# UK Footfall Index

This index shows the percentage change in visitors – or footfall (FF) – to retail environments in the United Kingdom between two different periods of time (for example, consecutive months or days). The FF data were provided by the Local Data Company and covers the period from July 2015 to date; the data refer to the number of people walking in front of, or into, a retail unit.

The FF index captures major seasonal changes, such as the end of the summer or the beginning of a new year, where people tend to spend less time in retail areas, especially when compared with the Christmas period for example. The UK FF index is not designed to isolate and observe local changes, or particular retail area characteristics, but rather it is designed to be representative of what is happening more generally in any given town in the UK.

The FF index is calculated as follows:

Equation 1

where b = Total footfall at period *b*, a = Total footfall at period *a*, *a≠b,* and *n* is an integer representing the distance between b and a, for example, if a=July 2016 and b=July 2017, n=12.

Quantities *b* and *a* depend on the aggregated FF counts at each location. There is a great heterogeneity in spatial-temporal distribution: we began with 9 locations in July 2015, a figure which rose to 791 by the end of January 2018. To accommodate these variations, a system of weights is applied to each location which enables making *b* and *a* statistically comparable to each other. The weighting procedure is explained in Appendix A. Along with the weighted system, the FF index also accounts for possible duplicated counts which may be generated by sensors in close proximity to each other. This is explained in more detail in section 3.3.1.

1. Monthly index

In the early months of the project, the number of cities and locations of the sensors were very limited. We initially only had sensors in 2 urban locations (London and Market Harborough), so the FF index recorded during this period cannot be considered to be representative of the whole country. However, by July 2016 the number of urban locations with sensors had reached 52. As this numerical increase was coupled with an increase in the diversity of locations, we have chosen to begin our analysis of the FF index from this point. Therefore, Figure 1 shows the FF index on a monthly aggregated basis for the period of July 2016 to January 2018, and in it we can clearly observe different trends across the year. For instance, the large drop in FF after the Christmas period is very apparent; conversely, the variance in FF over the summer months (July-August) is quite small (-2%).

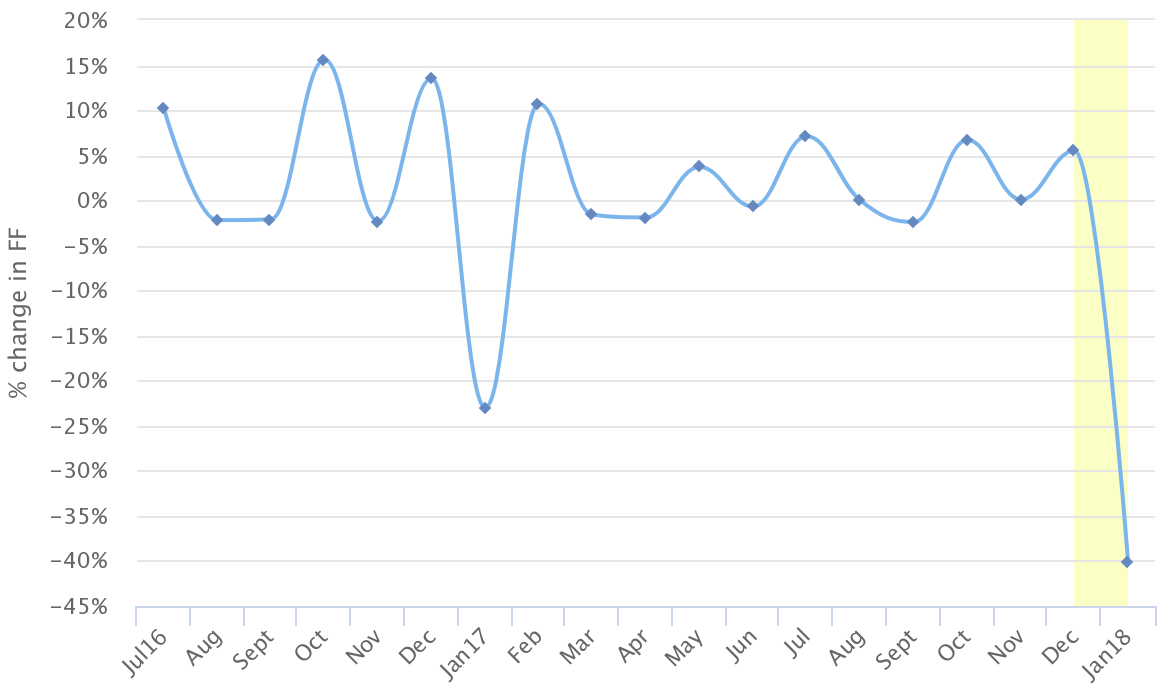


Figure 1. Monthly FF index. We can observe that there is a considerable reduction in footfall in the periods of Dec-Jan 2017 and Dec-Jan 2018 (-23% and -40% respectively), while the remaining months do not demonstrate changes larger than 15.5%.

1. Daily index

Figure 2 shows the FF index on a daily scale, measuring, for example, the difference between any given Sunday-Monday. The index clearly demonstrates the large change in FF between 25th and 26th of December when stores across the UK reopen on Boxing Day. Interestingly enough, the turnover of people on those days was larger in 2016 than in 2017. With this daily scale, we can detect the expected circadian rhythms found in urban areas: during weekdays the change in FF is relatively stable (most of the points in Figure 2 are around 0-25%), whereas the larger positive/negative changes correspond to Saturday-Sunday-Monday (the points oscillate between ±40%).

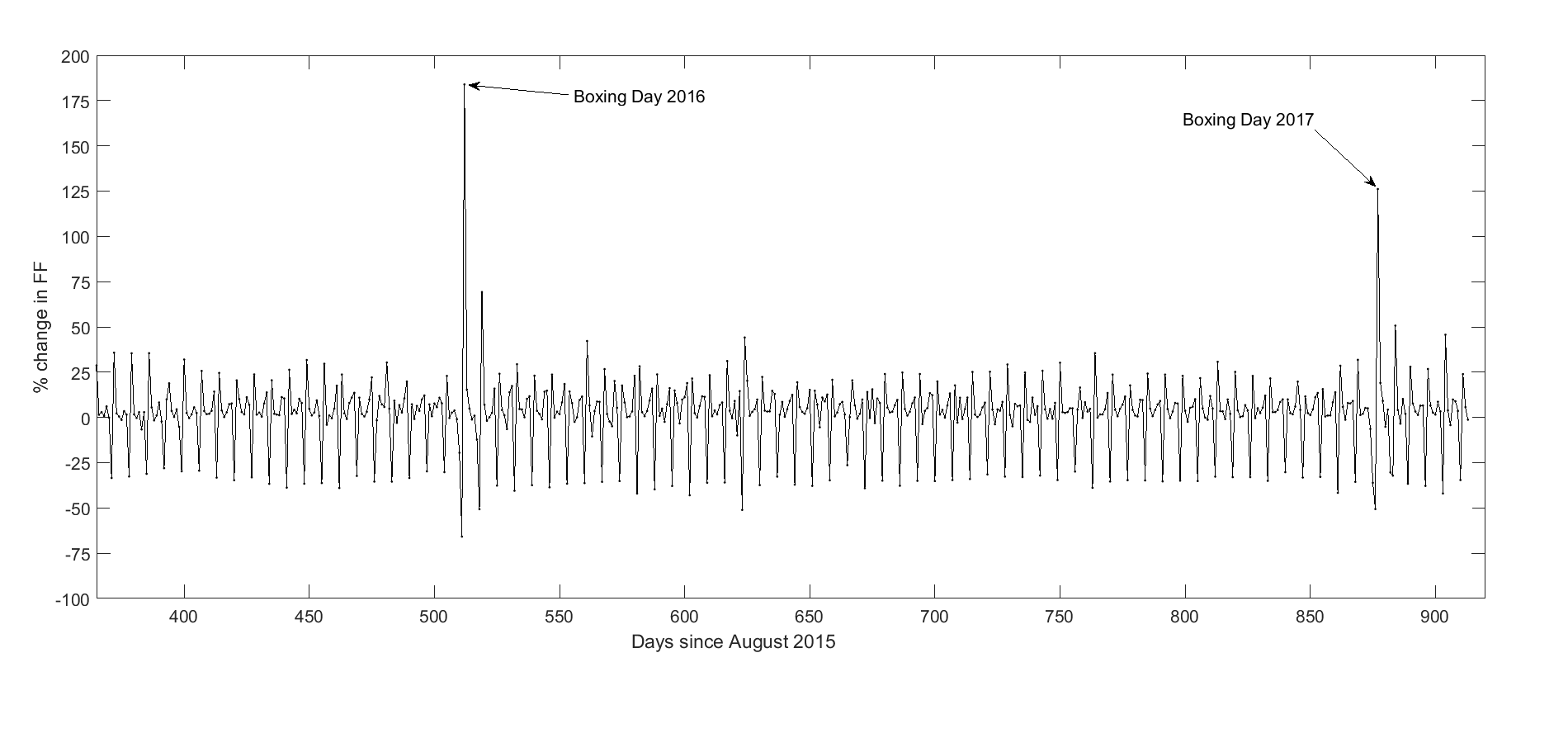
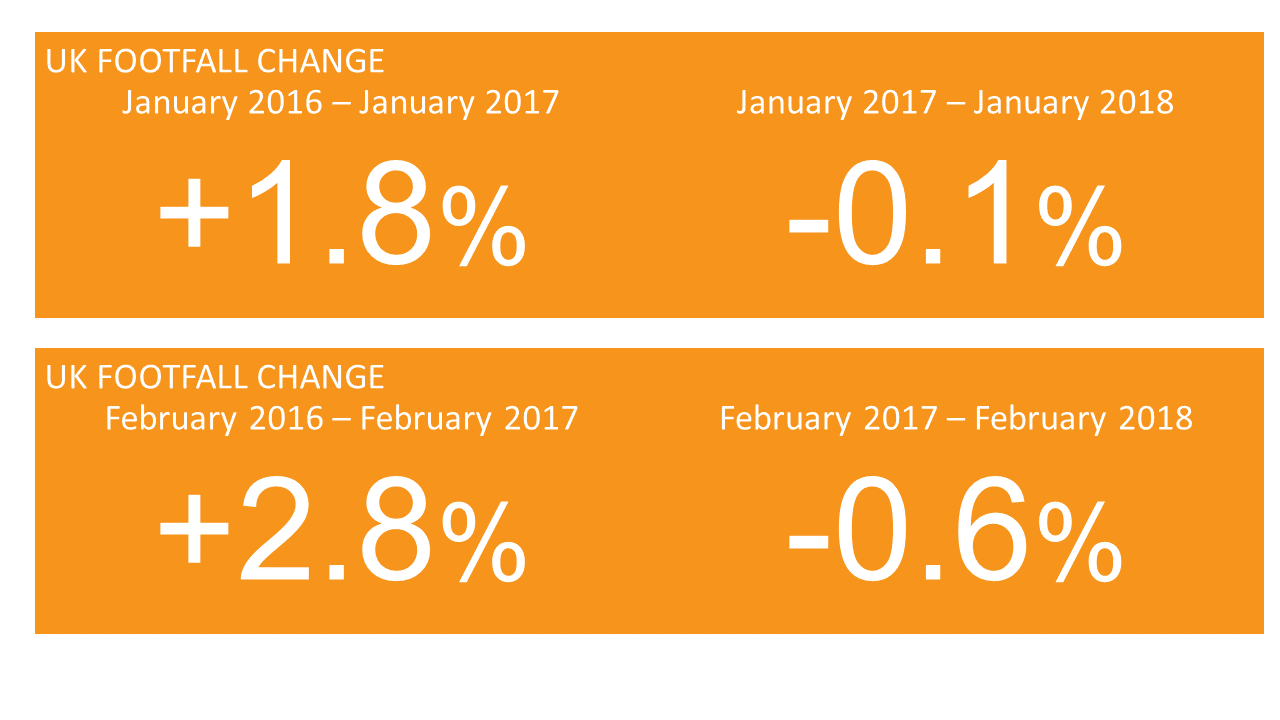


Figure 2. Daily FF index. The evident increase in FF on 26th December 2017/2018 is a confirmation that these data enable us to detect major shifts in FF throughout Great Britain.

For a clearer visualisation, we can also show the results as a single comparative number, or as a smooth version of Figure 2. For example:



In general, the figures show that there is a fall in FF in 2016/2017 when compared with the same periods in 2015/2016.

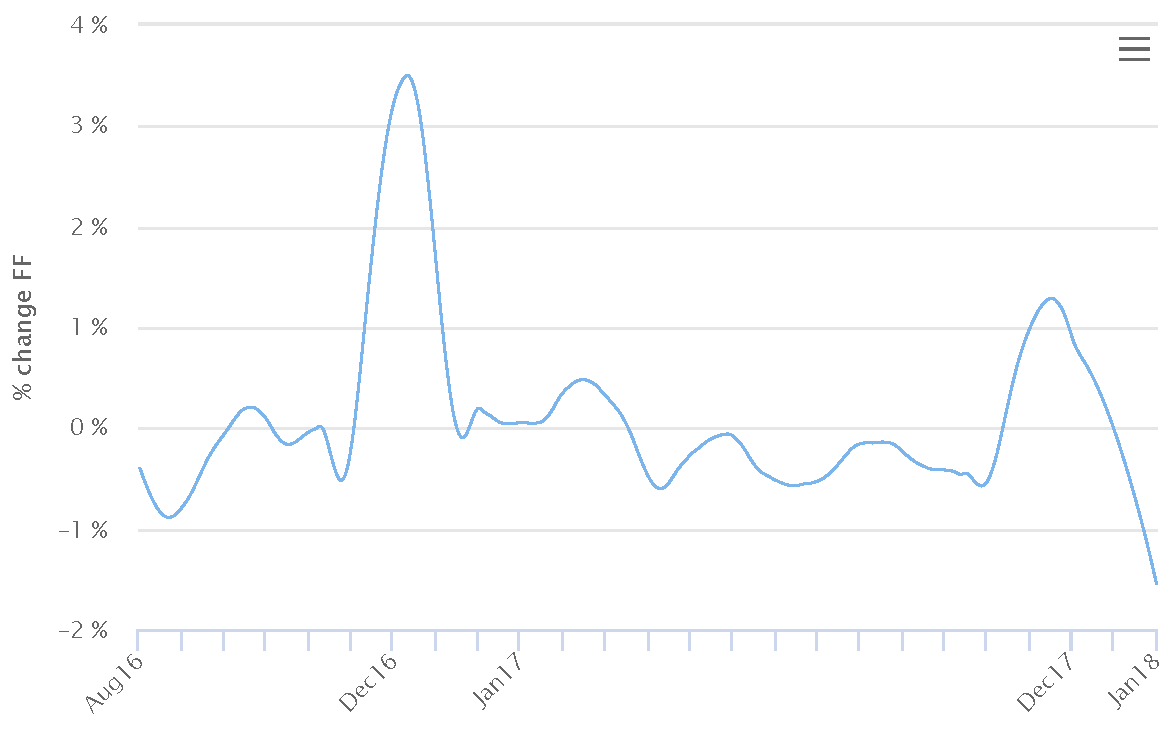


Figure 3. Smooth representation of Figure 2.

With the above smooth representation, we can more clearly see the large variance in FF between December and the rest of the year, as well as the consistency in FF between January–December 2017. This index also captures the decrease in FF in December 2016/2017 when compared with December 2015/2016.

The decrease in FF at the end of 2017 when compared with 2016 can be seen in more detail in Figure 4:

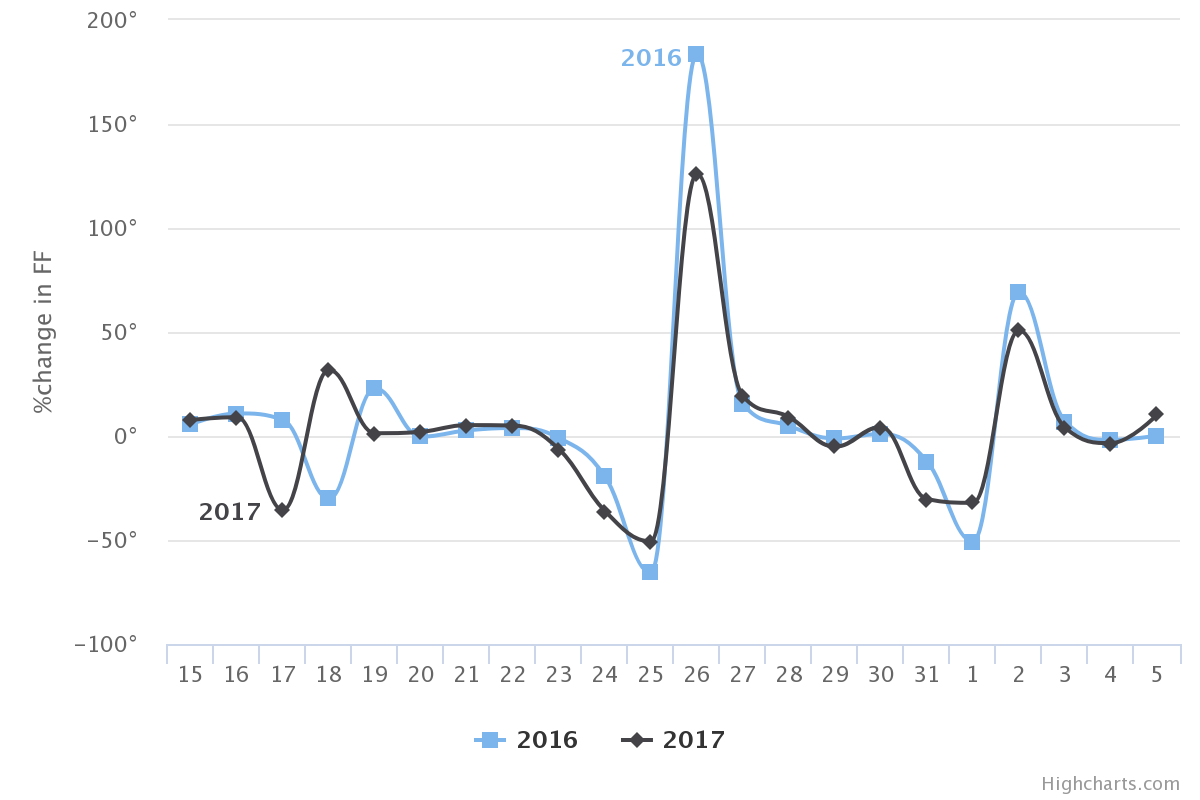


Figure 4. Comparison of the daily change in FF from 15th December 2016 to 5th January for the years 2016 and 2017. The 18th illustrates the difference between a Monday (2017) and a Sunday (2016). The 25th exhibits a similar drop in FF in both years, regardless of the fact that it was a Monday in 2017.

When comparing December 2016 with December 2017, we found a total FF change of -23%.

1. Representativeness

The fundamental question about this index is how representative it is for any retail centre in the UK and whether it contains any biases towards locations, types of retail areas, and other factors? We will address these questions below.

* 1. Locations

Since July 2016, 20 urban locations have accounted for 81% of the total FF in any given month, with London constantly contributing ~27% of that total. However, in our index, the bias towards these urban locations is compensated for by the weighted system explained in Appendix A. For example, a typical distribution of each sensor’s contribution to the total FF is shown in Figure 5. The bulk of the distribution – 80% of locations – is between 0-0.25%, which confirms that the FF index is not capturing the flow of people around any particular location.

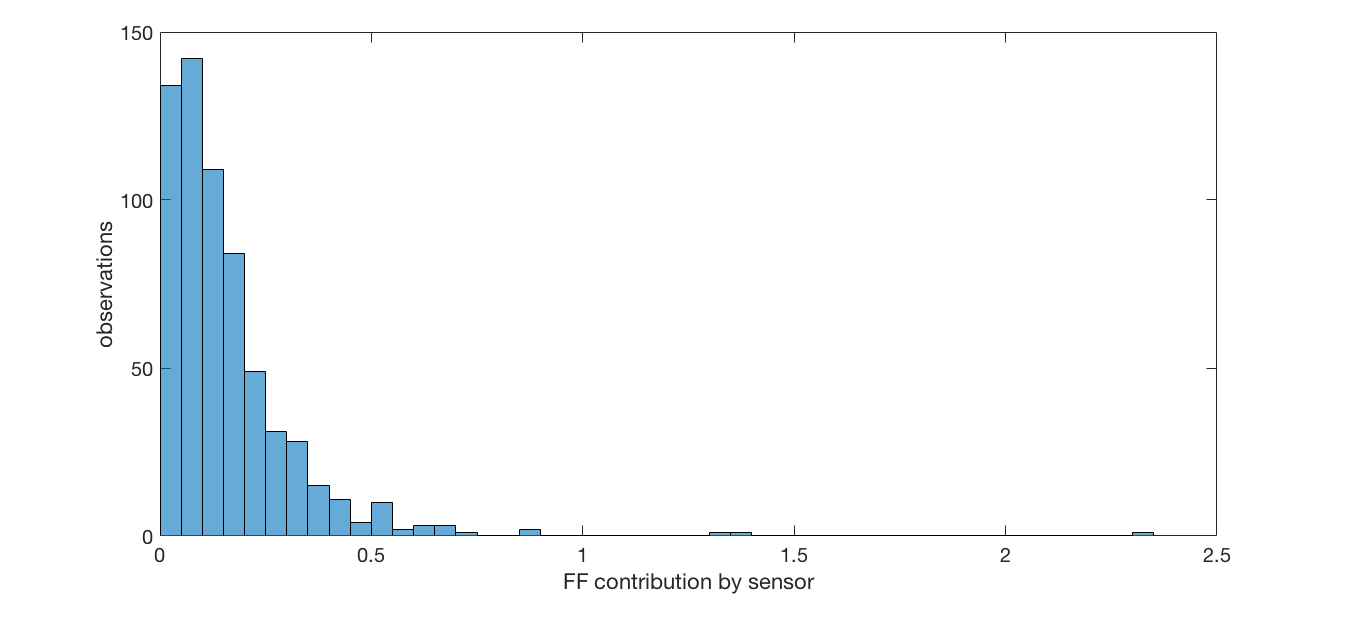


Figure 5. The probability distribution of the contribution to the total FF from each of the 631 locations operating in September 2017. As can be observed, only three sensors are contributing slightly more than 1% of the total.

* 1. Type of street

Once we established that the distribution of sensor locations was not causing undue bias, we then queried their position in respect of the type of street they are positioned on: i.e. are these sensors predominantly located in high streets, retail parks, shopping centres, etc.? Using the full address of the location of the sensors, we classified each location with its street type as derived from the OpenStreetMap highway tag definition.[[1]](#footnote-1) We found that the sensors are located in seven different types of areas: pedestrian/residential (55% of the locations), primary (16%), secondary (5%), tertiary (7%), service (9%), trunk (3%), and unclassified (5%). Consequently, the FF index is dominated by the pedestrian/residential street type (see Figure 6).

That said, not all pedestrian/residential streets are equal. For instance, sensors installed at shopping centres and retail parks are also classified as being located in pedestrian/residential streets, this is because people’s movements in shopping centres occur in a constrained environment with limited or no motor vehicles circulating. From January 2017 to date, 60% of the sensors are in pedestrian/residential locations. Within that figure, 20% of the sensors (approximately 90) are located in various shopping centres, and contribute just 11% of the total FF counted in January 2017. In short, our index is fundamentally measuring the FF on pedestrian/residential streets as opposed to shopping centres. In particular, our FF index is a measure of people walking on suburban and pedestrianised high streets.

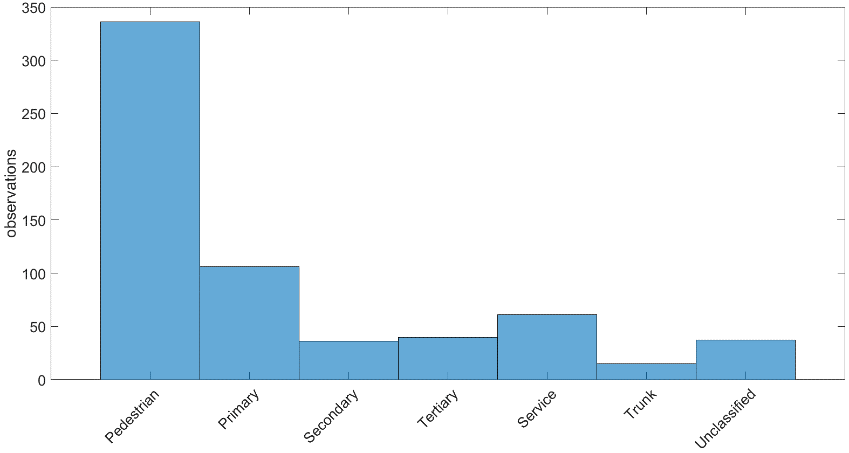


Figure 6. September 2017 probability distribution by type of street. The strong bias towards the pedestrian/residential street type is evident: 55% of sensors are located here, whereas the trunk type of location only contains 1% of sensors.

An interesting finding from this street type analysis is that each type of street has its own characteristic FF signal (FF as a function of time) and this signal does not depend on the type of business. In fact, shops of the same type can generate a whole variety of signals and values. From this, we inferred that the sensors are not capturing the FF generated by any particular business, but instead are capturing the FF produced more broadly by the characteristics of the areas in which the sensors are installed.

* 1. Type of shop

Finally, we explored if our index is biased towards a particular type of business. From July 2017 to January 2018, 105 shop types have been included in the sensor network, although not all shop types are present at each month. In fact, only 45% of shop types are included in the total 19 months, rising to 77% of shop types from March 2017. Consequently, the contribution of each shop type to the overall index varies from month to month, as it depends on the number of shops of that particular type (these counts are weighted as explained above). Before showing the actual FF per type of shop, we first need to address a bias in the counts generated.

* + 1. Nearest Neighbours sensors (NN)

We observed that when two (or more) sensors are in close proximity to one another (40 meters or less) they essentially measure the same local FF. In our current network, there are 264 pairs of sensors (representing 239 unique sensors) in this situation, distributed over 52 different urban locations with London having the greatest number of NN sensors (84). A special case is the Ridings Shopping Centre (Wakefield), where 20 of the 33 sensors installed are in close proximity to one another.

For these 264 pairs, we averaged the hourly counts from each sensor to obtain a single measure (we refer the interested reader to Appendix B for the details about this process). For example, sensors 547 and 276 are located in Bridge St, Chester, 22.3m apart from one another. As can be observed from Figure 7, the daily mean FF is positively correlated and has quite similar values, even though each shop is in a completely different retail category. We therefore assigned a new ID to the averaged signal and classified it as the category ‘Merged’, being sure to also remove the original FF counts for the merged sensors.

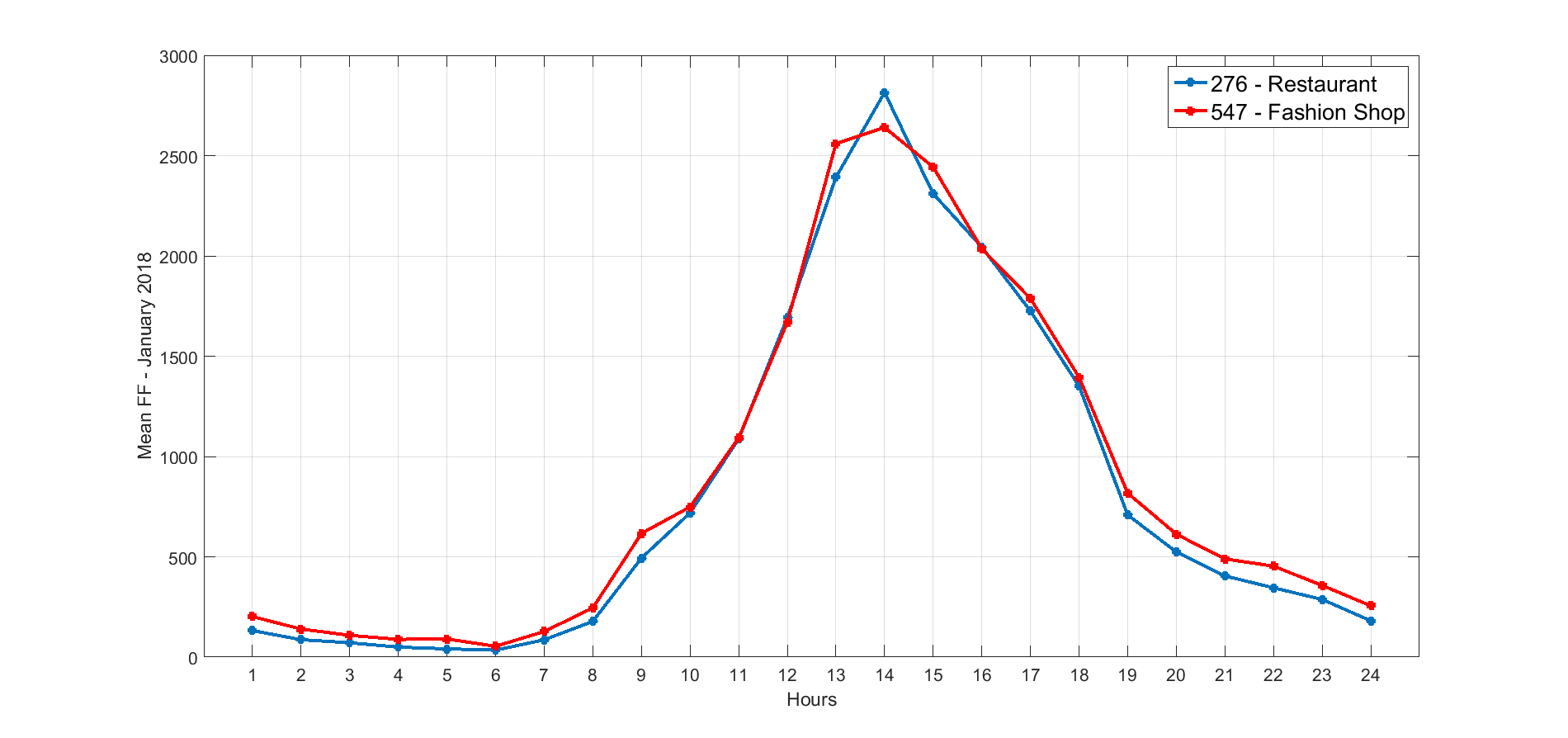


Figure 7. FF signal for a Restaurant and a Fashion shop located on a pedestrian street. The correlation matrix between these two variables has correlation coefficients of 0.99 [ref] reinforcing the evident interaction between these locations.

Figure 8 shows the probability distribution for each type of shop in terms of its particular contribution towards the FF index in each month since July 2016. All months share the same lognormal-like distribution, where the bulk is concentrated in the lower contributors. In general, between 70-90% of the total FF is generated by 95% of the shops, and only 7 types (Merged, Mobile Phones, Restaurant, Charity, Chocolatiers, Fashion, and Sports) account for the 10-30% of the total. The heterogeneity of shop types is an indicator that, as in the case of locations, these sensors are not capturing the FF outside a particular type of shop per se, but rather can be viewed as representative of a vast range of retail businesses.

The prevalence of the Merged category in 12 of the 19 months in Figure 8 provides another indication of how the FF measured is not tied to a particular type of shop, but rather to a cluster of types and/or the particular characteristics of a given retail area. If we break up the Merged category into its original types as in Figure 9, we still have the same type of distribution, but the location of the specific categories differ slightly. For instance, the category of shoe shops appear in the last three months in Figure 9, where originally there was only Merged (Figure 8). Understanding the reasons for this, and the precise evolution of the top contributors from month to month, could help us to more fully understand some of the particular variations of each location over and above the quantity of sensors in the retail location. This analysis is beyond the scope of this present work.

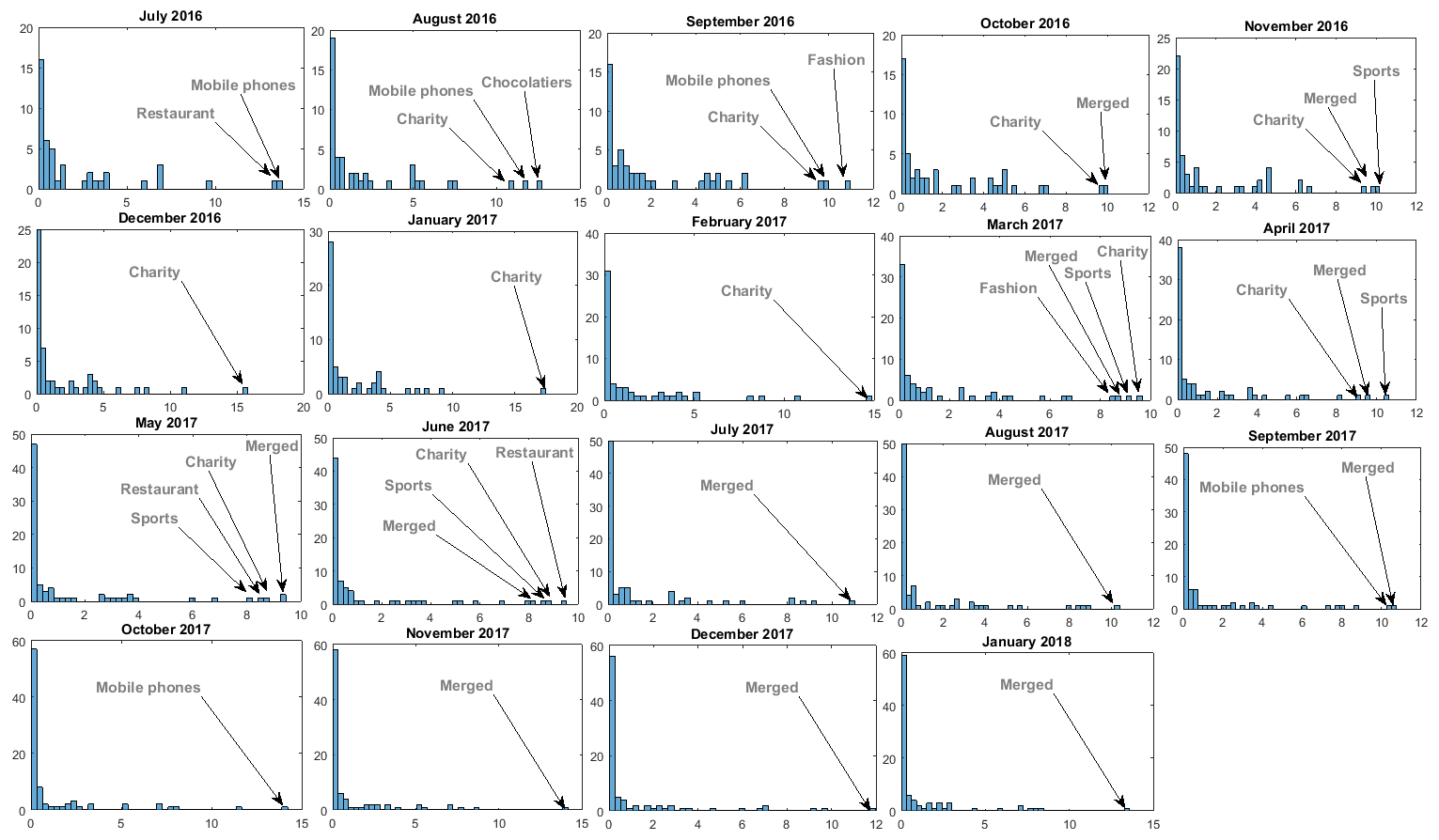


Figure 8. Distribution of FF by type of shop with the ‘Merged’ category.

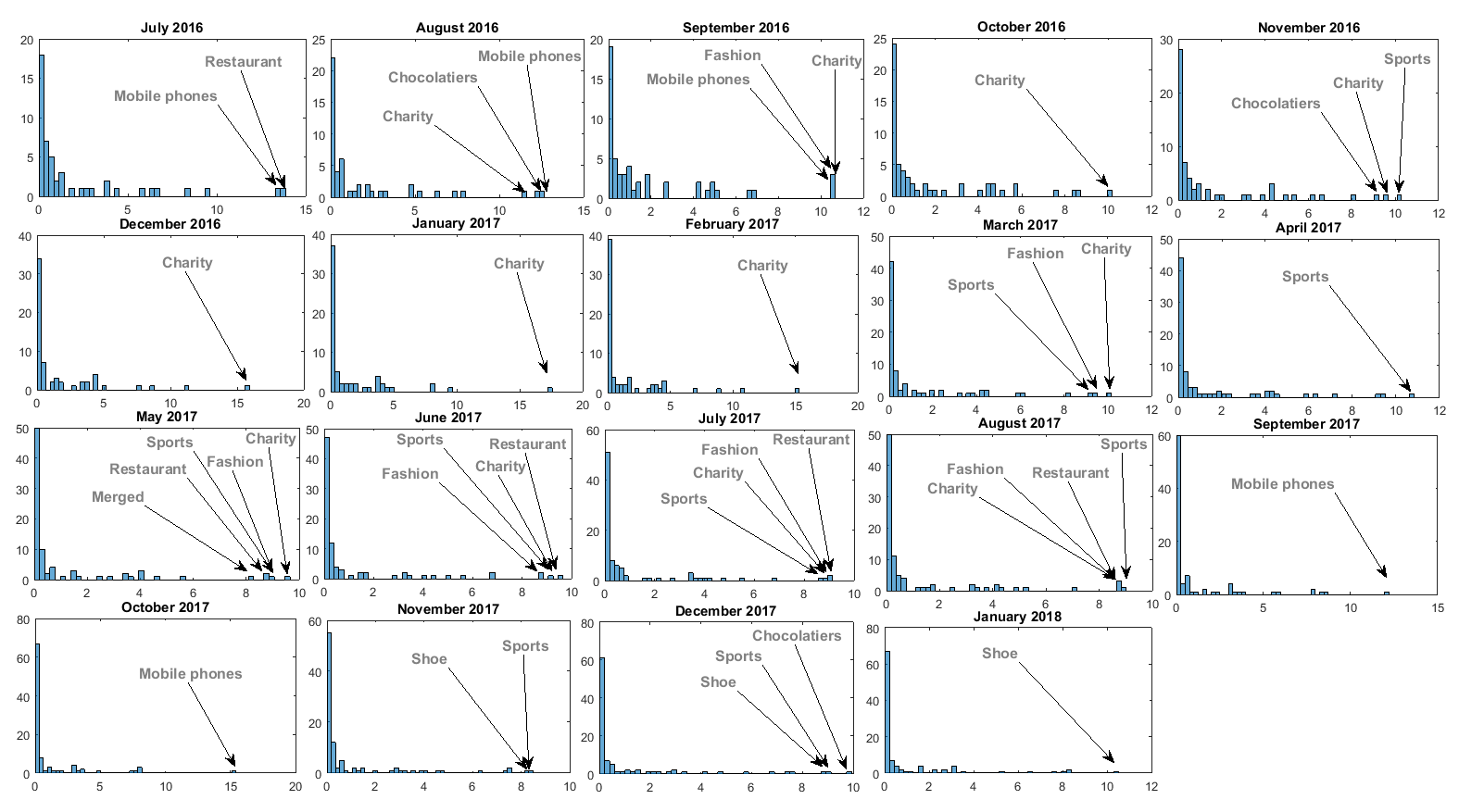


Figure 9. Distribution of FF type of shop without the ‘Merged’ category.

**Appendix A**

The formulation of the FF index (Equation 1) is quite simple: it is a comparison of the aggregated counts between the two different periods. The real challenge is the actual construction of the periods *b* and *a*, as the number of sensors between the two time periods is not necessarily the same in each case. In fact, even when the number of sensors is the same, the number of hours measured by the same sensor on two different days may be different. In principle, this discrepancy makes *b* and *a* incomparable, as an increase in FF between two different periods could be the result of having more sensors in one period than another. In order to adjust for this, the first measure we applied was to weight each period by the mean number of active sensors in that period. For example, if a = 10,000 (100 sensors) and b = 17,000 (130 sensors), the weighted counts will be a = 100, b = 130. An additional problem is that, as mentioned above, the number of active sensors is not constant in any given month. Figure A - 1(a) plots the number of active sensors every five minutes during December 2017. When we isolate a single day from this data (Figure A - 1(b)), we can observe the precise hours in which the number of sensors is constant.

|  |  |
| --- | --- |
| C:\2017\ldc\indicator\oneHour\figures\fiveMinSensors.png  a) | C:\2017\ldc\indicator\oneHour\figures\fiveMinWeek.png  b) |

Figure A - 1. in bon.

The large variation in the number of active sensors in a single day makes our first attempt to weight the FF counts by the mean number of sensors in a given month inaccurate. To fully normalise the counts across different periods, we need to weight the aggregated 1-hour counts by the number of the active sensors at each hour using the following method:

1. Let be the hourly counts at sensor at hour , and let be the number of active sensors at hour during a given period P, so the aggregated counts by hour in the a single period P, AH, are defined as

Equation A - 1

Basically, we are creating an array of 24 points, where each point is the weighted sum of each hour in P for all sensors with counts in that particular hour.

1. Finally, the total weighted FF in P is calculated as:

Equation A - 2

Equation A - 2 is the sum over all the weighted hours present in a single period.

In essence, what we are doing is normalising the counts by the number of sensors using the number of individual hours with data as a proxy for the number of sensors.

**Appendix B**

To detect the NN sensors, we created a distance matrix between all sensors in the network and extracted all pairs that are within a distance of 40m or less, obtaining 264 pairs. We performed a correlation analysis by pairs to investigate if the FF signal from both sensors followed the same pattern or not. To different degrees, all pairs exhibited a positive correlation. It is important to note that this correlation is only in terms of the signal, not in terms of the nominal value at each point in time.

To create an average FF value at each hour for the NN sensors, we proceeded as follows:

1. Case one. Both sensors have FF counts in the same of hours. The merged FF is just the mean FF at each hour from each pair. An example of these type of pairs is shown in Figure C.1, where we can observe how the FF signal of the two sensors is practically the same hour by hour both in shape and value.

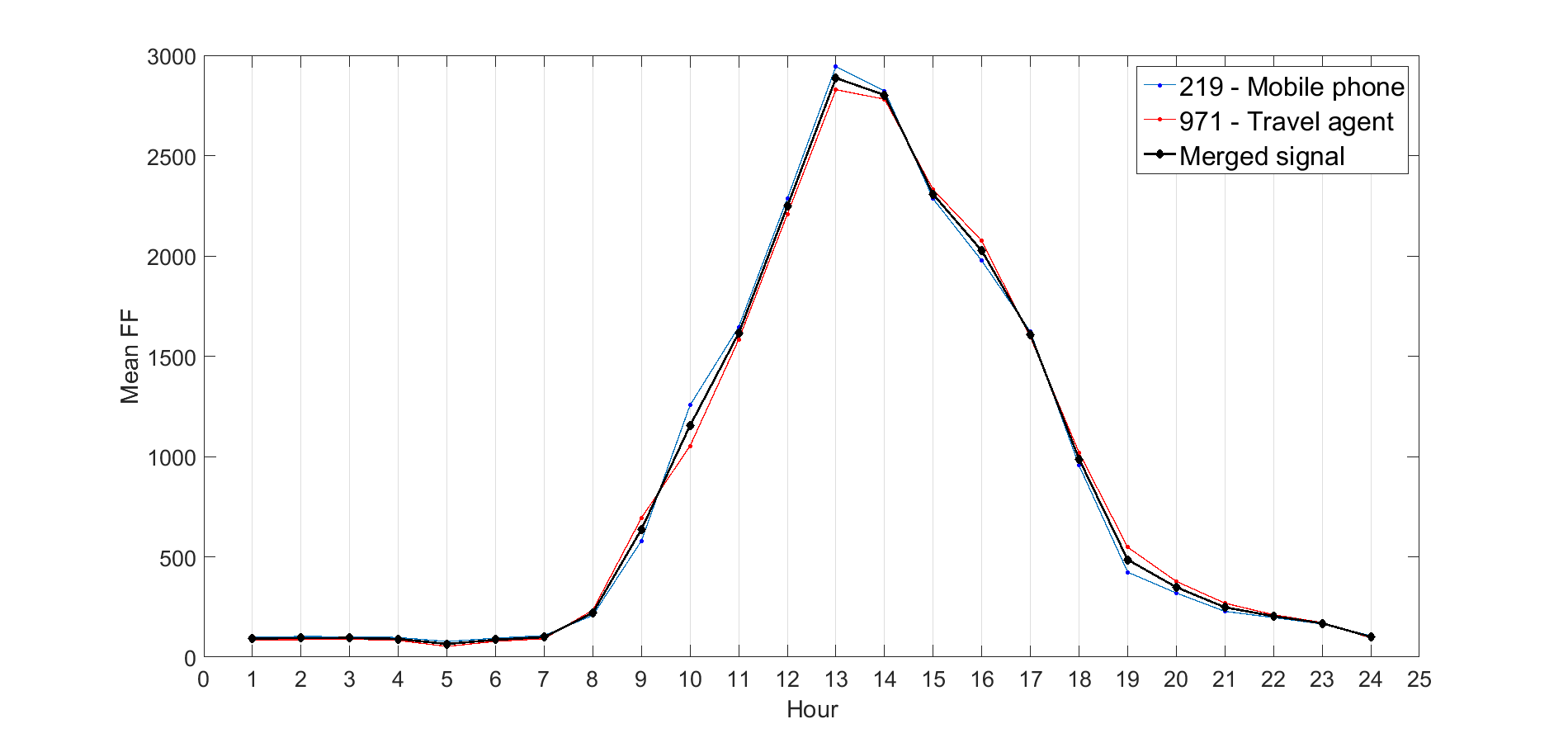


Figure C - 1. FF signals for two sensors (blue and red lines) at King Edward St, Hull, which are 30m apart. This street is a pedestrian one, and the “Bell shape” obtained, is a typical signature of this street type.

1. Case two. Both sensors are missing FF counts in the same hours. In this case, we only averaged between common hours and the rest of the hours were left out of the merging process. For example, sensors 200 and 473, both at the Ridings Shopping Centre (Wakefield) and separated by 20m, only have common counts for the hours of 08.00-19.00, so the merged signal reflects these 11 hours only. In this case, we deleted the FF counts for sensor 200, kept the FF from 1 to 7 for sensor 473, and provided the mean FF for the merged counts with a new ‘Merged’ ID.

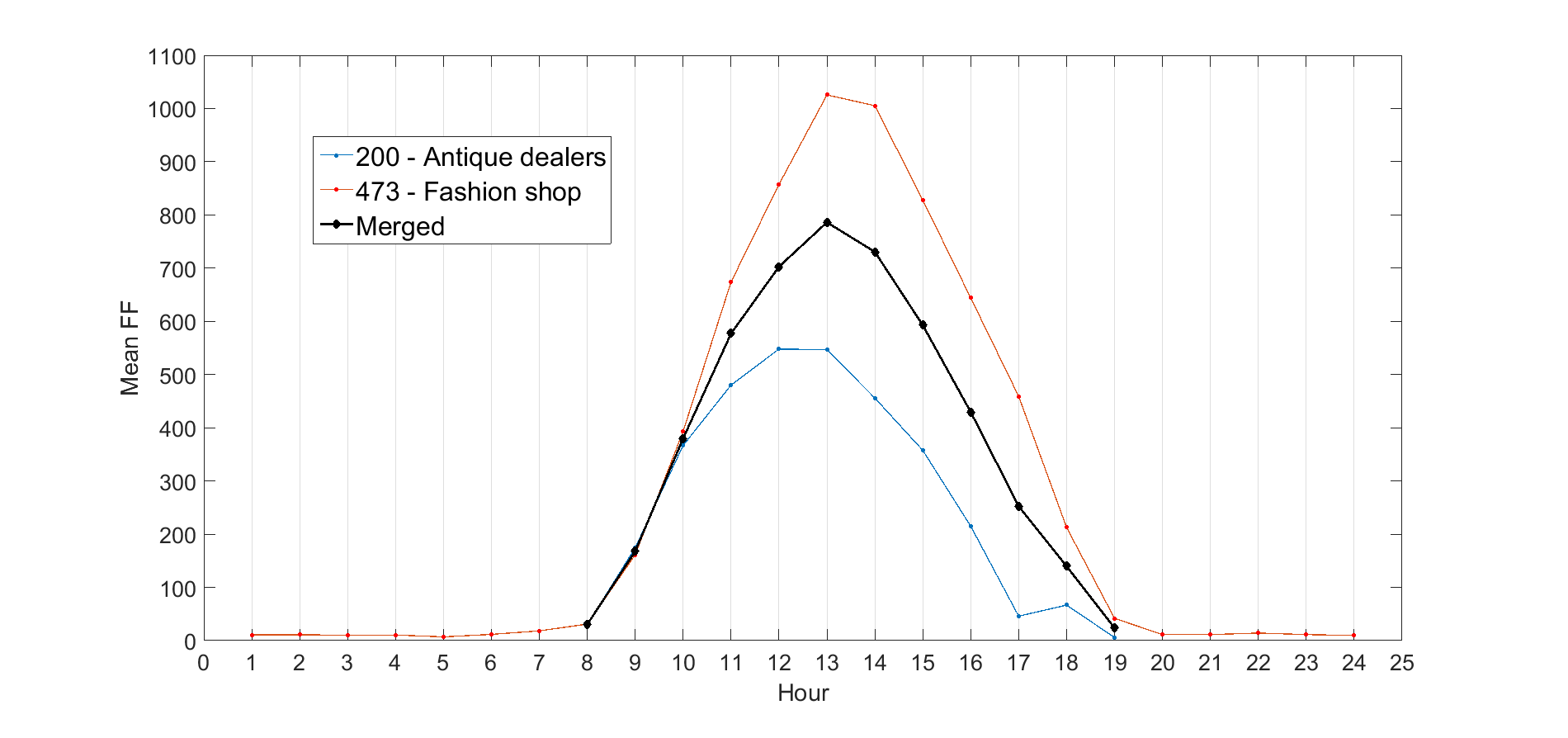


Figure C - 2. FF signals for two sensors (blue and red) and the mean FF (black) between them hour by hour.

Despite these adjustments, there are still some important concerns about the NN sensors which would hold more significance for measures not as broad as the FF index or for comparing FF between retail areas. For instance:

* We need to investigate the wide range of correlation coefficients found amongst these pairs. This could highlight micro differences across locations.
* The actual FF values. Long dwellers effect.
* Instead of calculating the mean, the max/min could be taken as the merged count. In fact, the min could be the actual FF for some locations.
* Currently, we calculate the mean between pairs, but there are also clusters of NN where three or more sensors are very close to each other. This needs to be further considered.

1. https://wiki.openstreetmap.org/wiki/Key:highway. [↑](#footnote-ref-1)