**Topic: Network Analytics**

**Instructions:**

**1. Business Problem**

* 1. **Objective**
  2. **Constraints (if any)**

**Using R and Python codes perform:**

**2. Data Pre-processing**

**2.1 Data cleaning, Feature Engineering etc.**

**3. Exploratory Data Analysis (EDA):**

**3.1. Summary**

**3.2. Univariate analysis**

**3.3. Bivariate analysis**

**4. Model Building**

**4.1 Build the Network model on the given data sets.**

**4.2 Perform the Network Analytics.**

**4.3 Model(s) Improvement (Try with different no. of clusters)**

**5. Result Share the benefits/impact of the solution - how or in what way the business (client) gets benefit from the solution provided.**

**6. problem statement information and data dictionary is given in next page.**

**Note:**

**The assignment should be submitted in the following format:**

* **R code**
* **Python code**
* **Code Modularization should be maintained**
* **Documentation of the modules (elaborating on steps mentioned above).**

**1) Use the flights\_hault and connecting routes datasets to find betweenness, closeness centrality, page rank, directed and undirected graphs.**

**Dataset: Flights\_hault**

**column names:**

**ID","Name","City","Country","IATA\_FAA","ICAO","Latitude","Longitude","Altitude","Time","DST","Tz database time"**

**Dataset: Connecting routes**

**column names:**

**"flights", " ID", "main Airport","main Airport ID", "Destination ", "Destination ID", "haults", "machinary"**

**2) Business Problem: Three data sets of social media networks are attached.**

**- plot circular graph**

**- Star Graph for the three data sets**

**- Check for its nodes and directed and undirected graphs**

**Draw your insights on the data.**