Big Data and Big Models for a better customer experience

Olga Feldman

Strategic Manager Analytics,
Transport for London

BIG DATA – The Art of the Possible, Cambridge

10 September 2015

TRANSPORT FOR LONDON
EVERY JOURNEY MATTERS

Today's presentation - agenda

1. TfL Customer Experience Analytics

2. TfL Models

3. Customer Experience Analytics Research

Transport for London: Our Purpose

'Keep cities working and growing and make life better'

Plan ahead to meet the challenges of a growing population

Unlock economic development and growth

Meet the rising expectations of our customers and users

TfL: Integrated body created in 2000 responsible for the Capital's transport system

What our customers want

Understand what we stand for

Excellent reliability and customer experience

Value for Money

Progress & Innovation

Trust

These principles guide the work that we do

London's Big Data

46 million bus journeys a week

8.6 million people rising to 10 million people by 2030

25 million London Underground journeys a week

London is 'Big', so our data is 'Big' too…

Data is gold dust

Data is at the centre of delivering better urban transport:

- **Open Data –** Information that we make freely available to power travel information apps and other tools

- **Big Data** – The analysis of one or more large data sets to reveal patterns or trends and enable action to be taken

Sources of Big Data

Oyster and contactless cards

Social media

Bus location data

Big Data

Traffic information

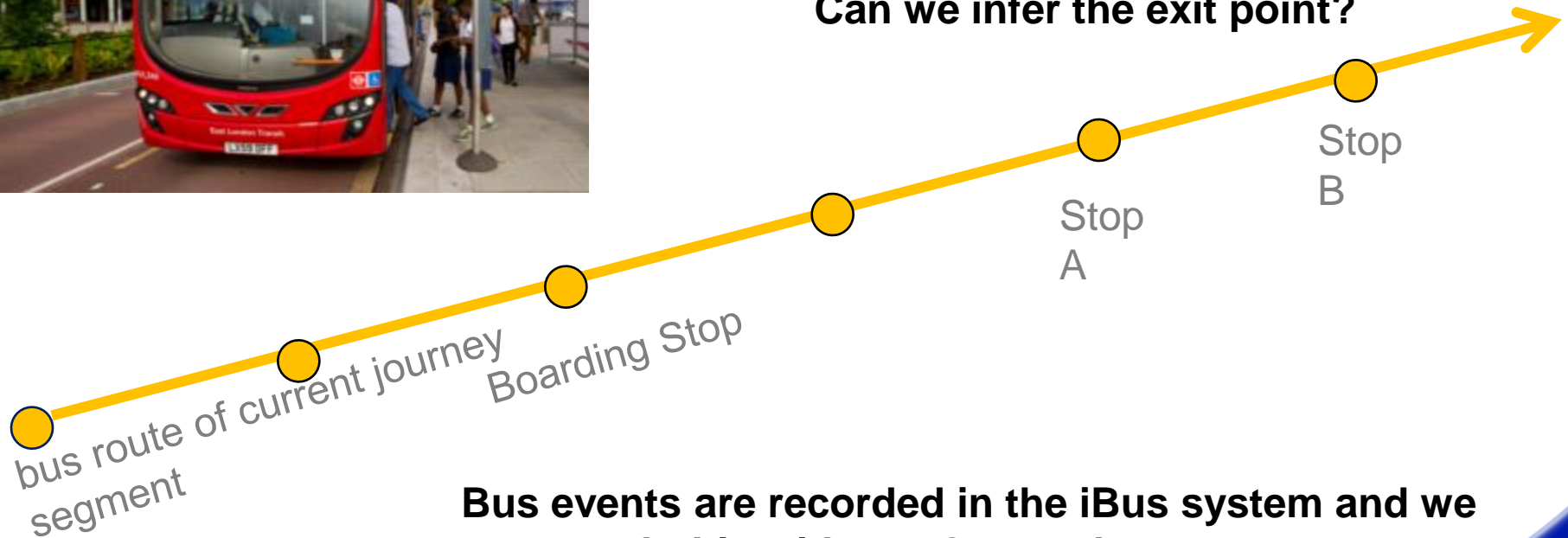Apps and mobile/WIFI data

Asset Data

Inferring destination for bus trips

A customer taps an Oyster card on the reader, which records the location and time

Can we infer the exit point?

Stop B

Stop A

Boarding Stop

bus route of current journey segment

Bus events are recorded in the iBus system and we can match this with our Oyster data

Where is the next tap?

From the location of the next tap (if there is one), we can infer where a customer alights

bus route R

Underground line

Station Y

Stop X

Stop B

If next trip begins at stop X, the current segment is inferred to end at stop A

Stop A

bus route of current journey segment

If next trip begins at station Y, the current segment inferred to end at stop B

Using our Big Data tools

	Time	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###
LEWISHAM CENTRE		0	0	4	2	3	3	3	2	3	4	1	3	8	0	5	10	12	0	0	6	0	4	3	8	1	5	7	2
LOCKMEAD ROAD		2	0	5	4	3	6	4	2	4	5	3	8	7	6	5	18	7	0	0	7	0	6	3	9	2	8	8	5
THE SQUIRRELS		2	0	5	4	3	6	4	2	5	5	3	7	7	6	5	18	8	2	0	7	0	6	3	9	3	8	8	2
ST MARGARET'S CHURCH / BRANDRAM ROAD		2	0	5	4	3	6	5	2	5	5	3	7	8	6	6	10	2	2	4	8	0	6	3	9	4	8	6	2
BLACKHEATH HOSPITAL		2	1	6	4	3	7	5	2	5	5	4	7	8	5	8	11	3	3	5	8	0	6	3	9	4	7	6	2
BLACKHEATH STATION #		2	1	6	3	4	11	7	2	6	11	8	11	8	7	10	18	9	7	8	9	0	9	5	12	6	14	4	5
ROYAL PARADE		2	2	6	3	4	12	7	1	6	11	8	12	9	8	10	19	10	7	7	12	0	12	5	14	5	14	4	3
MONTPELIER ROW		2	2	6	3	5	12	7	1	6	11	8	14	12	8	10	20	11	8	6	11	0	14	5	15	8	15	3	4
ST GERMANS PLACE		2	2	6	3	5	12	9	1	6	11	8	16	12	9	13	21	11	9	12	15	0	15	6	15	8	16	3	4
STRATHEDEN ROAD / SHOOTERS HILL ROAD		2	2	6	3	7	14	11	2	11	17	13	23	20	12	19	37	13	24	17	18	0	21	6	18	8	18	3	4
BLACKHEATH / ROYAL STANDARD		4	3	10	9	15	29	23	10	26	43	25	40	48	27	27	54	28	43	29	22	10	21	8	20	10	17	11	6
KIRKSIDE ROAD		4	6	11	11	15	33	26	11	32	43	27	43	50	34	28	54	29	43	32	23	10	24	9	22	8	17	11	6
WESTCOMBE PARK STATION #		4	7	12	11	16	33	26	12	32	43	27	45	50	37	30	54	31	43	32	24	10	24	10	21	9	18	11	7
WESTERDALE ROAD		4	7	12	11	16	32	27	11	33	43	26	45	50	38	30	54	28	43	31	23	11	23	10	21	8	17	10	5
SAINSBURY'S AT GREENWICH PENINSULA		5	7	16	16	18	30	43	18	36	43	34	46	50	40	44	56	34	43	32	22	11	27	9	24	11	18	11	5
ODEON CINEMA		5	8	15	16	19	29	46	25	38	42	36	47	50	40	43	54	40	43	31	22	11	27	10	24	11	18	11	5
MILLENNIUM VILLAGE SOUTH		5	8	15	18	19	29	47	25	39	42	37	47	50	40	43	56	40	43	31	22	11	26	11	22	11	19	11	5
MILLENNIUM VILLAGE / OVAL SQUARE		5	8	16	19	18	34	48	28	44	43	45	47	50	40	42	55	40	47	30	31	12	30	12	26	14	23	15	7
NORTH GREENWICH STATION <>		8	8	2	8	15	7	8	2	5	20	3	8	5	7	5	4	6	5	3	7	9	10	6	21	4	9	4	2
BOORD STREET		8	7	2	8	14	7	8	1	6	20	3	8	5	7	5	5	6	4	3	7	9	10	6	21	4	9	4	2
BLACKWALL LANE		8	8	3	8	14	7	8	0	9	16	3	9	4	8	9	6	6	3	3	7	11	9	6	24	3	12	5	2
MORDEN WHARF ROAD		8	8	3	8	14	7	9	0	9	16	3	9	4	8	9	6	6	3	3	7	11	9	6	24	3	12	5	2
BLACKWALL TUNNEL / EAST INDIA DOCK ROAD		8	6	3	7	12	7	6	0	6	13	3	9	5	8	9	7	6	3	3	7	9	8	6	23	4	11	5	2
ANDREW STREET		9	6	3	7	12	7	6	0	6	14	4	13	8	11	10	9	6	7	4	8	8	8	6	23	3	14	5	3
ST LEONARD'S WHARF		9	6	4	9	13	5	11	3	9	16	6	18	10	15	11	12	6	9	6	10	9	8	7	23	4	14	8	3
COVENTRY CROSS ESTATE		9	7	4	10	13	5	12	3	9	16	6	17	11	15	11	11	6	8	6	9	9	8	7	23	4	14	9	3
BROMLEY BY BOW STATION <>		5	6	4	7	12	4	9	1	7	17	4	10	15	10	9	7	5	6	4	9	9	7	8	22	4	14	6	3
GRACE STREET		5	6	4	7	12	4	9	1	7	17	4	10	15	10	9	7	6	6	4	8	9	7	8	22	4	13	5	3
BOW INTERCHANGE		5	5	2	6	9	4	7	1	7	10	2	8	11	9	8	6	6	6	4	8	7	7	7	18	5	12	4	3
MARSHGATE LANE		4	5	2	6	8	4	7	1	7	10	2	8	11	9	8	6	6	6	4	8	7	7	6	18	5	11	4	3
WARTON ROAD		4	6	9	9	15	10	9	3	13	17	12	21	9	18	20	8	3	6	4	12	9	6	8	17	6	11	7	3
STRATFORD HIGH ST STN [DLR] / CARPENTERS ROAD		4	6	10	10	16	9	7	3	13	19	14	24	12	21	23	10	3	7	4	12	9	8	10	16	6	10	10	3
STRATFORD BUS STATION <> # [DLR]		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0

We use our analysis to monitor congestion so that we can tailor our bus services where needed

We also use this for our bus route groupings so we can design good transfer points

Case Study: Major bridge closure

- Summer **2014:** Wandsworth Borough Council had to **close Putney Bridge** for emergency repair work.

- Bus services had to stop either side of bridge. People could walk or cycle across.

- We used Oyster taps and iBus location data to predict how many bus passengers affected

Analytics in Action – Putney Bridge

- Approx **40,000 unique Oyster** made **111,000 bus journeys** a week that the bridge

- Roughly **half** of these journeys **started or ended very close to the bridge**, so transfer necessary.

- **56,000** journeys crossed the bridge in the middle of a trip. These would require two bus trips, one either side of the bridge

- Result: arranged to offer transfer facilities so that customers would not be charged twice

- Sent **targeted emails** to provide customers with information about alternative routes **to minimise** the impact

Case Study: Automated fare refunds

Sometimes things go wrong and we refund fares

Revenue collection

Operational issues or customer mistakes

Automated refunds

Pattern matching

Case Study: Customer Segmentation at Stations

Baker Street - Cluster Sizes
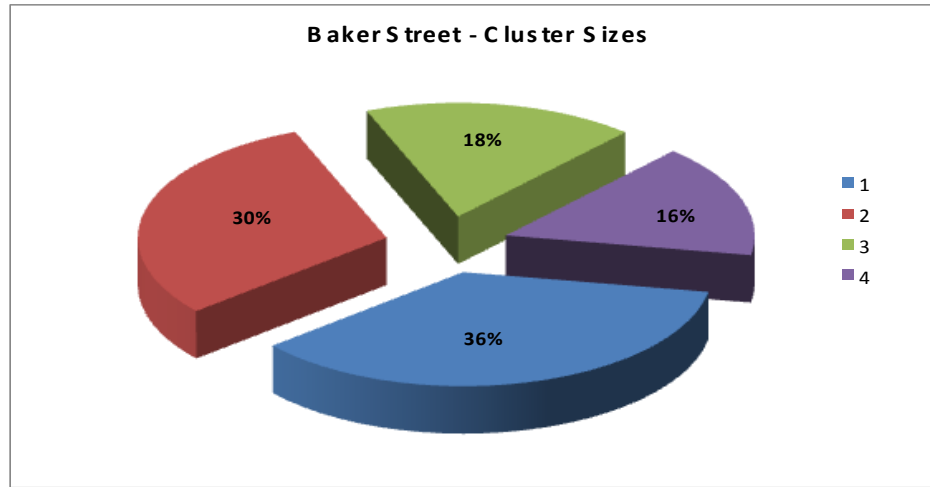
Cluster	Description	Median Start Time	Journeys / travelled day	Ratio unique stations / days travelled	# of Regular Days	# of Irregular Days	Cluster Sample Size
1	Regular Frequent User	08:11	2.5	0.7	11.6	4.6	36%
2	Occasional User (Resident)	12:53	1.9	1.3	2.5	3.6	30%
3	Irregular Frequent User	11:50	2.6	0.9	4.1	12.1	18%
4	Occasional User (Visitor)	11:01	3.4	3.1	1.9	2.0	16%

Our analysis of journey records helps us understand the types of customers who use our stations. This helps us plan ticket facilities, signage and commercial offering.

Case Study: Influencing travel demand
Analysis of travel patterns helps customers with flexibility

Leaflet

Whiteboard poster

Email

Announcement

London Underground

New or occasional customer at Mile End?

If you don't use this station often, you may like to know that the busiest time here is between 08:15 and 08:45.

If you are able to travel outside this time you could have a more comfortable journey.

MAYOR OF LONDON — TRANSPORT FOR LONDON — EVERY JOURNEY MATTERS

Know your travel options

The Central line at this station can become busy at peak times. You may benefit from a quicker journey into central London if you take the first District or Hammersmith & City line train.

| Mile End | Monument | Bank |
District — 10 minutes →

| Mile End | Liverpool Street |
Hammersmith & City — 9 minutes →

More than a billion journeys are made on the Tube each year, with almost one million journeys made on the Central line each day and the number is increasing as London continues to grow.

We are investing to improve the capacity and frequency of your Tube services but we know that at certain times and places the network can be very busy.

We are providing information on the busiest time at Mile End station as we know that over 60% of customers do not regularly travel from here. The busiest time at this station is between 08:15 and 08:45. If you are able to travel outside this time you could have a more comfortable journey.

To find out how TfL are reinvesting in transport, visit: tfl.gov.uk/campaign/reinvesting-in-transport

♻ recycle — Printed on recycled paper

🌐 tfl.gov.uk
ℹ 24 hour travel information 0343 222 1234*
✉ Sign up for email updates tfl.gov.uk/emailupdates
🐦 @TfLTravelAlerts

*Service and network charges may apply. See tfl.gov.uk/terms for details.

Highbury & Islington station

New or occasional customer at this station?

The busiest time here is between 0815 and 0830

TfL is investing to improve the capacity and frequency of Tube services but we know that at certain times and places the network can be very busy.

If you are able to travel outside this time you could have a more comfortable journey.

MAYOR OF LONDON — UNDERGROUND — TRANSPORT FOR LONDON — EVERY JOURNEY MATTERS

Are our new emails displaying well on your device? If not, allow images or view online

Home | Plan journey | Status update | Tube

TUBE — UNDERGROUND

Dear seed email recipient,

I am writing to share some new analysis on Oxford Circus station.

Every Tube station is different and research shows that over 75% of Oxford Circus customers do not use the station regularly and so may not know that the very busiest part of the peak time at this station is from 17:30 to 18:30. Anyone able to travel outside of this time could have a more comfortable journey.

More than a billion journeys are made on the Tube each year, with almost one million journeys made on the Central line each day. We know at certain stations and at certain times it can be very busy. That's why we are investing to improve the capacity and frequency of your Tube services.

For full details and to find out more about how we are reinvesting to improve your Tube, visit tfl.gov.uk/reinvesting-in-transport

We are carrying out a short 3 question survey on this email and we would welcome your views. To participate, please click the box below:

To feedback on this email, please click here

Yours sincerely,

Stuart Reid
Travel Demand Management Programme Director

---

Leytonstone – to be played from 0800 – 0830

"TfL is investing to improve the capacity and frequency of Tube services but we know that at certain times and places the network can be very busy.

The busiest time here is between 08:15 and 08:30.
If you are able to travel outside this time you could have a more comfortable journey."

Bethnal Green Results
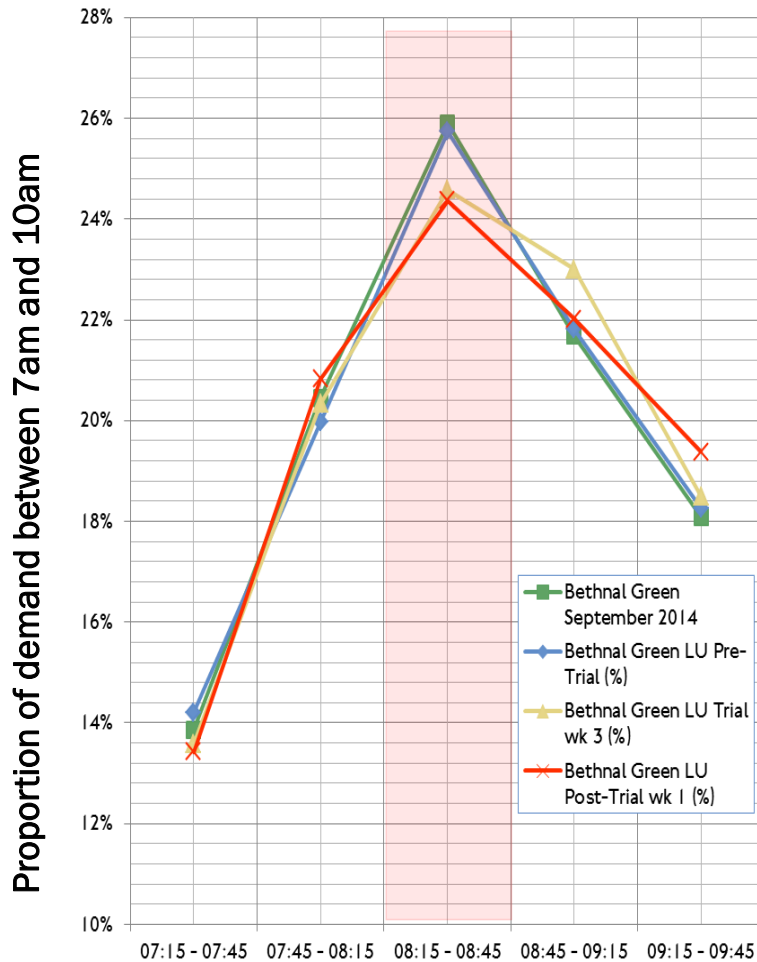
The results suggests change in passenger behaviour:

- Demand distribution over the peak period consistent between pre trial and September periods
- Approximate 5% shift in demand during target time period of 08:15 to 08:45.
- Total Peak Demand over the trial period relatively unchanged

TfL Transport Models

Why do we need transport models?

London 2050

TfL's suite of models

❖ London's population is now at around 8.6 million.

❖ The population is expected to exceed 10 million in the early 2030s, and to reach 11 million by 2050, with growth rising year on year in the interim.

Absolute change in population, 2011 to 2050 (thousands)

- 150 to 200
- 100 to 150
- 50 to 100
- 10 to 50
- 0 to 10

Strategic demand models - LTS

Travel cost →

← Planning data

Land-use transport interaction models - LonLUTI
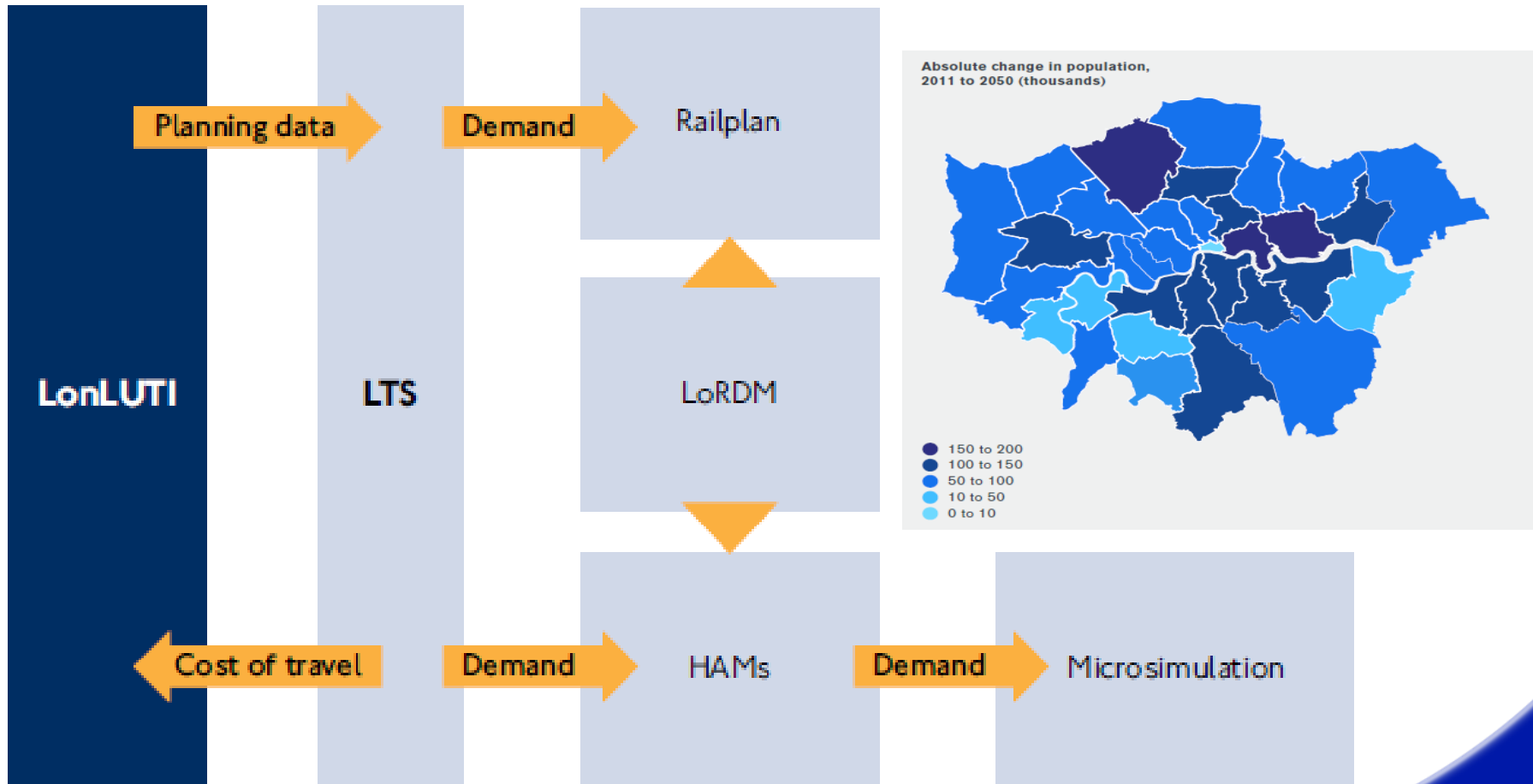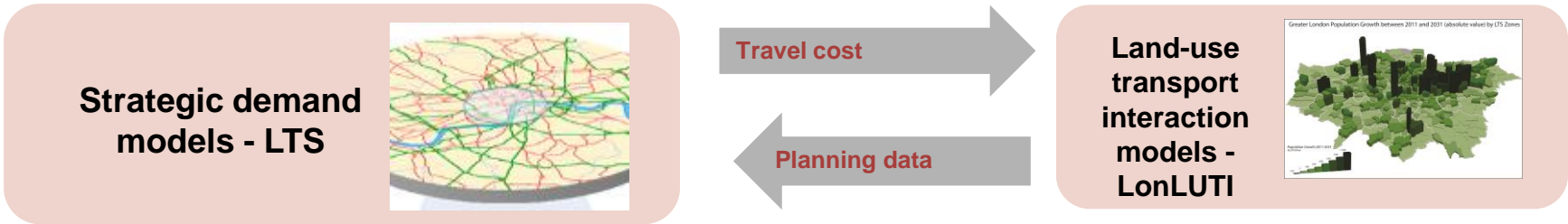
Demand modelling

What is LTS?

- **The London Transportation Studies Model (LTS)** is a strategic multi-modal model for London and its surrounding area.

- **A four-stage aggregate transport model**

- **Used to prepare forecasts of:**
  - growth in total travel
  - changes in travel patterns
  - the mode of transport chosen (car, public transport, walking and cycling) and
  - the routing of trips through the road and public transport networks.

The IM view – LTS

The world moves on

The IM view – LTS

Land-Use Model LonLUTI

Developers

Property market

Firms

Labour market

Residents

Product markets

Transport infra-structure suppliers

Transport market

Transport service suppliers
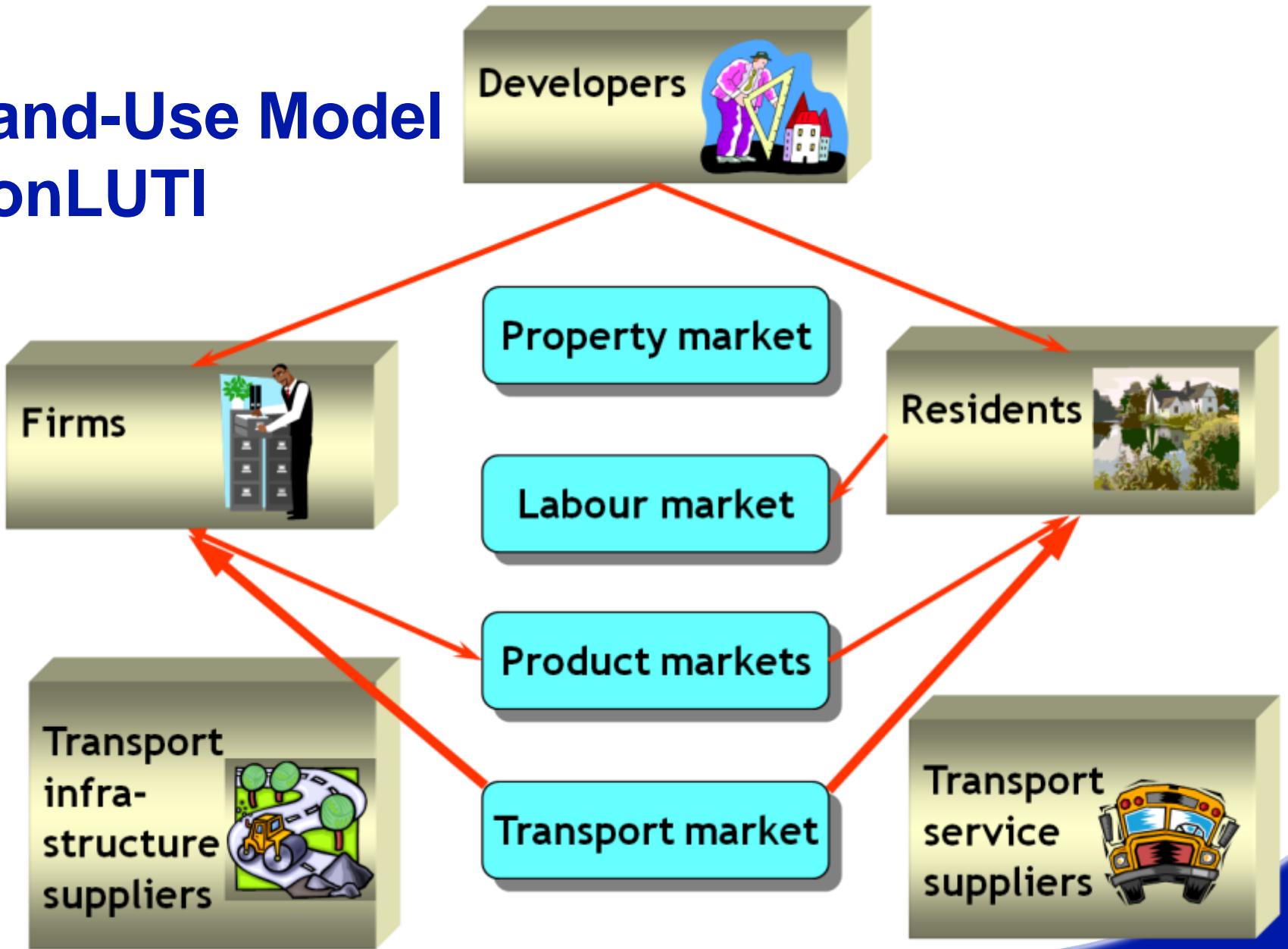
Railplan

LUL and DLR Crowding
RT027AC01 - 2016 AM Reference Case

Legend:
- < 1 standing / sqm
- 1 - 2 standing / sqm
- 2 - 3 standing / sqm
- 3 - 4 standing / sqm
- > 4 standing / sqm
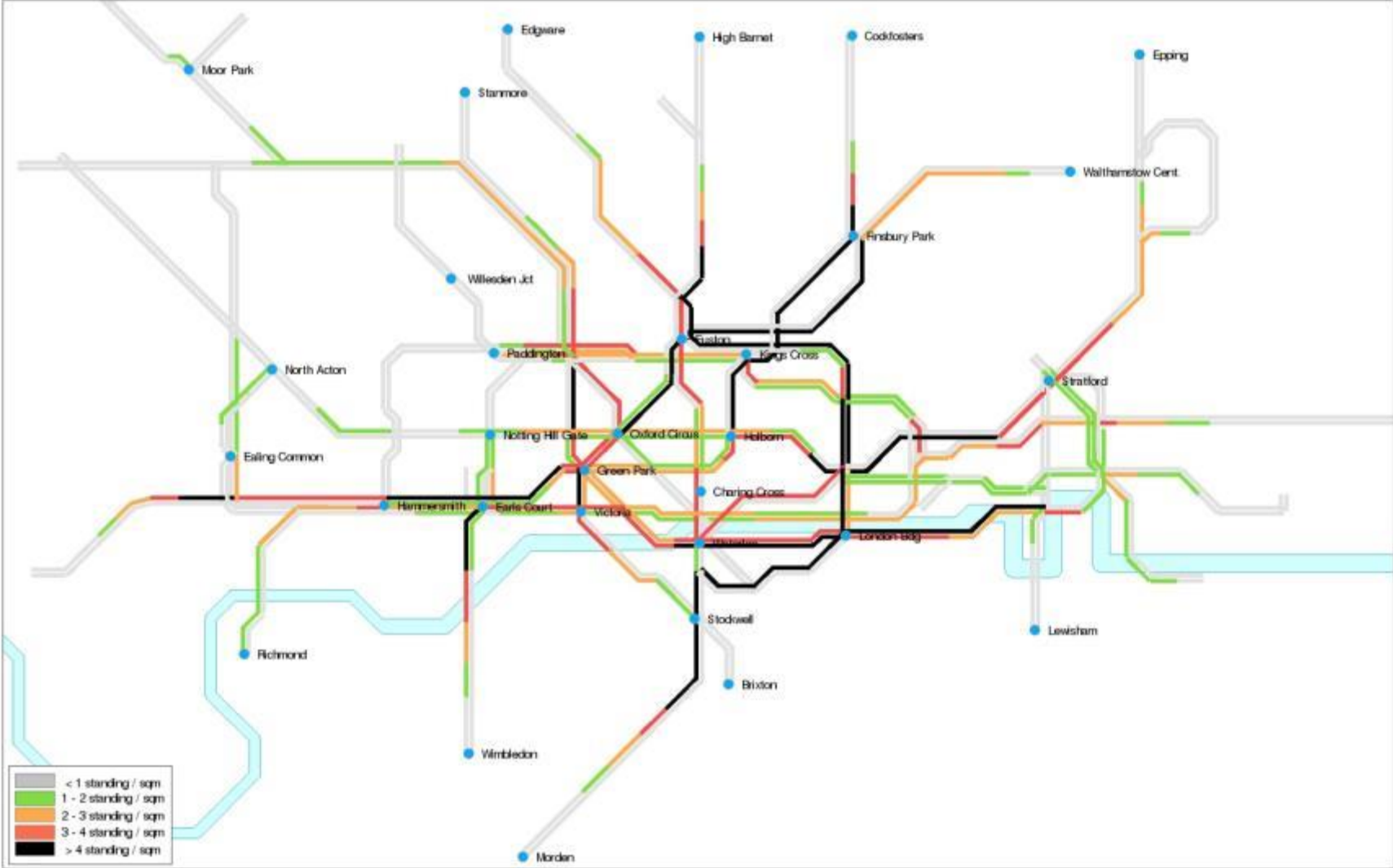
Note:
- Peak hour crowding (54% peak period demand)
- Standing density factor of 7 pax/sqm
- Includes reliability factor

Transport for London

5 sub-regional models - South– Detailed Area

Customer Experience Analytics Research

Customer Experience Analytics Research
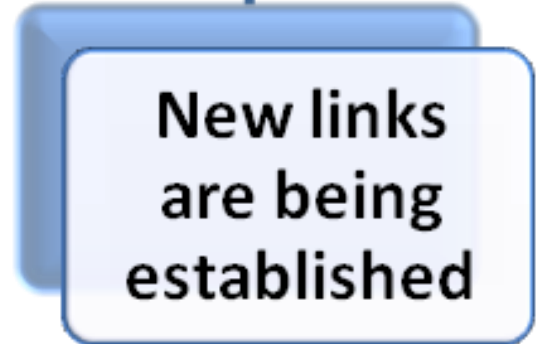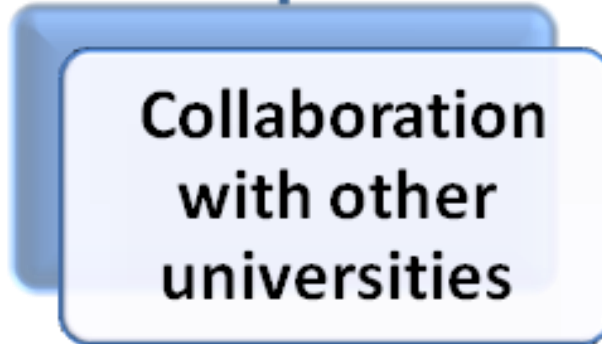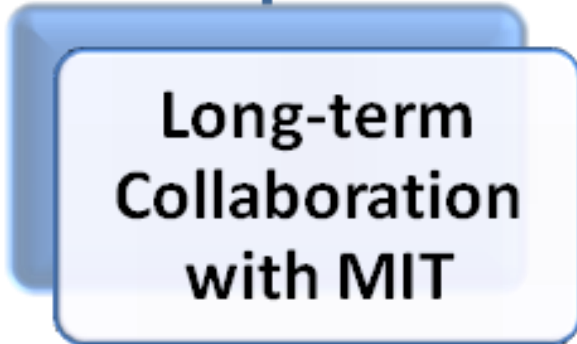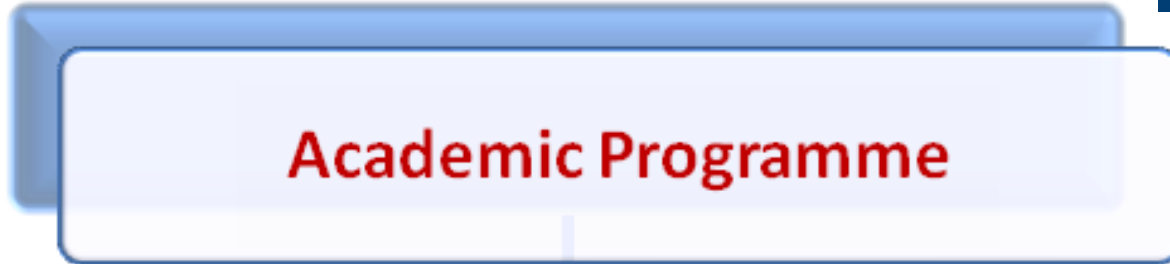
Research Topics
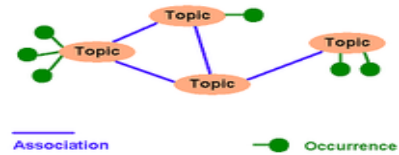
- **Travel Behaviour:** How can we better explain travel behaviour and urban dynamics? How these insights can be used in the improvements to existing transportation models, through both the development of new behavioural models and new modelling approaches?

- **Safety:** Innovations to reduce the number of accidents

- **Network Performance:** Integrating ticketing, bus, traffic congestion, and incident data for better performance of the bus and road networks

- **Information Provision:** Developing further personalised services for those customers who want tailored information

- **Predictive Analytics:** Predicting platform and train congestion at stations

- **Data:** Innovations around data mining tools and geo-spatial visualisations to bring data to life

THANK YOU

OlgaFeldman@tfl.gov.uk