Data warehouse Design

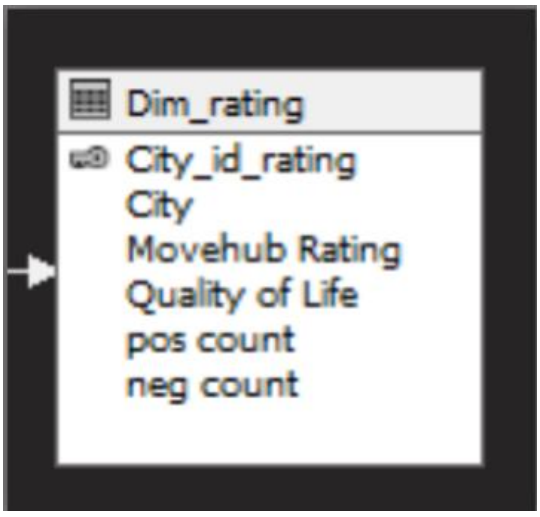
In this data warehouse we will be using four dimensions which are Dim_health, Dim_safety, Dim_living and Dim_rating.

Dim_health: This dimension has all the city measures related to health indexes and health rating of the cities. It includes city name, pollution rating, healthcare, climate and traffic commute index. From this we can analyze which city is better to live if health is considered.

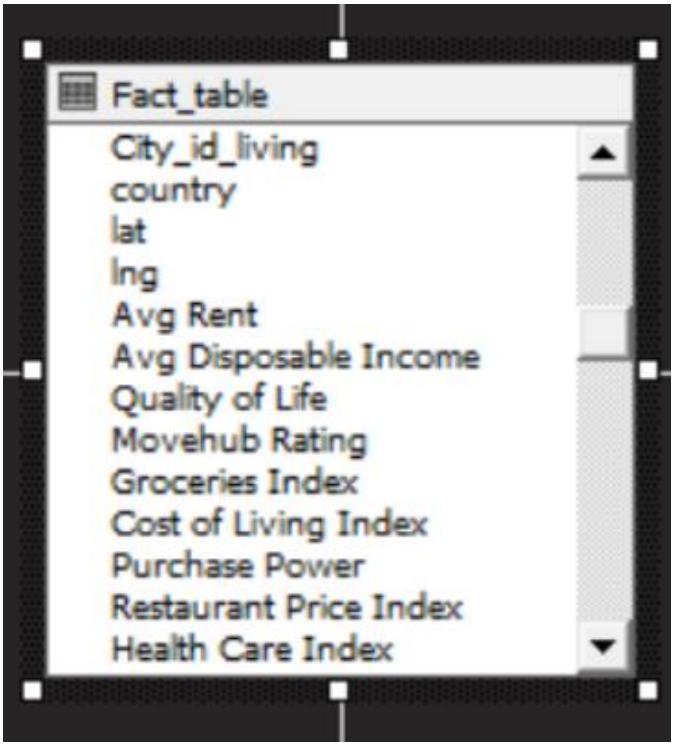
Dim_living: This dimension has the measures regarding the living life of the cities which includes the cost of living index, purchase power measures, average income and house rent in the cities, from this we can understand which city is better in terms of income vs expenditure.

Dim_safety: This dimension contains Safety rating of cities. Safety is an important aspect of any city. Crime ratings, geographical details of the cities and respective countries of the cities are in this dimension table.

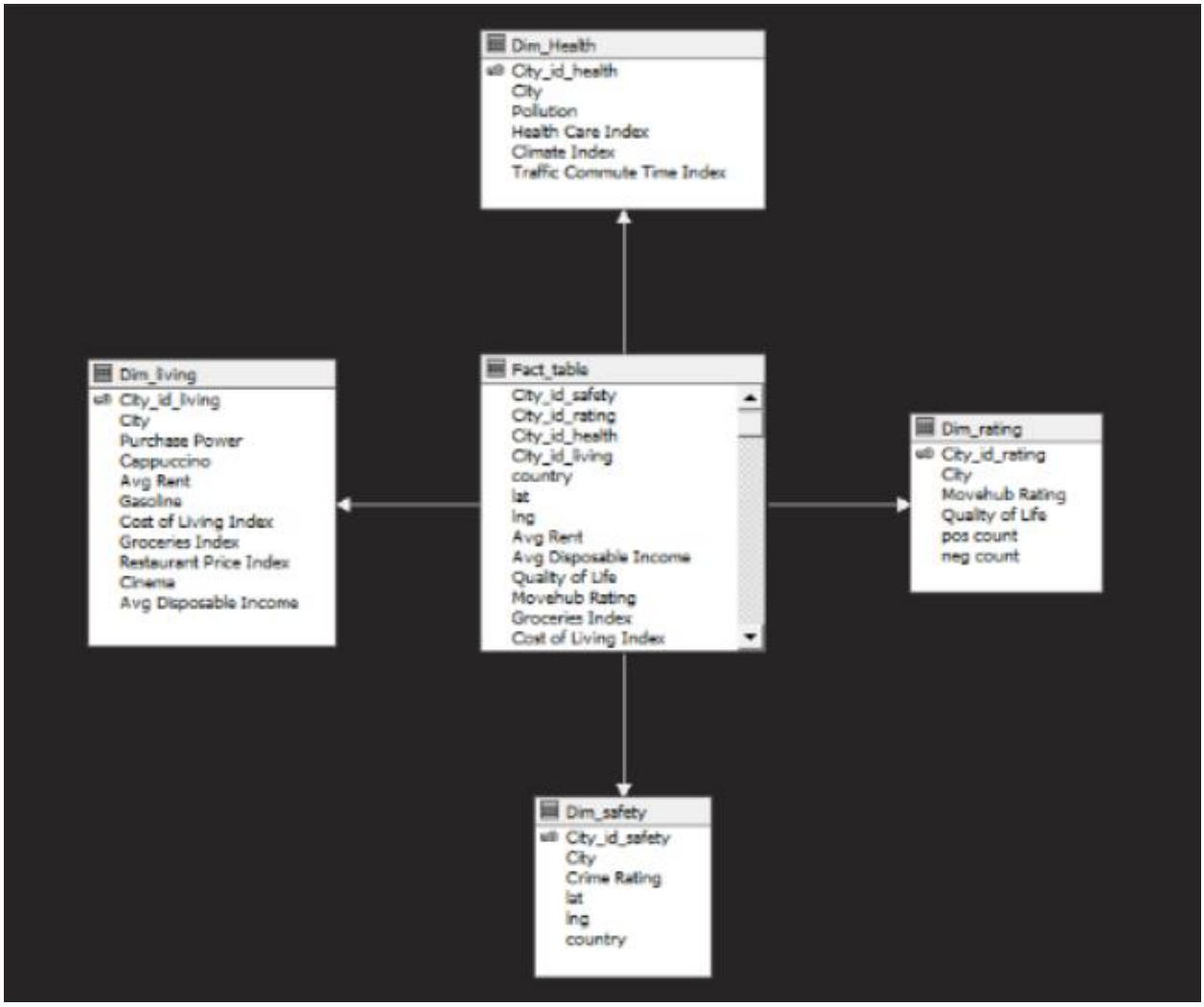
Dim_rating: Overall rating for the city measure is in this dimension. Positive and negative scores of the respective cities from the twitter are also added in this dimension.



Fact table: For our data warehouse all the dimension tables are connected to the fact table with the surrogate keys.



Star Schema: In this project star schema is used to connect fact tables with the dimensions table.



Star Schema for the project