Phase 2 Report

August 17, 2017

1   Detection of Circle ID Coordinates

After the creation of Circle IDs in previous works, now we detect the mid-coordinates for the loading/unloading areas which fall into the same circle, as the representative of *circle id coordinate*.

All coordinates are converted into Cartesian coordinates for each latitude and longitude. For each delivery area carries the same Circle ID:

$$X = \cos(Latitude) \times \cos(Longitude)$$
$$Y = \cos(Latitude) \times \sin(Longitude)$$
$$Z = \sin(Latitude) \tag{1}$$

Then central latitudes and longitudes are computed:

$$Central\ Longitude = \arctan(Average(Y), Average(X))$$
$$Central\ Square\ Root = \sqrt{Average(X) \times Average(X) + Average(Y) \times Average(Y)}$$
$$Central\ Longitude = \arctan(Average(z), Central\_Square\_Root) \tag{2}$$

Data Set: **circlecoordinates.csv**

2   Occupancy of Each Stop

For this analysis, we have divided the days into 30 minutes bins for working hours (From 8am to 8pm). In each time slot, the most crowded/popular loading/unloading areas and associated seccio censals are created. Some parameters are considered to calculate the average occupancy ratio for each delivery area:

- N : total number of days which have check-ins,

- TO: total occupancy in general (total occupancy in day_1 + ... total occupancy in day_N),

- AS: number of slots,

- *OD*: daily average occupancy level (TO/N),

- *OCCL*: average occupancy level per slot

The target delivery areas are the ones with more than or equal 70% average occupancy considering the slot numbers.

Hour slots	Capacity of parking lots	Number of delivery areas with more than 100% occupancy	Number of delivery areas with 70% ≤ occupancy ≤ 100%	The most problematic section	The section neighborhood
08:00 – 08:30	1, 2, 3	32	53	131 – 64	el Camp d'en Grassot i Gràcia Nova
08:30 – 09:00	1, 2, 4	37	70	27 – 50	la Sagrada Família
09:00 – 09:30	1, 2, 3	37	92	30 – 66	la Dreta de l'Eixample
09:30 – 10:00	1, 2, 3	35	113	10 – 27	el Barri Gòtic
10:00 – 10:30	1, 2, 3	36	124	102 – 40	Sant Gervasi - la Bonanova
10:30 – 11:00	1, 2, 3	35	120	33 – 77	la Dreta de l'Eixample
11:00 – 11:30	1, 2, 3, 4	38	119	140 – 34	el Guinardó
11:30 – 12:00	1, 2, 3	36	111	33 – 77	la Dreta de l'Eixample
12:00 – 12:30	1, 2, 3	36	115	33 – 77	la Dreta de l'Eixample
12:30 – 13:00	1, 2, 3	34	110	33 – 77	la Dreta de l'Eixample
13:00 – 13:30	1, 2, 3	34	106	33 – 77	la Dreta de l'Eixample
13:30 – 14:00	1, 2, 3, 4	35	86	33 – 77	la Dreta de l'Eixample
14:00 – 14:30	1, 2, 3, 4	37	103	33 – 77	la Dreta de l'Eixample
14:30 – 15:00	1, 2, 3	36	121	33 – 77	la Dreta de l'Eixample
15:00 – 15:30	1, 2, 3, 4	34	117	33 – 77	la Dreta de l'Eixample
15:30 – 16:00	1, 2, 3	34	103	33 – 77	la Dreta de l'Eixample
16:00 – 16:30	1, 2, 3	34	102	33 – 77	la Dreta de l'Eixample
16:30 – 17:00	1, 2, 3	33	92	33 – 77	la Dreta de l'Eixample
17:00 – 17:30	1, 2, 3	34	93	33 – 77	la Dreta de l'Eixample
17:30 – 18:00	1, 2, 3, 4	35	85	33 – 77	la Dreta de l'Eixample
18:00 – 18:30	1, 2, 3	34	77	208 – 42	el Parc i la Llacuna del Poblenou
18:30 – 19:00	1, 2, 3, 4	35	87	109 – 71	Sant Gervasi – Galvany
19:00 – 19:30	1, 2, 3	36	78	114 – 1	Vallcarca i els Penitents
19:30 – 20:00	1, 2, 3, 4	41	65	30 – 66	la Dreta de l'Eixample

For each hour slot, the delivery areas with more than or equal to 70% average occupancy (considering the number of slots) are the ones with low capacity. These delivery areas have only 1, 2 or 3 (a few of them, 4) parking slots.

Number of delivery areas with more than 100% represents the number of delivery areas with a high circulation by shorter visits, and these areas are highly used.

Number of the delivery areas with more than or equal to 70% and less than or equal to 100% represents the number of delivery areas which are popular by drivers.

The most problematic sections for the delivery areas (less than or equal to 100% or greater than or equal to 70%) are found by checking the frequency of the delivery areas. **To interpret**,

In hour slot 08:00 - 08:30,

- The delivery areas have 1, 2 or 3 parking slot capacities.

- 32 delivery areas are used by more than 100% capacity and it points out the fact that there is a high circulation.

- 53 delivery areas are used by more than or equal to 70% and less than or equal to 100% capacity and it points out the popular delivery areas.

- *Section 131-64* is the most problematic section which has 7 delivery areas in popular ones (53).

In the following table, we have the occupation percentage of the delivery areas which have the greatest number of parking slots in Barcelona.

Delivery Areas	Capacity	08:00 – 08:30	08:30 – 09:00	09:00 – 09:30	09:30 – 10:00	10:00 – 10:30	10:30 – 11:00	11:00 – 11:30	11:30 – 12:00	12:00 – 12:30	12:30 – 13:00	13:00 – 13:30	13:30 – 14:00
9478	19	12%	15%	18%	20%	23%	22%	20%	20%	18%	16%	15%	18%
9028	15	25%	30%	31%	35%	34%	33%	29%	27%	26%	23%	23%	21%
8743	15	19%	15%	17%	17%	18%	19%	18%	17%	16%	17%	17%	15%
10014	15	12%	14%	15%	15%	15%	16%	15%	17%	16%	15%	14%	13%

Delivery Areas	Capacity	14:00 – 14:30	14:30 – 15:00	15:00 – 15:30	15:30 – 16:00	16:00 – 16:30	16:30 – 17:00	17:00 – 17:30	17:30 – 18:00	18:00 – 18:30	18:30 – 19:00	19:00 – 19:30	19:30 – 20:00
9478	19	12%	11%	12%	13%	14%	13%	14%	17%	13%	12%	11%	9%
9028	15	18%	18%	17%	16%	16%	17%	17%	15%	13%	13%	12%	11%
8743	15	14%	14%	15%	14%	14%	16%	16%	14%	12%	11%	11%	9%
10014	15	13%	13%	13%	14%	13%	12%	11%	11%	11%	10%	10%	9%

It can be concluded that where we have availability problem at the delivery areas with 1, 2, 3 and 4 parking lots, the others with high number of slots have low occupancy levels.

In the file *occupancydene.xlsx*, all of the tramos, their occupancy levels, related seccio censals, and the activity types information are merged. Using a basic filtering process will be easy for PowerBI plotting. Since it is a big file, and there is no way to show everything here in the report, we would like to point out the different cases. We compare the general activity type distribution in the whole data with the activity type distribution of tramos which have more than the average value.

Activity Type Distribution in General

Activity Type 1 ● Activity Type 2 ● Activity Type 3 ● Activity Type 4
Activity Type 5 ● Activity Type 6
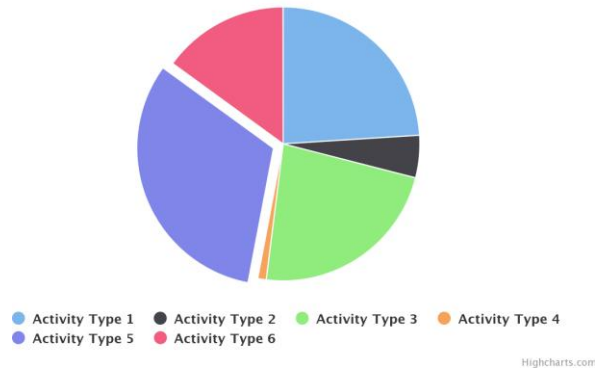
We checked the activity type distribution for each tramo in each hour slot. For each related activity type, we took the number of tramos who have more than average percentage.

	08-09	09-10	10-11	11-12	12-13	13-14	14-15	15-16	16-17	17-18	18-19	19-20
Activity Type 1	1035	936	808	722	687	859	1005	1110	1037	961	865	856
Activity Type 2	574	725	799	847	811	740	722	942	994	909	863	745
Activity Type 3	701	882	7871	726	792	839	986	1441	1357	1192	1045	846
Activity Type 4	715	940	1124	1087	1054	1057	1117	1163	1262	1192	1125	913
Activity Type 5	807	857	989	983	974	881	700	421	457	491	501	554
Activity Type 6	886	790	898	1025	1004	850	723	701	776	885	1032	1131
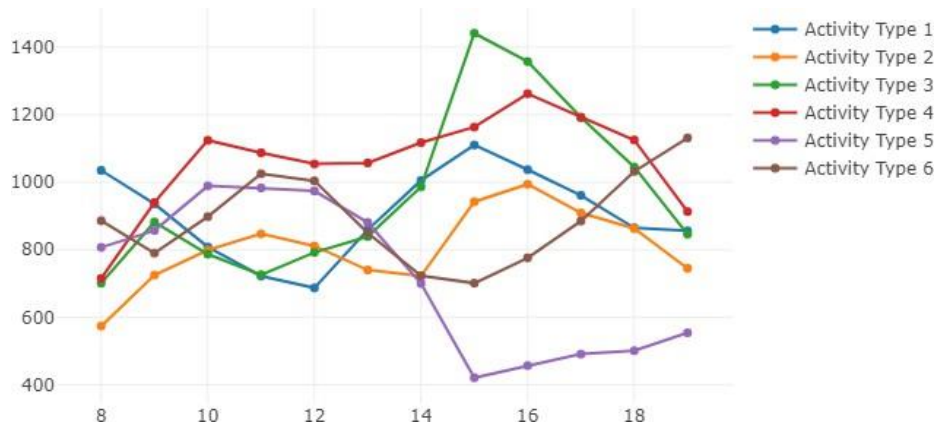
Figure 1: Unexpected High Behavior of Activity Types for Hour Slots

4

- From 08:00 to 09:00, Activity Type 1 has the greatest number of tramos which have more than 24% for the distribution.

- From 09:00 to 10:00, Activity Type 1 and 4 have the greatest number of tramos which have more than 24% and 1% (respectively) for the distribution.

- From 10:00 to 15:00, Activity Type 4 has the greatest number of tramos which have more than 1% for the distribution.

- From 15:00 to 17:00, Activity Type 3 has the greatest number of tramos which have more than 23% for the distribution.

- From 17:00 to 18:00, Activity Type 3 and 4 have the greatest number of tramos which have more than 23% and 1% (respectively) for the distribution.

- From 18:00 to 19:00, Activity Type 4 has the greatest number of tramos which have more than 1% for the distribution.

- From 19:00 to 20:00, Activity Type 6 has the greatest number of tramos which have more than 24% for the distribution.

Figure 1, shows that Activity Type 5 is the one that is close to the general type distribution, although it has the largest amount of check-ins in the data set.

For Seccio Censal, "33-77" (La Dreta de l'Eixample) is the most problematic seccio by the occupancy level for each hour slot (even in each 30 minutes slot).

3   Repeatability of Each Stop

The file is *user percentage seccio.csv*.

Each seccio censal is analyzed in order to observe their repeatability for the users. The following table is for the seccio censals which have repetition with more than 15% of the users. The rest can be found in the .csv file.

SECCIO CENSAL CODE	SECCIO CENSAL NAME	PERCENTAGE OF USERS WHO REPEATED
33-77	LA DRETA DE L'EIXAMPLE	23.6211 %
35-81	LA DRETA DE L'EIXAMPLE	19.6429 %
34-79	LA DRETA DE L'EIXAMPLE	19.4365 %
36-83	LA DRETA DE L'EIXAMPLE	16.9884 %
34-78	LA DRETA DE L'EIXAMPLE	16.0232 %
32-75	LA DRETA DE L'EIXAMPLE	15.8404 %
38-90	L'ANTIGA ESQUERRA DE L'EIXAMPLE	15.4583 %

For Activity Types:

SECCIO CENSAL CODE	SECCIO CENSAL NAME	ACTIVITY TYPE 1
121-78	LA VILA DE GRACIA	37.68844 %
127-51	LA VILA DE GRACIA	38.76712 %
112-89	EL PUTXET I EL FARRO	39.5543 %
70-70	SANTS-BADAL	42.15686 %
192-47	SANT ANDREU	44.87952 %
129-56	LA VILA DE GRACIA	48.20847 %

SECCIO CENSAL CODE	SECCIO CENSAL NAME	ACTIVITY TYPE 2
189-32	SANT ANDREU	13.74172 %
192-46	SANT ANDREU	13.88255 %
69-69	LA BORDETA	16.74528 %
193-59	LA SAGRERA	18.1818182 %
20-4	EL FORT PIENC	19.48399 %
194-61	LA SAGRERA	24.46809 %

SECCIO CENSAL CODE	SECCIO CENSAL NAME	ACTIVITY TYPE 3
133-79	EL CAMP D'EN GRASSOT I GRACIA NOVA	44.94311 %
193-56	LA SAGRERA	45.45455 %
193-59	LA SAGRERA	50 %
193-57	LA SAGRERA	51.16279 %
188-29	SANT ANDREU	51.86404 %
196-72	LA SAGRERA	52.94118 %

SECCIO CENSAL CODE	SECCIO CENSAL NAME	ACTIVITY TYPE 4
200-90	NAVAS	5.065295 %
76-109	SANTS	5.279503 %
76-113	SANTS	5.364152 %
190-34	SANT ANDREU	7.105767 %
190-36	SANT ANDREU	8.092369 %
193-56	LA SAGRERA	18.181818 %

SECCIO CENSAL CODE	SECCIO CENSAL NAME	ACTIVITY TYPE 5
15-41	LA BARCELONETA	47.23618 %
195-63	LA SAGRERA	48.21429 %
131-69	EL CAMP D'EN GRASSOT I GRACIA NOVA	49.33568 %
27-54	LA SAGRADA FAMILIA	52.52698 %
17-47	SANT PERE, SANTA CATERINA I LA RIBERA	52.76832 %
10-25	EL BARRI GOTHIC	56.17046 %

SECCIO CENSAL CODE	SECCIO CENSAL NAME	ACTIVITY TYPE 6
84-35	LES CORTS	27.0113 %
141-40	EL GUINARDO	27.25322 %
193-56	LA SAGRERA	27.27273 %
21-11	EL FORT PIENC	27.37931 %
207-41	EL CLOT	27.60671 %
195-64	LA SAGRERA	38.09524 %

It seems like half of the Activity Type 5 users have visited some specific seccios censals in their routes, whereas Activity Type 4 users have visited different seccio censals.

4   The Distance Between Stops

The distance between the stops could not be calculated since there is a problem with Google API limitation. It will be in a function later, but for now, we have the whole dataset which stands for the physical distance between tramos (pairwise).

5   The Radius of Possible Influence of the Commercial Activity Around Each Stops
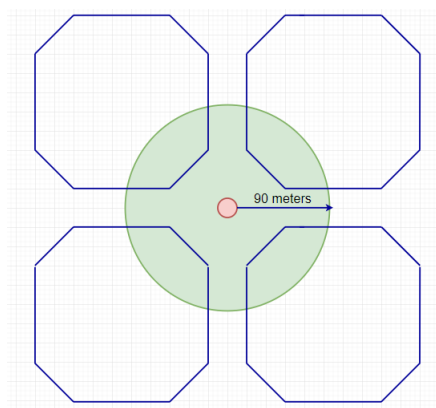
6   New Activity Types

In order to create the new activity types, the algorithms are built. Because of the year the data came from (2014), the information is not current and can be wrong. On the other hand, there are a lot of empty stores in the dataset. We only have 52241 out of 67117 businesses after the cleansing step. We can simply say that we have lost 22.16% of our data.

In total, we have 55296 users to be identified for the new activity types. However, we can only identify 7505 out of 55296 users (13.5%). It points out the need of a new data set, the current one for more accurate results.

The current algorithm is already fast and it does the comparison in one and half minutes, after the daily route computation is done.

FIRST AND LAST STOP ANALYSIS FOR THE NEIGBORHOODS ARE LOCATED CLOSE TO THE BORDER OF BARCELONA

There are 19 border neighborhoods in Barcelona, whereas there are only 8 neighborhoods we have data from. In this section, we analyze the first and the last check-ins for the daily routes of each user in order to see the deliveries come from out of Barcelona (or vice versa).

Neighborhood 13: La Marina de Port:
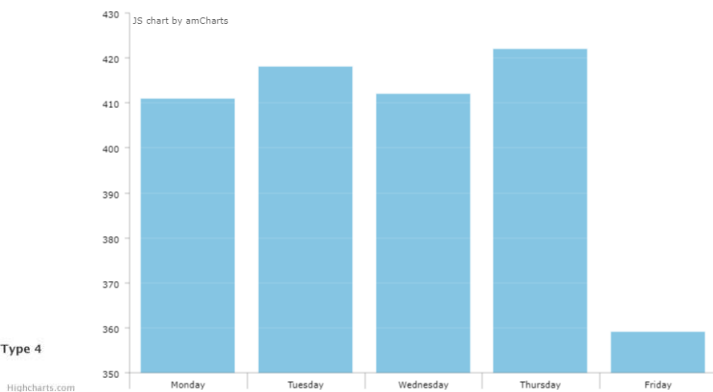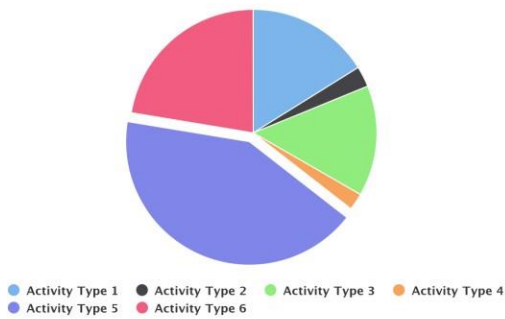
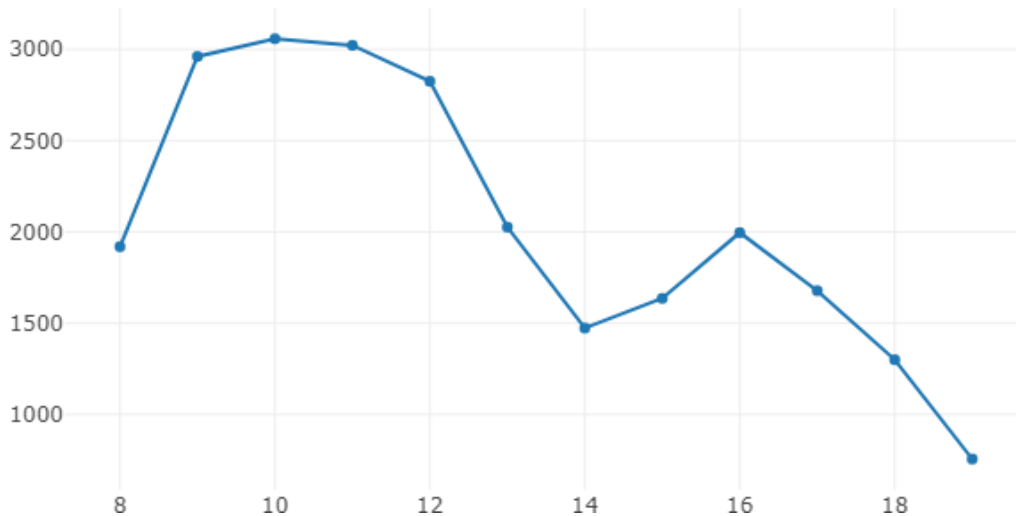Activity Type 5 (Transport), is the most common type in Neighborhood 13. From Monday to Thursday, there are more deliveries which occurred as the first or the last one in the daily routes. The hour distribution shows us that the drivers arrived this neighborhood (maybe from out of Barcelona) mostly between 9 am and 10 am.

Neighborhood 16: La Bordeta:
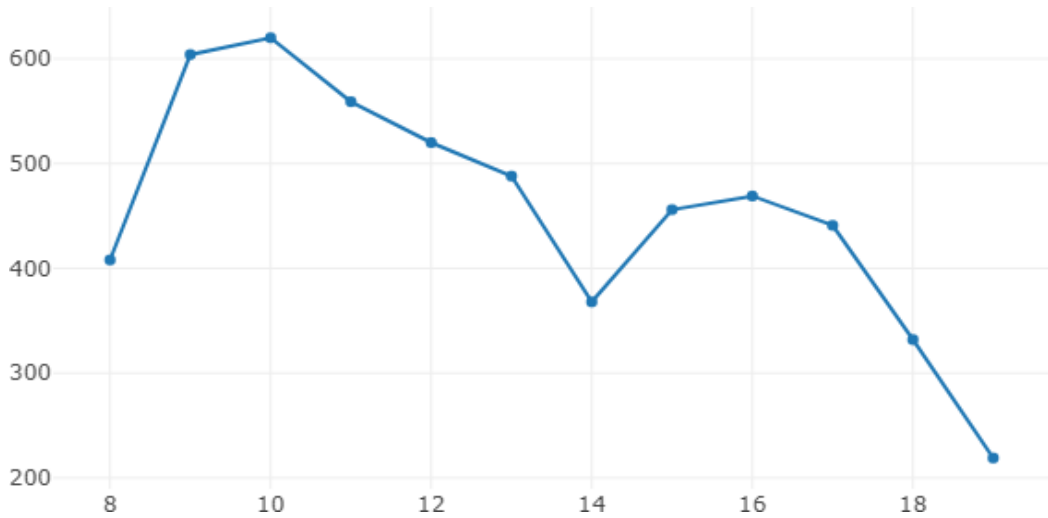
Activity Type Distribution for Neighborhood 16

- Activity Type 1 ● Activity Type 2 ● Activity Type 3 ● Activity Type 4
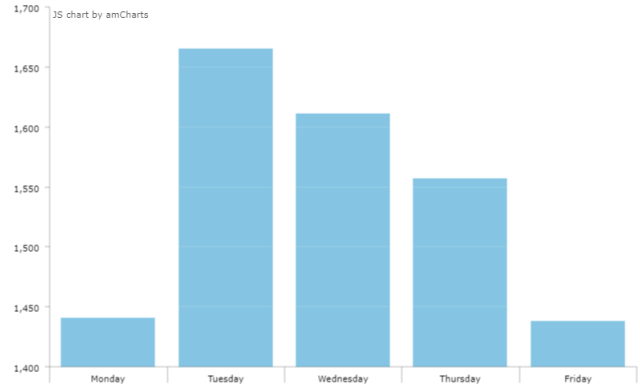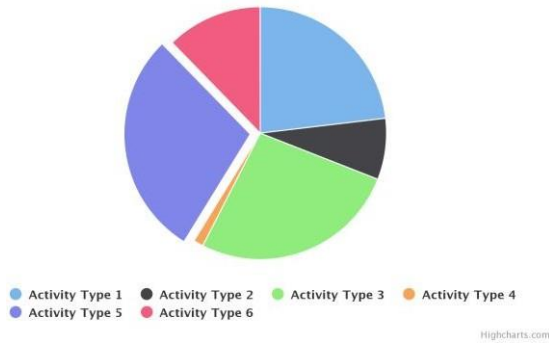- Activity Type 5 ● Activity Type 6

HOUR DISTRIBUTION IN NEIGHBORHOOD 16
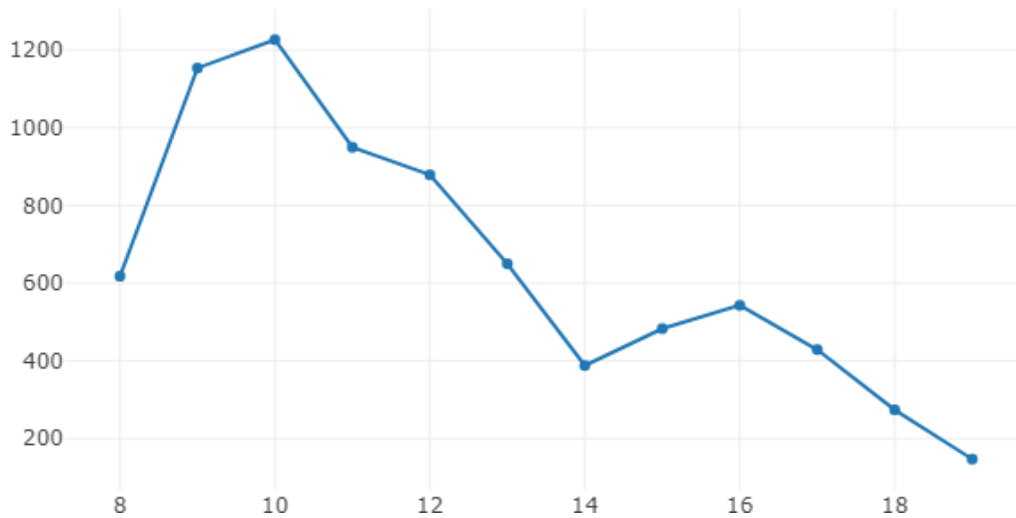
Activity Type 3 is the most common type in Neighborhood 16. From Monday to Thursday, there are more deliveries which occurred as the first or the last one in the daily routes. The hour distribution shows us that the drivers arrived this neighborhood (maybe from out of Barcelona) mostly between 9am and 11am, and they left between 3pm and 5pm.

Neighborhood 17: Sants-Badal

Activity Type Distribution for Neighborhood 17

- Activity Type 1
- Activity Type 2
- Activity Type 3
- Activity Type 4
- Activity Type 5
- Activity Type 6

Highcharts.com
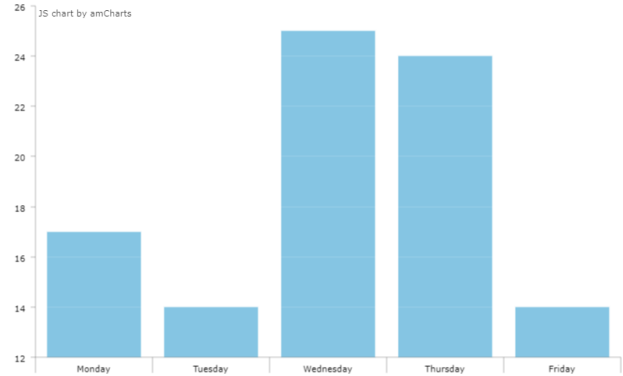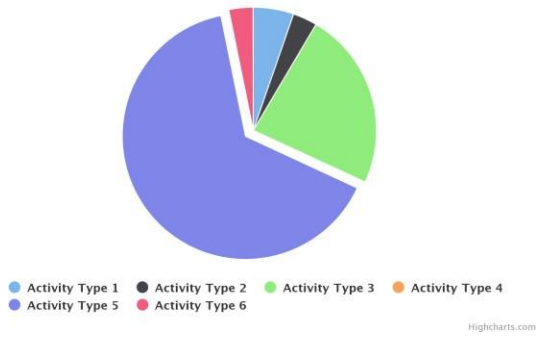
JS chart by amCharts
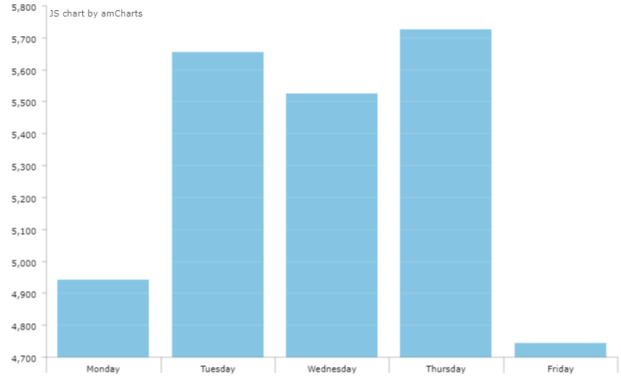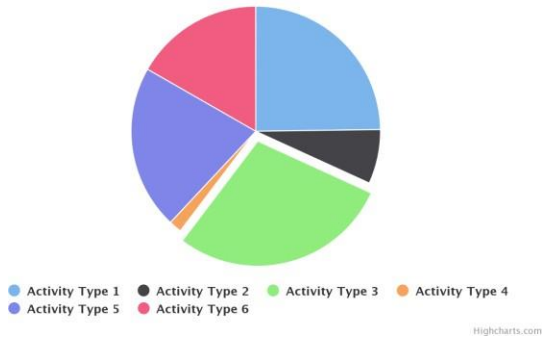
| HOUR DISTRIBUTION IN NEIGHBORHOOD 17 |

Activity Type 3 is the most common type in Neighborhood 16. From Tuesday to Thursday, there are more deliveries which occurred as the first or the last one in the daily routes. The hour distribution shows us that the drivers arrived this neighborhood (maybe from out of Barcelona) mostly between 10am and 11am, and they left between 4pm and 5pm.

Neighborhood 20: La Maternitat I Sant Ramon

Activity Type Distribution for Neighborhood 20

- Activity Type 1
- Activity Type 2
- Activity Type 3
- Activity Type 4
- Activity Type 5
- Activity Type 6

Highcharts.com

JS chart by amCharts

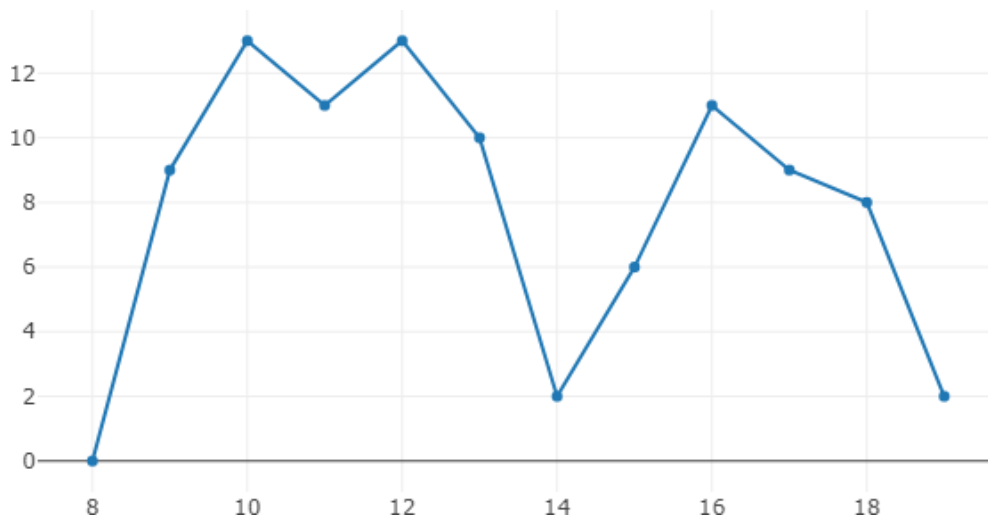| Monday | Tuesday | Wednesday | Thursday | Friday |
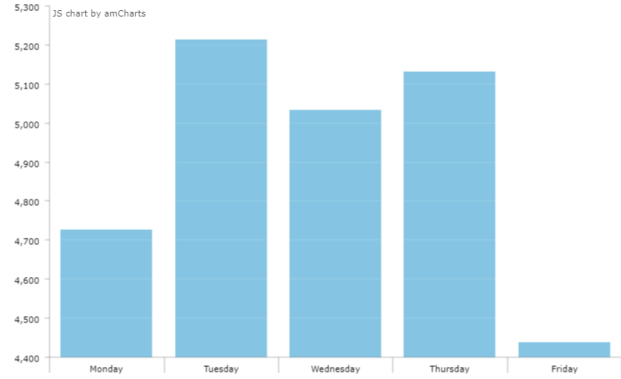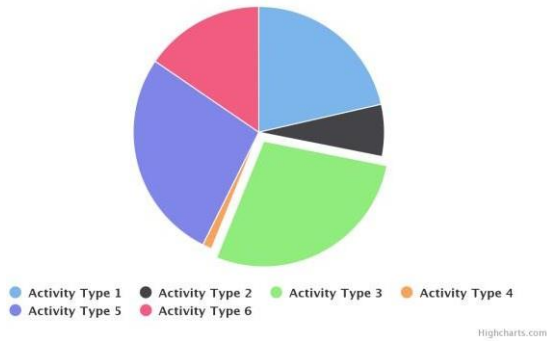
HOUR DISTRIBUTION IN NEIGHBORHOOD 20

Activity Type 3 is the most common type in Neighborhood 16. From Tuesday to Thursday, there are more deliveries which occurred as the first or the last one in the daily routes. The hour distribution shows us that the drivers arrived this neighborhood (maybe from out of Barcelona) mostly between 10am and 11am, and they left between 4pm and 5pm.

Neighborhood 21: Pedralles

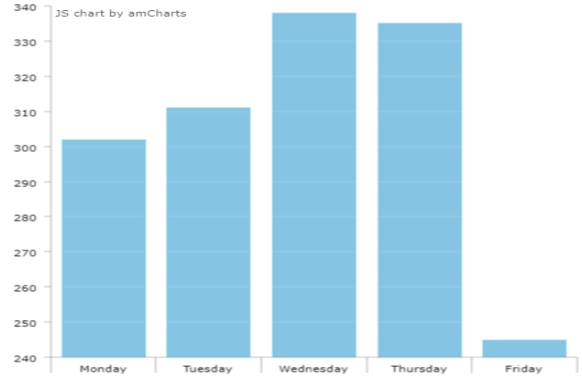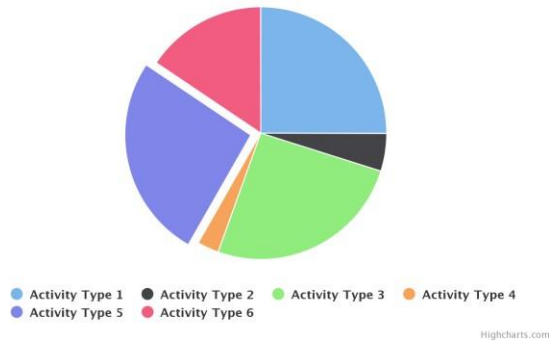Activity Type Distribution for Neighborhood 20

JS chart by amCharts

HOUR DISTRIBUTION IN NEIGHBORHOOD 21
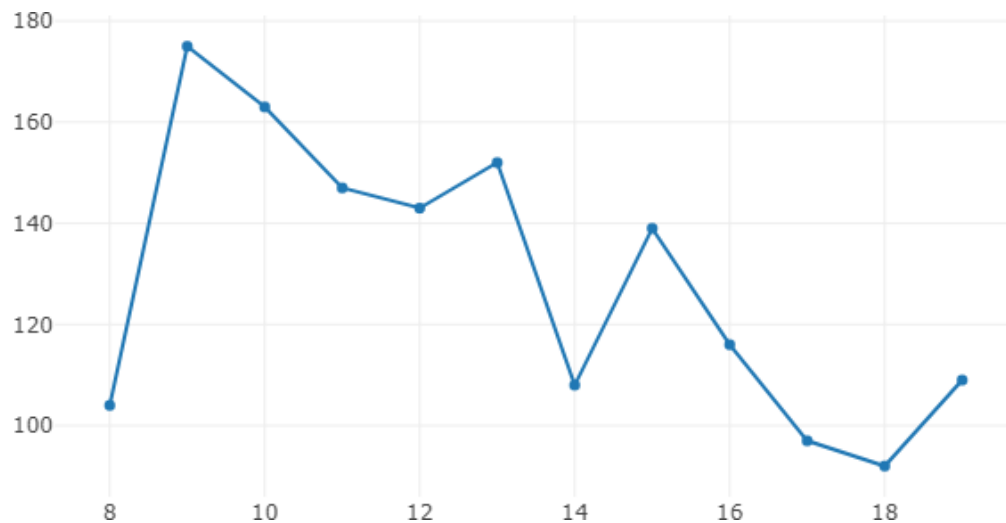

Activity Type 5 is the most common type in Neighborhood 16. From Tuesday to Thursday, there are more deliveries which occurred as the first or the last one in the daily routes. The hour distribution shows us that the drivers arrived this neighborhood (maybe from out of Barcelona) mostly between 10am and 11am, and they left between 4pm and 5pm.

Neighborhood 22: Sarria

Activity Type Distribution for Neighborhood 22

- ● Activity Type 1 ● Activity Type 2 ● Activity Type 3 ● Activity Type 4
- ● Activity Type 5 ● Activity Type 6

Highcharts.com

JS chart by amCharts

Monday, Tuesday, Wednesday, Thursday, Friday

| HOUR DISTRIBUTION IN NEIGHBORHOOD 22 |

Activity Type 5 is the most common type in Neighborhood 16. From Wednesday to Thursday, there are more deliveries which occurred as the first or the last one in the daily routes. The hour distribution shows us that the drivers arrived this neighborhood (maybe from out of Barcelona) mostly between 10am and 11am, and from 12am to 1pm, and they left between 4pm and 5pm.

Neighborhood 39: Sant Genis del Agudels
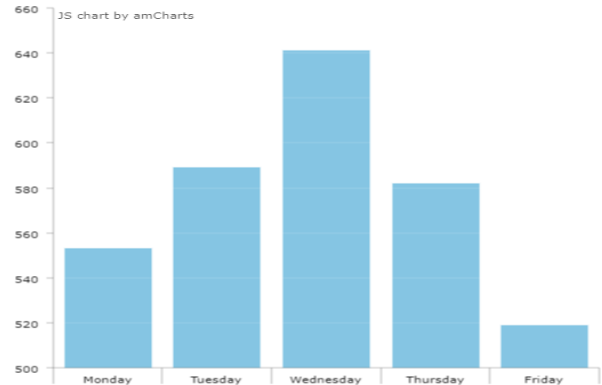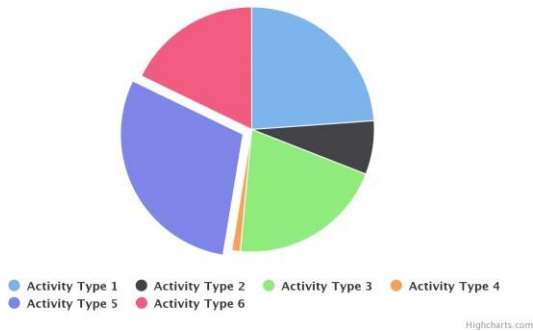

Activity Type Distribution for Neighborhood 39

- Activity Type 1
- Activity Type 2
- Activity Type 3
- Activity Type 4
- Activity Type 5
- Activity Type 6

Highcharts.com

JS chart by amCharts

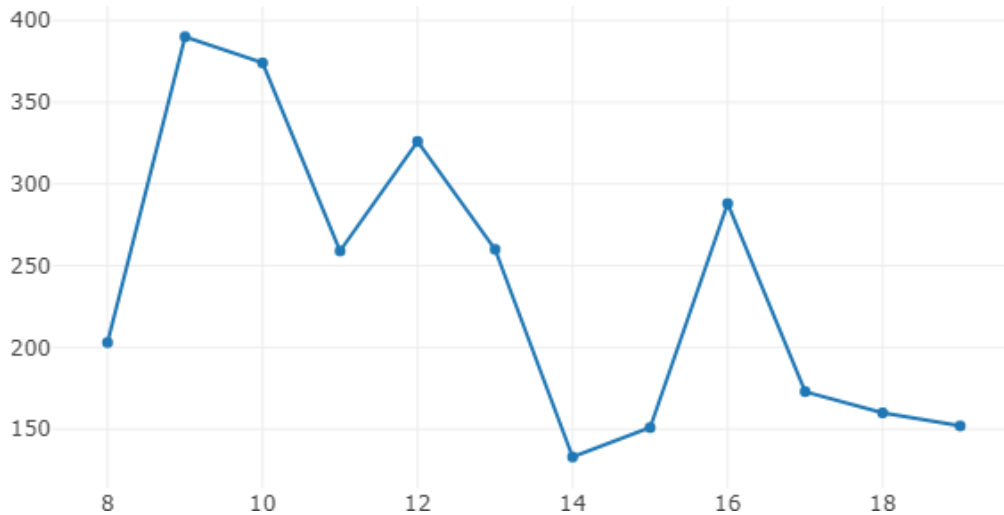HOUR DISTRIBUTION IN NEIGHBORHOOD 39

Activity Type 5 is the most common type in Neighborhood 16. From Monday to Thursday, there are more deliveries which occurred as the first or the last one in the daily routes. The hour distribution shows us that the drivers arrived this neighborhood (maybe from out of Barcelona) mostly between 9am and 10am, and they left between 3pm and 4pm.

Neighborhood 40: Montbau

Activity Type Distribution for Neighborhood 40

● Activity Type 1 ● Activity Type 2 ● Activity Type 3 ● Activity Type 4
● Activity Type 5 ● Activity Type 6

JS chart by amCharts

HOUR DISTRIBUTION IN NEIGHBORHOOD 40

Activity Type 5 is the most common type in Neighborhood 16. From Tuesday to Wednesday, there are more deliveries which occurred as the first or the last one in the daily routes. The hour distribution shows us that the drivers arrived this neighborhood (maybe from out of Barcelona) mostly between 9am and 10am, and they left between 4pm and 5pm.