

Data warehouse Design

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

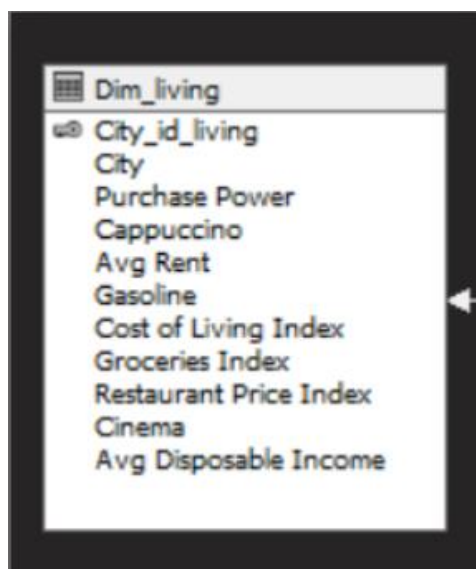
Dim_health: This dimension has all the city measures related to health indexes and health rating of the cities. It includes city name, pollution rating, healthcare, climate and traffic commute index. From this we can analyze which city is better to live if health is considered.



The screenshot shows a window titled "Dim_Health" with a list of attributes. The attributes are: City_id_health, City, Pollution, Health Care Index, Climate Index, and Traffic Commute Time Index. The window has a standard toolbar with icons for undo, redo, and search. The text "health_dimensions.PNG (196 x 127)" is visible at the bottom of the window.

Dim_Health
City_id_health
City
Pollution
Health Care Index
Climate Index
Traffic Commute Time Index

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.



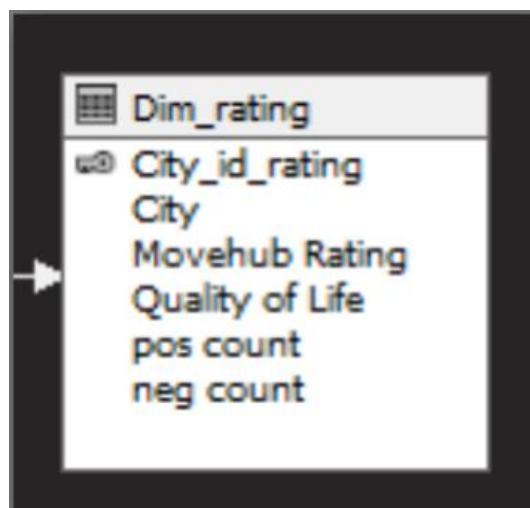
The screenshot shows a window titled "Dim_living" with a list of attributes. The attributes are: City_id_living, City, Purchase Power, Cappuccino, Avg Rent, Gasoline, Cost of Living Index, Groceries Index, Restaurant Price Index, Cinema, and Avg Disposable Income. The window has a standard toolbar with icons for undo, redo, and search. A white arrow points to the right side of the window.

Dim_living
City_id_living
City
Purchase Power
Cappuccino
Avg Rent
Gasoline
Cost of Living Index
Groceries Index
Restaurant Price Index
Cinema
Avg Disposable Income

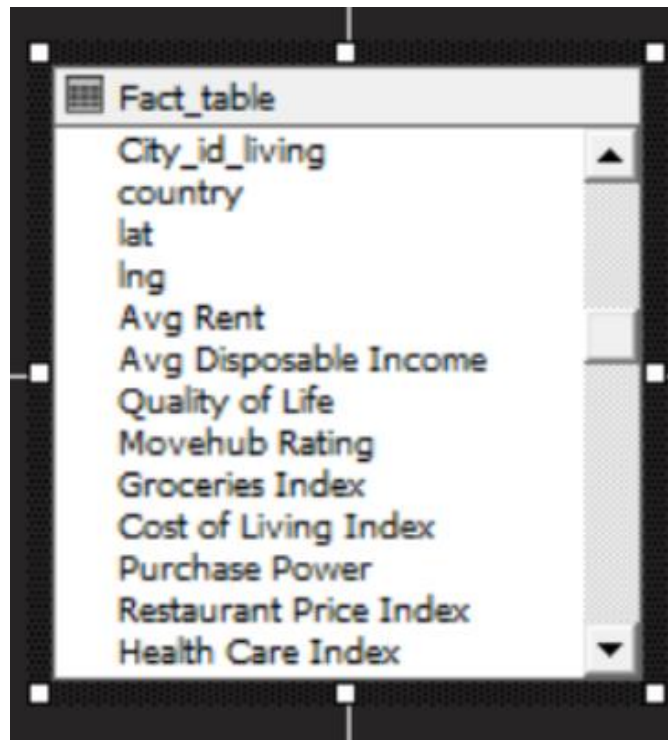
Dim_safety: This dimension contains Safety rating of cities. Safety is an important aspect of any city. Crime ratings, geographical details of the cities and respective countries of the cities are in this dimension table.



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.



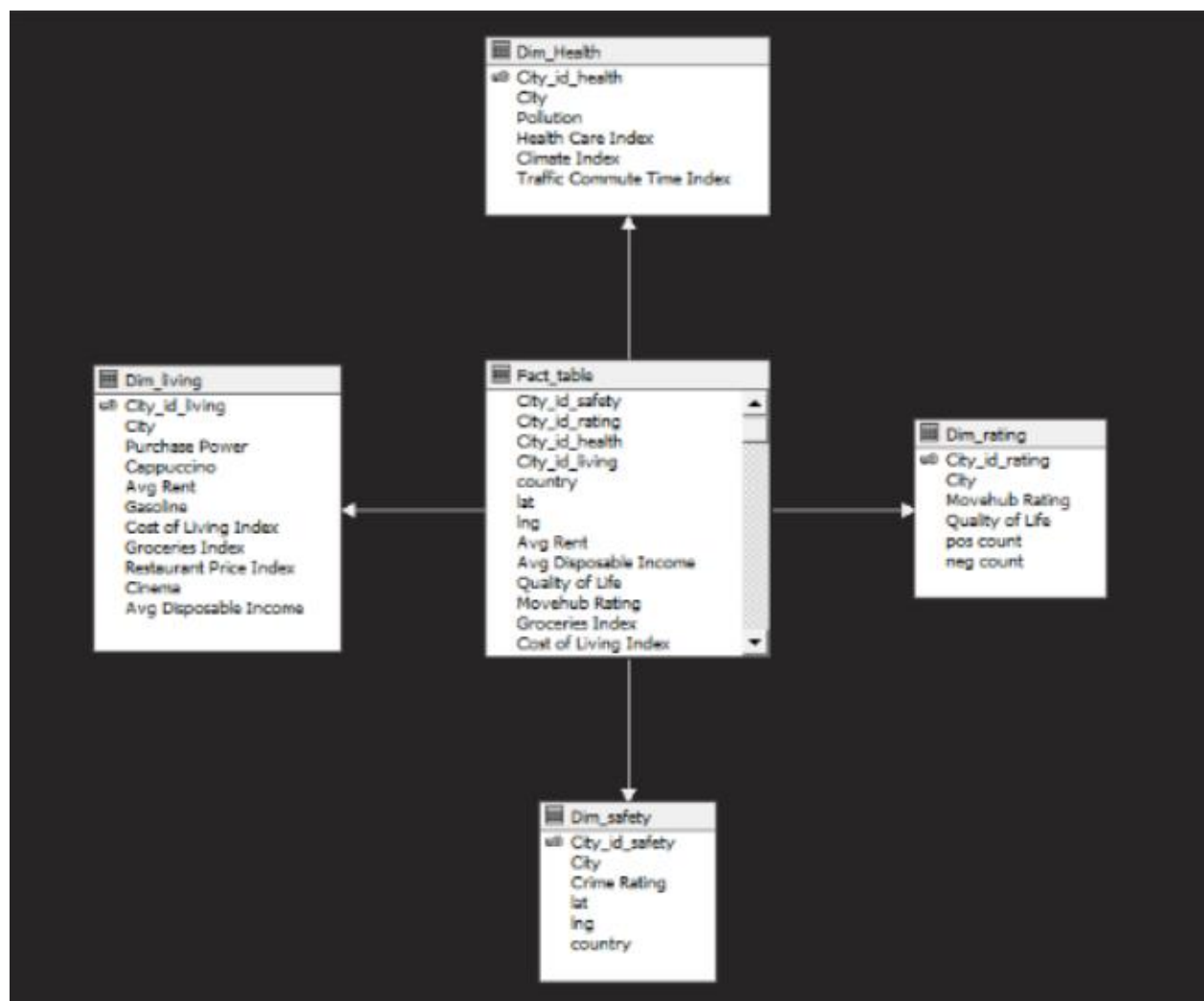
Fact table: For our data warehouse all the dimension tables are connected to the fact table with the surrogate keys.



A screenshot of a database table named 'Fact_table'. The table contains 14 rows of data. The first two rows are 'City_id_living' and 'country'. The remaining 12 rows are various indices and metrics: 'lat', 'lng', 'Avg Rent', 'Avg Disposable Income', 'Quality of Life', 'Movehub Rating', 'Groceries Index', 'Cost of Living Index', 'Purchase Power', 'Restaurant Price Index', and 'Health Care Index'. The table is displayed in a window with a scroll bar on the right.

Fact_table
City_id_living
country
lat
lng
Avg Rent
Avg Disposable Income
Quality of Life
Movehub Rating
Groceries Index
Cost of Living Index
Purchase Power
Restaurant Price Index
Health Care Index

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



Star Schema for the project