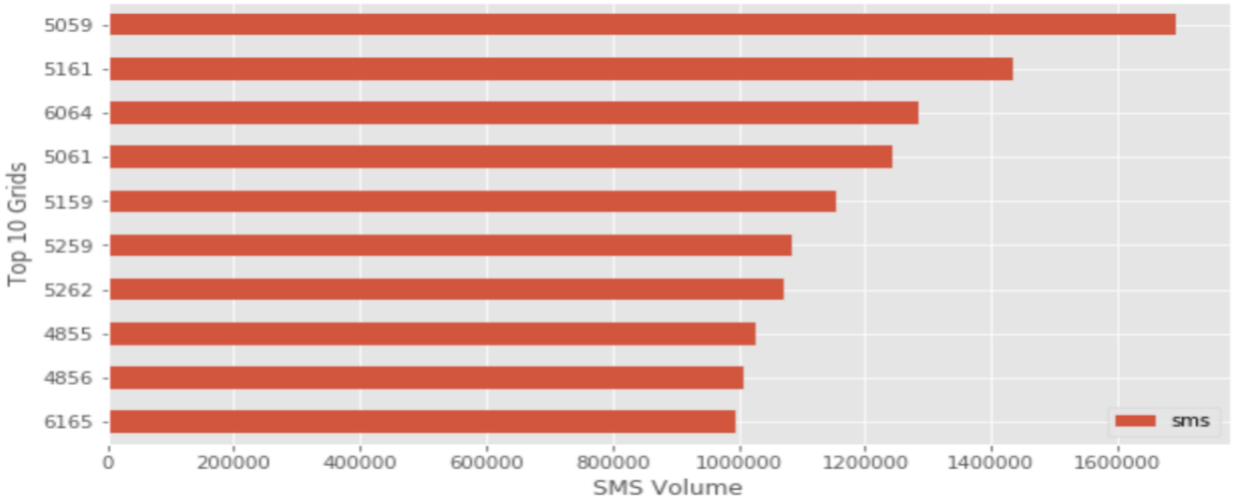
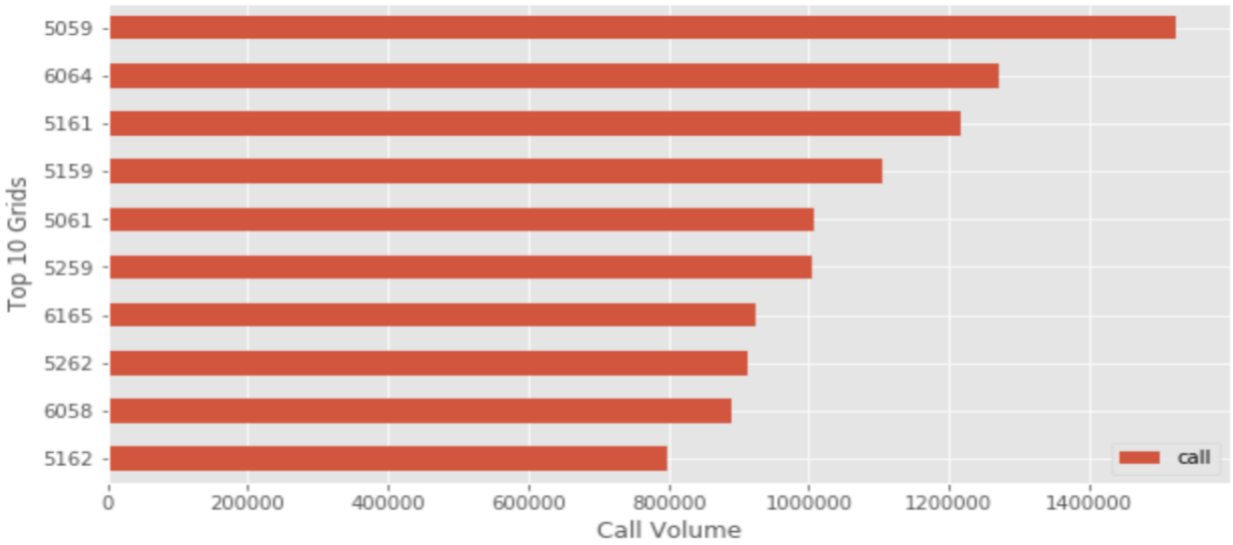
1. EXPLORATORY DATA ANALYSIS

In the network, SMS-In and SMS-Out utilizes the control channel, Call-In and Call-Out utilizes transmitted over voice channel and the internet is transmitted over broadband frequencies. Thus, we will use SMS (sum of SMS-In & SMS-Out), Call (sum of Call-In & Call-Out) and Internet activity for the analysis.





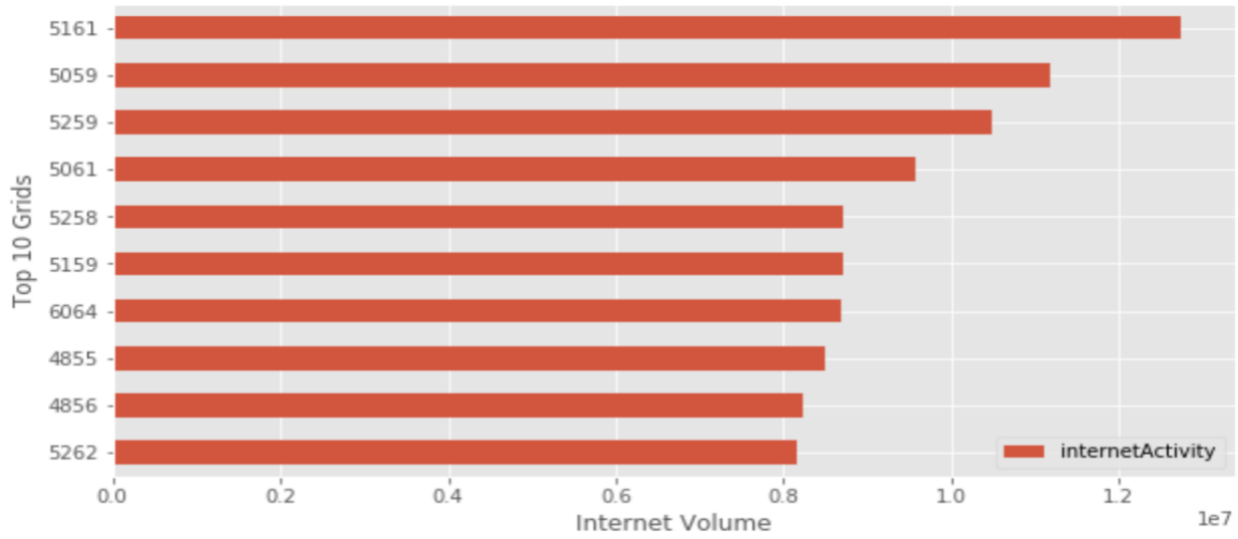
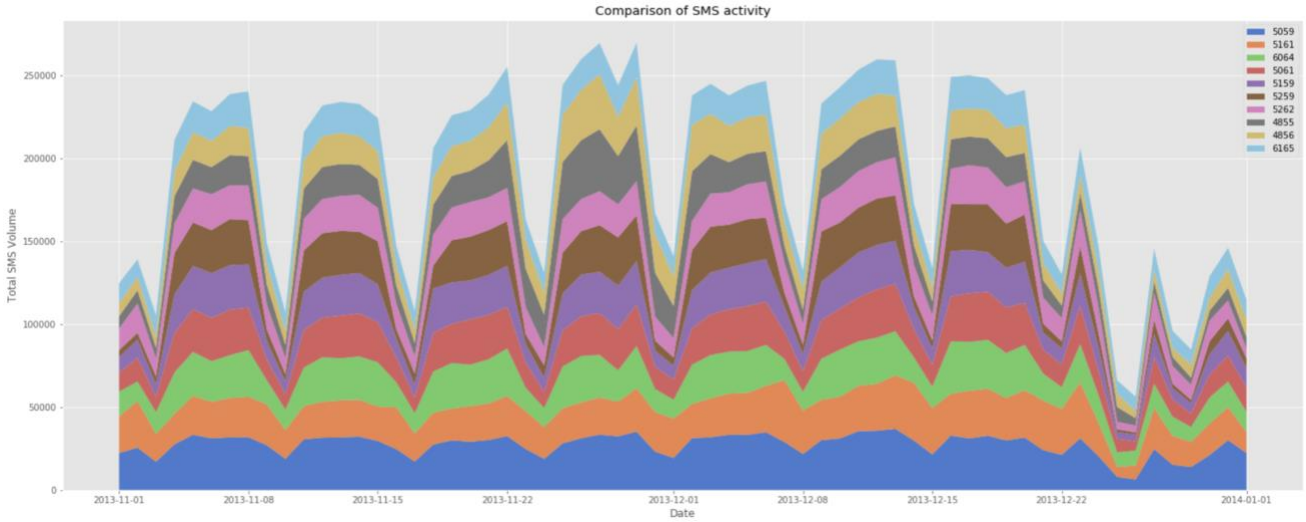
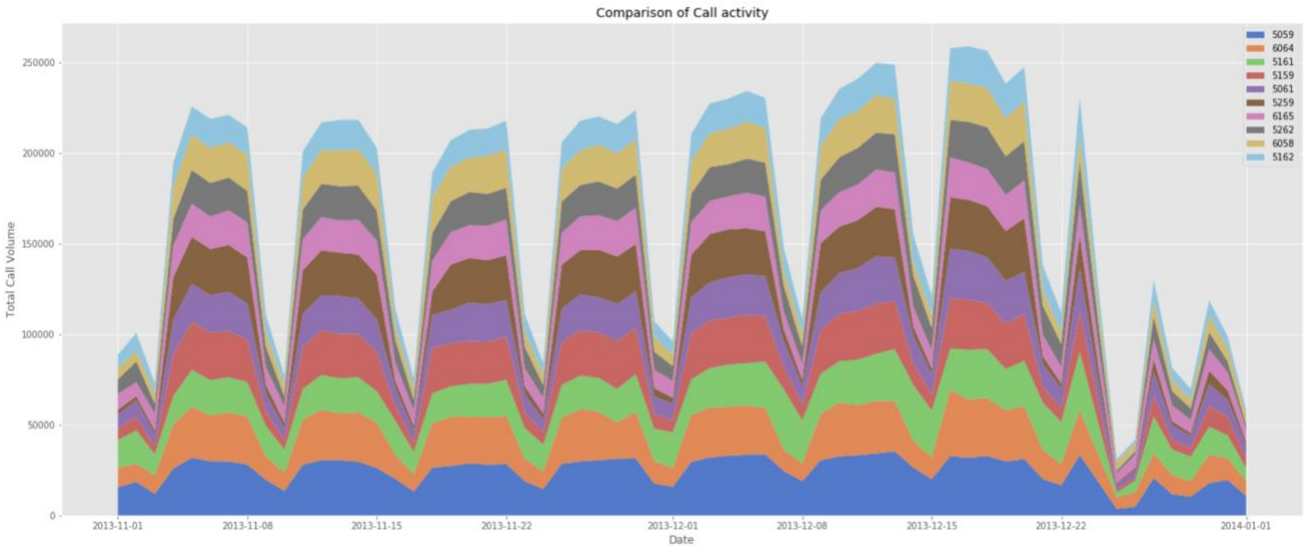


Fig 1: Horizontal bar plots showing top 10 grids with high total volumes in each telecommunication activity





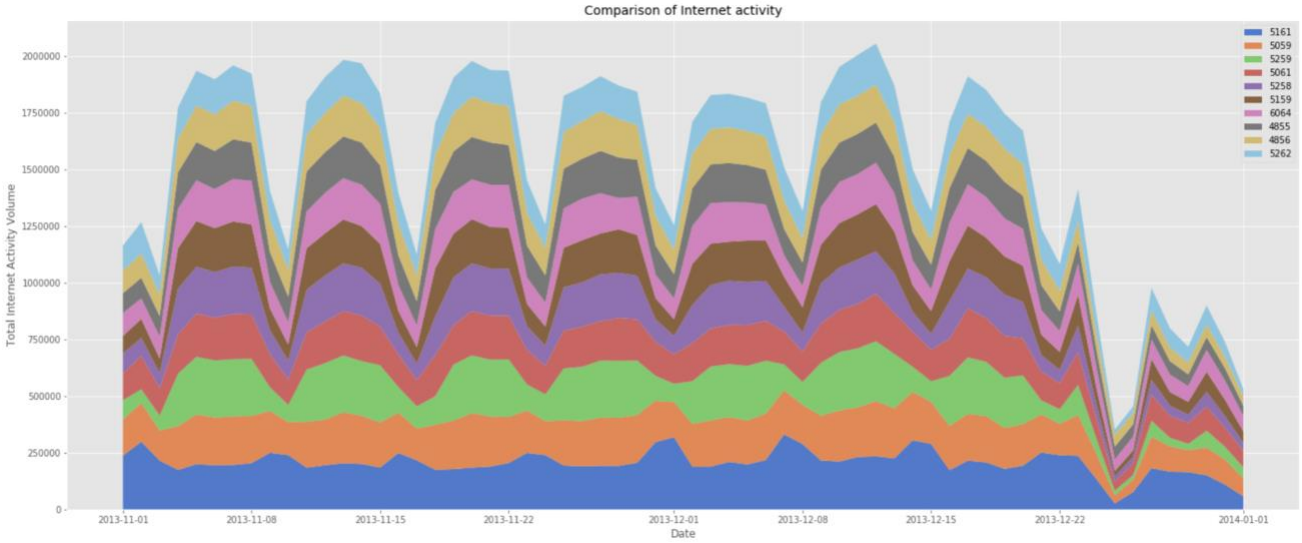


Fig 2: Stacked area plots showing comparison of top 10 grids daily pattern

Top 10 grids that experience high volumes for each of these activities for the 2 months are identified. From the above plots, we see that the top four grids have highly varying total volume and rest of the grids have almost same total volume for each activity. This is further verified by performing a set of one-way ANOVA tests as shown in Fig 7. P-value < 0.05 indicates mean values of the grids are not equal, P-value > 0.05 indicates means values of the grids are equal.

```
#Comparison of mean SMS values of top 4 grids
stats.f_oneway(daily5059.sms.to_list(), daily5161.sms.to_list(), daily6064.sms.to_list(), daily5061.sms.to_list())
```

```
F_onewayResult(statistic=15.6837152166561, pvalue=2.3305180034398526e-09)
```

```
#Comparison of mean SMS values of rest of the grids from the top 10 list
stats.f_oneway(daily5159.sms.to_list(), daily5259.sms.to_list(), daily5262.sms.to_list(), daily4855.sms.to_list(),
               daily4856.sms.to_list(), daily6165.sms.to_list())
```

```
F_onewayResult(statistic=1.0806731220881904, pvalue=0.37063204247098347)
```

```
#Comparison of mean Call values of top 4 grids
stats.f_oneway(daily5059.call.to_list(), daily6064.call.to_list(), daily5161.call.to_list(), daily5159.call.to_list())
```

```
F_onewayResult(statistic=8.245022535234758, pvalue=3.0159552209709387e-05)
```

```
#Comparison of mean Call values of rest of the grids from the top 10 list
stats.f_oneway(daily5061.call.to_list(), daily5259.call.to_list(), daily6165.call.to_list(), daily5262.call.to_list(),
               daily6058.call.to_list(), daily5162.call.to_list())
```

```
F_onewayResult(statistic=2.089333950123052, pvalue=0.06609465518655394)
```

```
#Comparison of mean Internet values of top 4 grids
stats.f_oneway(daily5161.internet.to_list(), daily5059.internet.to_list(), daily5259.internet.to_list(),
               daily5061.internet.to_list())
```

```
F_onewayResult(statistic=8.388137367527829, pvalue=2.5001501890290148e-05)
```

```
#Comparison of mean Internet values of rest of the grids from the top 10 list
stats.f_oneway(daily5258.internet.to_list(), daily5159.internet.to_list(), daily6064.internet.to_list(),
               daily4855.internet.to_list(), daily4856.internet.to_list(), daily5262.internet.to_list())
```

```
F_onewayResult(statistic=0.42971229199857125, pvalue=0.8278755021476163)
```

Fig 3: One-Way ANOVA tests verifying that, except the top 4 grids, rest have similar mean volumes

Location of these grids in the map shows that they are all from Duomo & Milano Centrale region.

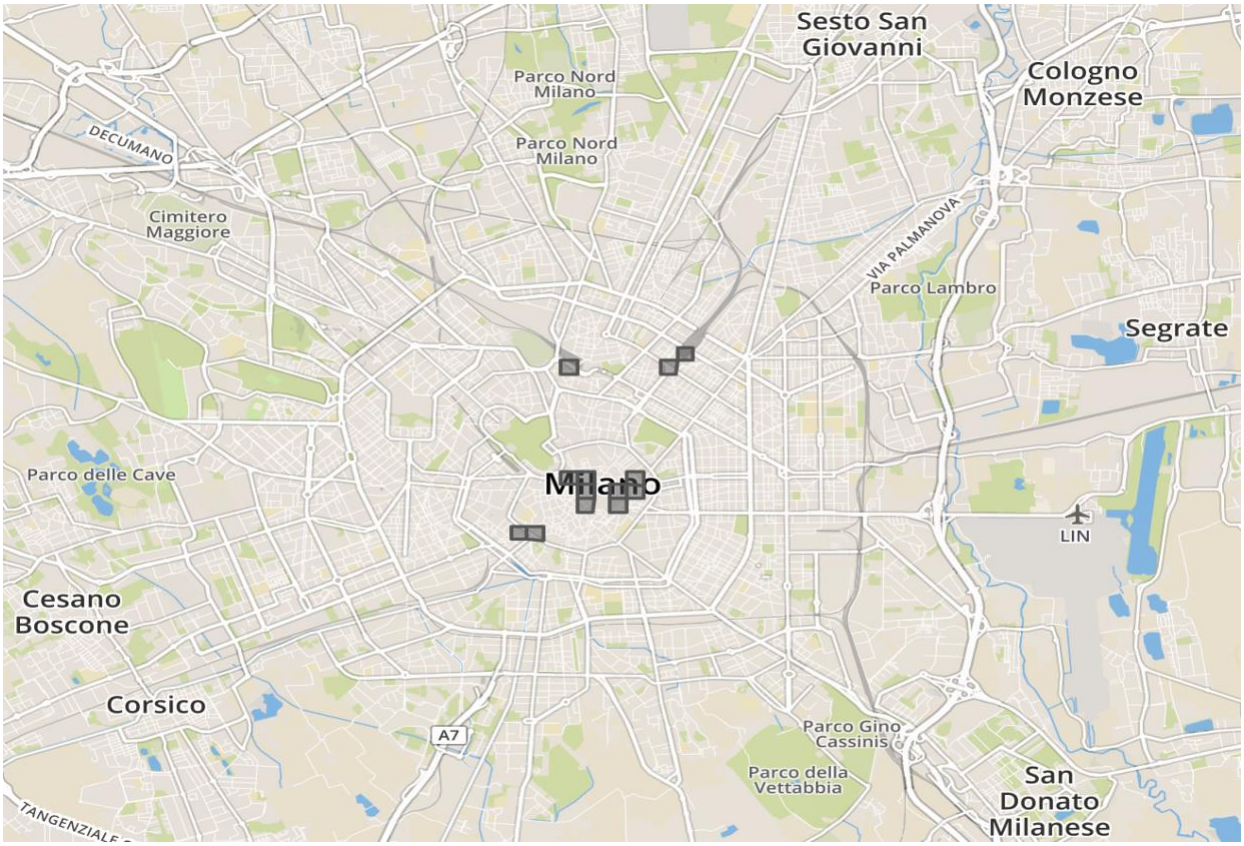


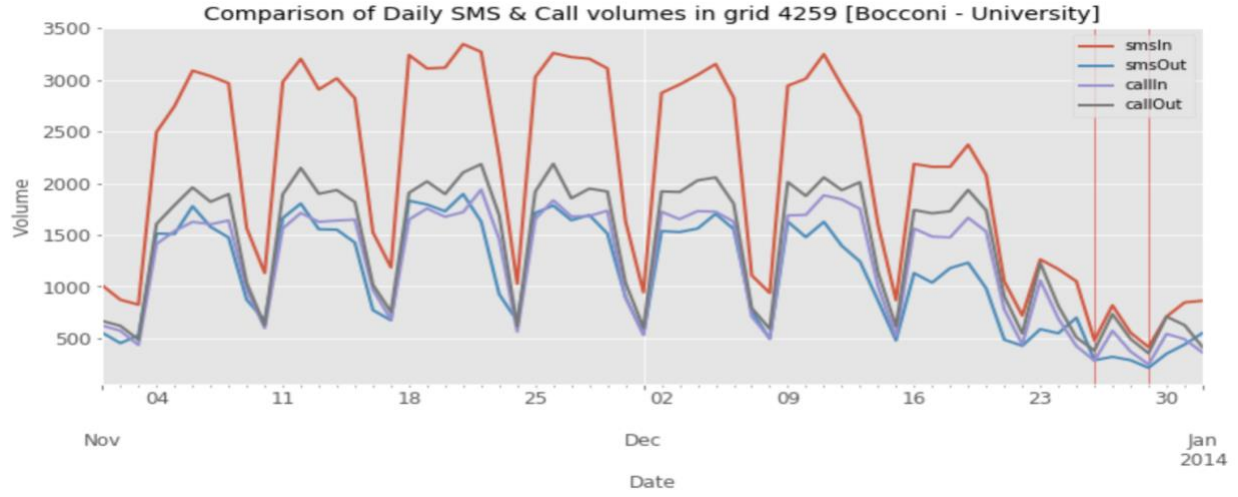
Fig 4: Location of top 10 high-volume grids in the city map

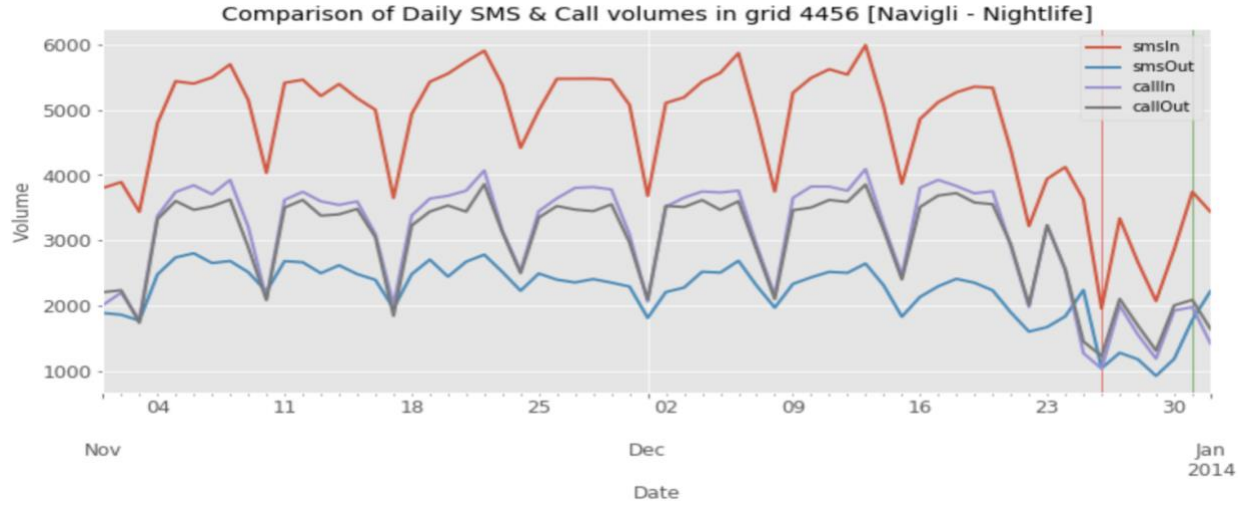
In Detail Analysis of grids from different sectors

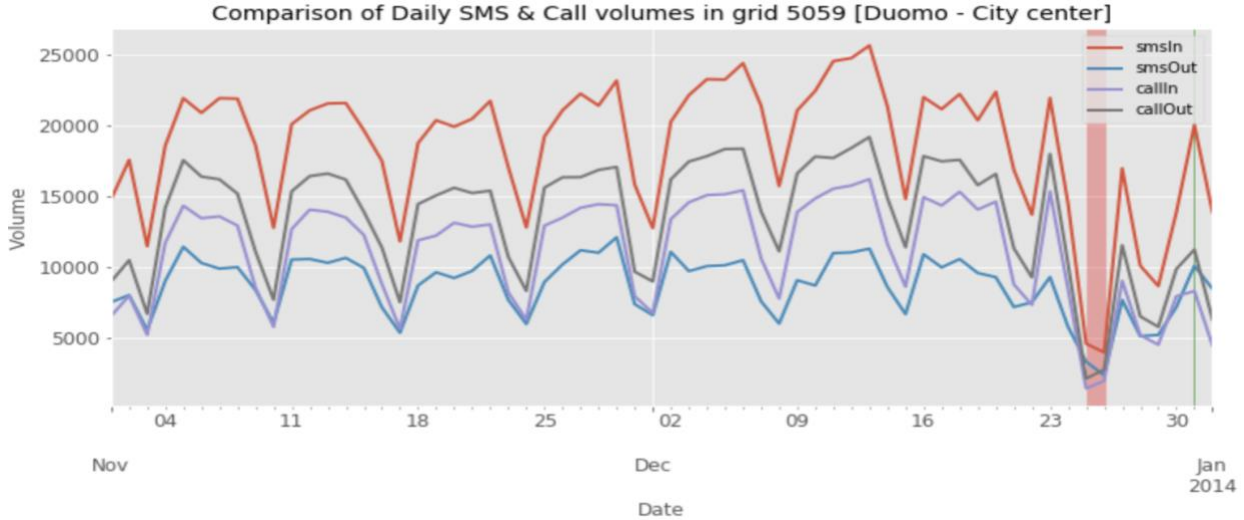
All the top contributing grids are from the city's center and mostly near transport hubs, they are expected to show similar behavior and most of these grids have approximately same mean values. In order to capture variations in the city's telecommunication activities, we will examine the following four grids that has markedly different behavioral signatures,

4259 - Bocconi, one of the most famous Universities in Milan
4456 - Navigli district, one of the most famous nightlife places in Milan
5059 - Duomo, the city center of Milan
5346 - Fiera, residential neighborhood of Milan

Daily Activity Plots:







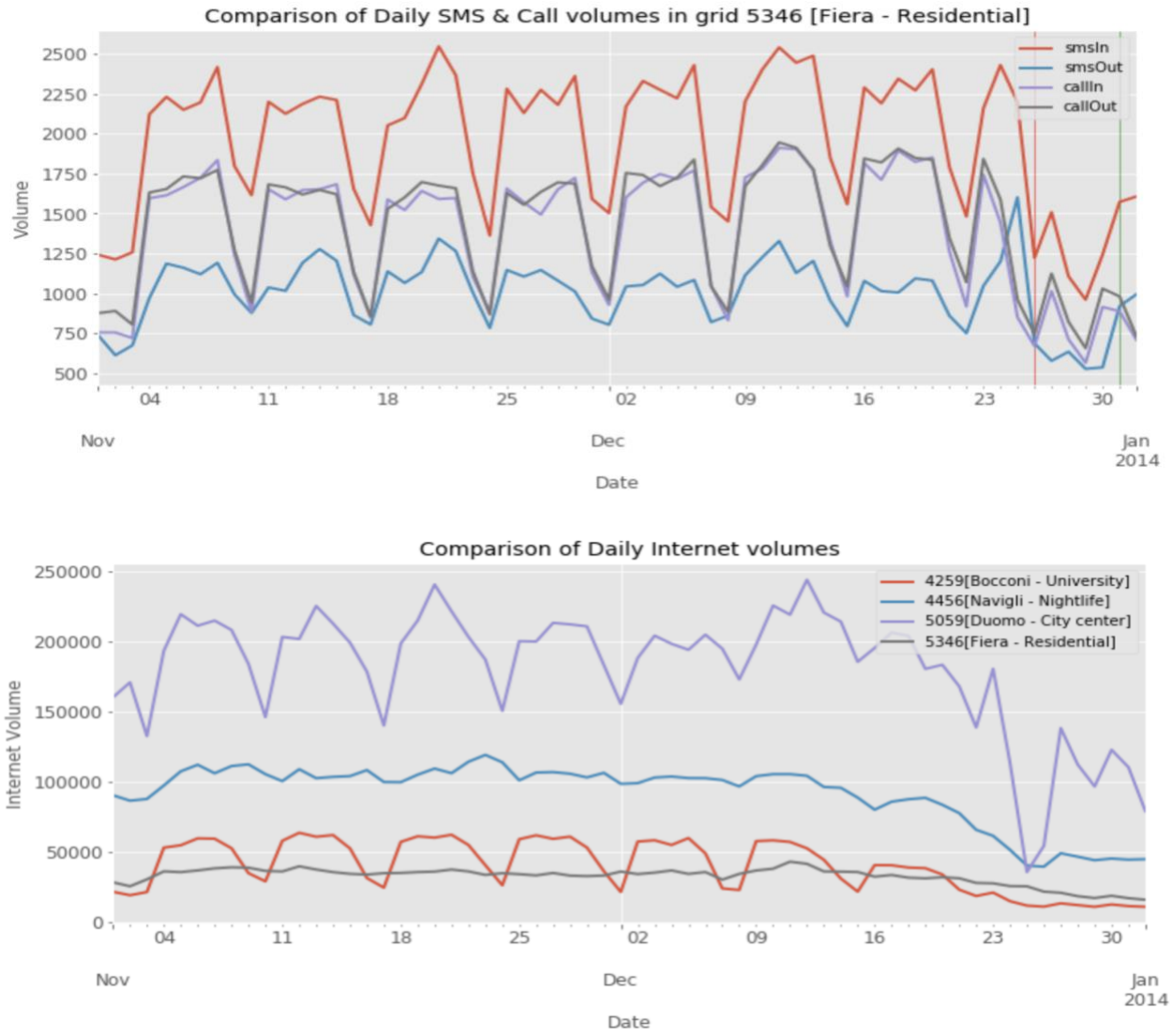


Fig 5: Time-series plot of daily telecommunication activities of the four grids

- All four grids have received high volumes of incoming SMS compared to other activities. Outgoing SMS has the least volume, almost equal amounts of Calls are made and received.
- Duomo has highest volume of all activities, followed by Navigli, then Bocconi and Fiera in the end. We can order the grids based on total volumes as,
- Duomo [city center] > Navigli [nightlife] > Bocconi [university] > Fiera [residential]
- All four grids exhibit seasonality in SMS & Call activities. In internet activity Navigli & Fiera doesn't show any seasonality. This may because of IoT, with many devices always being connected to the network.
- There is a drop in the volumes towards December end in all the plots [holiday season].

Hourly Activity Plots:

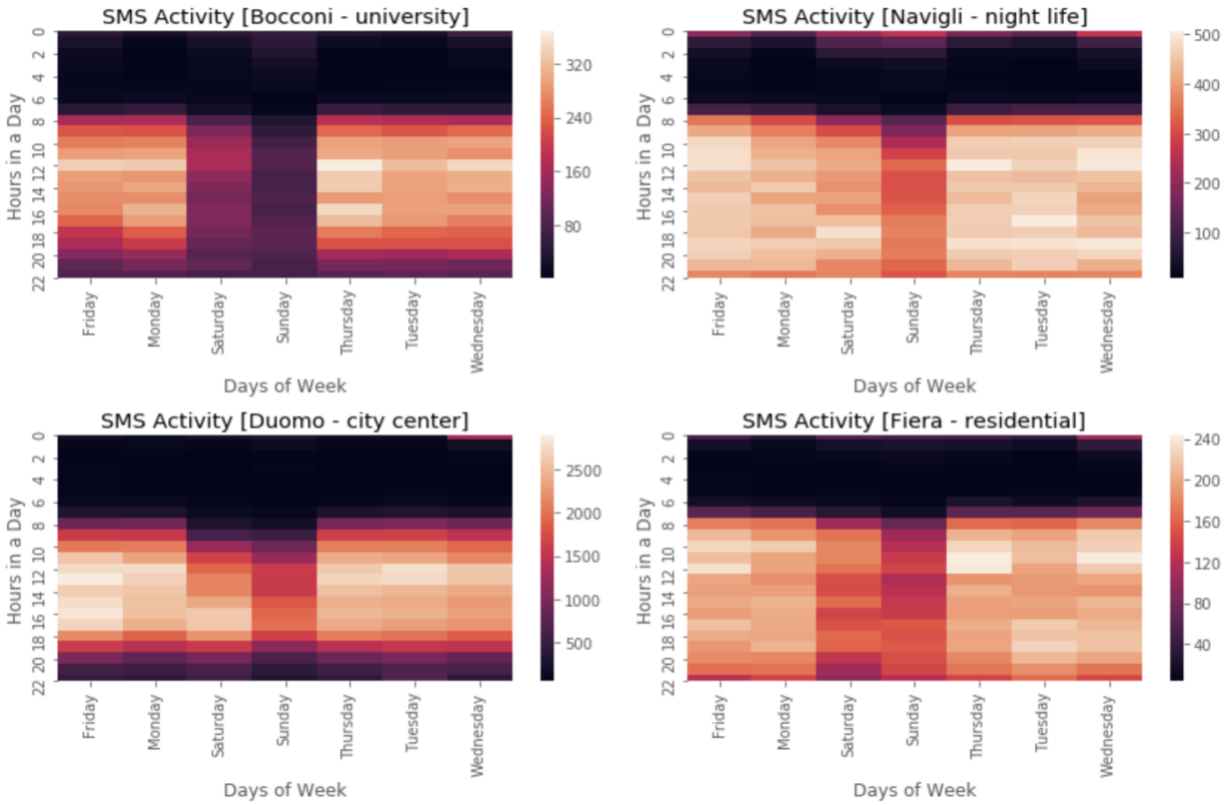


Fig 6: Heat map of hourly SMS activity of the four grids

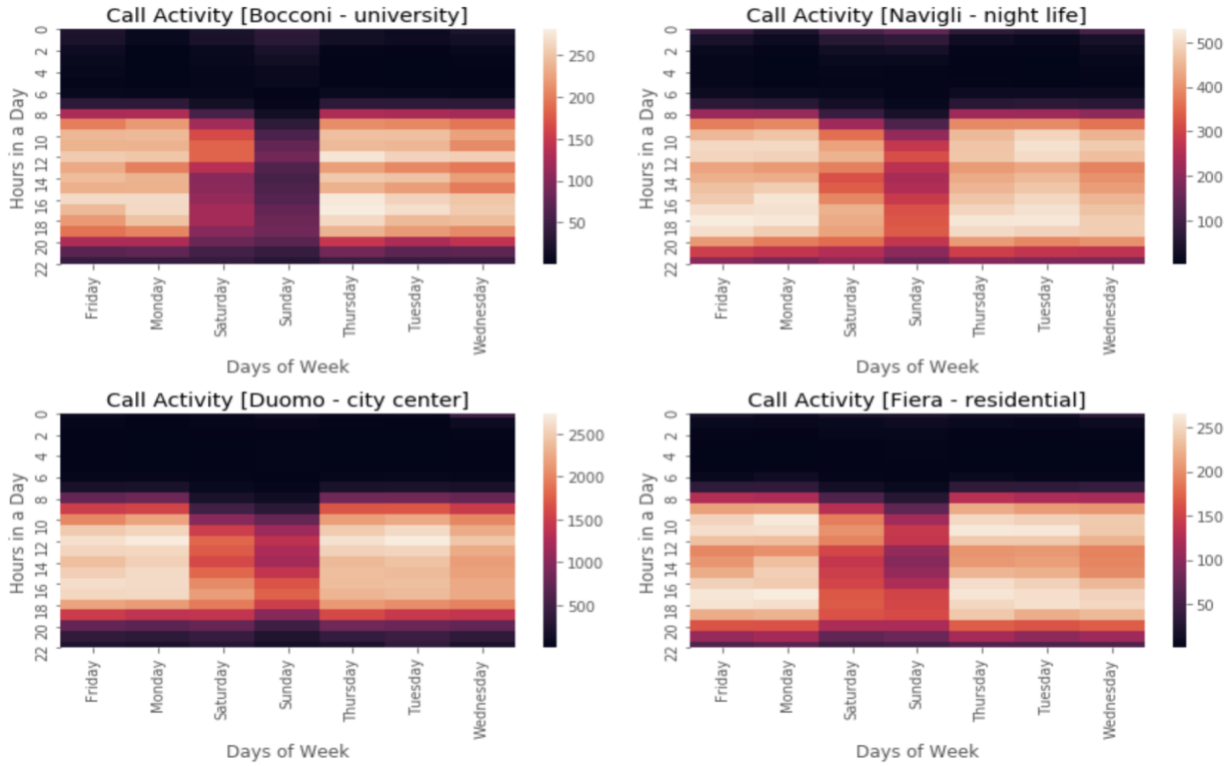


Fig 7: Heat map of hourly Call activity of the four grids

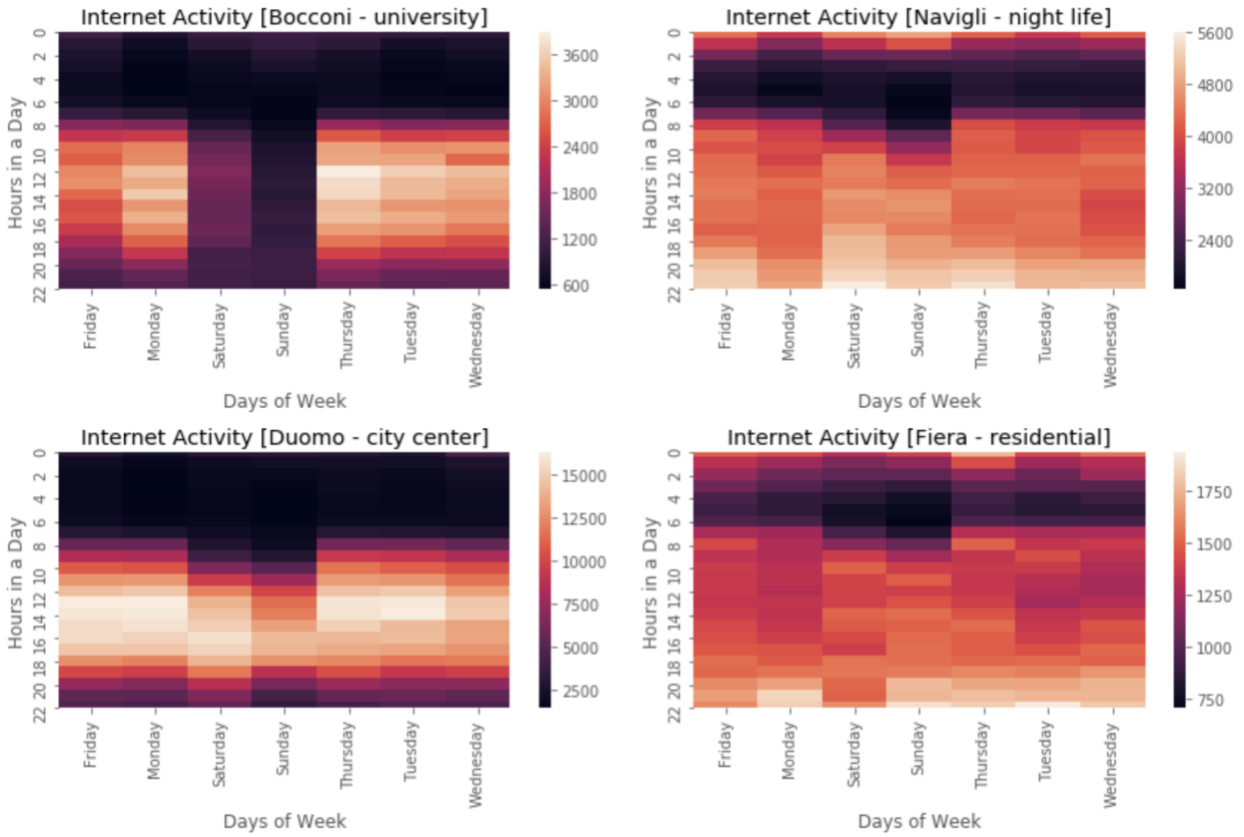


Fig 8: Heat map of hourly telecommunication activities of the four grids

Heat maps shows significant differences in behavior of the four grids,

SMS Activity:

- In general, there is less SMS activities during the weekends (Saturday & Sunday).
- Navigli, night life region shows SMS activities until 2 am in the night on weekend.
- Fiera & Navigli regions are very active from 7am till 10pm on all days.
- Duomo, city center has very less SMS activity from 10pm until 7am in the morning on weekdays and 9 am on weekends.
- Bocconi, university shows less SMS activity compared to others.
- Navigli, Duomo & Fiera shows a sudden bright region on Wednesday 12am, this must be due to New Eve falling on Tuesday.

Call Activity:

- Call activity has similar pattern as SMS, but lesser volumes.
- Surprisingly, there is no significant call activity on New Year eve. This shows how people are more connected via SMS and internet these days. Another possibility is that calls may have been made via internet.

Internet:

- Navigli & Fiera has internet activities almost all through the night. Even Bocconi, university shows some sparse activity after midnight on weekends.
- Duomo, city center although has the highest internet volumes, shows a steady pattern for all activities, 8am – 10pm on weekdays and 10am to 10pm on weekends.