Hypothetical Question: does crime and crime frequency effect building permits issued in St.Louis?

1. Source of data extract from:
   1. SLMPD (st. louis metro police department) crime csv for 12 months of 2018
      1. Combined all 12 months of 2018 into one dataframe
      2. Dropped: FlagCrime, FlagUnfounded, FlagAdministrative, FlagCleanup, LocationName, LocationComment
      3. Dropped all rows with ‘-1’ count value due to that crime being changed into a new one
         1. Example: assault reported originally, 2 months later victim dies, now converted to murder (preventing duplicate incidents from carrying over)
      4. Replaced all blank values with Null value
      5. Converted all column heads to lower case (makes it easier for pgadmin to join)
   2. City of St.Louis obtained csv for district numbers and their neighborhood name
      1. Converted all column heads to lower case (makes it easier for pgadmin to join)
      2. heads to lower case
   3. City of St. Louis obtained csv for building permits issues in all of 2018
      1. Shows number of permits, average day to issue permit, and dollar amount of all permits issued for each neighborhood
      2. Converted all column heads to lower case (makes it easier for pgadmin to join)
2. Import cleaned crime data frame, neighborhood/district number dataframe, and building permits dataframe to pgadmin database: stl\_crime6. **RELATIONAL DATABASE**
3. Now we have three tables in pgadmin and we need to have one database to query data to answer our hypothetical question.
   1. Under ‘Views’ in Schemas, create a hard database of all the three tables joined with the columns we are interested in (joined on neighborhood name, nei\_name). Big advantage of this is so that we can always go directly to the cleaned database without having to select and join the tables and our colleagues can easily just work directly with the database in Views.