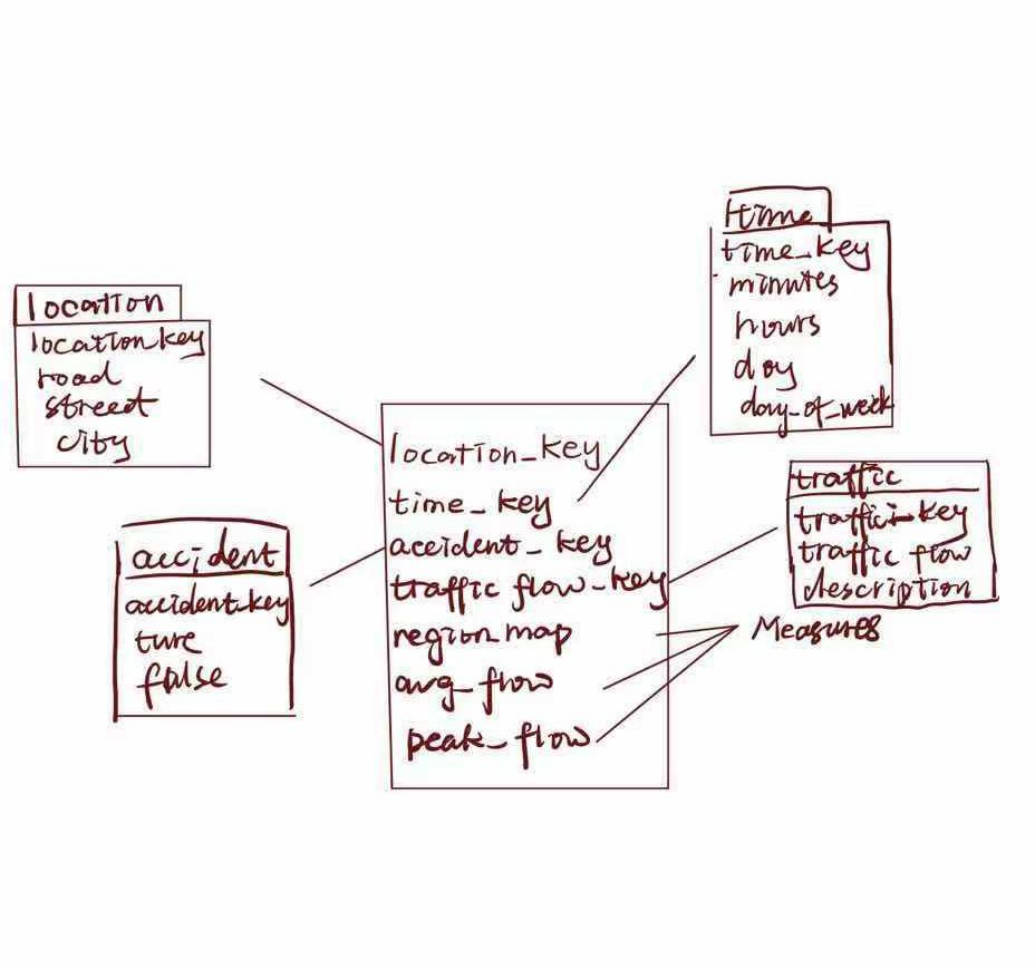
Question 1

Suppose that a city transportation department would like to perform data analysis on motorway traffic for the planning of a motorway construction based on the city traffic data collected at different hours every day.

1. Design a spatial data warehouse that stores the motorway traffic information so that people can easily see the average and peak time traffic flow by motorway, by time of day, and by weekdays, and the traffic situation when a major accident occurs.

Spatial data warehouse
Dimensions: location, time, accident, traffic flow
Measurements: region_map, avg_flow, peak_flow
Define cube trafficflow_star [location, time, accident, traffic flow]:
region_map = pointers to all motorway regions,
avg_flow=avg(traffic_flow),
peak_flow= max(traffic_flow)
Define dimension location as (location_key, road, street, city)
Define dimension time as (time_key, minutes, hours, day, day_of_week)
Define dimension accident as (accident_True/False)
Define traffic flow as(traffic flow_key, traffic flow description)



2. What information can we mine from such a spatial data warehouse to help city planners?

This data warehouse can demonstrate the average time of traffic flow (traffic intensities) and peak time of traffic flow according to the value of location user choose, the value of time, and when there is accident happens. By analyzing this spatial data warehouse, we can mine the information, such as when is the peak time of a certain location in the motorway, we can also detect changes and trends along a spatial dimension from the traffic jam location, and to help the city planners to change the traffic control rules. Furthermore, we can analysis the data warehouse to find the time period and location where the traffic accidents are likely to occur and the impact of the accidents will have to the traffic flow of the area nearby, in the way, the city planners can take measures to reduce accidents and solve the traffic jam problem when the accident happens.

3. This data warehouse contains both spatial and temporal data. Propose one mining technique that can efficiently mine interesting patterns from such a spatio-temporal data warehouse.

The mining technique classification can efficiently mine interesting patterns from such a spatio-temporal data warehouse. Analyse spatial objects and time, accident dimensions to derive classification schemes, such as decision trees in relevance to certain properties (location:location_1, time:7:00am-8:00am, accident_True:1) and divide the traffic flow into three bins: smooth, normal, jam to classy the traffic situation at a certain time and location.

Question 2

Traffic situations are often auto-correlated: the congestion at one motorway intersection may trigger the congestion in nearby motorway segments after a short period of time. Suppose we are given motorway traffic history data in Dublin, including road construction segment, traffic speed associated with motorway segment, direction, time, and so on. Moreover, we are given weather conditions by weather bureau in Dublin.

Design a data mining method to find high-quality spatio-temporal association rules that may guide us to predict what could be the expected traffic situation at a given motorway location.

In his question, we are provided the information of motorway segment, direction, time, and weather conditions. In this circumstance, we can firstly use Geographic Information Systems (GIS)-based data pre-processing method to integrate diverse data sets, extract spatio-temporal relationships, classify numeric data into ordinal categories, and encode spatio-temporal relationship data in tabular format for use by conventional (non-spatio temporal) association rule mining method.

Then, we can use Apriori algorithm for finding association rules. For the prediction using regression modeling methods might be suitable. Input output pairs denoted by (xi,yi) are approximated by a function of the form $y = f(x)$ and, for this motorway traffic case.