

## CS F364 Design & Analysis of Algorithms

### Lecture 11b

Topic: Relation between Spanning Tree problem and Data Clustering problem

#### Outline:

- Data Clustering:
  - General Definition of Data Clustering
  - Difficulty of Solving the Data Clustering Problem
  - Approximating the solution to the Data Clustering Problem
  - Hierarchical Agglomerative Clustering (HAC)
  
- HAC Algorithms and their relation to Minimum Spanning Trees
  - Graphs vs. Datasets
  - Clusters of vertices vs. Clusters of data points
  - Spanning trees vs. Clusters
  - HAC Algorithm vs. Computing Minimum Spanning Trees