

Parking Insights

Local Authority Data Reporting Pipeline

Martyn Minear, Yousef Majeed

18 May 2022

Project Overview

Project Aim: To use parking data to act as a barometer of economic recovery across England post Covid-19 lockdowns and to continue exploration of UK travel patterns.

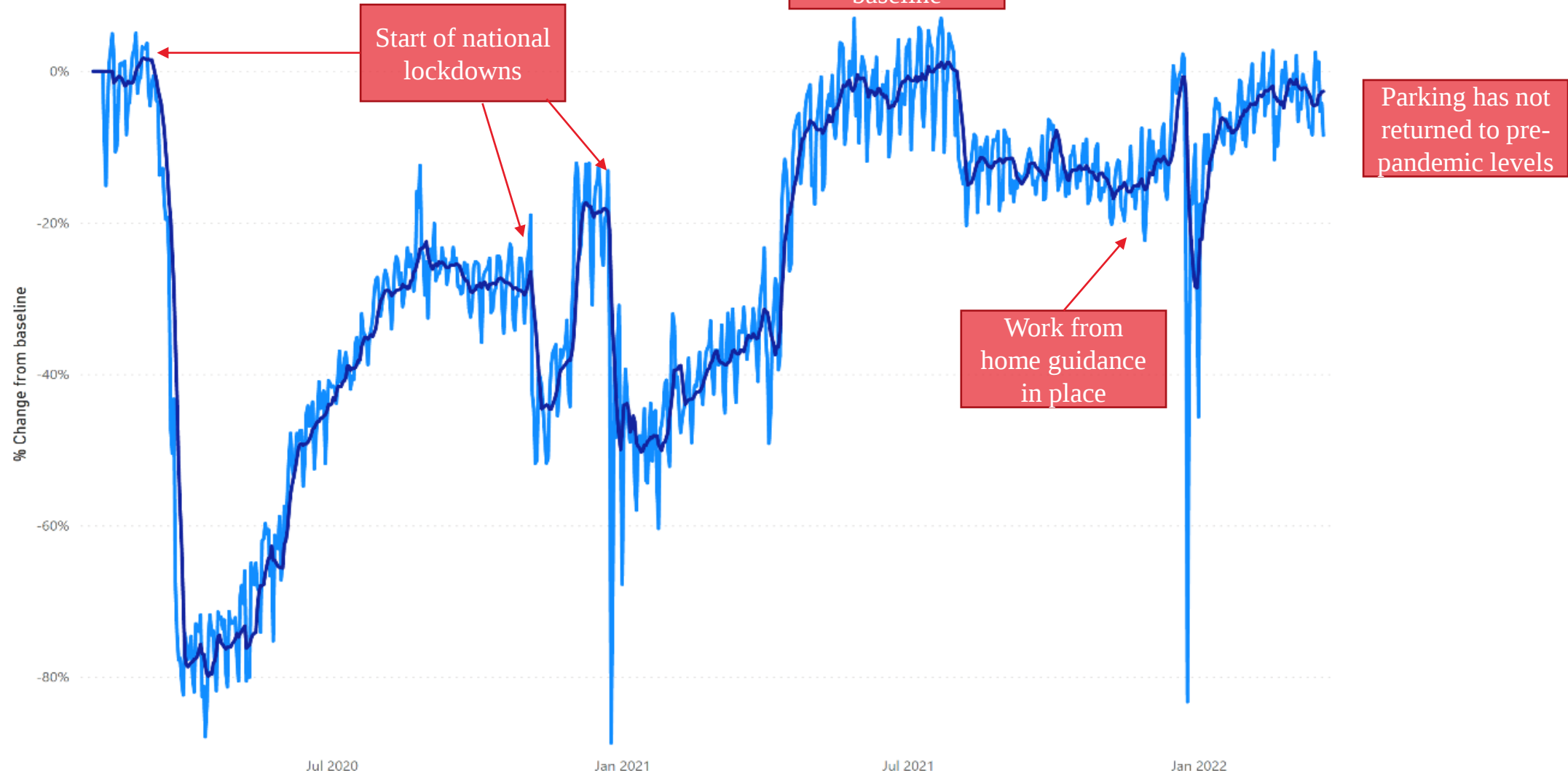
As part of the Transport Technology Forum (TTF), the Department for Transport's (DfT) Smarter Traffic Management Team and Arup have developed a transport data pipeline which consolidates parking data in England since the start of the COVID-19 pandemic as an extension of the Award Winning "DfT TTF Covid-19 Insights Work".

The parking data ingestion pipeline:

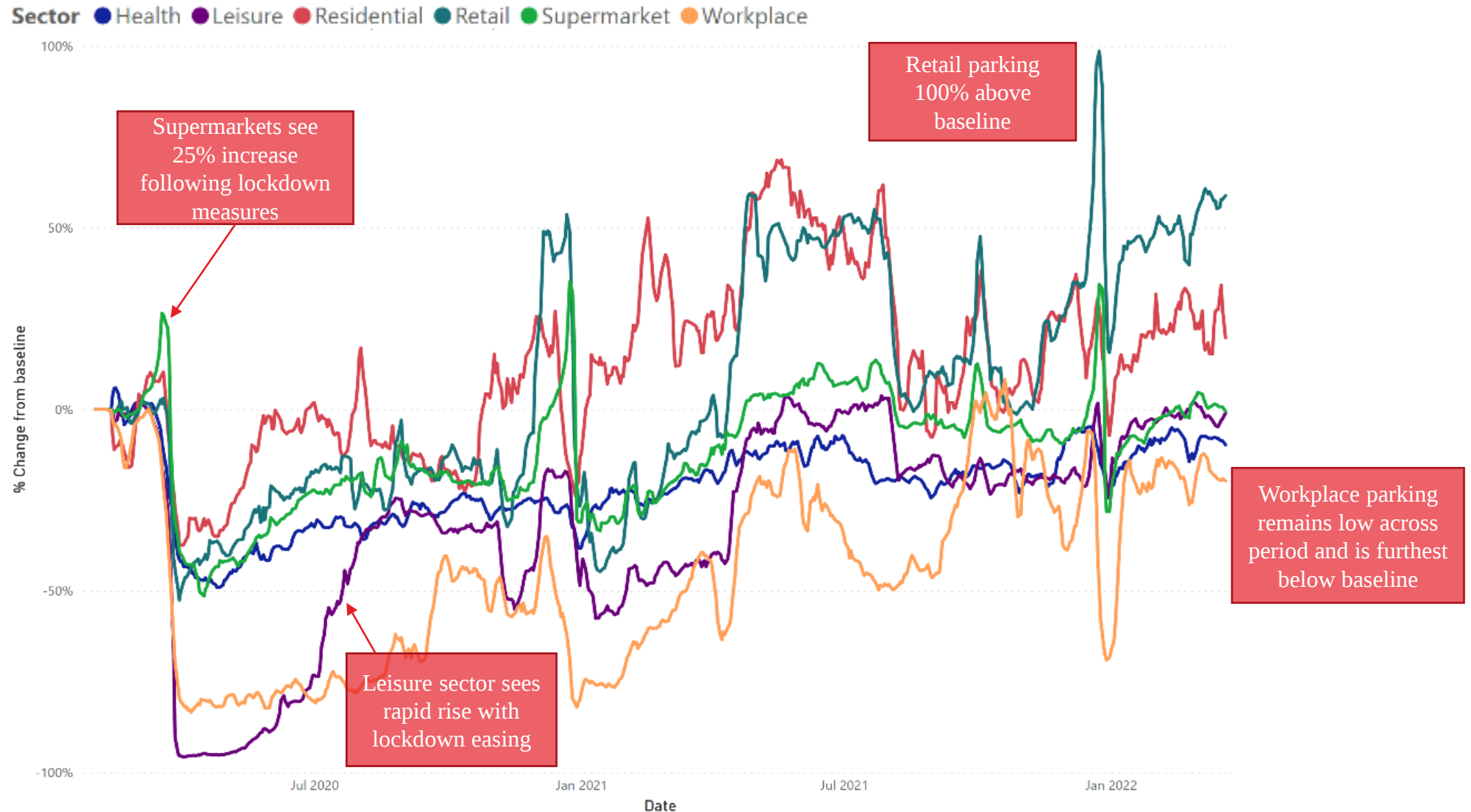
- Different metrics used for analysis (i.e. % occupancy or hourly incoming)
- Comparative analysis across time, regions and sectors nationwide
- Scripts created to support automation and ingestion of data
- Results to be used by HMT and published via TTF weekly digest

National Picture

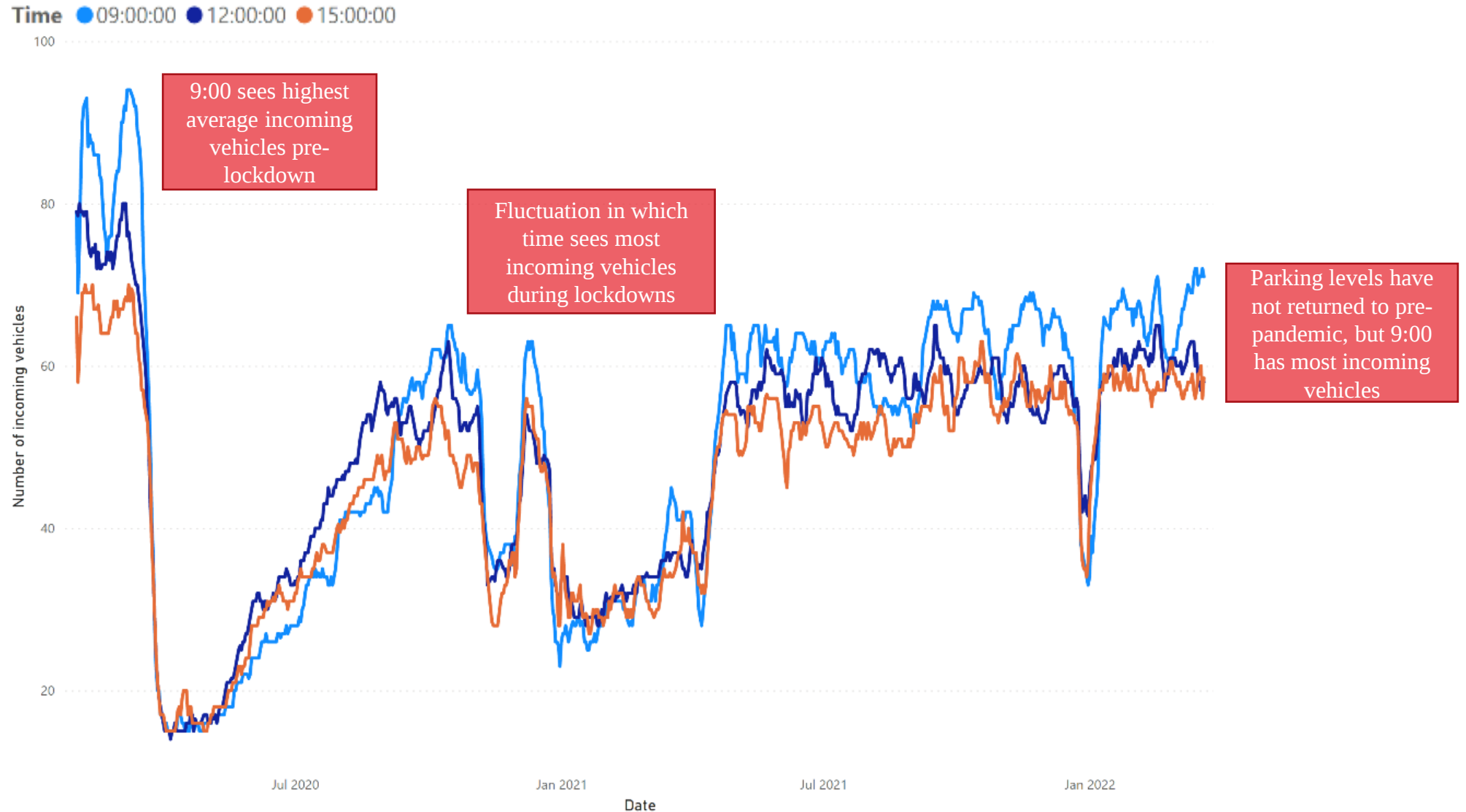
● Median Difference ● Rolling Average



Sector Comparison



Occupied Spaces National Picture



Experience Insights

Data Acquisition

Engagement with the data community can be slow

- Identifying correct contacts to provide data
- Establishment with British Parking Association
- 'Freedom of Information' requests

Goodwill from the transport community is essential for longevity

- Weekly Digests show the outcome of data provider efforts
- Good relationships with nationwide data suppliers
- Project relies on goodwill

Establish the data and analysis purpose

- Data requirements were not established at project start
- Parking data count types adds further complexity
- Pipeline allows more flexibility in data requests and schema

Experience Insights

Data Handling

Data can be provided through many methods

- Live data via API is optimal
- ‘Non-live’ data is at risk of human error
- Algorithms group, process and validate incoming data

Consistent schemas may not exist across datasets

- Multiple types of sensor data combined in a single data feed
- Pipeline handles multiple data metrics
- Create a consistent manner across sources and data types

Consistency of data is vital for analysis

- Data is not always spatially accurate or temporally complete
- Validation checklist to test and remove data from faulty or offline sensors
- Avoid skewing data aggregations

Manual data handling is not sustainable long-term

- Manual interventions required at project start
- Pipeline is a lower cost and longer term service
- Push for automated data provision from local authorities that manually handle data

Experience Insights

Data Coverage

Establishing a representative dataset is relies on engagement with providers

- Coverage only as widespread as authorities that have responded
- Target larger suppliers for greater representation
- Multiple datasets improves coverage and depth

Analysis of representative datasets will contain assumptions

- Normalise each dataset by the number of sensors in each area
- Importance of examining multiple datasets
- Greater confidence in conclusions

Live and historic datasets are required to monitor against baselines

- Availability of data is limited where counts are taken at less regular intervals
- Historic data from February 2020 needed as this is classed as the baseline