Assessment for CASA0002 – Urban Simulation

Network and SIM

Part I: London's underground resilience

1. Topological network

1.1. Centrality measures:

In the study of transportation networks, particularly complex systems such as the London Underground, it is essential to identify the most structurally and functionally important stations. Bloch, Jackson and Tebaldi (2023) claimed that centrality measures derived from graph theory provide a rigorous framework for quantifying the relative importance of nodes (stations) within a network. This section introduces and justifies the selection of three centrality measures: *degree centrality*, *betweenness centrality*, and *closeness centrality*, to characterize and identify crucial stations in the underground network.

i Degree Centrality

Definition and Equation Degree centrality is a fundamental measure that captures the number of direct connections a node has within the network. For an undirected graph $G = (V, E)$, where V is the set of nodes and E is the set of edges, the degree centrality $C_D(v)$ of a node $v \in V$ is defined as:

$$C_D(v) = \frac{\deg(v)}{N-1}$$

where $\deg(v)$ denotes the degree of node v, and N is the total number of nodes in the network. The measure is normalized by $N-1$ to facilitate comparison across networks of different sizes (Derrible (2012)).

Application in the Underground Context In the context of the London Underground, degree centrality reflects the number of other stations directly accessible from a given station, either through a single train line or by transfer Derrible (2012). Stations with high degree centrality, such as King's Cross St Pancras, typically serve multiple lines or act as interchange points. These stations play a key role in maintaining the local connectivity of the network.

Importance for Network Functioning According to Zhang, Pan and Lai (2021), high-degree stations are vital for passenger flow and route flexibility, enabling travelers to access multiple lines without significant detours. Disruption at such nodes may significantly impact service efficiency and passenger distribution, especially during peak hours (Su *et al.* (2023)).

ii Betweenness Centrality

Definition and Equation Betweenness centrality quantifies the extent to which a node lies on the shortest paths between other pairs of nodes. It is formally defined as:

$$C_B(v) = \sum_{i \neq j \neq v} \frac{\sigma_{ij}(v)}{\sigma_{ij}}$$

where σ_{ij} is the total number of shortest paths between nodes i and j, and $\sigma_{ij}(v)$ is the number of those paths that pass through node v(Freeman (1977)).

Application in the Underground Context Betweenness centrality highlights nodes that serve as transfer corridors or bridging stations within the Underground (Mussone and Notari (2025)). For example, stations such as Green Park or Oxford Circus often lie on the shortest routes connecting disparate parts of the network. These nodes may not have the highest degree but are essential for enabling efficient travel between many station pairs.

Importance for Network Functioning Liu *et al.* (2018) pinted out that nodes with high betweenness centrality represent potential bottlenecks or points of vulnerability. A failure or congestion at such stations would disrupt many origin-destination pairs, significantly affecting network-wide travel efficiency. Thus, betweenness centrality is instrumental in evaluating the resilience and structural robustness of the Underground.

iii Closeness Centrality

Definition and Equation Closeness centrality measures how near a node is to all other nodes in the network. For a node v, it is defined as the inverse of the average shortest path distance to all other nodes:

$$C_C(v) = \frac{1}{\sum_{v \neq t} d(v,t)}$$

where $d(v,t)$ denotes the length of the shortest path from node v to node t, and N is the number of nodes in the network (Bavelas (1950)).

Application in the Underground Context Closeness centrality identifies stations that are centrally located in terms of travel time or distance. Stations such as Bank or Holborn are often close to a large number of other stations and can be reached from many locations with relatively few transfers.

Importance for Network Functioning Stations with high closeness centrality enable efficient access to the rest of the network and are critical for minimizing overall travel times. They are often located in geographical and topological centers of the network and serve as ideal points for dispersing or collecting traffic.

Synthesis and Relevance

Together, these three centrality measures offer complementary insights into the importance of stations in the Underground:
- Degree centrality captures local importance and identifies hubs for immediate accessibility and transfers.
- Betweenness centrality identifies strategic connectors that hold the network together across routes.

- Closeness centrality highlights spatial efficiency, marking stations from which the network is most accessible.

By applying these measures in combination, one can systematically identify stations that are structurally and operationally indispensable to the Underground's performance and resilience. These metrics are crucial not only for understanding current usage patterns but also for planning targeted investments, maintenance priorities, and emergency response strategies (Bavelas (1950)).

The first 10 ranked nodes for each of the 3 measures are listed below:

	station_name	degree_t
0	Stratford	0.0225
1	Bank and Monument	0.0200
2	Baker Street	0.0175
3	King's Cross St. Pancras	0.0175
4	Waterloo	0.0150
5	Earl's Court	0.0150
6	Canning Town	0.0150
7	Liverpool Street	0.0150
8	Oxford Circus	0.0150
9	West Ham	0.0150

	station_name	betweenness_t
0	Stratford	23768.093434
1	Bank and Monument	23181.058947
2	Liverpool Street	21610.387049
3	King's Cross St. Pancras	20373.521465
4	Waterloo	19464.882323
5	Green Park	17223.622114
6	Euston	16624.275469
7	Westminster	16226.155916
8	Baker Street	15287.107612
9	Finchley Road	13173.758009

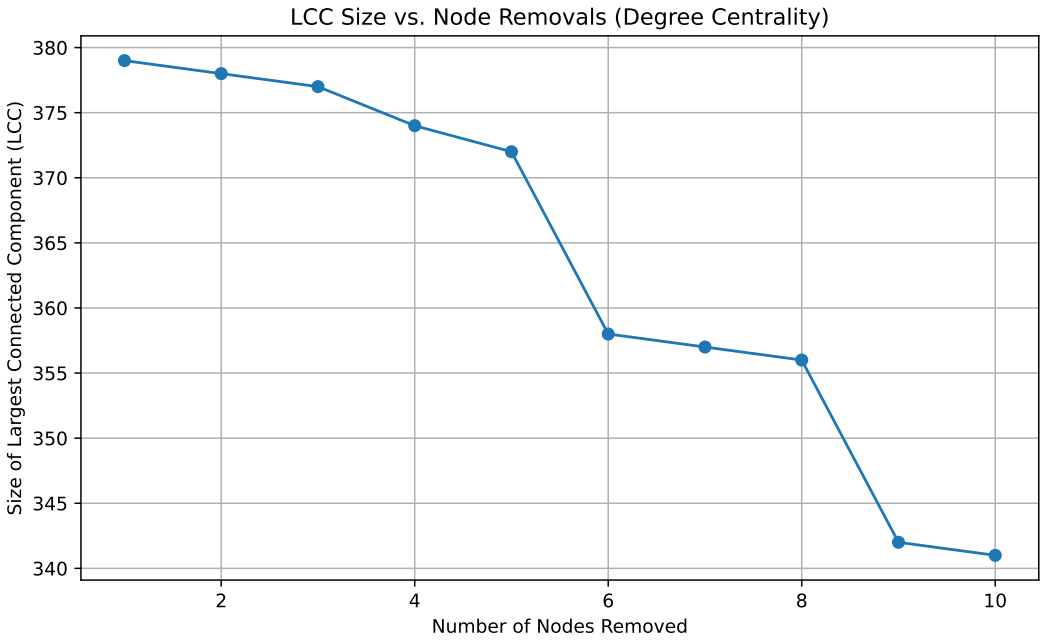
	station_name	closeness_t
0	Green Park	0.114778
1	Bank and Monument	0.113572
2	King's Cross St. Pancras	0.113443
3	Westminster	0.112549
4	Waterloo	0.112265
5	Oxford Circus	0.111204
6	Bond Street	0.110988
7	Angel	0.110742
8	Farringdon	0.110742
9	Moorgate	0.110314

1.2. Node removal

To evaluate the structural resilience of the urban rail network, we simulate sequential node removals guided by three classical centrality measures: degree, betweenness, and closeness. After each removal, the Largest Connected Component (LCC) is recalculated to assess the resulting fragmentation. Figure LCC Size vs. Node Removals by Centrality Measure presents a comparative line plot of LCC sizes as a function of the number of nodes removed under each centrality-based attack strategy.

Degree centrality

Step 1: Removed Stratford, LCC size: 379
Step 2: Removed Bank and Monument, LCC size: 378
Step 3: Removed King's Cross St. Pancras, LCC size: 377
Step 4: Removed Baker Street, LCC size: 374
Step 5: Removed Oxford Circus, LCC size: 372
Step 6: Removed Canning Town, LCC size: 358
Step 7: Removed Earl's Court, LCC size: 357
Step 8: Removed Waterloo, LCC size: 356
Step 9: Removed Willesden Junction, LCC size: 342
Step 10: Removed Green Park, LCC size: 341

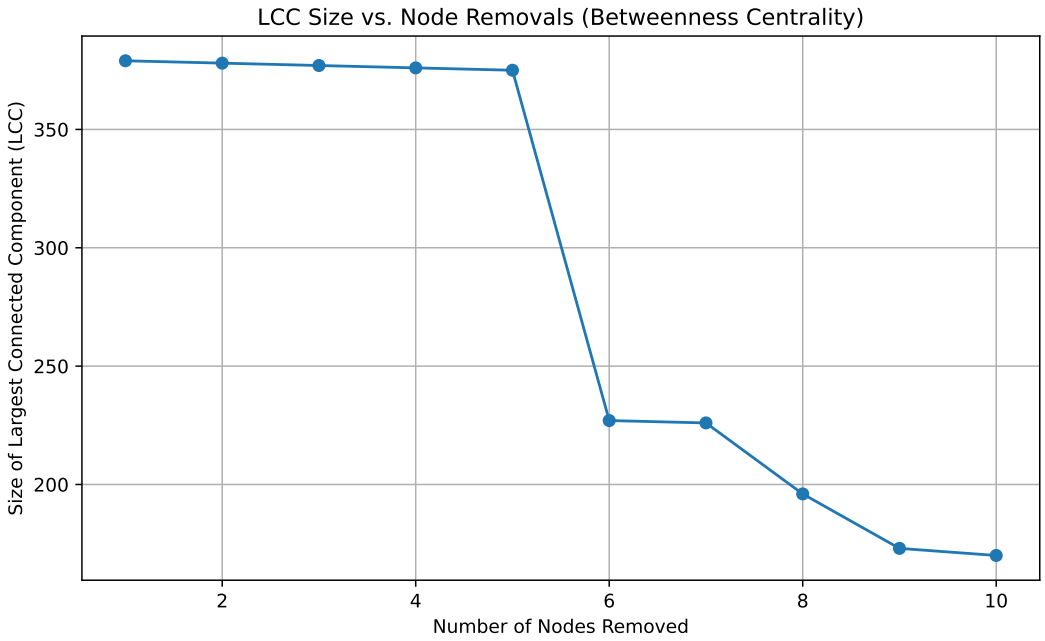


Resilience analysis complete.

Node removals based on degree centrality cause a slow and steady decline in the size of the largest connected component (LCC), suggesting that highly connected stations are not always structurally critical. The network remains largely intact after several removals, indicating resilience to local failures.

Betweenness centrality

Step 1: Removed Stratford, LCC size: 379
Step 2: Removed King's Cross St. Pancras, LCC size: 378
Step 3: Removed Waterloo, LCC size: 377
Step 4: Removed Bank and Monument, LCC size: 376
Step 5: Removed Canada Water, LCC size: 375
Step 6: Removed West Hampstead, LCC size: 227
Step 7: Removed Earl's Court, LCC size: 226
Step 8: Removed Shepherd's Bush, LCC size: 196
Step 9: Removed Euston, LCC size: 173
Step 10: Removed Baker Street, LCC size: 170

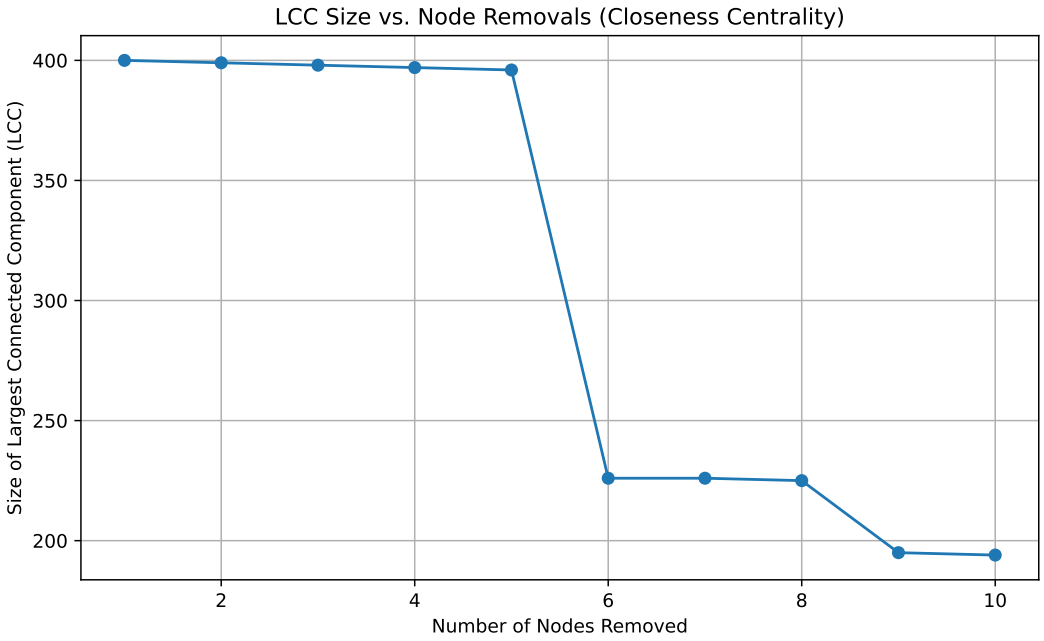


Resilience analysis complete.

Betweenness centrality leads to a sharp drop in LCC size after the removal of key bridging nodes. The sixth removal, corresponding to West Hampstead, causes fragmentation, showing that betweenness is effective at identifying structural weak points.

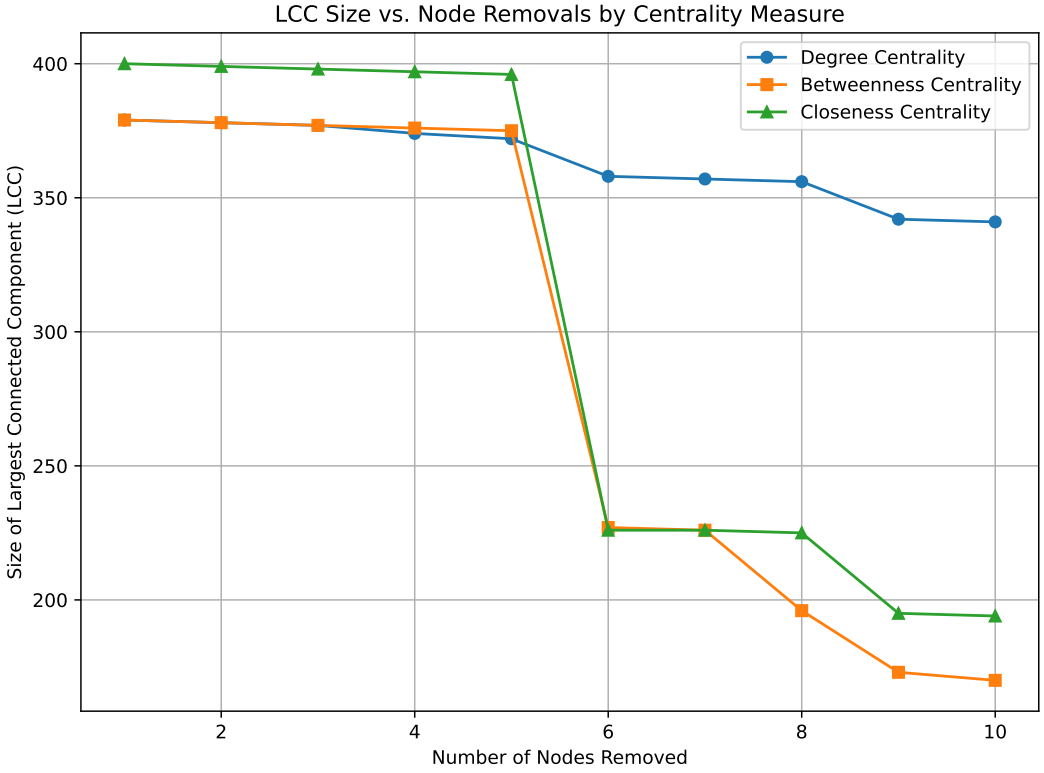
Closeness centrality

Step 1: Removed Green Park, LCC size: 400
Step 2: Removed King's Cross St. Pancras, LCC size: 399
Step 3: Removed Waterloo, LCC size: 398
Step 4: Removed Bank and Monument, LCC size: 397
Step 5: Removed West Hampstead, LCC size: 396
Step 6: Removed Canada Water, LCC size: 226
Step 7: Removed Stratford, LCC size: 226
Step 8: Removed Earl's Court, LCC size: 225
Step 9: Removed Shepherd's Bush, LCC size: 195
Step 10: Removed Oxford Circus, LCC size: 194



Resilience analysis complete.

Interestingly, closeness centrality shows a similar pattern, with station Canada Water triggering rapid decline. This overlap suggests that closeness can also capture structurally important hubs, especially when spatial and topological centrality coincide.



The results reveal that betweenness centrality is the most effective measure for identifying structurally critical nodes. The LCC size remains stable through the first five node removals but drops sharply between the fifth and sixth removal, from approximately 375 to 227. This indicates that betweenness-ranked nodes act as bridging stations, whose removal causes

major segmentation of the network. Interestingly, closeness centrality, while traditionally viewed as a measure of global efficiency rather than structural vulnerability, exhibits a similarly sharp drop. This suggests that some top closeness-ranked nodes may also serve as intermodal connectors with high structural importance.

In contrast, degree centrality, which prioritizes nodes with the highest immediate connections—leads to a more gradual reduction in LCC size. The persistence of a large connected component after several degree-based removals indicates that the network topology retains a certain level of redundancy. This echoes observations by Derrible and Kennedy (2010), who noted that metro systems often embed topological resilience through alternate routes and transfer options.

The sudden inflection in the betweenness and closeness curves suggests that a small subset of nodes holds disproportionate control over network connectivity. These are likely to include transfer hubs such as King's Cross St. Pancras or Bank, which serve as bottlenecks across multiple lines. This aligns with Freeman (1977) foundational theory of betweenness centrality, which emphasized the strategic role of nodes occupying critical positions in the shortest paths of the network. Furthermore, network vulnerability under targeted attacks, as demonstrated in the classic work by Albert, Jeong and Barabási (2000), is clearly visible here, especially when removals are guided by topological centrality rather than random selection. But the detailed inspection reveals that West Hampstead and Canada Water are the two stations responsible for this dramatic drop in LCC under both betweenness and closeness-based strategies. These stations function as crucial interchanges that connect multiple lines or subnetworks and their removal severs inter-line connectivity, leading to fragmentation.

Consistent with findings in the transportation literature, such as Freeman (1977), our results confirm that betweenness centrality is particularly suited for resilience analysis in transit networks. Its ability to capture inter-line dependencies and chokepoint behavior makes it a valuable tool for identifying high-risk stations. Closeness centrality, although less consistent, may still serve as a complementary indicator in identifying spatially central yet vulnerable nodes.

2. Flows: weighted network

2.1. Centrality of weighted network

	mode	station_origin	station_destination	Flow_Sum
0	DLR	Stratford International DLR	King George V	92
1	DLR	Stratford International DLR	London City Airport	206
2	DLR	Stratford International DLR	Pontoon Dock	148
3	DLR	Stratford International DLR	West Silvertown	107
4	DLR	Stratford International DLR	Woolwich Arsenal	954
...	...	...	...	...
62575	LU	Canning Town	Southwark	1174
62576	LU	Canning Town	Bermondsey	776
62577	LU	Canning Town	Canada Water	2553
62578	LU	Canning Town	North Greenwich	2316
62579	LU	Canning Town	Canary Wharf LU	2862

The most important station according to each these 3 measures in the weighted network and the computing process:

Degree Centrality:

```
deg_london_w = dict(G_lu.degree(weight='Flow_Sum'))

nx.set_node_attributes(G_lu,deg_london_w,'degree_w')

max_node_d = max(deg_london_w, key=deg_london_w.get)

print(f"The node with the highest degree centrality is: {max_node_d},\n"
    f"with a value of: {deg_london_w[max_node_d]}")
```

The node with the highest degree centrality is: Bank and Monument,
with a value of: 2685403

Betweenness centrality:

```
bet_london_w=nx.betweenness_centrality(G_lu,weight='Flow_Sum',normalized=False)

nx.set_node_attributes(G_lu,bet_london_w,'betweenness_w')

max_node_b = max(bet_london_w, key=bet_london_w.get)

print(f"The node with the highest betweenness centrality is: {max_node_b},\n"
    f"with a value of: {bet_london_w[max_node_b]}")
```

The node with the highest betweenness centrality is: West Hampstead,
with a value of: 31133.25

Closeness centrality:

```
clos_london_w=nx.closeness_centrality(G_lu, distance='Flow_Sum')

nx.set_node_attributes(G_lu,clos_london_w,'closeness_w')

max_node_c = max(clos_london_w, key=clos_london_w.get)

print(f"The node with the highest closeness centrality is: {max_node_c},\n"
    f"with a value of: {clos_london_w[max_node_c]}")
```

The node with the highest closeness centrality is: West Hampstead,
with a value of: 1.5774576506254367e-06

2.2. Largest OD flow

	station_origin	station_destination	mode	Flow_Sum	population
2837	Bank and Monument	Waterloo	LU	25031	185560
55995	Waterloo	Bank and Monument	LU	22897	157642
39164	Oxford Circus	Victoria	LU	14212	114051
54871	Victoria	Oxford Circus	LU	13442	137095

	station_origin	station_destination	mode	Flow_Sum	population
56020	Waterloo	Canary Wharf	LU	11735	157642
50141	Stratford	Liverpool Street	EZLLU	11017	218574
32687	London Bridge	Canary Wharf	LU	9945	124229
9088	Canary Wharf	Waterloo	LU	9790	109297
30048	King's Cross St. Pancras	Victoria	LU	8909	136921
2601	Bank and Monument	Canary Wharf	DLRLU	8797	185560

The largest OD flow is ('Bank and Monument', 'Waterloo')

Number of people affected: 185560

2.3. Optimal alternatives and Walking time analysis

In order to find the nearest station, we write code to identify alternative underground stations within walking distance (1 km) of the closed "Bank and Monument" station. It uses the geodesic distance to filter nearby stations and then downloads the local pedestrian street network via OSMnx. For each candidate station, it calculates the shortest walking route using the road network and estimates walking time based on a speed of 1.2 m/s (**based on UK Department for Transport – WebTAG guidance, Similarly, we also use this speed as the standard in the analysis of the second part.**). The results, including station name, path length, and walking time, are compiled into a sorted DataFrame, which is shown below.

	Station	PathLength_m	WalkTime_min
6	Cannon Street	405	5.6
5	Mansion House	579	8.0
0	Moorgate	818	11.4
7	Tower Hill	840	11.7
8	St. Paul's	856	11.9
1	London Bridge	910	12.6
2	Aldgate	961	13.3
3	Liverpool Street	971	13.5
4	Tower Gateway	1141	15.8

Thus, they should go to Canno Street station as an alternative, whcih will take them about 5.6 minutes, and the shorest path is shown below.

Part II: Spatial Interaction models

1. Models and calibration

1.1. Model introduction

Spatial interaction models are mathematical frameworks used to estimate flows between origins and destinations, factoring in both the attractiveness of destinations and the deterrent effect of distance or travel cost. Among these, urban gravity models specifically focus on predicting and analyzing the movement of people, goods, and services within urban and regional systems, applying principles similar to gravitational forces to explain spatial flows. Inspired by Newton's law of gravitation, these models assume that interaction between two locations is directly proportional to their "masses" such as population, employment and amenities, and inversely proportional to some function of the distance or cost between them (Wilson (1971)). Here, we introduce four commonly used gravity model variants: the *unconstrained*, *origin-constrained*, *destination-constrained*, and *doubly constrained models*.

i Unconstrained Gravity Model

The unconstrained gravity model estimates flows between two locations i and j as:

$$T_{ij} = k \cdot \frac{O_i^{\alpha} D_j^{\gamma}}{f(c_{ij})}$$

where:
- T_{ij} : predicted flow from origin i to destination j
- O_i : origin mass (e.g., population at i)
- D_j : destination mass (e.g., jobs or services at j)
- $f(c_{ij})$: deterrence function, typically increasing with cost c_{ij} (e.g., travel time or distance)
- k : normalization constant
- α : Exponent on the origin variable — adjusts how strongly origin capacity influences flow
- γ : Exponent on the destination variable — adjusts how strongly destination attractiveness influences flow

A common choice for $f(c_{ij})$ is an exponential or power function, such as:

$$f(c_{ij}) = c_{ij}^{\beta} \quad \text{or} \quad f(c_{ij}) = e^{-\beta c_{ij}}$$

where $\beta > 0$ is a deterrence parameter that governs sensitivity to distance.

In the practicals and assignment, we choose $f(c_{ij}) = c_{ij}^{\beta}$, and cost factor is often represented by distance d. So, we can get $f(c_{ij}) = d_{ij}^{\beta}$

Use Case - Exploratory Urban Interaction Analysis The unconstrained gravity model is particularly useful in exploratory analyses where no prior knowledge of total flows from origins or to destinations is available. It is commonly applied in early-stage urban studies or simulations to infer potential spatial interaction patterns based solely on population size and accessibility. For instance, it can be used to estimate hypothetical trade volumes or migration intensities between cities in the absence of empirical flow data, assuming that larger cities attract more flows and that distance acts as a frictional constraint. This approach has also been used in cross-national contexts, such as modeling international migration or tourism flows without border-specific data, suggested by Liu *et al.* (2014). However, the lack of constraints may limit its predictive accuracy in practical planning scenarios.

ii Origin-Constrained Gravity Model

This model constrains flows to match known total outflows from each origin:

$$T_{ij} = O_i \cdot \frac{D_j^{\gamma} f(c_{ij})}{\sum_j D_j^{\gamma} f(c_{ij})} \quad (\alpha = 1)$$

where

$$O_i = \sum_j T_{ij}$$

Here, O_i is fixed, and flows are allocated among destinations j based on their attractiveness D_j and proximity to i. The denominator ensures that the total outflow from each i matches O_i. thus

$$A_i = \frac{1}{\sum_j D_j^{\gamma} f(c_{ij})}$$

Use Case - Commuting and Travel Demand Estimation The origin-constrained model is most suitable when total flows from each origin like residential zone are known but their distribution among destinations is unknown. A prominent application is in commuting flow estimation, where the number of workers living in each neighborhood is given, but their specific job locations must be inferred based on job densities and travel costs. Transportation planners frequently employ this model to forecast travel demand from suburbs into urban cores, particularly in developing cities where detailed employment destination data may be lacking. It has also been adopted in health accessibility studies to evaluate patient flows from homes to medical facilities while holding origin population fixed (Wilson (1971)).

iii Destination-Constrained Gravity Model

This model is symmetrical to the origin-constrained version, ensuring the total inflows to each destination are fixed:

$$T_{ij} = D_j \cdot \frac{O_i^{\alpha} f(c_{ij})}{\sum_i O_i f(c_{ij})} \quad (\gamma = 1)$$

where

$$D_j = \sum_i T_{ij}$$

which is exactly the opposite of the previous one. D_j is fixed, and flows are allocated among origins i based on their features O_i and proximity to j. The denominator ensures that the total outflow from each j matches D_j. thus

$$B_j = \frac{1}{\sum_i O_i^{\alpha} f(c_{ij})}$$

Use Case - Facility Access or Service Catchment Modeling When destination capacities or inflow totals are fixed—such as the number of students a school can accept or the maximum number of patients a hospital can serve—the destination-constrained gravity model is appropriate. This model allocates demand across origins based on their relative proximity and population size while ensuring that no destination is overloaded. As indicated by Wills (1986), it is commonly used in health services planning, disaster response logistics, and educational catchment modelling. For example, a regional health department may use this model to estimate how many patients each hospital will receive from different neighborhoods, assuming each hospital has a capped intake capacity and patients prefer closer facilities.

iv Doubly Constrained Gravity Model

This model imposes constraints on both origins and destinations:

$$T_{ij} = A_i B_j O_i D_j f(c_{ij}) \quad (\alpha = \gamma = 1)$$

with balancing factors A_i and B_j ensuring the row and column totals match the known origin and destination flows:

$$A_i = \frac{1}{\sum_j B_j D_j f(c_{ij})} \quad and \quad B_j = \frac{1}{\sum_i A_i O_i f(c_{ij})}$$

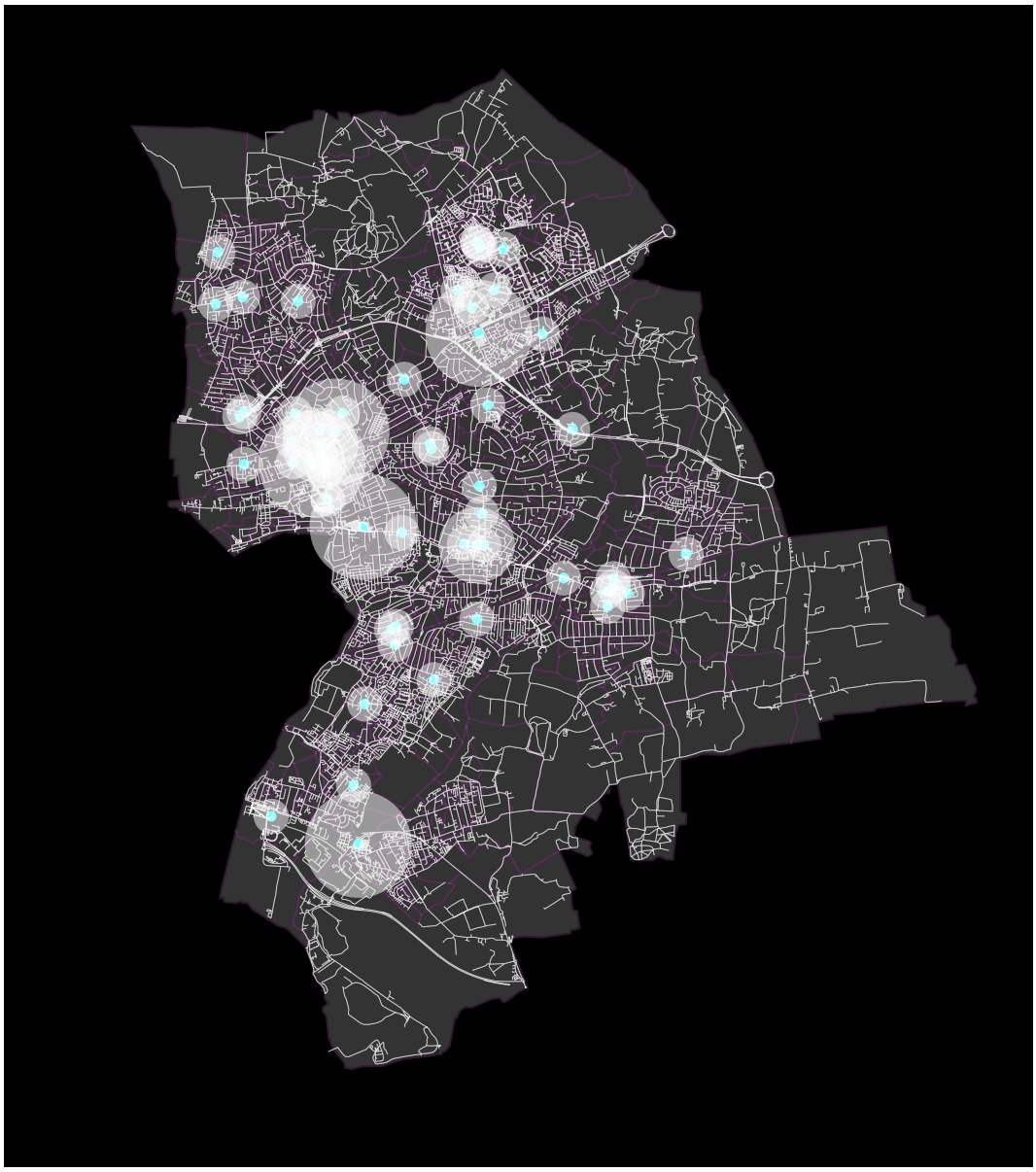
This formulation is usually solved iteratively (Furness method) and produces the most accurate fit when both marginals are known.

Use Case - Transport and Urban Systems Calibration The doubly constrained gravity model is the most data-intensive but also the most accurate among the gravity model family, making it ideal for calibrated modeling in well-studied urban systems. It is used when both the total outflows from each origin and the total inflows to each destination are known, such as in household travel surveys or census-based commuting datasets. This model is widely applied in regional transportation models, logistics routing, and land-use forecasting, particularly in metropolitan areas with reliable OD (origin-destination) matrices. For example, metropolitan planning organizations (MPOs) use this model to simulate peak-hour transit flows between neighborhoods and employment hubs, taking into account actual boarding and alighting totals (Boyce and Williams (2015)). Its iterative balancing procedure ensures that predicted flows match empirical margins closely, making it a standard in transport modeling software.

1.2. Model selection and calibration

The gravity model posits that the interaction T_{ij} between origin i (population location, in our case, represented by the centroid of OAs) and destination j (supermarket) increases with: - the attractiveness of the destination (supermarket area), and - the size of the origin (population), and decreases with the cost of travel (distance).

All the features make it ideal for modeling accessibility and shopping behavior, where people are drawn more to larger supermarkets but are deterred by greater distance. Since no actual flow data (e.g., shopper counts) are available, a constrained gravity model can be used. Origin-constrained model is most suitable, because the total number of trips from each origin is assumed to be equal to the population. Furthermore, through the γ and β values that have been calibrated in practical, we directly substitute the known parameters into the model and can directly calculate A_i of each origin. Finally, we could get T_{ij}.



```
distance_matrix.loc['E00011326', 'E00011326'] = 1
distance_matrix.loc['E00011710', 'E00011710'] = 1
```

For the two potential areas where supermarkets may be added, the coordinates of the origin and the proposed supermarket location coincide. This results in zero values in the distance matrix, which may cause computational issues. To prevent this, all zero distances are replaced with 1.

```
Average_Area = retail_sub['square_meters'].mean()

df_model.loc[df_model['Destination'] == 'E00011326', 'Area'] = Average_Area
df_model.loc[df_model['Destination'] == 'E00011710', 'Area'] = Average_Area
```

Assume that the size of both new supermarkets is the same and corresponds to the mean of all sizes.

	Origin	T_ij	Population	difference
0	E00011202	435.0	435	0.0

	Origin	T_ij	Population	difference
1	E00011203	326.0	326	-0.0
2	E00011204	326.0	326	-0.0
3	E00011205	376.0	376	0.0
4	E00011206	567.0	567	0.0
...	...	...	...	...
770	E00176973	544.0	544	0.0
771	E00176974	506.0	506	0.0
772	E00176975	197.0	197	0.0
773	E00176979	413.0	413	0.0
774	E00176984	293.0	293	-0.0

For each origin, the total estimated flow to all destinations matches exactly the total flow, which means $\sum_j T_{ij} = O_i$.

2. Scenarios

2.1. Location of new supermarket

The same model and method as in the previous section are applied here, with the only difference being the distance matrix used. In the following analysis, potential supermarket investment locations will be incorporated into the distance matrix to recalculate A_i respectively, and predict T_{ij} subsequently.
Check the origin constrained condition.

For E00011326:

	Origin	T_ij	Population	difference
0	E00011202	435.0	435	-0.0
1	E00011203	326.0	326	-0.0
2	E00011204	326.0	326	0.0
3	E00011205	376.0	376	-0.0
4	E00011206	567.0	567	0.0
...	...	...	...	...
770	E00176973	544.0	544	-0.0
771	E00176974	506.0	506	0.0
772	E00176975	197.0	197	-0.0
773	E00176979	413.0	413	0.0
774	E00176984	293.0	293	-0.0

For E00011710:

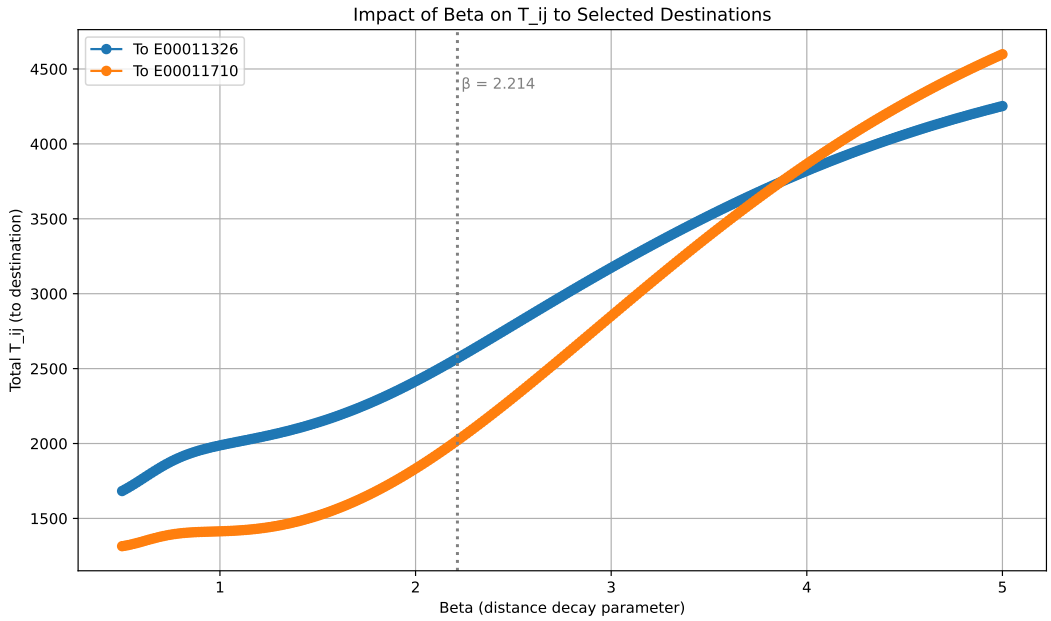
	Origin	T_ij	Population	difference
0	E00011202	435.0	435	0.0
1	E00011203	326.0	326	-0.0
2	E00011204	326.0	326	-0.0
3	E00011205	376.0	376	0.0
4	E00011206	567.0	567	0.0

	Origin	T_ij	Population	difference
...	...	...	...	...
770	E00176973	544.0	544	-0.0
771	E00176974	506.0	506	0.0
772	E00176975	197.0	197	0.0
773	E00176979	413.0	413	0.0
774	E00176984	293.0	293	0.0

Finally, we compare the total flow $\sum T_{ij}$ reach these two potential loactions and get the result:

Estimated customer flow for E00011326: 2569
Estimated customer flow for E00011710: 2021
Location E00011326 is better, with a higher estimated customer count.

2.2. Transport increasement



The parameter β in the gravity model captures how sensitive people are to distance when deciding where to travel—commonly referred to as the distance decay effect. A higher β means that people are more deterred by distance, so closer destinations attract disproportionately more flow. In contrast, a lower β implies weaker distance deterrence, allowing farther destinations to compete more equally (Liu *et al.* (2012)). In the graph, we observe that as β increases, total flows to both E00011326 and E00011710 rise, but the relative advantage shifts: E00011326 dominates at lower β values, while E00011710 overtakes it at higher β, suggesting the latter benefits more when local flows are emphasized. Assuming a sharp increase in transport efficiency, for instance, better connectivity or reduced travel time, the effective β would decrease, reducing the penalty of distance. Under this shift, E00011326 would likely regain its dominance, as more distant flows become viable and its broader appeal can outweigh proximity advantages.

On the ther hand, When β increases, the gravity model applies a stronger penalty to distance, meaning people become more likely to choose nearby destinations and less likely to travel

far. In practical terms, as β grows:
- Flows become more localized: distant destinations receive less flow, while nearby destinations gain a larger share.
- The influence of destination attractiveness (e.g., size or services) is suppressed by distance.
- The spatial interaction pattern becomes more fragmented, favoring dense clusters or locally dominant areas.

In the graph, as β increases, both destinations receive more flow—but E00011710 eventually surpasses E00011326 at $\beta \approx 3.8$. This suggests that E00011710 is more locally accessible or surrounded by denser origins, and thus benefits more in a high-β scenario. Conversely, E00011326 likely draws from a broader spatial base, which gets suppressed as β rises.

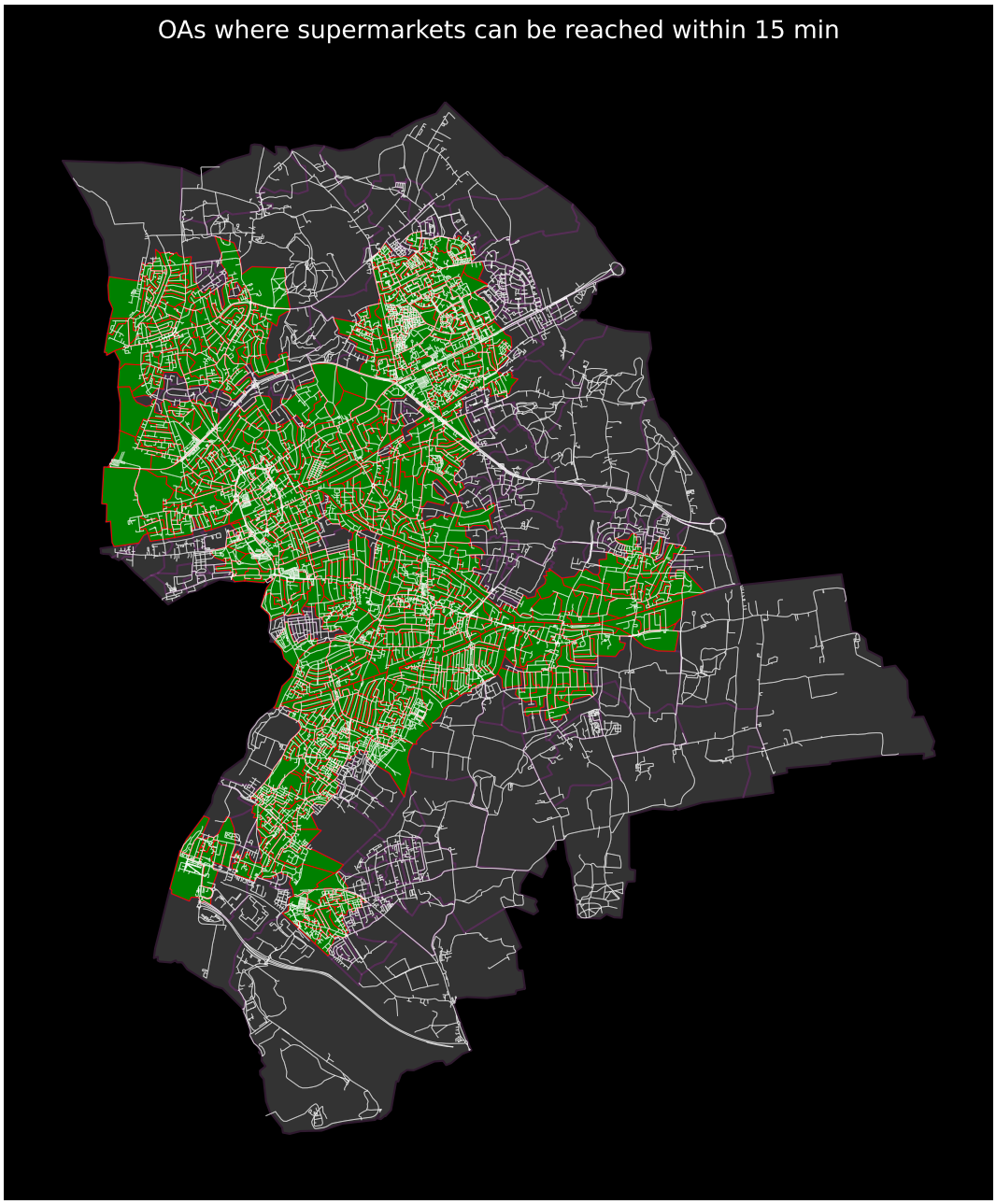
If β continues to grow, the system may converge toward a pattern where only the closest and most immediate destinations dominate, regardless of their absolute attractiveness—reflecting a highly frictional spatial system with limited mobility.

2.3. Extra - 15 min Circle

We evaluate pedestrian accessibility to retail locations within a 15-minute walking distance for each Output Area (OA). Assuming a walking speed of **72** meters per minute, we define a walk radius of 1,080 meters. For each OA, we use a street network graph (G) to generate an ego subgraph representing all nodes reachable within that distance. We Then check whether any retail locations, identified by their osmid, fall within this reachable area. The results are compiled into a DataFrame indicating whether each OA has access to at least one retail node, along with its population. Finally, we summarize the total number of OAs, those with access, and the covered and uncovered populations:

	Value
Total OAs	775
OAs with access to ≥1 supermarket	590
Total population	262051
Covered population	198931
Uncovered population	63120

The OAs that can walk to a supermarket within 15 minutes is shown in the figure below.



Due to space limitations, some parts of the code have been omitted, or may not be fully rendered in the PDF. The complete code can be found here(in the ‘Code’ folder, including all the qmd, ipynb and pdf files).

References

Albert, R., Jeong, H. and Barabási, A.-L. (2000) 'Error and attack tolerance of complex networks', *Nature*, 406(6794), pp. 378–382. Available at: https://doi.org/10.1038/35019019.

Bavelas, A. (1950) 'Communication Patterns in Task-Oriented Groups', *The Journal of the Acoustical Society of America*, 22(6), pp. 725–730. Available at: https://doi.org/10.1121/1.1906679.

Bloch, F., Jackson, M.O. and Tebaldi, P. (2023) 'Centrality measures in networks', *Social Choice and Welfare*, 61(2), pp. 413–453. Available at: https://doi.org/10.1007/s00355-023-01456-4.

Boyce, D.E. and Williams, H.C.W.L. (2015) *Forecasting Urban Travel: Past, Present and Future*. Edward Elgar Publishing.

Derrible, S. (2012) 'Network Centrality of Metro Systems', *PLOS ONE*, 7(7), p. e40575. Available at: https://doi.org/10.1371/journal.pone.0040575.

Derrible, S. and Kennedy, C. (2010) 'The complexity and robustness of metro networks', *Physica A: Statistical Mechanics and its Applications*, 389(17), pp. 3678–3691. Available at: https://doi.org/10.1016/j.physa.2010.04.008.

Freeman, L.C. (1977) 'A Set of Measures of Centrality Based on Betweenness', *Sociometry*, 40(1), pp. 35–41. Available at: https://doi.org/10.2307/3033543.

Liu, B. *et al.* (2018) 'Recognition and Vulnerability Analysis of Key Nodes in Power Grid Based on Complex Network Centrality', *IEEE Transactions on Circuits and Systems II: Express Briefs*, 65(3), pp. 346–350. Available at: https://doi.org/10.1109/TCSII.2017.2705482.

Liu, Y. *et al.* (2012) 'Understanding intra-urban trip patterns from taxi trajectory data', *Journal of Geographical Systems*, 14(4), pp. 463–483. Available at: https://doi.org/10.1007/s10109-012-0166-z.

Liu, Y. *et al.* (2014) 'Uncovering Patterns of Inter-Urban Trip and Spatial Interaction from Social Media Check-In Data', *PLOS ONE*, 9(1), p. e86026. Available at: https://doi.org/10.1371/journal.pone.0086026.

Mussone, L., Aranda Salgado and Notari, R. and (2025) 'A dynamic evaluation of an underground transportation system using image processing and centrality index computation', *Transportation Planning and Technology*, 48(3), pp. 506–535. Available at: https://doi.org/10.1080/03081060.2024.2403643.

Su, G. *et al.* (2023) 'Simulation-Based Method for the Calculation of Passenger Flow Distribution in an Urban Rail Transit Network Under Interruption', *Urban Rail Transit*, 9(2), pp. 110–126. Available at: https://doi.org/10.1007/s40864-023-00188-z.

Wills, M.J. (1986) 'A flexible gravity-opportunities model for trip distribution', *Transportation Research Part B: Methodological*, 20(2), pp. 89–111. Available at: https://doi.org/10.1016/0191-2615(86)90001-9.

Wilson, A.G. (1971) ‘A Family of Spatial Interaction Models, and Associated Developments’, *Environment and Planning A*, 3(1), pp. 1–32. Available at: https://doi.org/10.1068/a030001.

Zhang, R., Pan, J. and Lai, J. (2021) ‘Network Structure of Intercity Trips by Chinese Residents under Different Travel Modes: A Case Study of the Spring Festival Travel Rush’, *Complexity*, 2021(1), p. 1283012. Available at: https://doi.org/10.1155/2021/1283012.