Lec14 Intro to datasets

- *Zachary Karate Club* : 34 nodes and 78 edges

Lect15 Ingredients Network

(butter, sugar, flour, vannila)->**(oil, water)**->(Tomato, Onion, Pepper, garlic)

Lec16 Synonmy Network

Love->Friend->....... ; Enemy->Hate->....... Love X Hate. But there is a path from Love to Hate _(REASON : Degradation of Synonimity)

Lec17 Web Graph

- Webpage linking -> GOOGLE unleashed it! RAnking of search

Lec18 Social Network Datasets

- Who could be his next friend? - Facebook

Lec19 Real Work Network Datasets

Eg. of networks

1. Friendship (undirected)
2. Road(undirected)
3. Citation(directed)
4. Email(directed)
5. Collaboration/ coauthorship (undirected)

Data Formats

- CSV
- GML
- Pajek Net
- GraphML
- GEXF (By Gephi)

   csv -> .csv/.txt _
   -> edge list, adj list, weigted edge list

   gml_
   -> Graph Modelling Language

Without attributes

```
graph [

    node [
        id a
        label 'Node A'
    ]
    node [
        id b
        label 'Node b'
    ]
    edge [
        source a
        target b
        label 'E1'
    ]

]
```

With attributes

```
graph [

    heirarchic 1
    directed 1
    node [
        id a
        label 'Node A'
        graphics[
            x 0
            y 0
        ]
    ]
    node [
        id b
        label 'Node b'
        graphics[
            x 20
            y 30
        ]
    ]
    edge [
        source a
        target b
        label 'E1'
        LabelGraphics[
        text 'Happy New Year'
        ]
    ]

]
```

Pajeck Net

- uses .net/.paj extension

Pajek Net Format : Uses .NET extension

```
*Vertices 82670       *arcs
1 "entity"            4244 107
2 "thing"             4244 238
3 "anything"          4244 4292
4 "something"         4247 107
5 "nothing"           4248 1
6 "whole"             4248 54
```


Pajek Net Format

```
*Vertices 9
*Edges
1 2
1 9
2 9
2 3
2 8
3 8
3 4
4 5
4 7
5 7
5 6
6 4
```

Pajek Net Format with attributes

```
*arcs
4244 107 5
```

Real-World Network Data Sets   13 / 16

NPTEL

GraphML

GraphML Format: Uses .graphml extension

```
<?xml version="1.0" encoding="UTF-8"?>
<graphml xmlns="http://graphml.graphdrawing.org/xmlns"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://graphml.graphdrawing.org/xmlns
http://graphml.graphdrawing.org/xmlns/1.0/graphml.xsd">
<graph id="G" edgedefault="undirected">
<node id="n0"/>
<node id="n1"/>
<edge id="e1" source="n0" target="n1"/>
</graph>
</graphml>
```

Real-World Network Data Sets   14 / 16

NPTEL

GEXF

Lec20 How to download datasets

- Google Network datasets

(snap.stanford.edu/data, uci network data, konect, sonet for grahML datasets)