Project 1: 4G Mobile Core Capacity Analysis

Team No. 4
Tadeo Aguilar
Carlos Palomino
Federico López

Brief Introduction

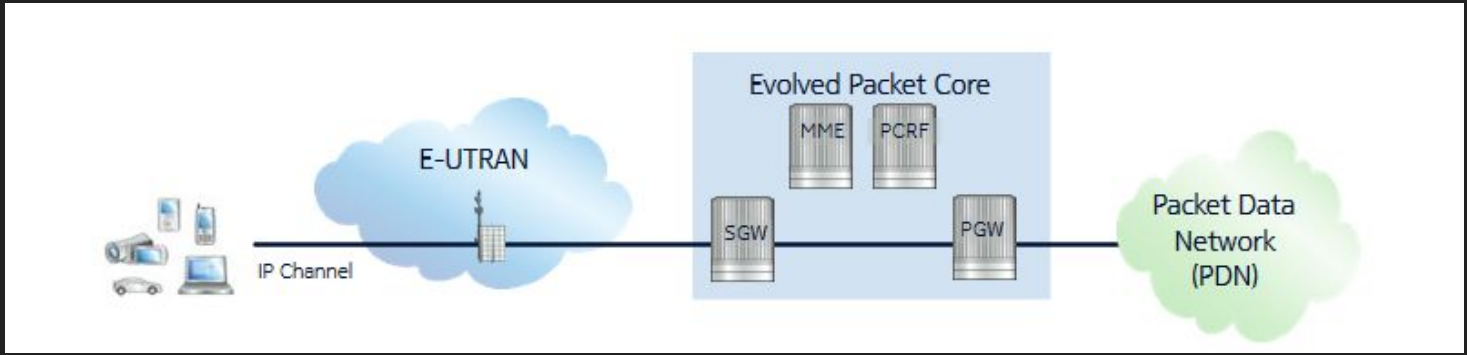
Project Proposal

- Our project is to uncover patterns in the Capacity KPIs of a 4G Core Network. We'll examine relationships between the KPIs; network capacity use, days and times of day when peaks are presented; capacity forecasts; and the related questions

Finding Data

- We are going to use real data from a 4G mobile operator in Mexico; the data consists of 6 Excel files containing capacity measured data during one week for 2 elements of the Core Network: MME and SAEGW

4G Network Architecture



- It is composed of Access Network (E-UTRAN), Core Network (Evolved Packet Core) and Packet Data Network (typically the Internet)
- In this project we are analyzing the Core Network: MME and SAEGW (combined SGW + PGW)

Excel files description

1. *mme.xlsx* : containing KPIs measurements for MME, there is a total of 6 MMEs
2. *mme_cpu_mem.xlsx* : containing CPU and memory KPIs measurements for MMEs
3. *saegw.xlsx* : containing KPIs measurements for SAEGW, there are 11 SAEGWs in total
4. *saegw_bearers.xlsx* : containing bearers measurements for SAEGW; it is separated as it is measured in a different frequency basis
5. *saegw_cpu_mem.xlsx* : containing CPU and memory KPIs meas. for SAEGW
6. *max_capacity.xlsx* : containing maximum capacity information for each element

KPIs description

The KPIs that we will analyze are the following:

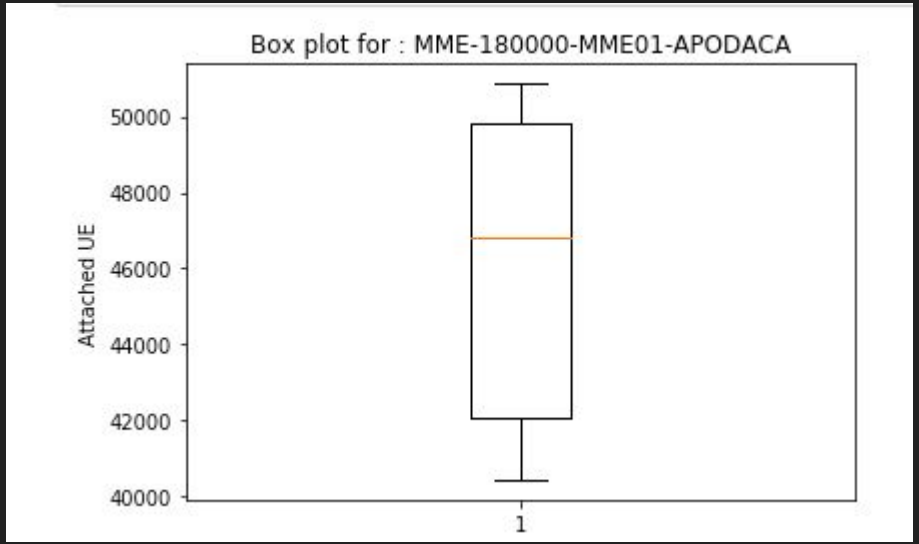
- MME:
  - ATTACHED UE: Attached user equipments
  - BEARERS: Quantity of data channels
  - AVECPUUSAGE: Average CPU utilization in the current interval. Unit: % Range: 0-100
  - PEAKCPUUSAGE: Peak CPU utilization in the current interval. Unit: % Range: 0-100
- SAEGW:
  - THROUGHPUT: Data rate measured in gigabits per second
  - BEARERS: Quantity of data channels
  - MAXCPUUTILIZATION: Max CPU utilization in the current interval. Unit: % Range: 0-100

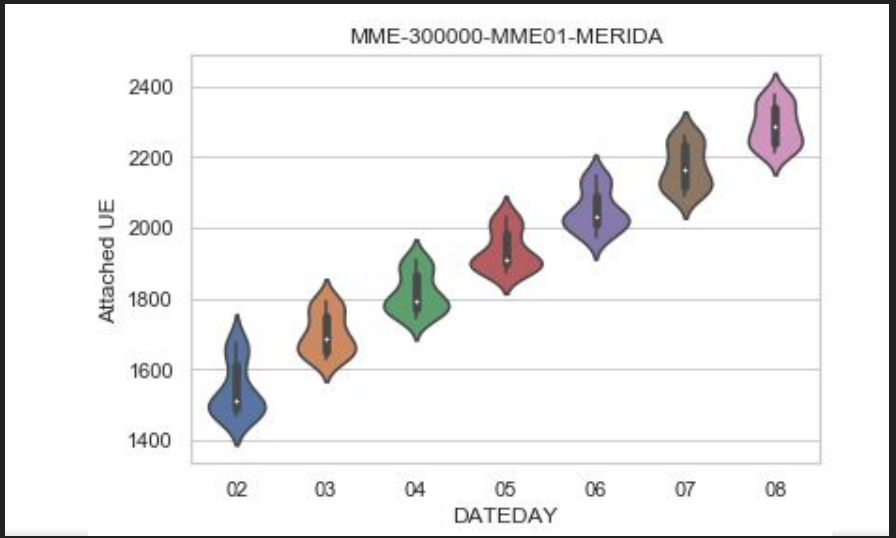
Methodology

Business Understanding	Data questions	Data Cleansing	Data Preparation	Data Analysis
Meeting with data owner to understand the files and Kpis	-Descriptive -Diagnostic -Predictive	-Outlier identifications -Empty and Nan treatment	-DF Manipulation -Regression tests -Api integration	-Data Visualizations -Presentation

Are there any outlier in KPI.s

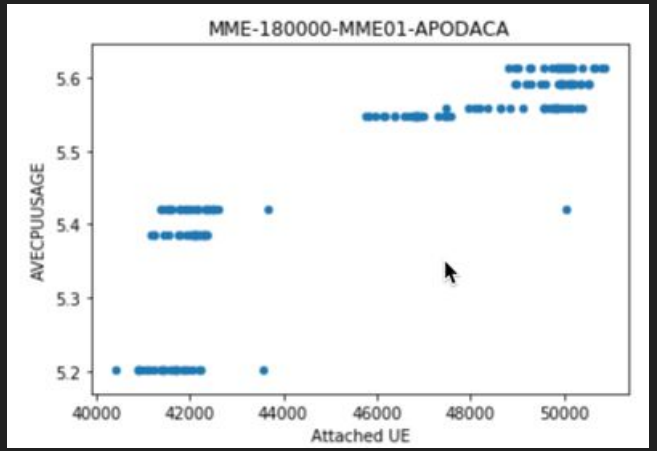
No

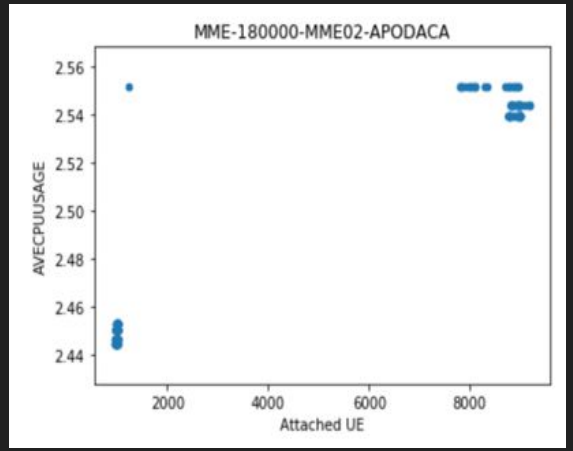


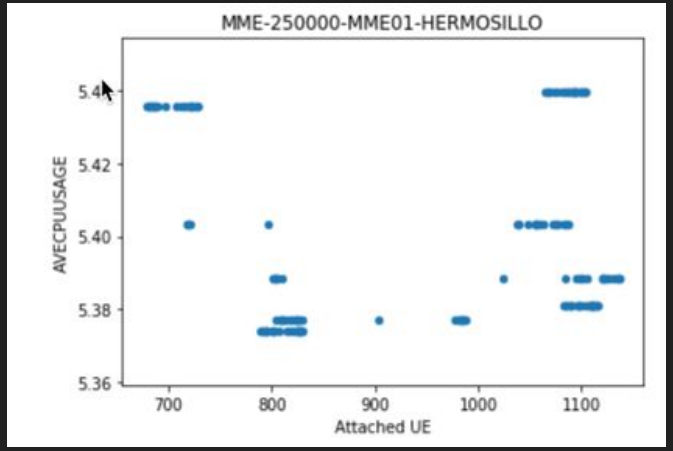


Is there a relationship in Attached UE and Average CPU Usage

There is no visible relationship in this variables

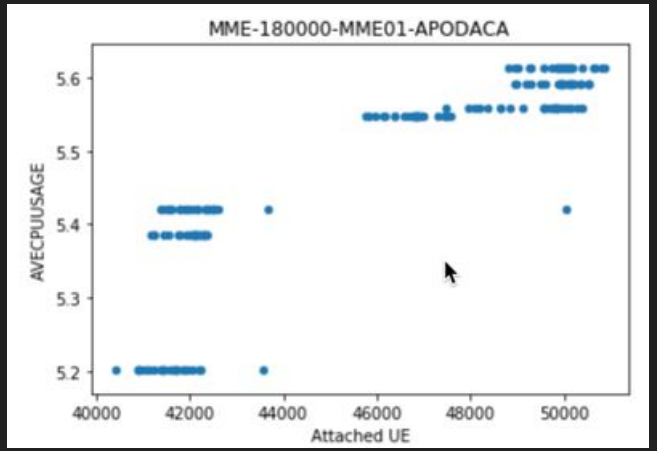


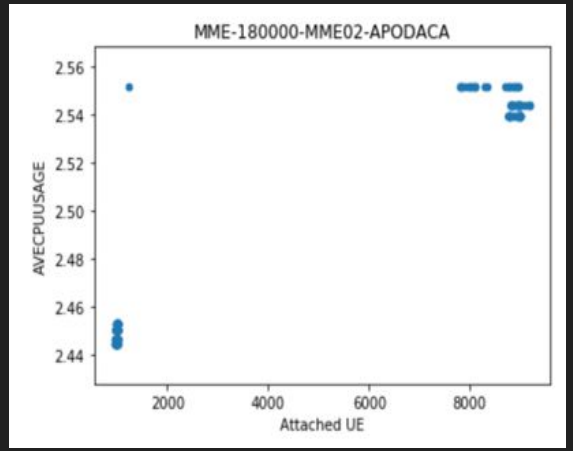


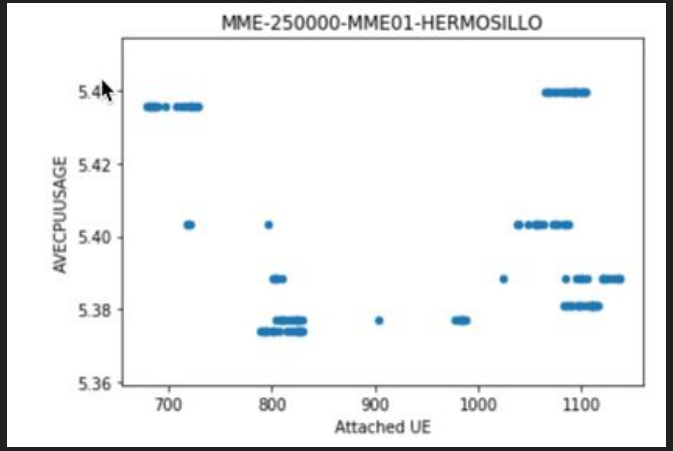


Is there a relationship in Attached UE and Average CPU Usage

There is no visible relationship in this variables

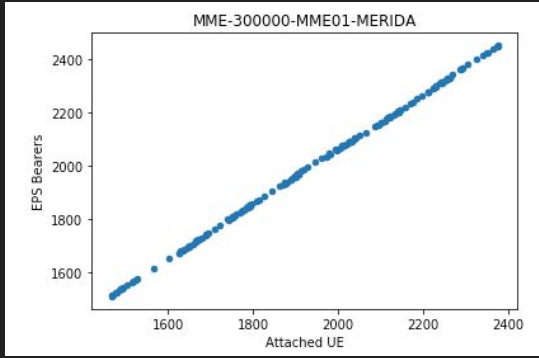


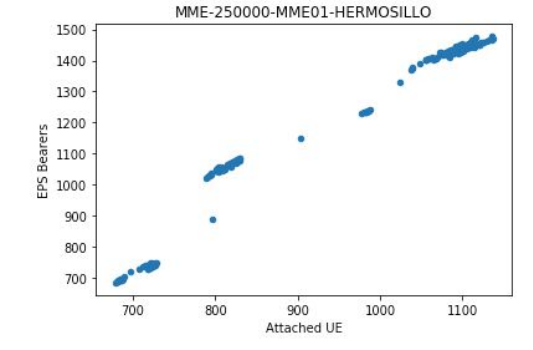


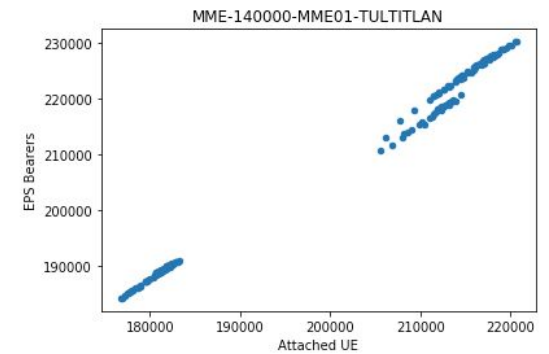


Is there a relationship in Attached UE and Bearers and other KPIs

There is a visible relationship in attached UE and EPS Bearers

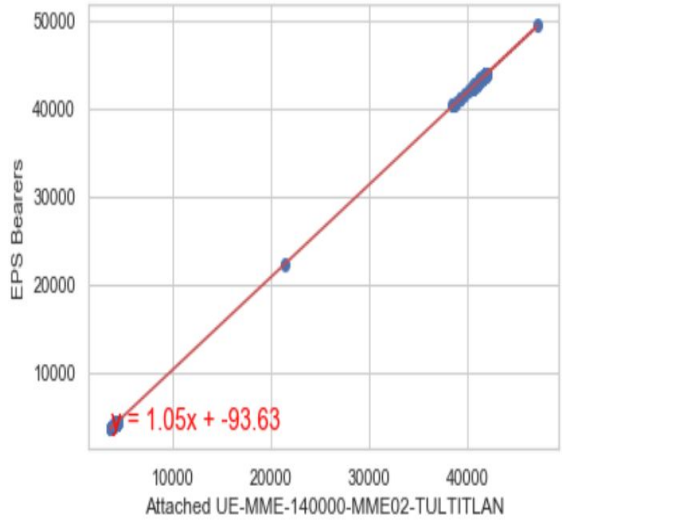




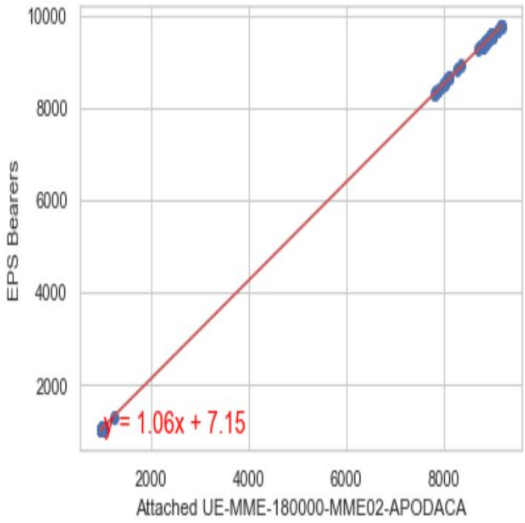


Regression test

The r-squared for MME-140000-MME02-TULTITLAN is: 0.9999911278807705

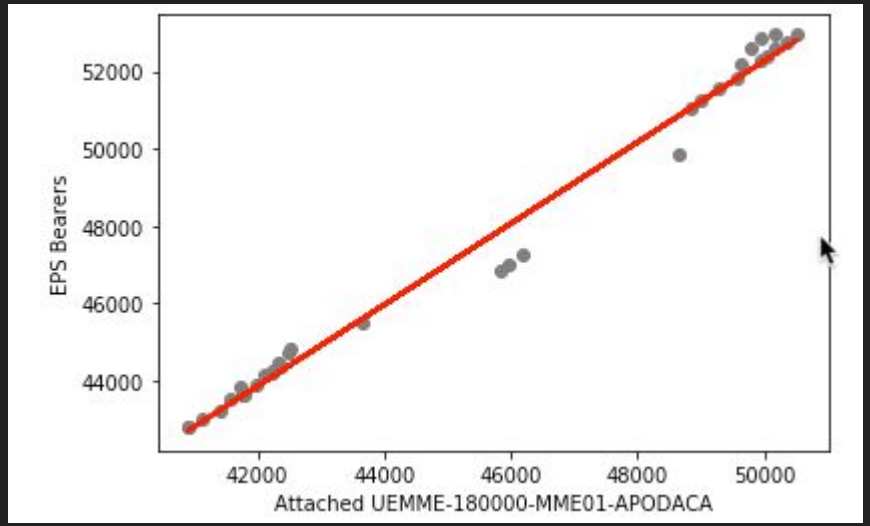


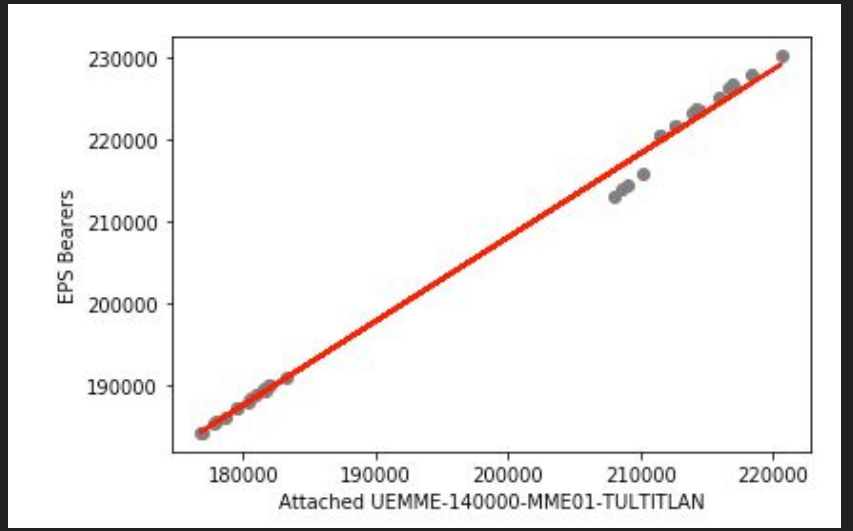
The r-squared for MME-180000-MME02-APODACA is: 0.9999932268958474



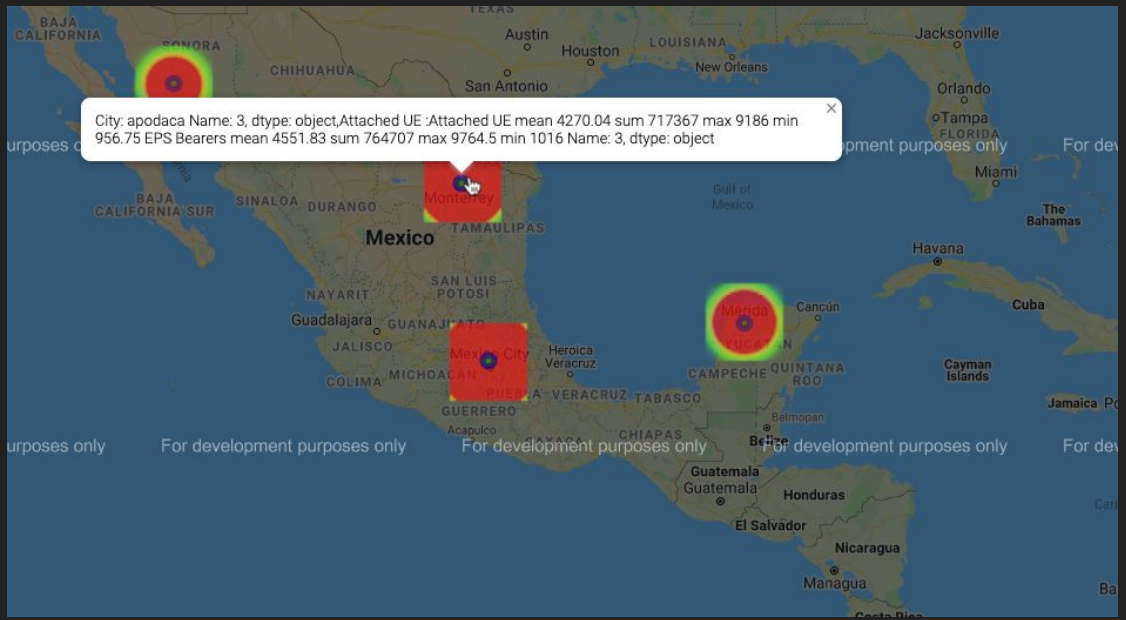
Can we predict the Bearers if we know the Attached UE.

Yes. What if these trends continues?





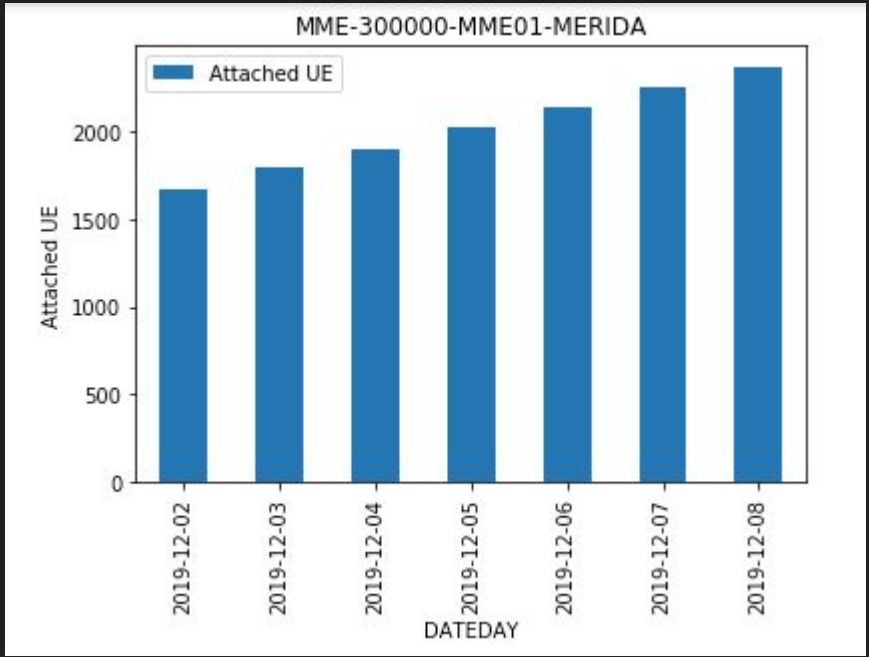
Can we identify in a Map the cities with the highest demand.



	MME	city	lat	lng	Attached UE				EPS Bearers	
					mean	sum	max	min	mean	sum
0	MME-140000-MME01-TULTITLAN	tultitlan	19.639212	-99.166865	200128.092262	33621519.50	220623.75	176803.25	208229.254464	34982514.75
1	MME-140000-MME02-TULTITLAN	tultitlan	19.639212	-99.166865	19927.809524	3347872.00	47176.00	3601.25	20773.763393	3489992.25
2	MME-180000-MME01-APODACA	apodaca	25.776468	-100.185874	45980.971726	7724803.25	50866.00	40431.00	48048.677083	8072177.75
3	MME-180000-MME02-APODACA	apodaca	25.776468	-100.185874	4270.041667	717367.00	9186.00	956.75	4551.828869	764707.25
4	MME-250000-MME01-HERMOSILLO	hermosillo	29.072967	-110.955919	945.319940	158813.75	1137.25	678.50	1199.958333	201593.00
5	MME-300000-MME01-MERIDA	merida	20.967370	-89.592586	1928.985119	324069.50	2375.75	1467.25	1989.684524	334267.00

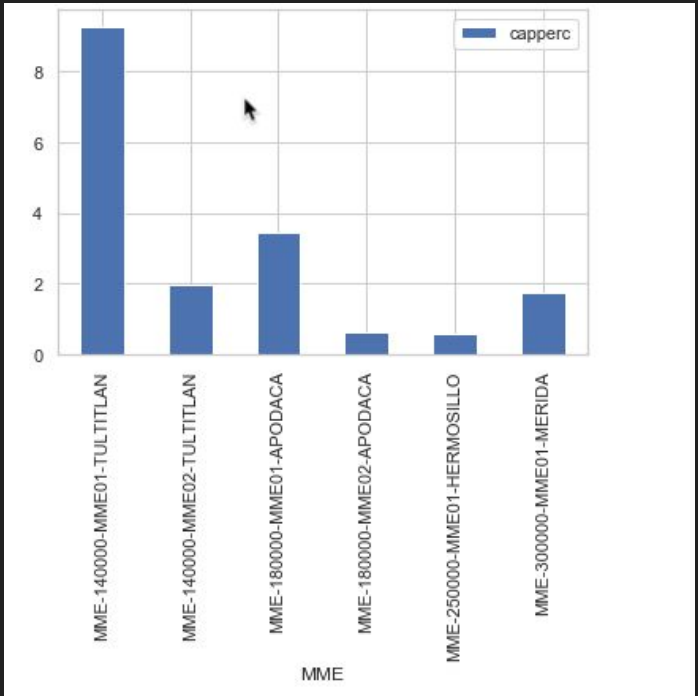
6 rows × 32 columns

What is the day with highest demand by MME



	MME	DATEDAY	Attached UE
0	MME-140000-MME01-TULTITLAN	2019-12-08	220623.75
1	MME-140000-MME02-TULTITLAN	2019-12-08	47176.00
2	MME-180000-MME01-APODACA	2019-12-08	50866.00
3	MME-180000-MME02-APODACA	2019-12-08	9186.00
4	MME-250000-MME01-HERMOSILLO	2019-12-08	1137.25
5	MME-300000-MME01-MERIDA	2019-12-08	2375.75

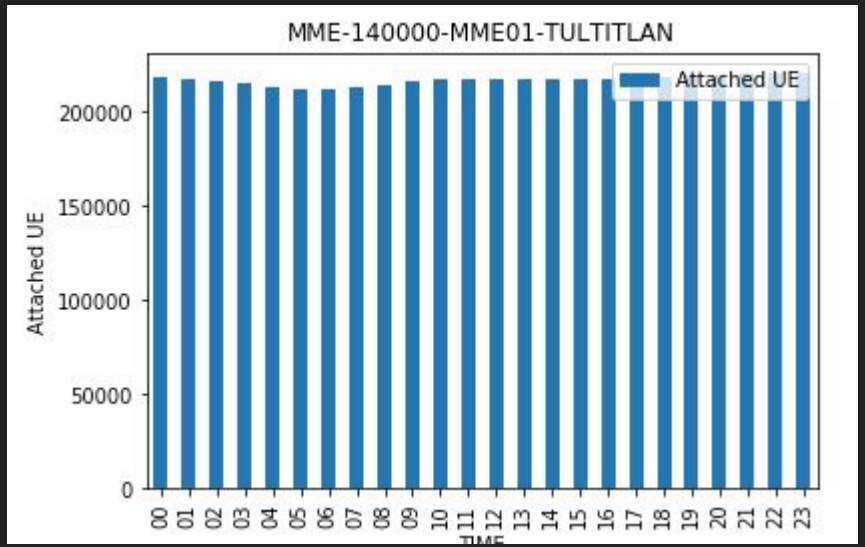
What is the utilization percentage by MME?

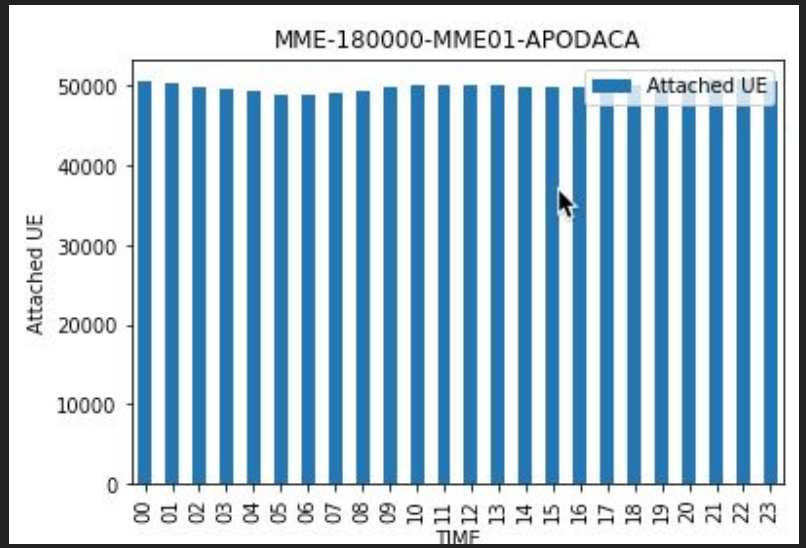


	MME	DATEDAY	capperc
0	MME-140000-MME01-TULTITLAN	2019-12-02	9.01%
1	MME-140000-MME01-TULTITLAN	2019-12-03	9.18%
2	MME-140000-MME01-TULTITLAN	2019-12-04	9.23%
3	MME-140000-MME01-TULTITLAN	2019-12-05	9.27%
4	MME-140000-MME01-TULTITLAN	2019-12-06	8.89%
5	MME-140000-MME01-TULTITLAN	2019-12-07	7.66%
6	MME-140000-MME01-TULTITLAN	2019-12-08	7.70%

What is the time with highest demand by MME

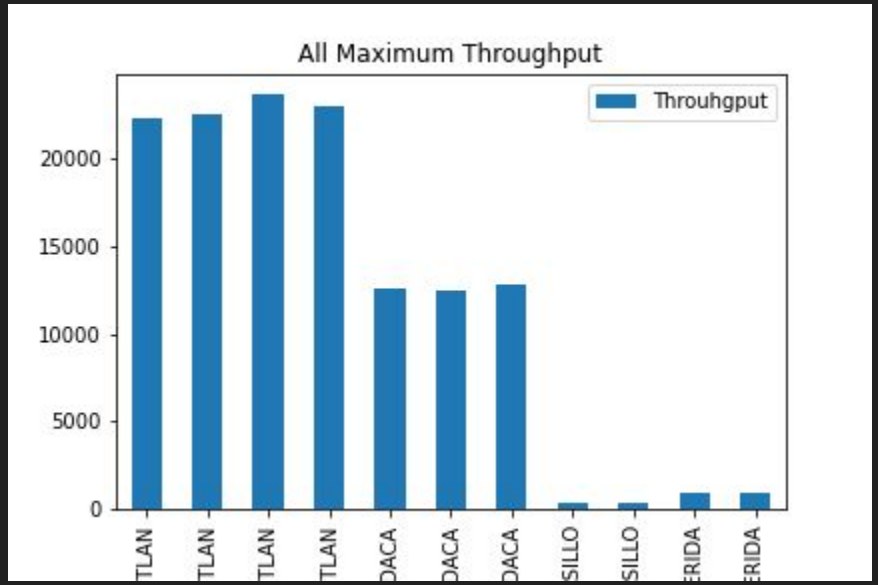
Equipments are turned off during nights.





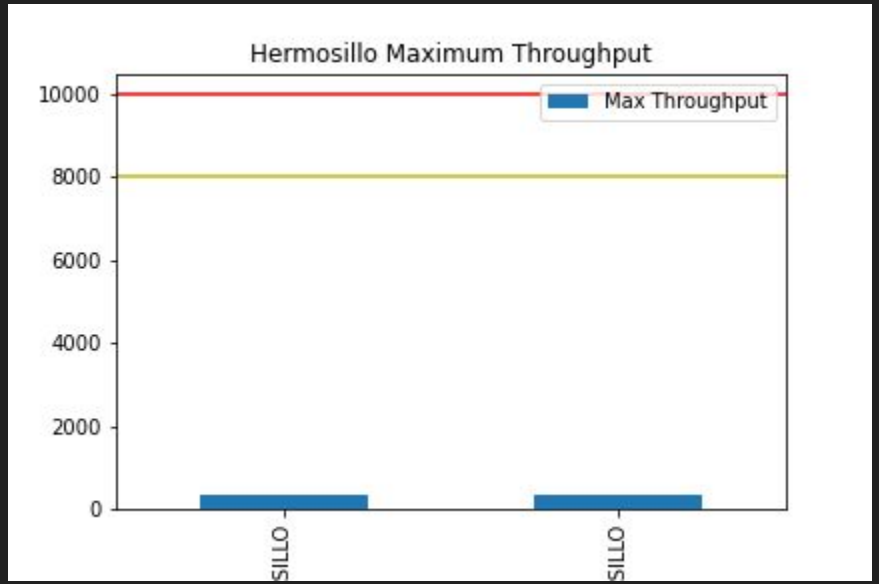
Are there any SAEGW in critical load?

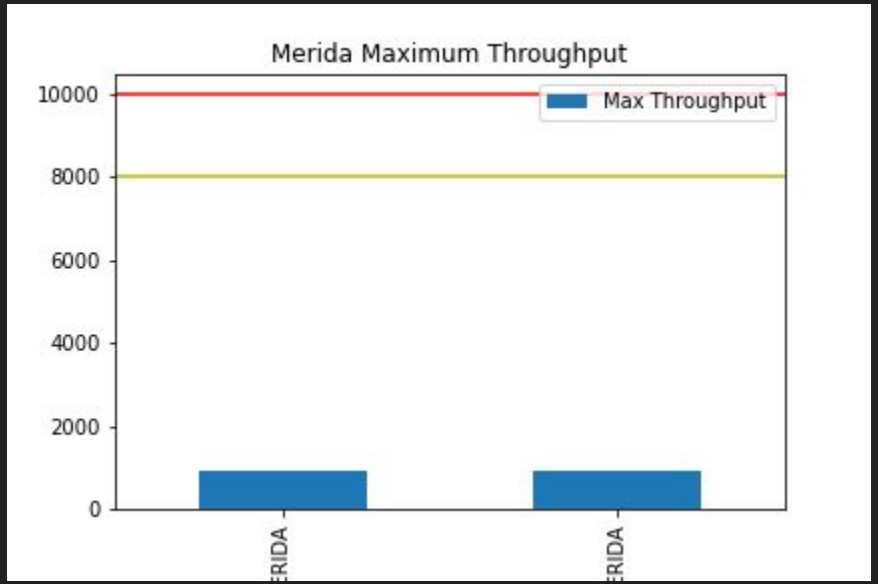
Knowing that a SAEGW at more than 80% load is considered critical.



Are there any SAEGW in critical load?

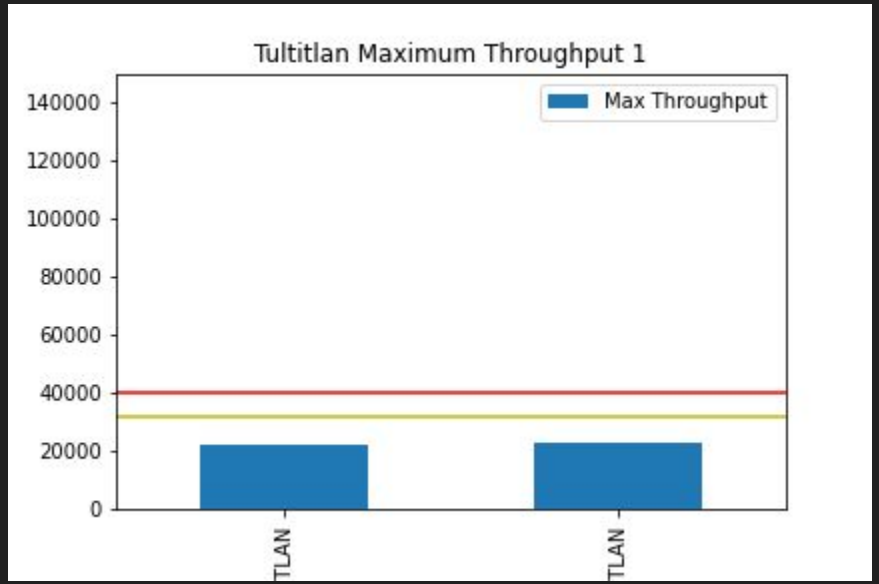
Knowing that a SAEGW at more than 80% load is considered critical.

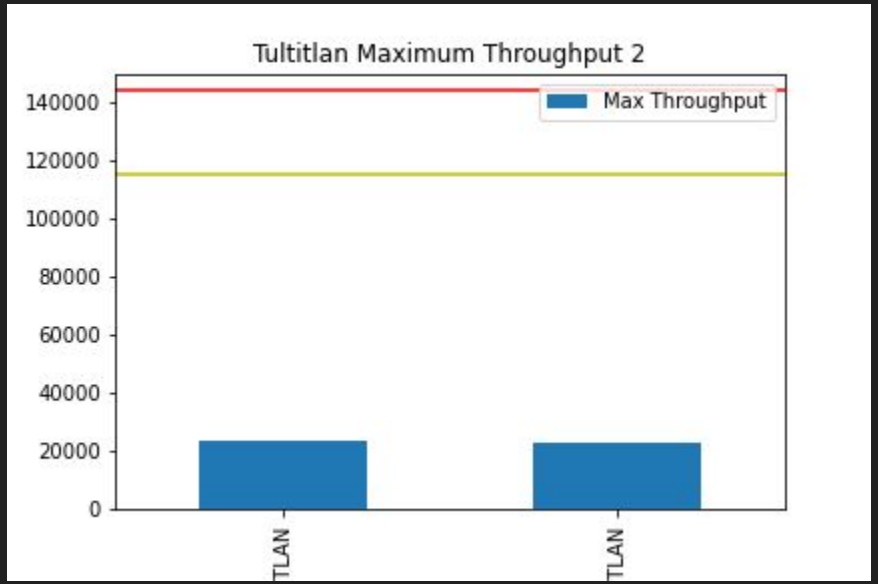




Are there any SAEGW in critical load?

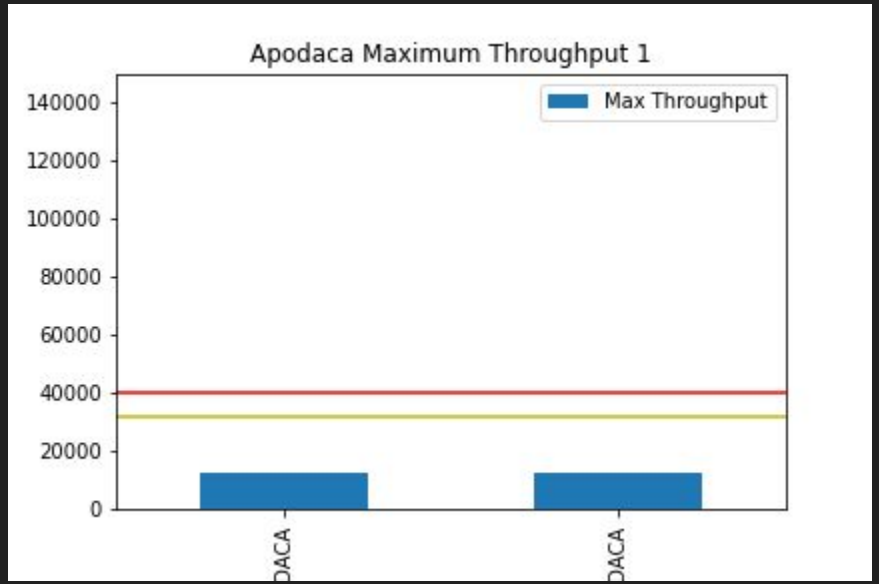
Knowing that a SAEGW at more than 80% load is considered critical.

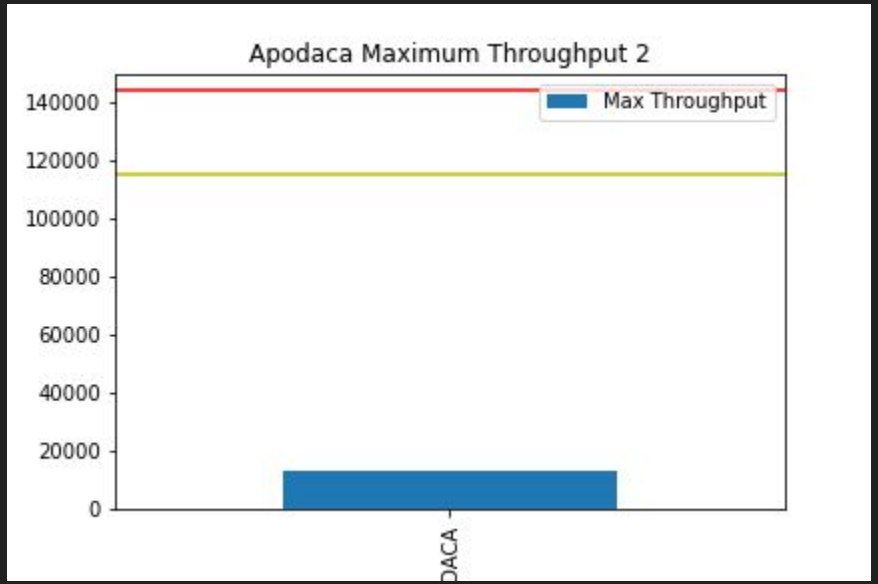




Are there any SAEGW in critical load?

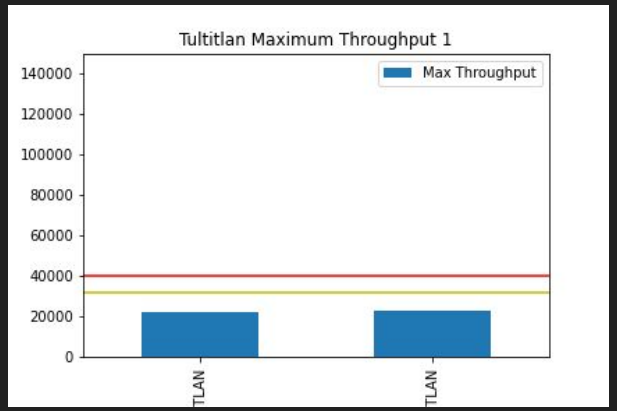
Knowing that a SAEGW at more than 80% load is considered critical.

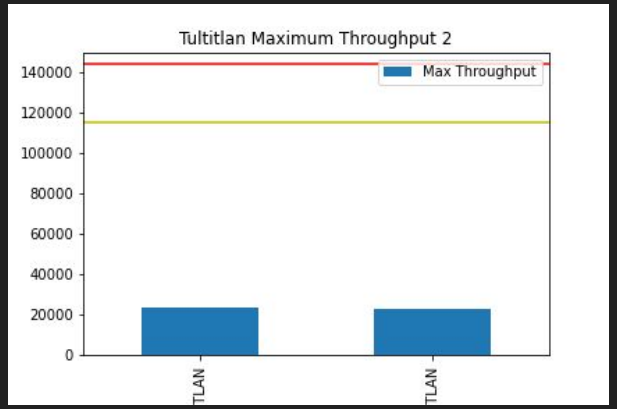




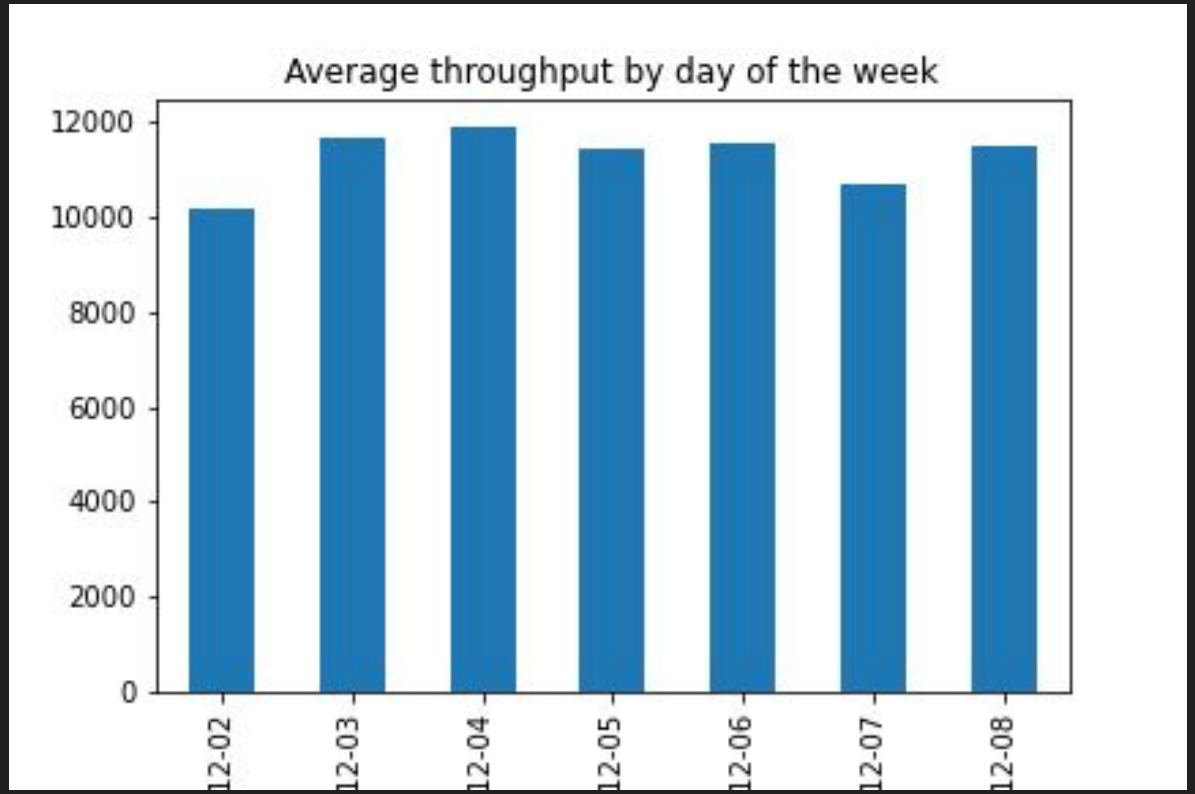
If there are any SAEGW way below critical, could we afford to shut one down?

	SAEGW	Throughput
0	140000-SAE01-TULTITLAN	22271.29
1	140000-SAE02-TULTITLAN	22482.07
2	140000-SAE03-TULTITLAN	23684.76
3	140000-SAE04-TULTITLAN	23028.11

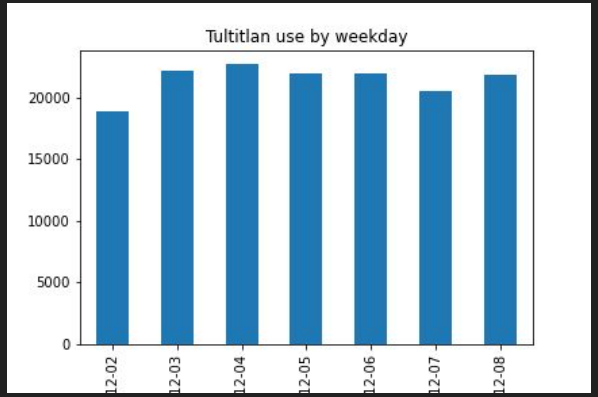


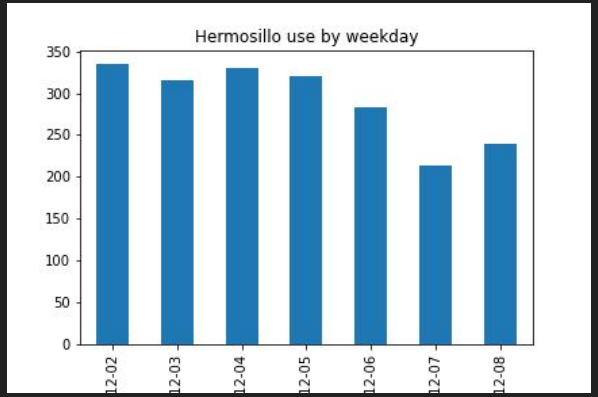


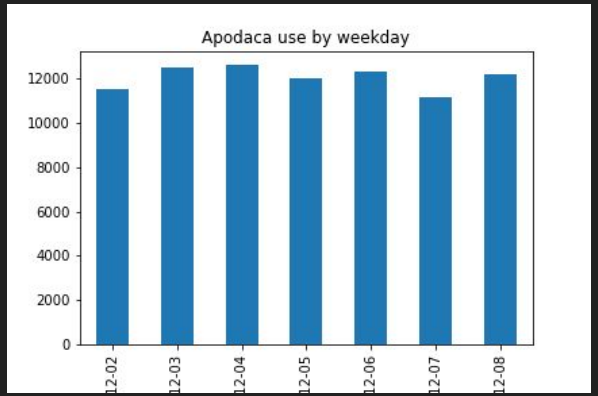
What day of the week gets the most SAEGW traffic?

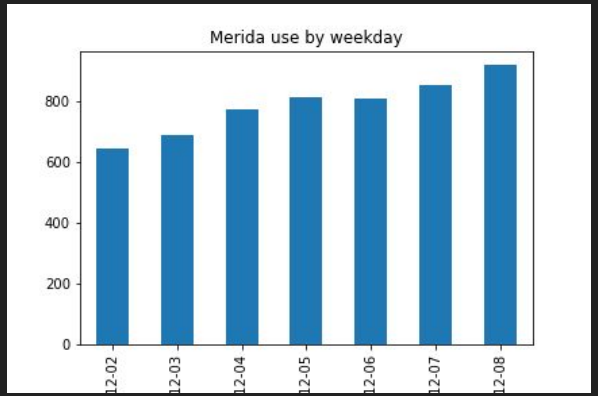


What day of the week gets the most SAEGW traffic?









Conclusions

- MME
  - Data is free of outliers
  - The most significant relationship in KPI’s is between Attached UE and Bearers
  - Bearer behavior can be predicted based on attached UE
  - Number of attached US grows everyday
  - Tultitlan is the city with the highest demand in Users
- SAEGW
  - There aren't any SAEGW above their critical load.
  - We could shut down many towers based on critical load, yet for redundancy we have to keep many of them
  - Every city consumes data differently throughout the week.