```
import pandas as pd

df = pd.read_csv('/route_detail.csv')

df.head()
```

	Unnamed: 0	route_id	stop_id	stop_name
0	0	1	1	THIRUVOTRIYUR
1	1	1	2	THIRUVOTRIYUR TEMPLE
2	2	1	3	THANGAL
3	3	1	4	ANNA NAGAR
4	4	1	5	ROYAPURAM P.S

Data cleaning

```
# removing unwanted columns
df = df.iloc[:,1:]
df.head()
```

	route_id	stop_id	stop_name
0	1	1	THIRUVOTRIYUR
1	1	2	THIRUVOTRIYUR TEMPLE
2	1	3	THANGAL
3	1	4	ANNA NAGAR
4	1	5	ROYAPURAM P.S

Top 10 Stop which have high bus traffic

```
stat = df.groupby(by=["stop_name"])["route_id"].count().sort_values(ascending=False)
stat.iloc[:10]
```

	route_id
stop_name	
BROADWAY	144
M.G.R.CENTRAL	117
CONCORDE	110
SAIDAPET	102
TAMBARAM	86
THIRUMANGALAM	81
VANDALUR ZOO	79
PERUNGALATHUR	74
VANDALUR GATE	74
M.G.R.KOYAMBEDU	73

dtype: int64

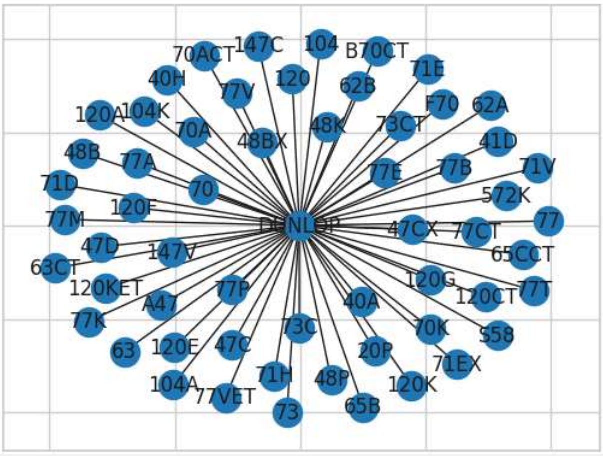
Visualize data

- Found network graph will best suit for the usecase

```
import matplotlib.pyplot as plt
import networkx as nx
```

Visualizing Buses passing through a stop

```
sample = df[df["stop_name"]=="DUNLOP"]
G = nx.from_pandas_edgelist(sample, "route_id", "stop_name")
nx.draw_networkx(G)
plt.show()
```



Data manipulation to calculate next stop

```
df.head()
df["next_stop"] = df.groupby(by=["route_id"])["stop_name"].shift(-1)
df.loc[:50]
```

	route_id	stop_id	stop_name	next_stop
0	1	1	THIRUVOTRIYUR	THIRUVOTRIYUR TEMPLE
1	1	2	THIRUVOTRIYUR TEMPLE	THANGAL
2	1	3	THANGAL	ANNA NAGAR
3	1	4	ANNA NAGAR	ROYAPURAM P.S
4	1	5	ROYAPURAM P.S	CLIVE BATTERY
5	1	6	CLIVE BATTERY	PARRYS
6	1	7	PARRYS	M.G.R.CENTRAL
7	1	8	M.G.R.CENTRAL	P.OR & SONS
8	1	9	P.OR & SONS	WESLEY H.S
9	1	10	WESLEY H.S	Y.M.I.A
10	1	11	Y.M.I.A	MANDAVELI
11	1	12	MANDAVELI	A.M.S.HOSPITAL
12	1	13	A.M.S.HOSPITAL	ADYAR O.T.
13	1	14	ADYAR O.T.	ADYAR DEPOT
14	1	15	ADYAR DEPOT	THIRUVANMIYUR
15	1	16	THIRUVANMIYUR	NaN
16	101	1	THIRUVOTRIYUR	THRUVOTRIYUR TEMPLE
17	101	2	THRUVOTRIYUR TEMPLE	THANGAL
18	101	3	THANGAL	ANNA NAGAR
19	101	4	ANNA NAGAR	ROYAPURAM P.S
20	101	5	ROYAPURAM P.S	CLIVE BATTERY
21	101	6	CLIVE BATTERY	PARRYS
22	101	7	PARRYS	M.G.R.CENTRAL
23	101	8	M.G.R.CENTRAL	DASAPRAKASH
24	101	9	DASAPRAKASH	TAYLORS ROAD
25	101	10	TAYLORS ROAD	AMINJIKARAI
26	101	11	AMINJIKARAI	NADUVANKARAI
27	101	12	NADUVANKARAI	ARUMBAKKAM
28	101	13	ARUMBAKKAM	NERKUNDRAM
29	101	14	NERKUNDRAM	MADURAVOYAL
30	101	15	MADURAVOYAL	VAANAGARAM
31	101	16	VAANAGARAM	VELAPPANCHAVADI
32	101	17	VELAPPANCHAVADI	KUMUNANCHAVADI
33	101	18	KUMUNANCHAVADI	POONAMALLEE
34	101	19	POONAMALLEE	NaN
35	102	1	BROADWAY	SECRETARIAT
36	102	2	SECRETARIAT	CHEPAUK
37	102	3	CHEPAUK	Q.M.C
38	102	4	Q.M.C	FORESHORE ESTATE
39	102	5	FORESHORE ESTATE	A.M.S.HOSPITAL
40	102	6	A.M.S.HOSPITAL	ADYAR O.T.
41	102	7	ADYAR O.T.	INDIRA NAGAR WATER TANK
42	102	8	INDIRA NAGAR WATER TANK	SRP TOOLS
43	102	9	SRP TOOLS	KANDANCHAVADI
44	102	10	KANDANCHAVADI	THORAPPAKKAM TEA SHOP
45	102	11	THORAPPAKKAM TEA SHOP	MOOTAKARAN CHAVADI
46	102	12	MOOTAKARAN CHAVADI	KARAPAKKAM
47	102	13	KARAPAKKAM	SHOZHANGANALLUR P.U.OFFICE
48	102	14	SHOZHANGANALLUR P.U.OFFICE	KUMARAN NAGAR
49	102	15	KUMARAN NAGAR	SEMMANCHERI
50	102	16	SEMMANCHERI	NAVALUR

- Assigning next stop for last stop as route_id (Bus number) for easy visualization later

```
df.loc[df["next_stop"].isna(), "next_stop"] = df.loc[df["next_stop"].isna(), "route_id"]
df.loc[:50]
```

	route_id	stop_id	stop_name	next_stop
0	1	1	THIRUVOTRIYUR	THIRUVOTRIYUR TEMPLE
1	1	2	THIRUVOTRIYUR TEMPLE	THANGAL
2	1	3	THANGAL	ANNA NAGAR
3	1	4	ANNA NAGAR	ROYAPURAM P.S
4	1	5	ROYAPURAM P.S	CLIVE BATTERY
5	1	6	CLIVE BATTERY	PARRYS
6	1	7	PARRYS	M.G.R.CENTRAL
7	1	8	M.G.R.CENTRAL	P.OR & SONS
8	1	9	P.OR & SONS	WESLEY H.S
9	1	10	WESLEY H.S	Y.M.I.A
10	1	11	Y.M.I.A	MANDAVELI
11	1	12	MANDAVELI	A.M.S.HOSPITAL
12	1	13	A.M.S.HOSPITAL	ADYAR O.T.
13	1	14	ADYAR O.T.	ADYAR DEPOT
14	1	15	ADYAR DEPOT	THIRUVANMIYUR
15	1	16	THIRUVANMIYUR	1
16	101	1	THIRUVOTRIYUR	THRUVOTRIYUR TEMPLE
17	101	2	THRUVOTRIYUR TEMPLE	THANGAL
18	101	3	THANGAL	ANNA NAGAR
19	101	4	ANNA NAGAR	ROYAPURAM P.S
20	101	5	ROYAPURAM P.S	CLIVE BATTERY
21	101	6	CLIVE BATTERY	PARRYS
22	101	7	PARRYS	M.G.R.CENTRAL
23	101	8	M.G.R.CENTRAL	DASAPRAKASH
24	101	9	DASAPRAKASH	TAYLORS ROAD
25	101	10	TAYLORS ROAD	AMINJIKARAI
26	101	11	AMINJIKARAI	NADUVANKARAI
27	101	12	NADUVANKARAI	ARUMBAKKAM
28	101	13	ARUMBAKKAM	NERKUNDRAM
29	101	14	NERKUNDRAM	MADURAVOYAL
30	101	15	MADURAVOYAL	VAANAGARAM
31	101	16	VAANAGARAM	VELAPPANCHAVADI
32	101	17	VELAPPANCHAVADI	KUMUNANCHAVADI
33	101	18	KUMUNANCHAVADI	POONAMALLEE
34	101	19	POONAMALLEE	101
35	102	1	BROADWAY	SECRETARIAT
36	102	2	SECRETARIAT	CHEPAUK
37	102	3	CHEPAUK	Q.M.C
38	102	4	Q.M.C	FORESHORE ESTATE
39	102	5	FORESHORE ESTATE	A.M.S.HOSPITAL
40	102	6	A.M.S.HOSPITAL	ADYAR O.T.
41	102	7	ADYAR O.T.	INDIRA NAGAR WATER TANK
42	102	8	INDIRA NAGAR WATER TANK	SRP TOOLS
43	102	9	SRP TOOLS	KANDANCHAVADI
44	102	10	KANDANCHAVADI	THORAPPAKKAM TEA SHOP
45	102	11	THORAPPAKKAM TEA SHOP	MOOTAKARAN CHAVADI
46	102	12	MOOTAKARAN CHAVADI	KARAPAKKAM
47	102	13	KARAPAKKAM	SHOZHANGANALLUR P.U.OFFICE
48	102	14	SHOZHANGANALLUR P.U.OFFICE	KUMARAN NAGAR
49	102	15	KUMARAN NAGAR	SEMMANCHERI
50	102	16	SEMMANCHERI	NAVALUR

Sample data creation - Dunlop stop

```
sample_routes = df[df["stop_name"]=="DUNLOP"]["route_id"]
sample = df[df["route_id"].isin(sample_routes)]
sample.head()
```

	route_id	stop_id	stop_name	next_stop
204	104	1	REDHILLS	AYURVEDHA ASHARAMAM
205	104	2	AYURVEDHA ASHARAMAM	KAVANGARAI
206	104	3	KAVANGARAI	SCREW FACTORY
207	104	4	SCREW FACTORY	SURAPEDU
208	104	5	SURAPEDU	KALLIKUPPAM

- **Except for base stop all other stop should be unique to route, So bus route can be visualized clearly without intersection between stops**

```
# sample["derived_next_stop"] = sample.apply(lambda row: row["next_stop"] if(row["next_stop"]=="DUNLOP") else row["route_id"]+row["next_stop"], axis=1)
# sample = sample.drop("derived_next_stop",axis=1)
sample.loc[~(sample["next_stop"]=="DUNLOP") & ~(sample["next_stop"]==sample["route_id"]),"derived_next_stop"] =  sample["route_id"]+ "-" +sample["next_
sample.loc[(sample["next_stop"]=="DUNLOP") | (sample["next_stop"]==sample["route_id"]),"derived_next_stop"] = sample["next_stop"]
sample.loc[~(sample["stop_name"]=="DUNLOP"),"derived_stop_name"] = sample["route_id"]+ "-" + sample["stop_name"]
sample.loc[sample["stop_name"]=="DUNLOP","derived_stop_name"] = sample["stop_name"]
sample.head()
```

```
<ipython-input-43-5519beb56ca2>:3: SettingWithCopyWarning: 


A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

<ipython-input-43-5519beb56ca2>:5: SettingWithCopyWarning: 


A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
```

	route_id	stop_id	stop_name	next_stop	derived_next_stop	derived_stop_name
204	104	1	REDHILLS	AYURVEDHA ASHARAMAM	104-AYURVEDHA ASHARAMAM	104-REDHILLS
205	104	2	AYURVEDHA ASHARAMAM	KAVANGARAI	104-KAVANGARAI	104-AYURVEDHA ASHARAMAM
206	104	3	KAVANGARAI	SCREW FACTORY	104-SCREW FACTORY	104-KAVANGARAI
207	104	4	SCREW FACTORY	SURAPEDU	104-SURAPEDU	104-SCREW FACTORY
208	104	5	SURAPEDU	KALLIKUPPAM	104-KALLIKUPPAM	104-SURAPEDU

Creating Graph

```
G = nx.from_pandas_edgelist(sample, "derived_stop_name", "derived_next_stop", create_using=nx.DiGraph())
print(nx.is_directed(G))
```

```
True
```

coloring based on bus number

```
!pip install distinctipy
```

```
Requirement already satisfied: distinctipy in /usr/local/lib/python3.10/dist-packages (1.3.4)
Requirement already satisfied: numpy>=1.16.3 in /usr/local/lib/python3.10/dist-packages (from distinctipy) (1.26.4)
```

```
from distinctipy import distinctipy

routes = sample["route_id"].unique()
n=len(list(routes))
# print(list(route))

reserved_color = [(1,0,0),(0,1,0),(0,0,1),(0.8,0.6,0.3)]

# generate N visually distinct colours
colors = distinctipy.get_colors(n,reserved_color)
# print(colors)
route_color_map = { route: color for route, color in zip(routes,colors)}
# print(route_color_map)
```

color edges based on route

```
edge_color = []
for u,v in G.edges():
    route_id = None
    u_list = u.split("-")
    v_list = v.split("-")
    if len(u_list) > 1:
        route_id = u_list[0]
    elif len(v_list) > 1:
        route_id = v_list[0]

    edge_color.append(route_color_map[route_id])
# print(edge_color)
```

Color node based on stops

```
stops = sample["stop_name"].unique()
n=len(list(stops))

print(n)

reserved_color = [(1,0,0),(0,1,0),(0,0,1)] + edge_color

# generate N visually distinct colours
colors = distinctipy.get_colors(n,reserved_color)
# print(colors)
stop_color_map = { stop: color for stop, color in zip(stops,colors)}
# print(stop_color_map)
```

```
200
```

```
node_color =[]
root_node = None
labels = {}
for node in G:
    if node == "DUNLOP":
        # current stop
        root_node = node
        node_color.append((0,1,0)) #green
        labels[node] = node
    elif node in list(sample["route_id"]):
        node_color.append((1,0,0)) #red
        labels[node] = node
    else:
        labels[node] = sample.loc[sample["derived_stop_name"]== node,"stop_id"].values[0]
        node_color.append(stop_color_map[node.split("-")[1]]) #route wise color
# print(root_node)
# print(labels)
```

- use graphviz radical layout
- add these details in legend
- increse edge width

Creating Legend for the graph

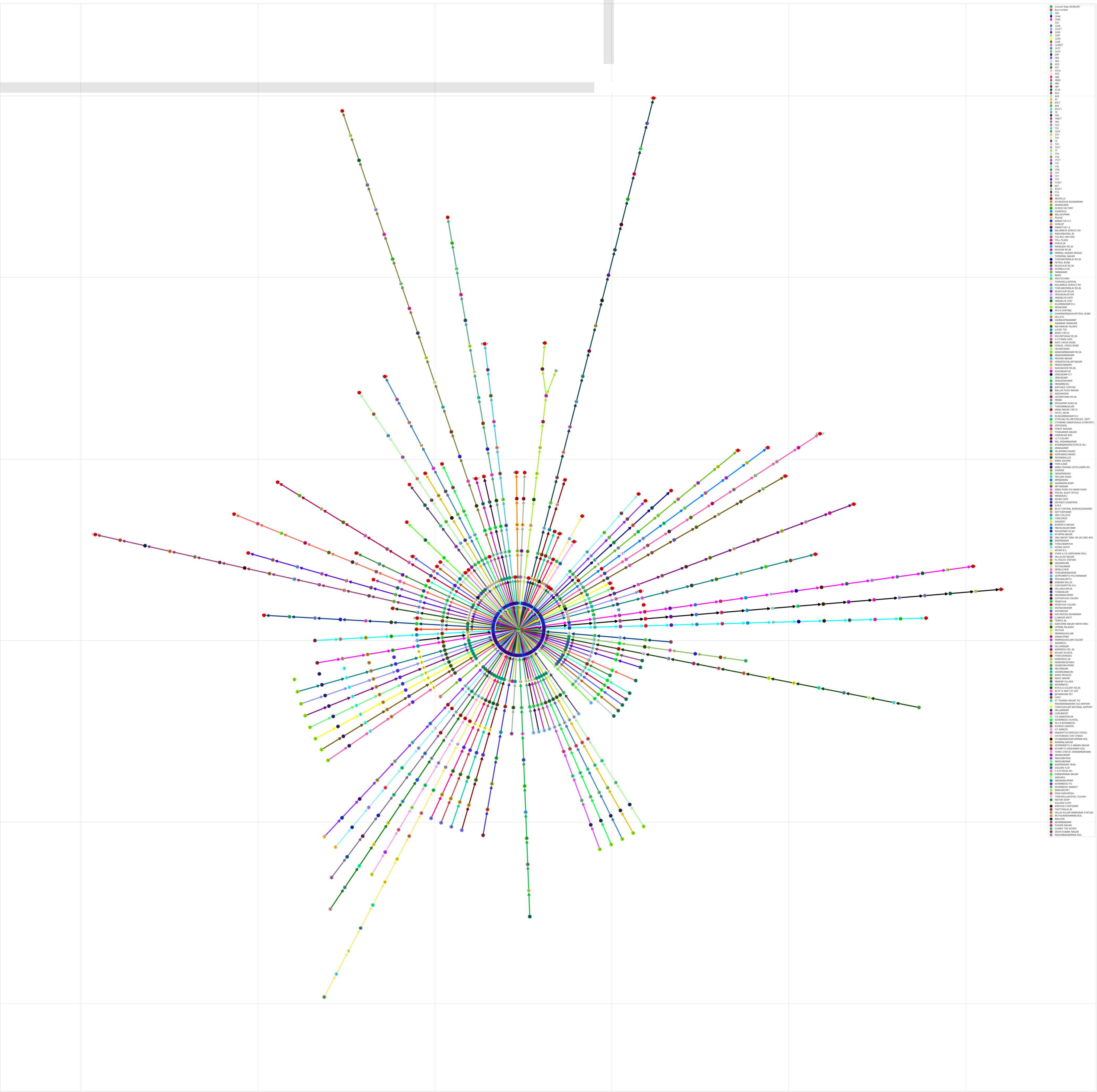
```
from matplotlib.lines import Line2D

legend_elements = [
    Line2D([0], [0], marker='o', color='w', label='Current Stop (DUNLOP)',markerfacecolor='g', markersize=15),
    Line2D([0], [0], marker='o', color='w', label='Bus number',markerfacecolor='r', markersize=15),
]
route_legend = [
    Line2D([0], [0], marker='d', color='w', label= bus_num,markerfacecolor=route_color_map[bus_num], markersize=15) for bus_num in route_color_map
]
stop_legend = [
    Line2D([0], [0], marker='o', color='w', label= stop,markerfacecolor=stop_color_map[stop], markersize=15) for stop in stop_color_map
]
legend_elements = legend_elements + route_legend + stop_legend
```

Drawing Graph

```
plt.figure(figsize=(100,100))
pos=nx.nx_pydot.graphviz_layout(G,prog="twopi",root=root_node)
nx.draw_networkx(G, pos=pos,edge_color=edge_color, node_color=node_color, arrows=True
    arrowsize= 25, width=5, labels=labels)
```

```
plt.legend(handles=legend_elements, loc='upper right')
```


Current Stop (DUNLOP)
Bus number
104A
104K
120A
120CT
120E
120F
120G
120K
120KET
147C
147V
20P
40A
40H
41D
47C
47CX
47D
48B
48BX
48K
48P
572K
62A
62B
63CT
65B
65CCT
70A
70ACT
70K
71D
71E
71EX
71H
71V
73C
73CT
77A
77B
77CT
77E
77K
77M
77P
77T
77V
77VET
A47
B70CT
F70
S58
REDHILLS
AYURVEDHA ASHARAMAM
KAVANGARAI
SCREW FACTORY
SURAPEDU
KALLIKUPPAM
PUDUR
AMBATTUR O.T.
DUNLOP
AMBATTUR I.E.
NALAMBUR SERVICE RD
MADURAVOYAL JN
TVS MCC MOTORS
TOLL PLAZA
PORUR JN
MANGADU RD JN
KOVOOR RD JN
PAMMAL ADAYAR BRIDGE
THENDRAL NAGAR
THIRUNEERMALAI RD JN.
PETROL BUNK
MUDICHUR RD JN.
IRUMBULIYUR
TAMBARAM
AVADI
POLYTECHNIC
THIRUMULLAIVOYAL
NOLAMBUR SERVICE RD
THIRUNEERMALAI RD.JN.
MUDICHUR RD.JN.
PERUNGALATHUR
VANDALUR GATE
VANDALUR ZOO
KILAMBAKKAM B.S.
BROADWAY
M.G.R.CENTRAL
DHARMAPRAKASH/PETROL BUNK
KELLEYS
SAYANI/AYNAVARAM
KAMABAR ARANGAM
NATHAMUNI TALKIES
LUCAS TVS
AVADI CIRCLE
KOILPATHAGAI RD.JN.
H.V.F.MAIN GATE
AAFS CROSS ROAD
VENGAL CROSS ROAD
VEERAPURAM
ARAKKAMBAKKAM RD.JN.
ARAKKAMBAKKAM
PERIYAR NAGAR
VENKATACHALAM NAGAR
PANDESWARAM
KADHAVOOR RD.JN.
KILKONDAIYUR
ORAGADAM O.T.
ORAGADAM
VENGATAPURAM
MENAMBEDU
AIRFORCE STATION
NALLUR PUDU NAGAR
KADHAVOOR
VEERAPURAM RD JN.
MORAI
MOGAIPAIR ROAD JN.
THIRUMANGALAM
ANNA NAGAR CIRCLE
HOTEL ARUN
NUNGAMBAKKAM R.S.
STERLING RD./METROLOG. DEPT.
UTHAMAR GANDHISALAI (CORP.OFF)
VIDYODAYA
PONDY BAZAAR
THIAGARAYA NAGAR
VINAYAGAR KOIL
I.C.F.COLONY
MEL AYNAMBAKKAM
AYNAMBAKKAM(3STATUE JN.)
VANAGARAM
VELAPPANCHAVADI
KUMUNANCHAVADI
POONAMALLEE
ANNA SQUARE
TRIPLICANE
ANNA POORNA HOTEL/DAMS RD
EGMORE
DASAPRAKASH
TAYLORS ROAD
AMINJIKARAI
KAVARAIPALAYAM
PATTABIRAM
ANNA ROAD P.O./DAMS ROAD
POSTAL AUDIT OFFICE
MANDAVELI
ADYAR GATE
DEFENCE QUARTERS
D.M.S
JN.OF CENTRAL AVENUE&ERIKARAI
KOTTURPURAM
ENG.COLLEGE
CONCORDE
SAIDAPET
BHARATHI NAGAR
MAHALINGAPURAM
MOGAPPAIR RD.JN.
AYYAPPA NAGAR
2ND WATER TANK ON SECOND AVE.
AYAPPAKKAM
THIRUVANMIYUR
ADYAR DEPOT
ADYAR B.S.
VIVEK & CO (KRISHNAN KOIL)
VALLALAR NAGAR
P1.POLICE STATION
DASAMAHAN
PUTHAGARAM
NEMILICHERI
THIRUNINDRAVOOR
VEPPUMPATTU PULIYAMARAM
PERUMALPATTU
KANDAN KOLLAI
CHEVVAIPETTAI R.S.
VELLAKULAM JN.
THANDALAM
KALYANAKUPPAM
CHITHATHUR COLONY
PERATHUR
PERATHUR COLONY
VISHNUVAKKAM
KIZHANOOR
AAYURVEDA ASHRAMAM
J.J.NAGAR WEST
TEMPLE JN.
KARISHMA NAGAR WATER BRG
UPPARA PALAYAM
POTHUR
PAMMADUKULAM
IRANKUPPAM
PAMMADUKULAM COLONY
KARIMEDU
VILLIVAKKAM
KONIMEDU RD. JN.
KOLADI SCHOOL
THIRUVERKADU
KONEMEDU JN.
KARAYANCAHVADI
SENNEERKUPPAM
MELPAKKAM
GOVARDANAGIRI
AVADI MOSQUE
NASIC NAGAR
PAKKAM VILLAGE
KOYAMBEDU
M.M.D.A.COLONY RD.JN.
JN OF III AND 1ST AVE
JAFFARKHAN PET
CIPET
ST THOMAS MOUNT PO
MEENAMABAKKAM OLD AIRPORT
THIRUSOOLAM NATIONAL AIRPORT
PALLAVARAM
CHROMEPET
T.B.SANATORIUM
KOYAMBEDU SCHOOL
M.G.R.KOYAMBEDU
KILPAUK GARDEN
ICF ANNEXE
ANAIKATTUCHERY(DIV STAGE)
CHITHAKADU (DIV STAGE)
VILINJAMBAKKAM AMMAN KOIL
KAMARAJ NAGAR
VEPPAMPATTU E.WARAN NAGAR
ATHIPATTU KRISHANER KOIL
THREE STATUE (AYANAMBAKKAM)
VAANAGARAM
MADURAVOYAL
NERKUNDRAM
AYAPPAKKAM TNHB
GOLDEN FLAT
C.R.P.CROSS RD.
KANNIAMMAN NAGAR
KARUKKU
MADANAKUPPAM
KOYAMBEDU P.S
KOYAMBEDU MARKET
BANGARUPET
POOCHIATHIPEDU
THIRUMULLAIVOYAL COLONY
RATION SHOP
GOLDEN FLATS
AYATHUR CHATHIRAM
THOTTIKALAI JN.
VELLAI KULAM AMBEDKAR STATUJN
MUTHUMARIAMMAN KOIL
AYALOOR
KOYAMBAKKAM
PUSHPA NAGAR
GUINDY TVK ESTATE
DEVIK ESWARI NAGAR
KADUMBADIAMMAN KOIL