

# The Roads of California – Project Plan

## Group Members

Basil Mashal	015223500	basil.mashal@helsinki.fi
Hanna Holtdirk	015222789	hanna.holtdirk@helsinki.fi
Jimi Hytönen	014720402	jimi.hytonen@helsinki.fi

## Description

We are going to analyze the road network of California, USA, and its relation to the different points of interest (hospitals, airports, ...).

To do this, we are going to deal with the following questions and preliminary analysis tasks:

1. Question: What is the general structure of the road network?  
Analysis tasks: visualization of the data (roads + points of interest), connectivity of the graph, degree distribution, distribution of path length, clustering coefficient, centrality
2. Question: Can we find out where the (big) cities are from the road network and the points of interest?  
Analysis tasks: community detection, points of interest in/near different communities
3. Question: Can we learn to make predictions about the placement of roads and points of interest?  
Analysis tasks : machine learning for link (road) prediction, classify points of interest regarding category, for a given subgraph (city) learn to predict what buildings (e.g. hospital) there have to be and where they would be placed

## Timeline

Week 5 - Project plan, set up project, visualization of the data

Week 6 - Structure of the network, communities (question 1 and 2)

Week 7 - Machine learning/predictions (question 3)

Week 8 - Presentation, project report