Table of Contents

ENERGY INFORMATICS FOR PUBLIC POLICY IN QATAR

Introduction:

According to AMIA [1], informatics is the science of how to use data, information, and knowledge to improve human health and the delivery of health care services. In public policy, it has hosts of application which are not restricted to just the medical profession. One such field is the field of energy which is part and parcel of our everyday life. But the focus of public informatics officer dealing with energy is not to ensure profits for the utility companies; that happens to be business related application of informatics that comes under the heading of energy business analytics. Instead, the goal of these professionals is to keep in check the energy consumption of the populace of which they are responsible for. The reason their work is imperative is because of two factors: the additional cost of each marginal unit of energy consumed gets translated collectively into the fuel imports which means increasing dependence on fossil fuels (in short, dependence on imports); secondly, each unit of energy consumed adds to the gross carbon footprint which is responsible for the climate change phenomenon.

Energy Efficiency Importance:

Global climate change poses a major threat to the health, economic prospects, and basic food and water sources of billions of people[2]. Negative effects of global climate change are already occurring, such as extreme weather events and reductions in global food supply[2]. While most people acknowledge this, they often forget that it is the very actions of the humans that has caused global warming leading to global climate change; this imputation has been well corroborated by reputed organizations such as NASA [3].

However, the umbrella of energy has a myriad of sectors thereby making it practically impossible for one individual (public policy maker) to address all the sectors at once; instead, they should commence by taking stock of the building sector first as they are at the epicenter of energy consumption. This is because they generate approximately 40 % of annual global GHG emissions [4]! Their eminence has attracted a lot of political and scientific pressure, which is why over the past 15 years, the evaluation of energy demand and use in buildings has become increasingly acute [5]. The endeavors that are being undertaken to curtail energy usage (and thus GHG emissions) therefore are myriad and diverse: these efforts are mostly technological or legislative that deal with the implementation of those technologies. For e.g., technical energy experts are developing cutting edge technologies, such as IoT [6] or Energy Star Equipment [7] to integrate in buildings; similarly, policy makers for e.g., are developing policies to successfully actualize these nifty tools inside the buildings: An example of this is the policy framework for Building Retro-Commissioning [8], Demand Response [9], etc. But before the latter can devise policies, they must take stock of the status quo first. To go about that they must learn data analytics techniques to make the most of the big data that is available to them.

Building Retrofitting & Qatar's Energy Policy:

This study showcases how informatics helps in benchmarking big data accrued by utility companies. Precisely, this is about exploring in depth, which kinds of buildings -among residential buildings – happen to have the highest potential for energy efficiency implementation. Therefore, all the work that is done in this project, will be to find avenues that could be shortlisted for the adoption of numerous techniques in energy sector that bring about energy efficiency (such as retrofitting, etc.). Investing in energy efficiency is analogous to killing two birds with one stone as it curtails GHG emissions and results in major savings (in thousands of dollars) with a relatively quick payback. And the reason why residential buildings were a target of this study was because these efforts will save money of the homeowners which aligns with the job of a public servant, which is to do good for the public.

To go about this task, actual energy consumption data of all residential buildings in Qatar was collected – spanning from 2017 to 2019 - from a utility company in Qatar called Kahramaa. It is the sole transmission and distribution system owner and operator for the electricity and water sector in Qatar. The reason for the selection of Qatar of the region of focus was because of the availability of authentic, well-recorded, big data and secondly, because Qatar is a country that heavily depends upon oil for its electricity generation. More importantly, Qatar's domestic energy consumption per capita is considered excessively high when compared to other developed nations[10] . In 2014 energy consumption reached 17 MWh per capita with 35 tonnes of carbon dioxide emissions per capita each year (International Energy Agency, 2018). Unlike the US, energy efficiency cannot be achieved in Qatar at the expense of electricity price hikes driven by oil prices because oil is not only abundant but electricity is provided to Qataris free of charge [11]. Such unique energy policies of Qatar, along with the fact that increasing population growth and income of Qataris will further enhance the rate of consumption, behoove energy policy makers to explore ways to bring about energy efficiency in such a unique and complex milieu of energy.

While there certainly are ways to motivate Qataris to buy or adopt renewable energy, they do not guarantee success. The cost of electricity would usually demoralize excessive energy consumption in typical households. Conversely, this is not the case for Qatari residents since electricity is free for so many. However, what will do wonders is retrofitting in buildings that are/will get flagged for bad energy performance. Corroborating from this paper 'Energy consumption and conservation practices in Qatar—A case study of a hotel building' [10] which asserts that retrofitting measures can save up to 49.3% energy compared to non-retrofitting measures (that can save 6.5 percent energy as per simulation results), it is established that one of the primary ways of curbing energy usage is through supplanting obsolete, deprecated equipment (such as lights, fans, HVAC systems, etc.). The replacement of inefficient electrical appliances and the establishment of stringent energy efficiency standards will also galvanize the populace there towards energy conservation practices. This to some extent solves another predicament which one of the paper [12] highlighted that access to free electricity had made Qatari residents resistant to change. This paper also maintained that since Qatar is bereft of energy policies incentivizing renewable energy adoption and no apprehension of any kind of economic ramifications of the energy overuse has made energy conservation a challenge in Qatar. Which is why this study becomes ever more important because it would serve as a rudimentary step in the quest of achieving energy efficiency pragmatically.

Use of Informatics:

The data available for this study was in an excel sheet format which comprised of 0.1 million rows, each denoting a distinct residential customer obtaining electricity and water from Qatar. Each customer was categorized into a heading called 'Customer Type' which entailed four categories, namely: Government Owned, Regular Customer, Qatari Owners, and Rented Premises by Qatari's. Their dwellings were categorized into premises names which denoted the types of residential buildings such as Villas, Houses, Flats, Beach Cabin, Villa, etc. The dwellings had also been assigned spatial coordinates in the data which could have been used for spatial analysis. However, since the geography of the country is such that it experiences similar weather conditions uniformly, obviating the need thereof.

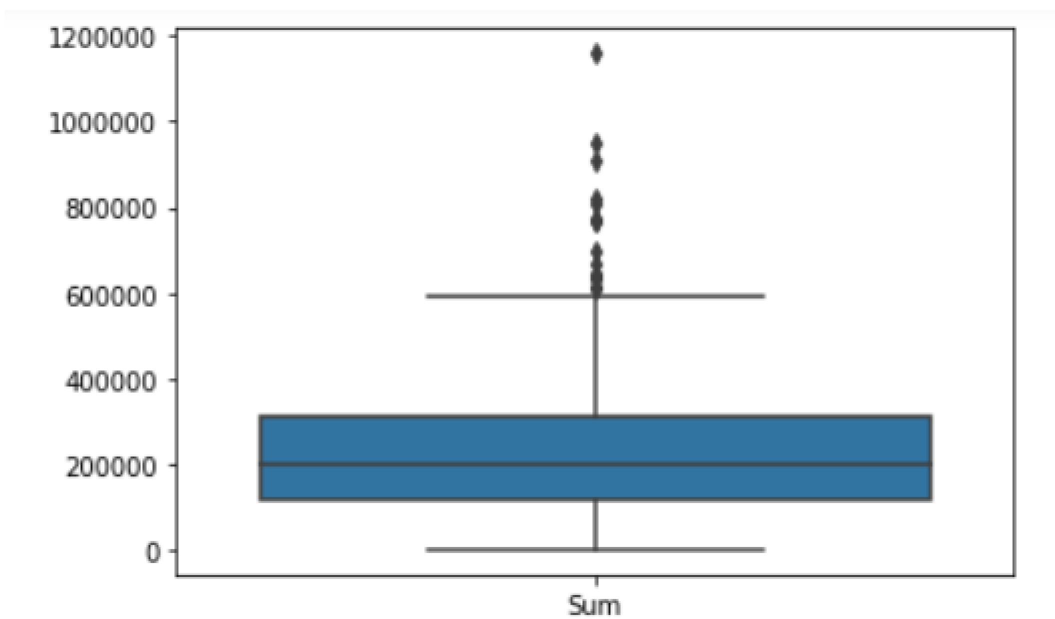
Of particular importance was the monthly consumption data (in kWhs) of each customer starting from January 2017 till December 2019. The entire work of this project revolves around the aforementioned attributes which were manipulated using two popular data analysis software called Python and Stata. It is worth noting that simply excel software would not have been enough for data analysis as it required preprocessing of the data – which would not have been done conveniently without using Python. Data for each year –

in monthly segments was recorded as single sheets for which joining was conducted in Python. Since the pivot tables of excel was not adequate to perform customized group-by functions – wherein each category was explored, compared, and analyzed in detail – Python Pandas came in clutch.

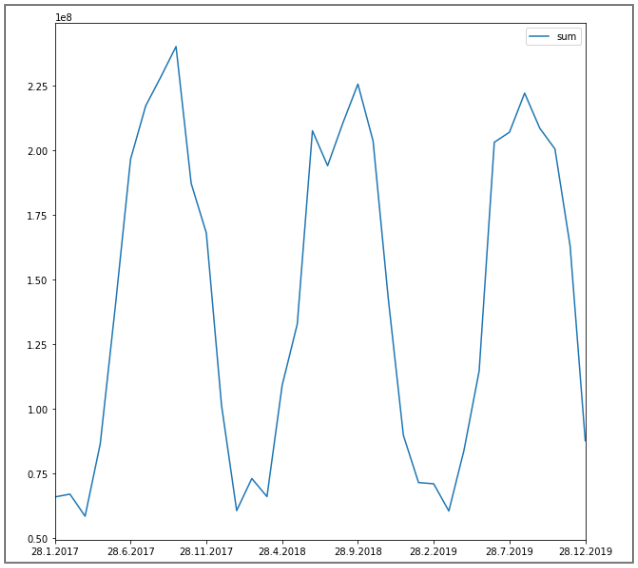
Thus, data was aggregated to give sum of all 36 months which was used as a target variable for most of the analysis, including hypothesis testing and regression analysis. This process of using informatics upon the data had three objectives: first to benchmark; secondly, to run regression; and thirdly, to conduct hypothesis tests (t-tests) to validate whether certain category of buildings is highly discrepant of energy consumption and whether it is worth concentrating energy efficiency practices therein or not.

⇨ **Benchmarking:**

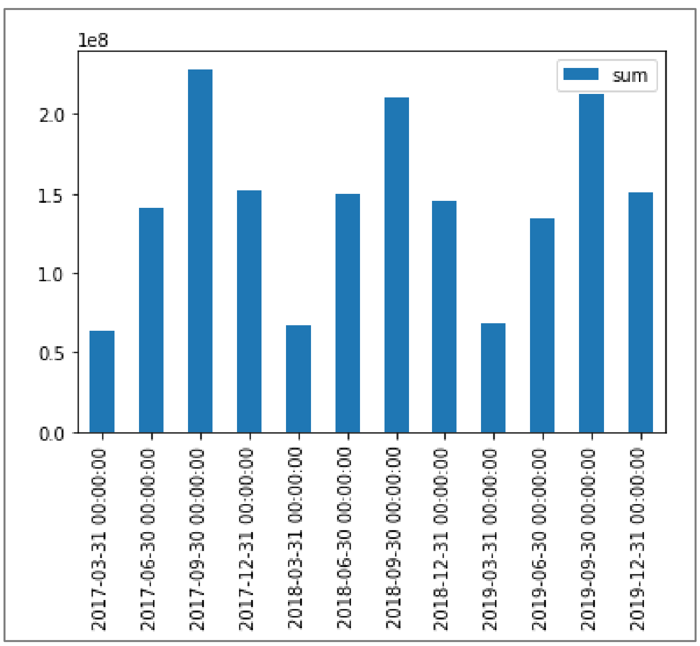
Before any benchmarking is done, it is essential to check for outliers as they can skew the mean consumption metric which is the most preponderant



metric used in this project to analyze different cohorts together. It was observed that the highest consuming month in all 36 months was the month of August in 2017 as shown in the diagram:



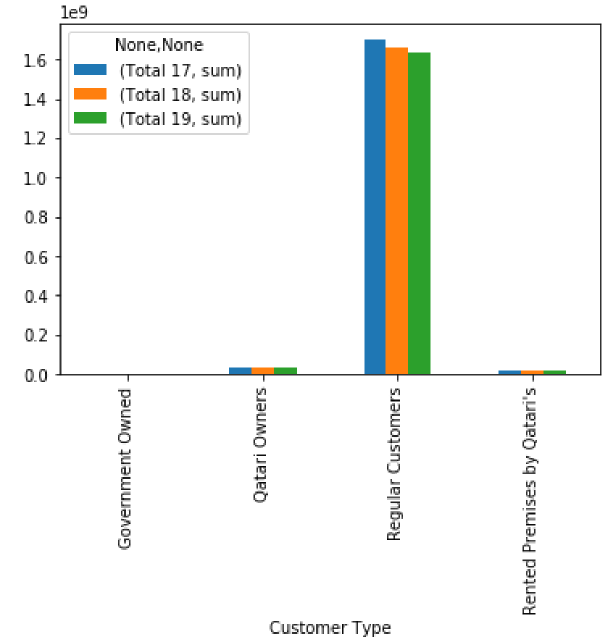
What is good to see is that overall peak energy consumption has been decreasing each August. What is also observed is the anomaly of increase in consumption in January 2017 and 2018, just before a quick reduction in February after which it goes on to hit a peak in August as usual. These two anomalies detected through simple visualization might go on to lead to energy savings if explored further. Unfortunately, with just one excel file, it is impossible to further delve into it.

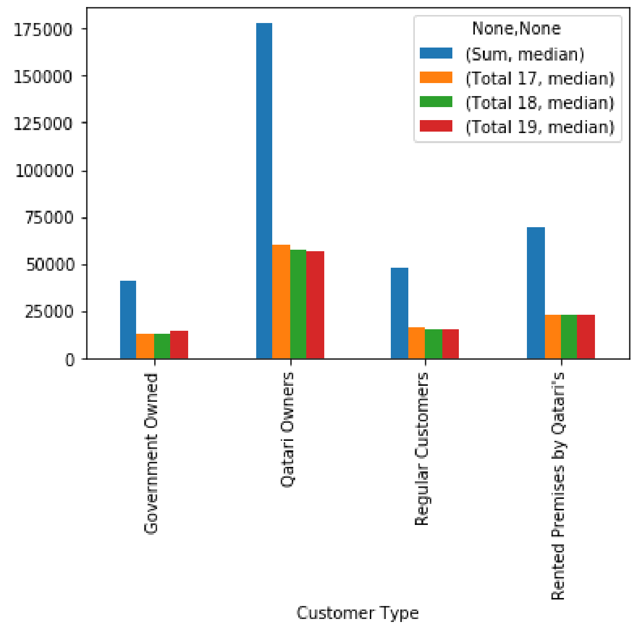


Informatics was also used to resample monthly data into quarters as they represent a holistic picture as shown in the left:

After exploring simple descriptive and detecting anomalies, aggregation was performed at Customer Type level to observe the trends of energy consumption between different types of residential owners.

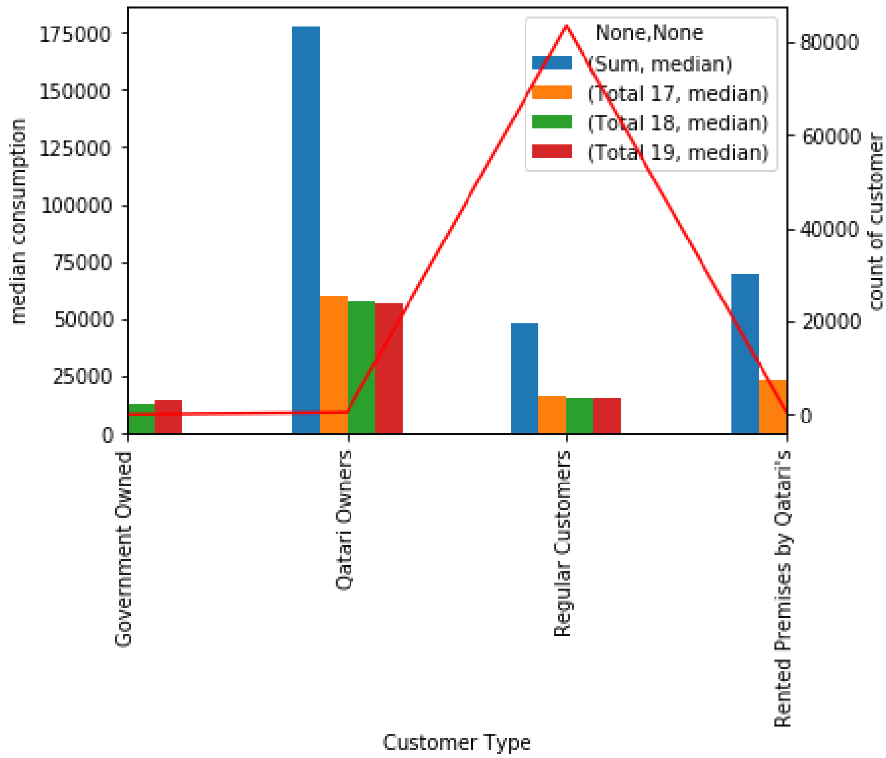
Upon doing so, two graphs were plotted each whereof shows a different picture. With the first diagram, one would thing that residential customers are the guilty party but upon using median values, true picture is painted:



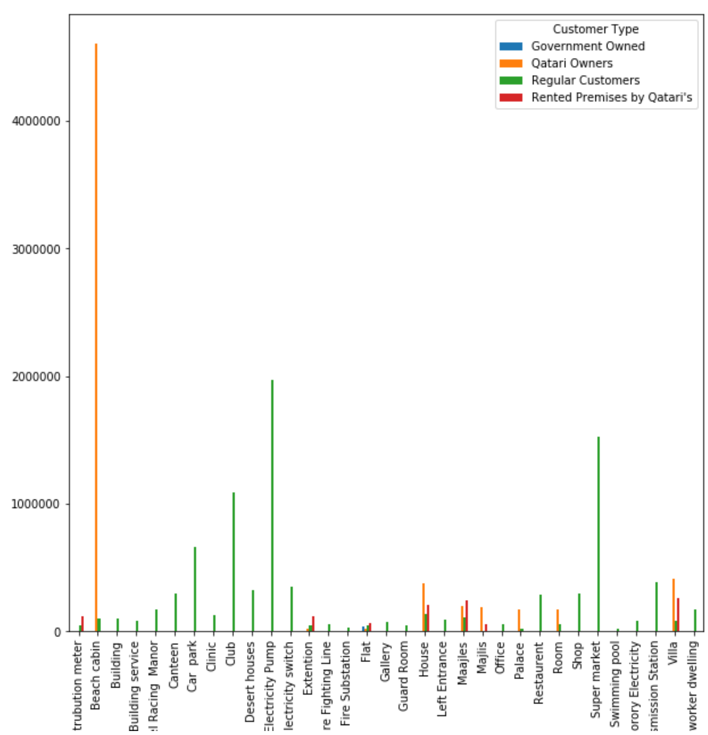


When median yearly consumption of all four groups is compared it becomes evident that Qatari owners have the highest median consumption, then come the Rented Premises by Qatari's whereafter, Regular Customers appear. This trend further becomes clear with the following dual axis graph:

It is seen that a few hundred Qatari and Rented Premises owners have a relatively higher median consumption whereas Regular Customers fare low in this bar.

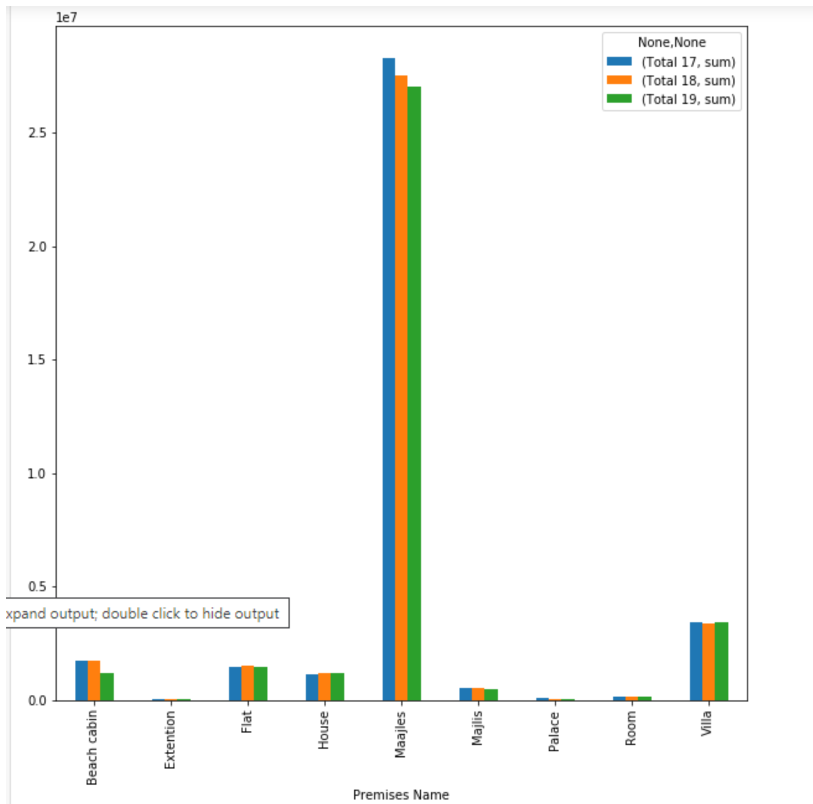


As it became clear that these groups are worth exploring, a pivot table was created to check for the premises of each customer type, that are consuming inordinate consumption. After which it was seen that Beach



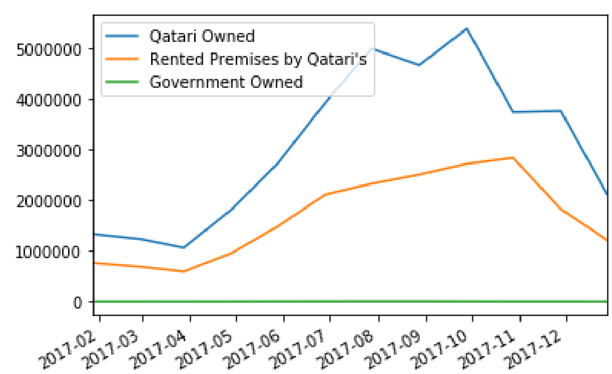
Cabin, House, and Villa had the highest median consumption in Qatari Owner's buildings. It became clearer that Qatari Owners should be made the object of focus for hypothesis as they showed anomalous values.

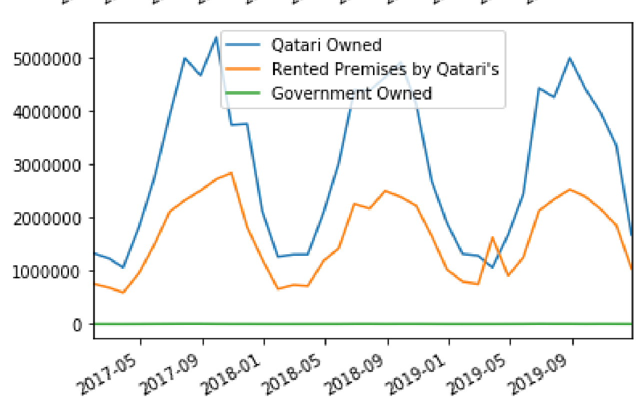
Thus, the data frame was filtered for Qatari Owners and now the total consumption of each of their premise was analyzed over the span of three years giving this figure: As evident in the figure, Maajles have the highest energy consumption followed by the Villas for Qatari Owners. This leads us to our first hypothesis.



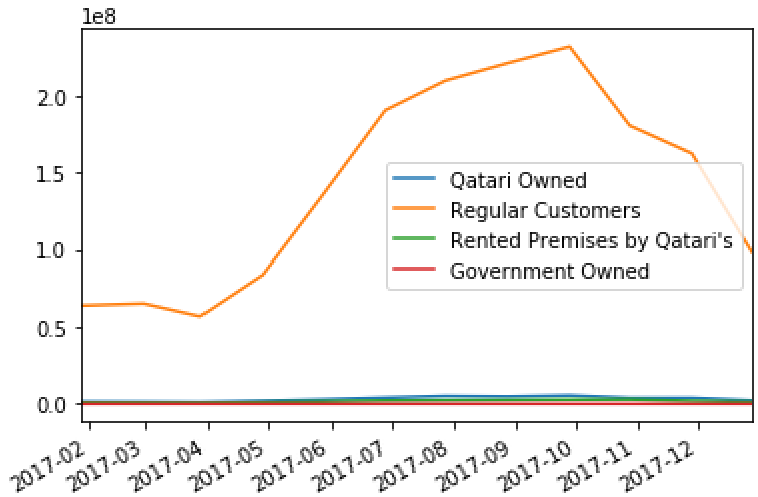
⇨ Hypothesis Testing I:

The first hypothesis was conducted between Qatari owner's majlis and Regular Customer's Majlis. A two tailed t test was performed between the two groups; there were 55 Majlis buildings for regular customers whereas there were 345 such buildings of Qataris. The t test gave a significance value of ***p =0.0015*** which rejects the null hypothesis that there is no difference between the two. Now, statistically it can be asserted using informatics that the energy consumption over the span of three years between Qatari owner's majlis and Regular Customer's majlis is significantly different making it worthy to be considered for energy efficiency practices. Upon analyzing Qatari owned buildings further by comparing their time series graph with respect to rented premises by Qatari's, it is found that the energy consumption pattern throughout the entire year is not as straightforward for Qatari owned buildings as it is for rented premises by Qatari's: After mid-year end, the energy consumption pattern for Qatari buildings is fickle in 2017 whereas Rented premises happen to increase until October after which there is a steady decline. Similar trend is experience for all three years denoting that Qatari buildings have anomalous energy consumption behavior.





On a side note, it should be kept in mind that the above visualizations exclude for Regular customers to better compare the two groups together. This is one of the critical component of data visualization where one visualization would fail to show the reality whereas another would show something starkly different. Upon including regular customers, timeline for the rest of the customer types essentially becomes linear line which is why it became necessary to weed them out.



Nonetheless, upon comparing Qatari Majlis to Regular customer's majlis, it becomes manifest that not only their energy consumption is substantially high, but it also follows the same fickle trajectory, characterized with ups and downs as for overall energy consumption pattern of all Qatari owned buildings, compared to Regular customers.

⇨ Regression:

Moreover, a bivariate regression analysis was done between customer type and sum of energy consumption to validate the hypothesis between these categories. To do this, since regression using categorial independent variables is much more convenient and meaningful in Stata, so it was employed. Upon conducting this test, it was found that in this model, relative to Qatari Owners, Regular Customer owned buildings had in average 171 MWh lesser consumption ($p < 0.001$) and Rented Premises by Qataris had on average, 143Mwh lesser consumption ($p<0.001$). This model was significant ($p < 0.001$) and accounted for 1.7 percent variation in the total energy consumption in the span of three years for residential buildings in Qatar.

```
regress Sum ib2.customertype

      Source |       SS           df       MS      Number of obs   =    84,666
-------------+----------------------------------   F(3, 84662)     =    488.94
       Model |  1.4147e+13         3  4.7156e+12   Prob > F        =    0.0000
    Residual |  8.1654e+14    84,662  9.6447e+09   R-squared       =    0.0170
-------------+----------------------------------   Adj R-squared   =    0.0170
       Total |  8.3069e+14    84,665  9.8114e+09   Root MSE        =     98207

----------------------------------------------------------------------------------------------
                         Sum |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-----------------------------+----------------------------------------------------------------
                customertype |
            Government Owned  |  -190499.4   69592.27    -2.74   0.006    -326899.7    -54099.1
               Qatari Owners  |          0  (base)
           Regular Customers  |  -171722.6   4566.915   -37.60   0.000    -180673.7   -162771.5
Rented Premises by Qatari's  |  -142933.1   5940.395   -24.06   0.000    -154576.2   -131289.9
                             |
                       _cons |   231534.4   4554.257    50.84   0.000     222608.1    240460.7
----------------------------------------------------------------------------------------------
```

After this, another regression analysis was done; this time it was a nested series of independent categorical variables named 'customer type' and 'premises name'. Results of multivariate regression analysis revealed that the variable customer type was no longer a significant predictor of the sum of energy consumption of buildings and none of the customer types had a significantly lower energy consumption than Qatari owners' buildings, on average, when controlling for the premises type of buildings. However, among the premise's names, all building premises had a significantly lower energy consumption than Beach Cabin on average ($p < 0.001$), when controlling for the customer type, signifying that it is worthwhile to explore the differences in premises types of buildings. This will lead to those type of premises that need the most attention for building retrofits. Overall, the model was significant ($p < 0.001$) and accounted for 7.8 % variation in total energy consumption of residential buildings in Qatar between 2017 and 2019.

One important point to understand about regression is that it is highly possible for a model to have high correlation, yet it is terrible. Conversely, it is possible that a model accounts for miniscule percent of variation in a dependent variable yet is worthwhile.

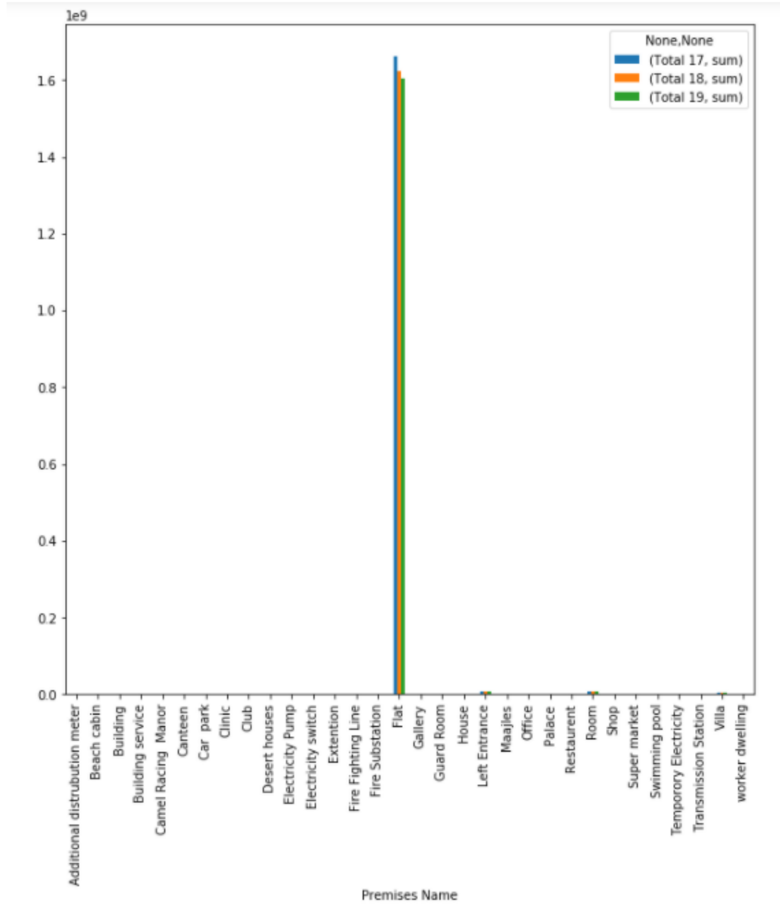
```
. regress Sum i.customertype ib2.pname

      Source |       SS           df       MS      Number of obs   =    84,666
-------------+----------------------------------   F(35, 84630)    =    205.13
       Model |  6.4960e+13        35  1.8560e+12   Prob > F        =    0.0000
    Residual |  7.6573e+14    84,630  9.0479e+09   R-squared       =    0.0782
-------------+----------------------------------   Adj R-squared   =    0.0778
       Total |  8.3069e+14    84,665  9.8114e+09   Root MSE        =     95121

----------------------------------------------------------------------------------------------
                         Sum |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-----------------------------+----------------------------------------------------------------
                customertype |
            Government Owned  |          0  (base)
               Qatari Owners  |   95868.24   67657.63     1.42   0.156    -36740.17    228476.7
           Regular Customers  |   17855.89   67261.21     0.27   0.791    -113975.5    149687.3
 Rented Premises by Qatari's  |   39545.34   67362.88     0.59   0.557    -92485.38    171576.1
                             |
                       pname |
Additional distrubution me..  |   -2256615   71433.83   -31.59   0.000     -2396625    -2116605
                 Beach cabin  |          0  (base)
                    Building  |   -2212013   95190.98   -23.24   0.000     -2398587    -2025440
            Building service  |   -1831607   72341.94   -25.32   0.000     -1973396    -1689817
          Camel Racing  Manor  |   -2144894   116555.9   -18.40   0.000     -2373343    -1916446
                     Canteen  |   -2015979   116555.9   -17.30   0.000     -2244428    -1787531
                    Car  park  |   -1653214   116555.9   -14.18   0.000     -1881663    -1424766
                      Clinic  |   -2187671   116555.9   -18.77   0.000     -2416120    -1959223
                        Club  |     396586   86909.93     4.56   0.000     226243.2    566928.8
               Desert houses  |   -1994030   116555.9   -17.11   0.000     -2222479    -1765582
            Electricity Pump  |  -342962.3   95190.98    -3.60   0.000    -529535.9   -156388.8
          Electricity switch  |   -1963432   116555.9   -16.85   0.000     -2191881    -1734984
                   Extention  |   -2260989   71179.87   -31.76   0.000     -2400501    -2121477
          Fire Fighting Line  |   -2215230   79667.73   -27.81   0.000     -2371378    -2059082
             Fire Substation  |   -2284282   116555.9   -19.60   0.000     -2512731    -2055834
                        Flat  |   -2257730   67360.31   -33.52   0.000     -2389756    -2125705
                     Gallery  |   -2251349   79667.73   -28.26   0.000     -2407497    -2095201
                  Guard Room  |   -2207249   82458.13   -26.77   0.000     -2368866    -2045632
                       House  |   -2101681   68765.62   -30.56   0.000     -2236461    -1966901
               Left Entrance  |   -2120128   67866.84   -31.24   0.000     -2253146    -1987109
                     Maajles  |   -2151065   67460.42   -31.89   0.000     -2283287    -2018843
                      Majlis  |   -2188200   75252.16   -29.08   0.000     -2335694    -2040706
                      Office  |   -2259889   86909.93   -26.00   0.000     -2430231    -2089546
                      Palace  |   -2256759   95120.56   -23.73   0.000     -2443195    -2070323
                  Restaurent  |   -2027240   95190.98   -21.30   0.000     -2213813    -1840666
                        Room  |   -2210671   67651.28   -32.68   0.000     -2343267    -2078075
                        Shop  |   -2022891   116555.9   -17.36   0.000     -2251340    -1794443
                Super market  |  -789034.3   116555.9    -6.77   0.000     -1017483   -560585.6
               Swimming pool  |   -2295283   116555.9   -19.69   0.000     -2523732    -2066835
       Temporory Electricity  |   -2228512   116555.9   -19.12   0.000     -2456961    -2000064
        Transmission Station  |   -1934098   116555.9   -16.59   0.000     -2162547    -1705650
                       Villa  |   -2144508   67734.52   -31.66   0.000     -2277267    -2011749
             worker dwelling  |   -2159826   77751.85   -27.78   0.000     -2312219    -2007433
                             |
                       _cons |    2298765   95191.24    24.15   0.000      2112191     2485339
----------------------------------------------------------------------------------------------
```

⇨ Hypothesis Testing II:

When comparing between different premises types, it was seen that flats had the highest consumption in all of Qatar. This is because there were around 85000 residential buildings in total, of which approximately 83000 were flats. Of the 83000 flats, there are 82891 flats owned by regular customers and the rest are owned by Qataris. Thus, a t-test was conducted to see for difference in the average total consumption of electricity between 2017 and 2019. The two tailed hypotheses gave a *p value = 0.49* which did not reject the null hypothesis that there is no significant difference between energy consumption in flats owned by both types of customers.



From policy perspective, this would serve as a unification point which can be used to implement standard energy efficiency protocol in all flats in Qatar thereby reducing energy consumption substantially. But when flats were analyzed further, it was found that median energy consumption in flats is highest for Rented Premises by Qatari's customer type. This anomaly means there is a possibility that both customer type's flats happen to consume energy significantly different than each other. Thus, a one tailed t-test was conducted between the two which gave a value of $p<0.001$ denoting that there is a significant difference in

	Sum		
	sum	median	count
Customer Type			
Government Owned	82070	41035	2
Qatari Owners	4344571	20992	73
Regular Customers	4889665606	48321	82891
Rented Premises by Qatari's	47065389	66186	615

the electricity consumption of flats owned by regular customers and those living in rented premises of Qataris.

⇨ Hypothesis Testing III:

To further bolster hypothesis II, another hypothesis was done using the following methodology:

	Sum
Area Name	
Al-Mansoura - Ben Derhum	490687459
Old Airport	334352081
Al-Mamoura - Ain Khaled - Al-K	293039239
Al-Wkeer	274656471
Najmaa	274461391
...	...
Rawda Rashed	474273
Musaimeer	405833
Abu-Thlouf - Al-Zobara	215644
Al-Jemelea	59375
Al-Jassra	35140

- First the area which had the highest energy consumption in Qatar was found:
- Then it was assessed which premises dominated Al-Mansoura – Ben Derhum

		Sum			
		mean	median	count	sum
Premises Name	**Customer Type**				
Flat	Government Owned	4.103500e+04	41035.0	2	82070
	Qatari Owners	9.826000e+04	98260.0	1	98260
	Regular Customers	6.562089e+04	57699.0	7405	485922677
	Rented Premises by Qatari's	8.578918e+04	84836.0	22	1887362
House	Regular Customers	5.106800e+04	51068.0	1	51068
Left Entrance	Regular Customers	1.731025e+05	32679.5	4	692410
Room	Regular Customers	5.375471e+04	56349.0	7	376283
Super market	Regular Customers	1.527587e+06	1527587.0	1	1527587
Villa	Regular Customers	4.974200e+04	49742.0	1	49742

- It turned out that this area was also dominated by flats (7430 flats)

- Of these 7430 flats, 7405 were owned by regular customers whereas, 22 were those that were rented out by Qataris.

- A one tailed t test was conducted between flats owned by regular customers in this area vs flats rented out by Qataris in this area

- *P value* turned out to be 0.007 denoting that there is significant increase consumption of energy for rented dwellings with respect to regular customer's flat in this area. This hypothesis compounds the last hypothesis.

Informed Policy Recommendations:

- **Recommendation 1:**
  - Based upon the tests and analysis above, it can be said that the discrepancy between different premises of buildings is significant enough to be given heed to. Since the approach of this project for bringing about energy efficiency is through retrofitting of old buildings with energy star (or any energy efficiency equipment), the above results can be used to state that flats must be subjected to a thoroughgoing process of energy audits and retrofits. Now, the spatial coordinates included in excel file will come into play as spatial mapping will help in locating the distribution of residential flats throughout Qatar.

- **Recommendation 2:**
  - Using nested regression modeling it can be maintained that the differences in energy consumption between different premises for different customer types are significant. Building upon the results of first hypothesis, as Qatari owned majlis have a significantly greater energy consumption than majlises owned by regular customers, it becomes necessary to implement stricter energy conservation protocol in their dwellings. It was also seen that the energy consumption pattern in a calendar year for Qatari Majlises was characterized by ups and downs; the anomalous decrease in their consumption in peak summer can be translated into an opportunity for partaking in 'Demand Response' with Kahramaa utility. Since the peak seasons require utility companies to fire up their peaker-plants to fulfill

energy demands, if premises like Qatari Majlises can be explored and brought in this program, the accumulated energy reduction during this season will indirectly prevent firing up extra peaker-plants; this will consequently save copious amounts of carbon emissions.

⇨ **Recommendation 3:**

- Last recommendation, based upon the results above, is intended to serve as a stepping-stone to energy efficiency in Qatar: It builds upon the analysis' methodology of finding and ranking areas that are most energy intensive. Hence, the recommendation that can be given to Kahramaa is to start with top 3 energy intensive areas and start retrofitting in the buildings showing exorbitant energy consumption. For e.g., the hypothesis done between flats of rented Qatari dwellings and regular customers dwellings in Al-Mansoura – Ben Derhum serves as a starting point. By devising energy conservation policies aimed precisely at the former type, the difference in energy consumption, which is statistically significant between the two can be reduced thereby saving energy as well.

Data Limitations:

While the data did seem comprehensive given its extent and precision, it had several limitations which makes it necessary to acknowledge certain caveats of the work done in this project. Usually, it is a norm to consider certain building related variables to perform an even-handed comparison of the buildings lest the comparison turns into what is typically referred as a comparison of apples and oranges. Since there were only few independent variables of the individual customers, controls were not implemented as comprehensively as they should have. This takes a significant toll on the validity of the analysis. Although the results of the analysis in this project are mostly statistically significant, one should learn the effect of controlling for independent variables while doing correlation or significance tests:

In this study, initially, through regression it was found that the variable customer type was a significant predictor of the aggregate energy consumption of residents. However, when this variable was controlled for the premise's names, it was observed that customer type did not remain statistically significant in predicting energy consumption. This serves as an alarm bell because it is likely that consideration of factors such as building square footages, vintage, and heating/cooling degree days would sway results to a different direction.

Unfortunately, this dataset was bereft of any such data of individual buildings adding shades of spuriousness to the analysis conducted. While this work was able to distinguish the categories that were peculiarly consuming more energy than their counterparts, it cannot be said that the analysis was evenhanded as it failed to consider the square footage of the buildings. For e.g., it is highly possible that the reason why Qatari Majlis clock higher monthly consumption than Regular Customer's Majlis is due to a higher covered area of the former. Thus, when the data gets normalized by this variable, it might end up denoting that Regular Customer's are, de facto, having a higher energy usage intensity which would turn the tables in our analysis.

Another speculation that can be used to criticize the analysis' results is the example of controlling for the building vintage to analyze residential building's energy consumption. It is highly plausible that Qatari owned buildings or Qatari rented dwellings are old construction with energy intensive equipment installed. For example, instead of state-of-the-art inverter air conditioners, window ACs are installed which are responsible for a higher median monthly consumption than Regular Customers. This becomes even more plausible given the fact [13] that population increase in Qatar has been steep over the past two decades giving rise to the construction of newer buildings that essentially would come into the domain of regular customers.

Moreover, this study has been predicated on the prospect of retrofitting in buildings. However, there is no data available that can be used to check which of the shortlisted building has undergone energy audit or retrofitting already. If such data is available it would streamline the work of this project effectively and perhaps, will also obviate much of the work that has been done in this project, hitherto.

The building vintage and renovation information is so indispensable that entire policies can be formulated based upon the findings from it. For e.g., with the limited information that was present, this study will perhaps suggest retrofits erroneously in buildings that are relatively new or have undergone renovation; however, these two variables (vintage and renovation information) will help in assessing which of the buildings are mandated to undergo retrofits given their construction year.

Similarly, with the goal of doing a just comparison, it was necessary to have information related to cooling degree days in Qatar for all three years. While it is possible that it might have only caused slight discrepancy in the results because climate of Qatar is almost uniform throughout the land, this kind of analysis in a different country/region would almost always require cooling or heating degree days to normalize energy consumption value; otherwise, the analysis would not hold any meaning.

While this project can be criticized in numerous ways, it must be remembered what the objective of the study was. With the goal of reducing energy consumption to curtail carbon emissions, it would have been quite meaningful to have data broken down by fuel sources and the consumption for each client, respectively. This is vital for the project since each fuel type has a different carbon emission value which must be analyzed. Eventually the metric that becomes a tool of comparison is 'carbon emissions per kWh' which then dictates whether a policy should be devised related to certain building stock or whether the current building stock can comply with a specific energy and environment policy (for e.g., reducing carbon emissions by 80 percent, etc.) with the existing energy efficiency tools.

For all the critique that one can come up with this study's approach, it was worthwhile analyzing building cohorts that were consuming most energy. This is because large buildings can now be shortlisted for getting contracted for the demand response program. Demand response provides an opportunity for consumers to play a significant role in the operation of the electric grid by reducing or shifting their electricity usage during peak periods in response to time-based rates or other forms of financial incentives [9]. However, this step is a secondary step, albeit its vitality. The reason being that Qatar's energy policy is yet to provide incentive/pricing policies related to demand response [14]. Nonetheless, once demand response is fully actualized, this study will prove essential. And this would further get complemented if the pricing data becomes available for each customer. This is another limitation of the study.

Lastly, one of the most important aspect that this study could not take advantage of was the aspect of machine learning to predict energy consumption. One of the basic approaches of using machine learning in such cases, where there is no data except time series of energy consumption, is to perform K-mean clustering. It is a form of unsupervised machine learning which is perfect for such cases, as the observations (not requiring or having any labels) can be sampled into cohorts to give noteworthy patterns (clusters) of energy consumption within a time frame. This would have been done with ease had data in segments of 15-minute consumption been available which would have been resampled to give accumulative energy consumption for each hour in 24 hours, for the span of three years (2017 to 2019). This kind of data is readily available in US and other developed nations because of smart metering technology in the world. However, this data set only possessed 36 readings of energy consumption (i.e., monthly consumption) which would not have sufficed for this kind of machine learning approach. The next step in this study is to collect 15-minute interval data for all customers of Kahramaa which would eventually lead to the detection of noticeable energy consumption patterns of residents of Qatar, in particular. Eventually these patterns can be analyzed to compare with patterns of residents from places that are known for energy efficiency to follow suit.

As the future steps of this study are laid down, the role of machine learning for energy policy should be underscored. As electricity markets involves a lot of scruples (in the shape of planning to meet ever increasing and fluctuating demands with margins of milliseconds), it would be a huge helping tool for the industry if reliable machine learning models can be made to predict energy consumption. This will serve as a benchmark for stakeholders involved in electricity markets leading to a more reliable and convenient process of electricity generation, transmission, and distribution. In terms of energy wastage, ML will help in deciding which plants can be shut off without the need of providing possible ancillary support (something which leads to gratuitous energy expenditure and carbon footprint). Thus, it would be useful to incorporate aspects of ML in the scope of this work.

Conclusion:

All in all, Qatar is a country where construction of new buildings upon stringent energy standards and building code will reduce more energy than renovating older buildings to conform to current standards [13]. Which means policy makers must assess which buildings are obsolete enough to be demolished and which ones should be commissioned. And while the policy environment of Qatar is not conducive for renewable incorporation, what remains a pragmatic approach is to implement renovations along latest technological standards to achieve maximum conservation [13]. Which is something that can be done using the informatics tool and the approach used in this study.

Bibliography:

[1] AMIA, "No Title." https://www.amia.org/fact-sheets/what-informatics.

[2] L. Steg, G. Perlaviciute, and E. van der Werff, "Understanding the human dimensions of a sustainable energy transition," *Front. Psychol.*, vol. 6, no. June, pp. 1–17, 2015, doi: 10.3389/fpsyg.2015.00805.

[3] NASA, "No Title." https://climate.nasa.gov/evidence/.

[4] A. 2030, "No Title." https://architecture2030.org/buildings_problem_why/.

[5] E. Delzendeh, S. Wu, A. Lee, and Y. Zhou, "The impact of occupants' behaviours on building energy analysis: A research review," *Renew. Sustain. Energy Rev.*, vol. 80, no. September 2016, pp. 1061–1071, 2017, doi: 10.1016/j.rser.2017.05.264.

[6] D. Sembroiz, S. Ricciardi, and D. Careglio, "Chapter 10 - A Novel Cloud-Based IoT Architecture for Smart Building Automation," in *Intelligent Data-Centric Systems*, M. Ficco and F. B. T.-S. and R. in I. D.-C. S. and C. N. Palmieri, Eds. Academic Press, 2018, pp. 215–233.

[7] E. Star, "No Title." https://www.energystar.gov/productfinder/.

[8] C. Buildings, "SEE Action Series : Energy Audit and Retro-Commissioning Policies for," 2012.

[9] O. of electricty-E. Gov, "No Title." https://www.energy.gov/oe/activities/technology-development/grid-modernization-and-smart-grid/demand-response.

[10] N. Ayoub, F. Musharavati, S. Pokharel, and H. A. Gabbar, "Energy consumption and conservation practices in Qatar - A case study of a hotel building," *Energy Build.*, vol. 84, pp. 55–69, 2014, doi: 10.1016/j.enbuild.2014.07.050.

[11] A. Meier, M. Darwish, and S. Sabeeh, "Complexities of saving energy in Qatar," *ECEEE 2013 Summer Study Proc.*, 2013.

[12] W. Al-Marri, A. Al-Habaibeh, and M. Watkins, "An investigation into domestic energy consumption behaviour and public awareness of renewable energy in Qatar," *Sustain. Cities Soc.*, 2018, doi: 10.1016/j.scs.2018.06.024.

[13] A. Kamal, S. G. Al-Ghamdi, and M. Koç, "Role of energy efficiency policies on energy consumption and CO2 emissions for building stock in Qatar," *J. Clean. Prod.*, 2019, doi: 10.1016/j.jclepro.2019.06.296.

[14] İ. Ş. Bayram and M. Koç, "Demand side management for peak reduction and PV

integration in Qatar,” 2017, doi: 10.1109/ICNSC.2017.8000100.