

COMP40370 – Data Mining

Practical 7

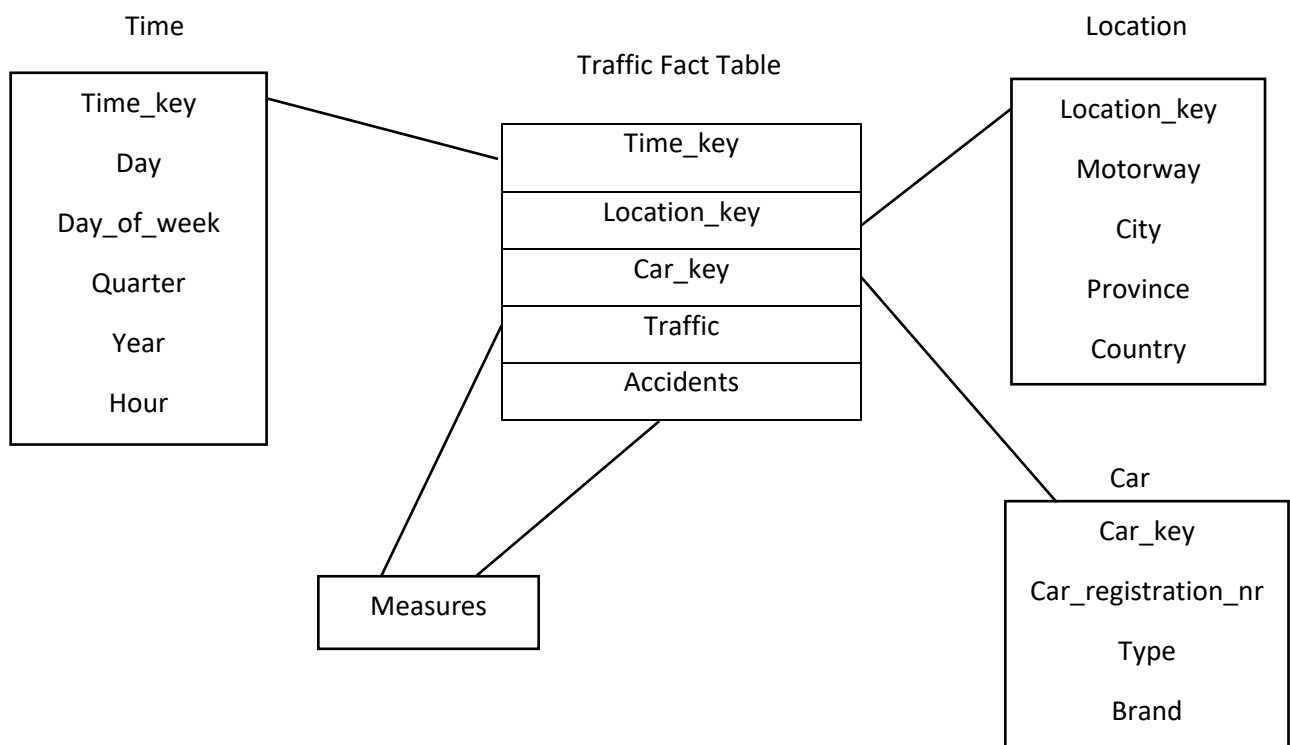
Weronika Wolska

17301623

Question 1:

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.

Star Schema:



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

We could mine which locations have heavy traffic and at which hours. We could also mine which places (or even specific cars) are likely to have an accident occur.

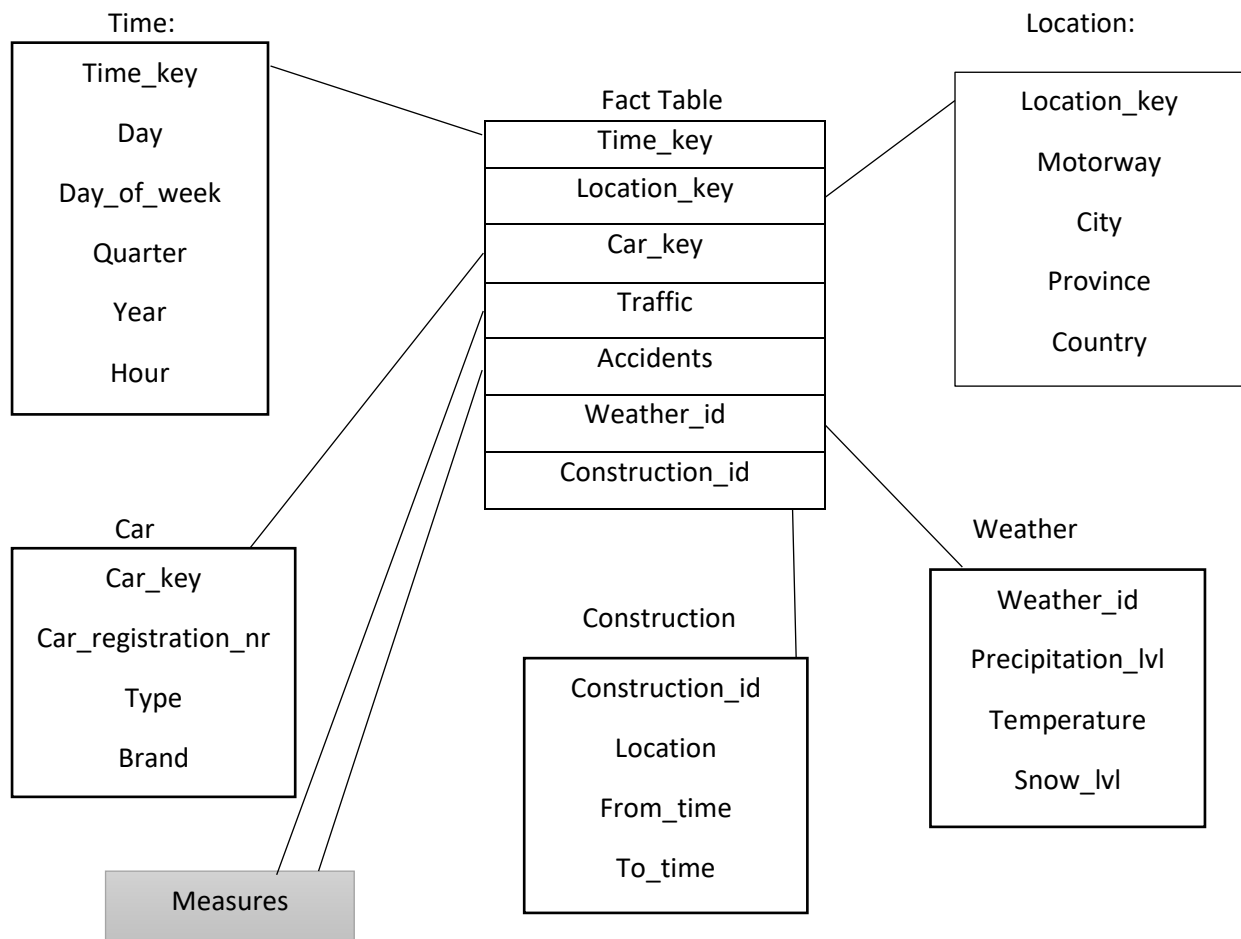
3. This data warehouse contains both spatial and temporal data. Propose one mining technique that can efficiently mine interesting patterns from such a spatiotemporal data warehouse.

DFS Mining

Question 2:

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

Redefine Star Schema:



You could use k-means clustering to find association rules. For example, to find traffic based on location and time, we could define a point as $x=\text{location}$ and $y=\text{time}$, then split the data into clusters to find which locations have heavy traffic at any particular time. We could expand this even further to add factors like weather and construction to get a more accurate representation of traffic. The bigger the data set, the better predictions we would get from these clusters.