Topic: Network Analytics

Instructions:

Please share your answers filled in-line in the word document. Submit code separately wherever applicable.

Please ensure you update all the details:

Name: _Nukala AyyappaBharthwaj_ Batch ID: _DSWDMCOS 21012022_

Topic: Network Analytics

Grading Guidelines:

1. An assignment submission is considered complete only when correct and executable code(s) are submitted along with the documentation explaining the method and results. Failing to submit either of those will be considered an invalid submission and will not be considered for evaluation.

2. Assignments submitted after the deadline will affect your grades.

Grading:

Ans	Date			Ans	Date
Correct	On time	A	100		
80% & above	On time	B	85	Correct	Late
50% & above	On time	C	75	80% & above	Late
50% & below	On time	D	65	50% & above	Late
		E	55	50% & below	
Copied/No Submission		F	45		

- **Grade A: (>= 90):** When all assignments are submitted on or before the given deadline.
- **Grade B: (>= 80 and < 90):**
  - When assignments are submitted on time but less than 80% of problems are completed.

    (OR)
  - All assignments are submitted after the deadline.

- **Grade C: (>= 70 and < 80):**
  - When assignments are submitted on time but less than 50% of the problems are completed.

    (OR)
  - Less than 80% of problems in the assignments are submitted after the deadline.

- **Grade D: (>= 60 and < 70):**
  - Assignments submitted after the deadline and with 50% or less problems.

- **Grade E: (>= 50 and < 60):**
  - Less than 30% of problems in the assignments are submitted after the deadline.

    (OR)
  - Less than 30% of problems in the assignments are submitted before the deadline.

- **Grade F: (< 50):** No submission (or) malpractice.

Hints:

1. Business Problem

1.1. What is the business objective?

1.1. Are there any constraints?

2. Work on each feature of the dataset to create a data dictionary as displayed in the below image:

Name of Feature	Description	Type	Relevance
ID	Customer ID	Quantitative, Nominal	Irrelevant, ID does not provide useful information

2.1 Make a table as shown above and provide information about the features such as its data type and its relevance to the model building. And if not relevant, provide reasons and a description of the feature.

3. Data Cleaning

4. Model Building

4.1. Perform network analytics on the given datasets.

4.2. Briefly explain the model output in the documentation.

5. Write about the benefits/impact of the solution - in what way does the business (client) benefit from the solution provided?

Problem Statement: -

There are two datasets consisting of information for the connecting routes and flight halt. Create network analytics models on both the datasets separately and measure degree centrality, degree of closeness centrality, and degree of in-between centrality.

- Create a network using edge list matrix (directed only**).**
- Columns to be used ***in R*:**

Flight_halt=c("ID","Name","City","Country","IATA_FAA","ICAO","Latitude","Longitude","Altitude","Time","DST","Tz database time")

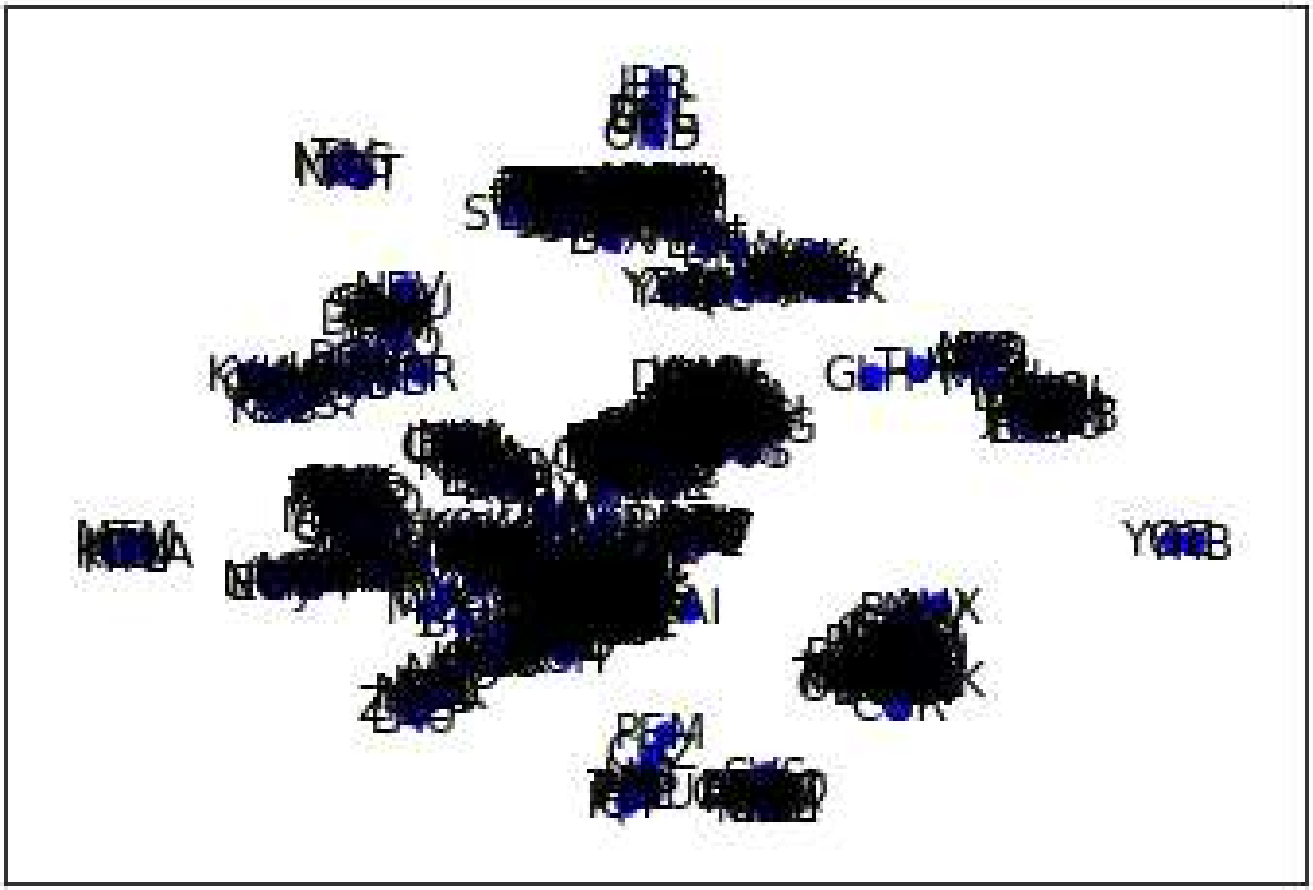
connecting routes=c("flights", " ID", "main Airport", "main Airport ID", "Destination ","Destination ID","haults","machinary")

connecting routes

	A	B	C	D	E	F	G	H	I
1	2B	410	AER	2965	KZN	2990		0	CR2
2	2B	410	ASF	2966	KZN	2990		0	CR2
3	2B	410	ASF	2966	MRV	2962		0	CR2
4	2B	410	CEK	2968	KZN	2990		0	CR2
5	2B	410	CEK	2968	OVB	4078		0	CR2
6	2B	410	DME	4029	KZN	2990		0	CR2
7	2B	410	DME	4029	NBC	6969		0	CR2
8	2B	410	DME	4029	TGK	\N		0	CR2
9	2B	410	DME	4029	UUA	6160		0	CR2
10	2B	410	EGO	6156	KGD	2952		0	CR2
11	2B	410	EGO	6156	KZN	2990		0	CR2
12	2B	410	GYD	2922	NBC	6969		0	CR2
13	2B	410	KGD	2952	EGO	6156		0	CR2
14	2B	410	KZN	2990	AER	2965		0	CR2
15	2B	410	KZN	2990	ASF	2966		0	CR2
16	2B	410	KZN	2990	CEK	2968		0	CR2
17	2B	410	KZN	2990	DME	4029		0	CR2
18	2B	410	KZN	2990	EGO	6156		0	CR2
19	2B	410	KZN	2990	LED	2948		0	CR2
20	2B	410	KZN	2990	SVX	2975		0	CR2
21	2B	410	LED	2948	KZN	2990		0	CR2
22	2B	410	LED	2948	NBC	6969		0	CR2
23	2B	410	LED	2948	UUA	6160		0	CR2
24	2B	410	MRV	2962	ASF	2966		0	CR2
25	2B	410	NBC	6969	DME	4029		0	CR2
26	2B	410	NBC	6969	GYD	2922		0	CR2
27	2B	410	NBC	6969	LED	2948		0	CR2
28	2B	410	NBC	6969	SVX	2975		0	CR2
29	2B	410	NJC	2972	SVX	2975		0	CR2

connecting routes

Graph :

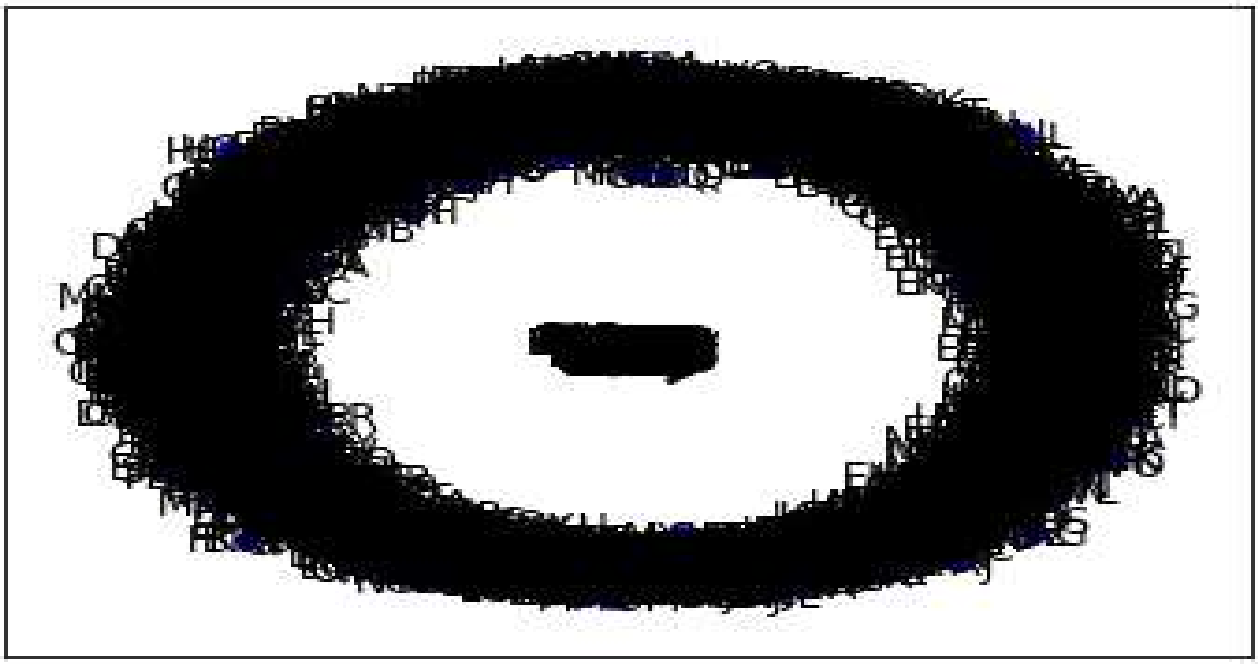


Average Clustering value is 0.2539

Flight_hault1

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	1	Goroka	Goroka	Papua Nev	GKA	AYGA	-6.08169	145.3919	5282	10	U	Pacific/Port_Moresby		
2	2	Madang	Madang	Papua Nev	MAG	AYMD	-5.20708	145.7887	20	10	U	Pacific/Port_Moresby		
3	3	Mount Ha	Mount Ha	Papua Nev	HGU	AYMH	-5.82679	144.2959	5388	10	U	Pacific/Port_Moresby		
4	4	Nadzab	Nadzab	Papua Nev	LAE	AYNZ	-6.56983	146.7262	239	10	U	Pacific/Port_Moresby		
5	5	Port More	Port More	Papua Nev	POM	AYPY	-9.44338	147.2201	146	10	U	Pacific/Port_Moresby		
6	6	Wewak Int	Wewak	Papua Nev	WWK	AYWK	-3.58383	143.6692	19	10	U	Pacific/Port_Moresby		
7	7	Narsarsua	Narssarss	Greenland	UAK	BGBW	61.16052	-45.426	112	-3	E	America/Godthab		
8	8	Nuuk	Godthaab	Greenland	GOH	BGGH	64.19092	-51.6781	283	-3	E	America/Godthab		
9	9	Sondre Str	Sondrestro	Greenland	SFJ	BGSF	67.01697	-50.6893	165	-3	E	America/Godthab		
10	10	Thule Air B	Thule	Greenland	THU	BGTL	76.5312	-68.7032	251	-4	E	America/Thule		
11	11	Akureyri	Akureyri	Iceland	AEY	BIAR	65.65999	-18.0727	6	0	N	Atlantic/Reykjavik		
12	12	Egilsstadir	Egilsstadir	Iceland	EGS	BIEG	65.28333	-14.4014	76	0	N	Atlantic/Reykjavik		
13	13	Hornafjorc	Hofn	Iceland	HFN	BIHN	64.29556	-15.2272	24	0	N	Atlantic/Reykjavik		
14	14	Husavik	Husavik	Iceland	HZK	BIHU	65.95233	-17.426	48	0	N	Atlantic/Reykjavik		
15	15	Isafjordur	Isafjordur	Iceland	IFJ	BIIS	66.05806	-23.1353	8	0	N	Atlantic/Reykjavik		
16	16	Keflavik In	Keflavik	Iceland	KEF	BIKF	63.985	-22.6056	171	0	N	Atlantic/Reykjavik		
17	17	Patreksfjo	Patreksfjo	Iceland	PFJ	BIPA	65.55583	-23.965	11	0	N	Atlantic/Reykjavik		
18	18	Reykjavik	Reykjavik	Iceland	RKV	BIRK	64.13	-21.9406	48	0	N	Atlantic/Reykjavik		
19	19	Siglufjordu	Siglufjordu	Iceland	SIJ	BISI	66.13333	-18.9167	10	0	N	Atlantic/Reykjavik		
20	20	Vestmanna	Vestmanna	Iceland	VEY	BIVM	63.4243	-20.2789	326	0	N	Atlantic/Reykjavik		
21	21	Sault Ste M	Sault Saint	Canada	YAM	CYAM	46.485	-84.5094	630	-5	A	America/Toronto		
22	22	Winnipeg S	Winnipeg	Canada	YAV	CYAV	50.05639	-97.0325	760	-6	A	America/Winnipeg		
23	23	Shearwate	Halifax	Canada	YAW	CYAW	44.63972	-63.4994	167	-4	A	America/Halifax		
24	24	St Anthony	St. Anthon	Canada	YAY	CYAY	51.39194	-56.0831	108	-3.5	A	America/St_Johns		
25	25	Tofino	Tofino	Canada	YAZ	CYAZ	49.08222	-125.773	80	-8	A	America/Vancouver		
26	26	Kugaaruk	Pelly Bay	Canada	YBB	CYBB	68.53444	-89.8081	56	-7	A	America/Edmonton		
27	27	Baie Come	Baie Come	Canada	YBC	CYBC	49.1325	-68.2044	71	-5	A	America/Toronto		
28	28	Bagotville	Bagotville	Canada	YBG	CYBG	48.33056	-70.9964	522	-5	A	America/Toronto		
29	29	Baker Lake	Baker Lake	Canada	YBK	CYBK	64.29889	-96.0778	59	-6	A	America/Winnipeg		

flight_hault1



Problem statement

There are three datasets given (Facebook, Instagram, and LinkedIn). Construct and visualize the following networks:

- circular network for Facebook
- star network for Instagram
- star network for LinkedIn

Create a network using an adjacency matrix (undirected only). The snapshots of those datasets are given below:

Facebook

	A	B	C	D	E	F	G	H	I
1	1	2	3	4	5	6	7	8	9
2	0	1	0	0	0	0	0	0	1
3	1	0	1	0	0	0	0	0	0
4	0	1	0	1	0	0	0	0	0
5	0	0	1	0	1	0	0	0	0
6	0	0	0	1	0	1	0	0	0
7	0	0	0	0	1	0	1	0	0
8	0	0	0	0	0	1	0	1	0
9	0	0	0	0	0	0	1	0	1
10	1	0	0	0	0	0	0	1	0
11									

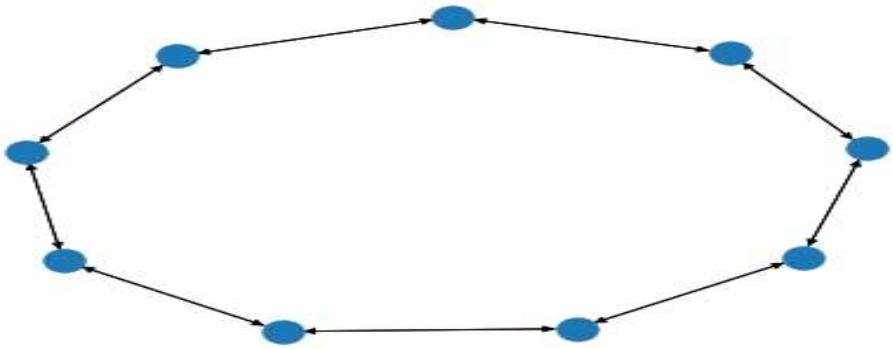
Instagram

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	1	2	3	4	5	6	7	8	9	10	11	12	13	
2	0	1	1	0	0	0	0	0	0	0	0	0	0	
3	1	0	1	1	0	0	0	0	0	0	0	0	0	
4	1	1	0	1	0	0	0	0	0	0	0	0	0	
5	0	1	1	0	0	0	0	0	0	0	0	0	1	
6	0	0	0	0	0	1	1	0	0	0	0	0	0	
7	0	0	0	0	1	0	1	1	0	0	0	0	0	
8	0	0	0	0	1	1	0	1	0	0	0	0	0	
9	0	0	0	0	0	1	1	0	0	0	0	0	1	
10	0	0	0	0	0	0	0	0	0	1	1	0	0	
11	0	0	0	0	0	0	0	0	1	0	1	1	0	
12	0	0	0	0	0	0	0	0	1	1	0	1	0	
13	0	0	0	0	0	0	0	0	0	1	1	0	1	
14	0	0	0	1	0	0	0	1	0	0	0	1	0	
15														
16														
17														
18														

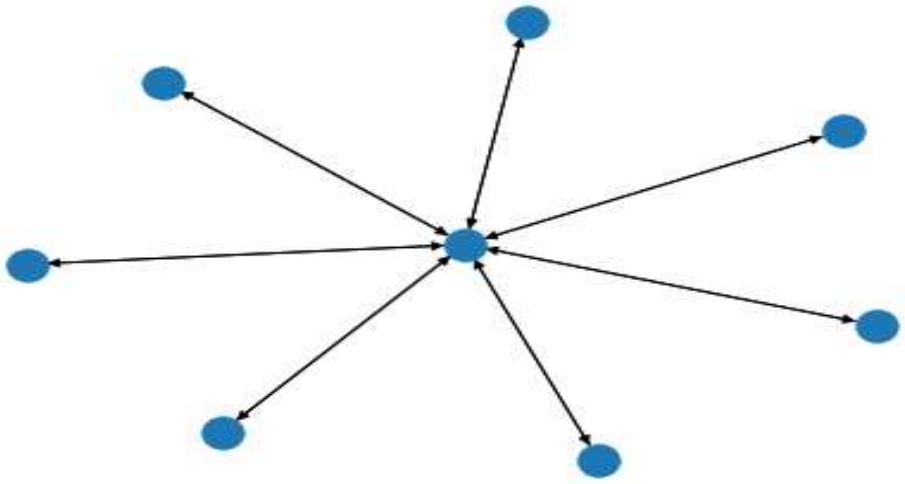
LinkedIn

	A	B	C	D	E	F	G	H	I	J
1	1	2	3	4	5	6	7	8		
2	0	1	1	1	1	1	1	1		
3	1	0	0	0	0	0	0	0		
4	1	0	0	0	0	0	0	0		
5	1	0	0	0	0	0	0	0		
6	1	0	0	0	0	0	0	0		
7	1	0	0	0	0	0	0	0		
8	1	0	0	0	0	0	0	0		
9	1	0	0	0	0	0	0	0		
10										
11										
12										
13										

Facebook Plot :



Instagram Plot:



Twitter Plot: