	Format Length	Element Definition	Examples of Valid Values
id (PK)	ID	Int	11	The crime code used to identify the crime	900
name	Crime name	Varchar	255	The name of the crime	Pickpocket
category_id (FK)	Crime Category ID	Int	11	The ID used to identify the category of the crime	Personal Theft

crime_categories

Physical Element Name	Logical Element Name	Format Type	Format Length	Element Definition	Examples of Valid Values
id (PK)	ID	Int	11	The ID of the crime category	0,1
Name	Name	Varchar	255	The name of the crime category (Defined by Uniform Crime Reporting (UCR)	Pickpocket
Type	Type	Varchar	255	The type of the crime (defined by UCR)	Violent, property

districts table

Physical Element Name	Logical Element Name	Format Type	Format Length	Element Definition	Examples of Valid Values
id (PK)	Reported district ID	Int	11	The crime code used to identify the district the crime was committed in	1385
area_id (FK)	Area ID	Int	11	The ID of the area the district is located in	13

areas_2018

Physical Element Name	Logical Element Name	Format Type	Format Length	Element Definition	Examples of Valid Values
id (PK)	ID	Int	11	The code used to identify one of 21 areas in Los Angeles.	13
name	Area name	Varchar	255	The 21 Geographic Areas or Patrol Divisions are also given a name designation that references a landmark or the surrounding community that it is responsible for. For example 77th Street Division is located at the intersection of South Broadway and 77th Street, serving neighborhoods in South Los Angeles.	Hollywood
violent_crimes	Violent crimes	Int	11	The number of violent crimes reported in 2018	1654
property_crimes	Property crimes	Int	11	The number of property crimes in reported in 2018	673
arrests	Arrests	Int	11	The number of arrests made in 2018	1234

descents table

Physical Element Name	Logical Element Name	Format Type	Format Length	Element Definition	Examples of Valid Values
id (PK)	Descent ID	Int	11	The ID used to identify an individual's descent	1,2
char_code	Character code	Char	1	The single letter code used to identify the descent	W, H
name	Descent Name	Varchar	255	The name of the descent	White, Hawaiian
population	Population	Int	11	The 2010 population of the descent category In Los Angeles	12789

premises table

Physical Element Name	Logical Element Name	Format Type	Format Length	Element Definition	Examples of Valid Values
id (PK)	Premise ID	Int	11	The code used to identify the premise of the crime	501
Name	Premise name	Varchar	255	The type of structure, vehicle, or location where the crime took place.	SINGLE FAMILY DWELLING