May 03, 2021

Exploratory Data Analysis of SACFA filing data of Telecom Service Providers and Model Fitting.

VJ CHRISTOPHER

JOINT WIRELESS ADVISER, WPC WING
Regional Licensing Office, Mumbai: 400067

Abstract

SACFA Siting is a term associated with the telecom network by which the base station transmitter or any telecom tower is validated through a clearing process instituted by WPC Wing, Department of Telecommunications. This is an online process done through an inter-ministerial coordination committee. The SACFA clearance is mandatory before the site can be deployed for operation in a network, by a telecom service provider.

In this paper, the SACFA filing data for the period 2014 to 2020 by the Telecom Service Providers(TSP) is studied to conduct an Exploratory Data Analysis(EDA) and then to understand its correlation with the AGR, Subscribers and Spectrum holding of the same TSP in this period. The OLS regression is conducted to fit a linear model to the data in order to find the association. The regression shows that that there is a linear association between the dependent and independent variables. Using the fitted model, an attempt is made to predict the SACFA filing for a future period.

The outcome of the study establish that the SACFA filing figure is a reflection of a TSP's revenue (AGR) and the subscribers it is holding in an LSA and so this figure can be considered as a key parameter of a TSP that represents it resilience and efficacy in that LSA.

Contents:

1. Introduction:-SACFA(Standing Advisory Committee Frequency Allocations) is a semi statutory body in WPC wing of Department of Telecom (DoT) that oversees the tower(mast) height clearance for installation of new antennas through an inter departmental coordination mechanism. The clearance not only ensures that the new antenna does not create radio interference to already existing antennas in its close proximity, but also examine the impact on the nearest airports (both civil and defence) such that the antenna tower does not obstruct the flight path and disturb the air traffic navigation to and fro from those airports. After obtaining the clearance from the SACFA, the antenna can be installed by the network operator and keep the site ready for operation.

The number of such SACFA applications filed is an indicator for the spread of the network of telecom service providers (TSP). In this project a comparison is made among 8 TSP s that operated in the period from 2014 to 2020 through an analysis of their SACFA filings. An Exploratory Data Analysis (EDA) of the SACFA record is conducted followed by a linear regression to see if the SACFA filing numbers has any association with the Revenue and Subscribers of the TSP as well as the Spectrum holding of the company. The later term, 'Spectrum holding' used in this paper is meant only the quantum acquired by the TSP in that particular year and should not be confused with the total spectrum holding of that TSP.

2. Objective: - To explore whether any of the factors, namely (i) Spectrum allocation to a TSP in each of the LSA, (ii) the aggregate gross revenue (AGR) from a LSA and (iii) the subscribers in that LSA influence the filing of SACFA applications in a particular LSA for the period 2014 to 2020. The objective is to find if any of the three explanatory variables has any association with the response variable, the number of SACFA applications.

3. Background: -It is quite rational to assume that the revenue earned by a TSP is used to expand its network for more geographical reach. This can result in more number of subscribers which in turn may cause the network to be expanded further with an increase in the number of SACFA filings. The Spectrum acquired by the TSP is another factor that could set off SACFA filings. This theory is to be tested in this paper. These points are elaborated below.

When spectrum is allocated to a TSP either through market methodologies or otherwise, there are some conditions to be fulfilled by the TSP in the subsequent years after the allocation. The spectrum is acquired by paying its price, it is not sensible for the company as well as the Government to waste the resources like Spectrum, but to put the same to optimum use. The new spectrum acquired will be used either to complement the working Base Transmitting Station(BTS) or will be used to deploy additional BTS depending on the network planning of the TSP. Either way,

the network will be upgraded with additional capacities.

It is a regulatory requirement that the any new BTS should get the SACFA clearance under the full siting category if it is a new tower or should file application for additional antenna category, if the spectrum is likely to be used under a new BTS antenna in an existing tower. There is a third possibility that the spectrum will be utilized using the existing BTS and antenna, say, with a revised network planning that there may not be any requirement for any SACFA clearance at all.

The important contractual obligation for a TSP is the roll out obligations mandated in the Unified License while acquiring spectrum. The roll out conditions make it mandatory to meet the urban as well as the rural deployment of network covering the minimum geographical area as assigned in the contract. Therefore, spectrum allocation can result in obligatory filing of SACFA applications in particular area.

The second factor is the AGR, the revenue earned from the particular LSA. This is the actual earning of the TSP. It is a reasonable assumption that a part of the revenue earned can be diverted to be used for expanding the network for improving the quality of services in an area. This could produce a demand for more BTS in that LSA and that would finally result in more number of SACFA applications. This points to a linear relation within a LSA that increase in AGR would give rise to a proportional increase in the number of the SACFA applications.

The third factor is the number of subscribers in a particular LSA. If you want to serve more number of consumers, then that would generate a demand for more spectrum and more number of BTS. The consumers continue to demand quality service which could be provided only with the above resources. Therefore, this can eventually cause the growth in the number of SACFA applications filed in a LSA. This variable is therefore selected as one of the explanatory variable.
It may be worth noting that all the variables selected are quantitative. Regression is the natural tool for testing relation between quantitative response variable and quantitative explanatory variable.

4. The research question:-

The hypothesis is that the number of SACFA filing by a TSP in a particular LSA is proportional to the spectrum acquired, the AGR of the company from that LSA and the number of subscribers in that LSA, in the context of the above argument. This paper is examining if the hypothesis can be proved from the SACFA data available during the period 2014 to 2020. Therefore, the research question can be formulated as follows:

"In a telecom service area, the number of SACFA filing by a particular telecom service provider in a year is correlated to the spectrum assigned to the TSP, the AGR of the TSP and the subscribers existing in that LSA "

The Null hypothesis is that the SACFA filing has no relation with any of the three explanatory variables. The alternate hypothesis is the SACFA filing is correlated to (AGR, Subscribers and Spectrum).
The OLS regression will be used to test this hypothesis using the available data

5. Data Collection: - The period of study was chosen 2014 to 2020 because of the ready availability of the data from the TRAI website for this period. Further, Reliance Jio Infocomm Ltd (referred to as JIO subsequently), the major telco in terms of the customer base came to the market after the auction of 2014.JIO was holding the Broadband spectrum (2300 MHz band) that was acquired in the year 2010 in 3G auction. But the technology was not matured to provide the service using the 2300 MHz spectrum until the VOLTE (voice over LTE) was cleverly deployed by JIO to commence LTE service in India. This was an important step by JIO that kick started LTE service straightaway unlike the other TSP who migrated from its lower cousins 2G and 3G. Although JIO launched the VOLTE in the year 2016, the SACFA filing has started before that period as seen from the data. The period chosen also cover the growth of 3G and LTE networks and also a period of downward trend in the use of 2G.

The key performance indicator report for each quarter of the year was downloaded from the website. From this report, the relevant indicators were extracted –the AGR data, Subscribers data, the Revenue data and the call drop data – for the period. This data was pre-processed for each of the TSP for each Service area. Then it was converted into yearly data. As for the spectrum data, the spectrum is distributed only through auction since the year 2012 onwards. This data is also available in the public domain at https://dot.gov.in. This was also downloaded, filtered and pre-processed to fit into the format for this paper.

In addition to the auction, spectrum has also been traded in the year 2016 and 2017. This was also taken into consideration for this project.

The Spectrum variable is processed on the logic that the spectrum won in a particular year would be used for SACFA filing after a time period of formalising the spectrum allotment by the WPC Wing. After each auction, WPC also conducts spectrum harmonisation exercise to have contiguous spectrum among the telcos. Therefore, it will be prudent to assume from a facile perspective that the spectrum acquired in a particular year would be utilized for SACFA filing in the next year. Though this assumption may not be held true always, but to a larger extent, the SACFA filing would follow this pattern. With this assumption, the following attribution has been assumed in the data collected:

Spectrum assigned in 2012 and 2013 (only 800 MHz): to be attributed to 2014 SACFA filing.

Spectrum assigned in 2014: to be attributed to 2015 SACFA filing.

Spectrum assigned in 2015: to be attributed to 2016 SACFA filing.

*Spectrum assigned in 2016: ½ to be attributed to 2017 SACFA filing

: ½ to be attributed to 2018 SACFA filing.

Spectrum assigned in 2017: to be attributed to 2018 SACFA filing.

Spectrum assigned in 2018: to be attributed to 2019 SACFA filing.

(*The spectrum auctioned in November 2016 has been assigned only in latter half of 2017. Therefore, the attribution has been divided equally into 2017 and 2018. There has not been any auction or trading in the year 2019 or 2020, the spectrum quantity during these years is zero only)

The number of SACFA applications filed was obtained from the WPC records. The raw data were cleaned and sorted by each TSP, LSA and finally tabulated year-wise.

As a summary, the following data is available for the EDA exercise:

(i)Telecom Service Providers: Total 8.

1. Aircel
2. Bharti Airtel.
3. Idea
4. Reliance Communication Ltd.
5. Tata communication
6. Jio
7. Vodafone
8. Vodafone Idea.

(The Aircel and Dishnet data were merged and renamed as Aircel. The Reliance Comm and Reliance telecom data merged and renamed as Reliance. The Tata Communication and Tata Teleservices Maharashtra Ltd (TTML) were also merged and renamed as Tata. The Idea, Vodafone and Vodafone Idea are treated three distinct TSP s.)

(ii)The Service Areas (LSA):

The 22 LSA are falling into 4 group as Metro, Group A, Group B and Group C service areas.

Metro= [Mumbai, Delhi, Kolkata]

Class A= [Andhra Pradesh, Gujarat, Karnataka, Maharashtra, Tamil Nadu]

Class B= [Haryana, Kerala, Madhya Pradesh, Punjab, Rajasthan, UP (East), UP (West), West Bengal]

Class C= [Assam, Bihar, Himachal Pradesh, Jammu & Kashmir, North East, Orissa]

6. The limitation of the study: - All the operators (TSP) who played major role in the telecom field in India were selected for the study. There are 8 TSP s selected for the study period from Jan 2014 to Dec 2020. In 2014, there were few more TSP s in addition to the eight selected like -MTS, Videocon, Telenor etc. However, these TSPs were either merged or quit the telecom business in India slowly. Therefore, they were not considered. But companies like Aircel and TATA had provided service for quite substantial period and had a considerable geographical coverage till recent time. When we come to 2020, there are only three service Providers left-Bharti, Jio and Vodafone Idea. Due to this reason, consistent presence of all TSP s during the whole period is not guaranteed, but in general the data chosen is expected to assist the study in such a way that this study will be able to extract patterns, if any, as evidence to prove the hypothesis.

The AGR data and subscriber data in the TRAI records are maintained in three months quarterly period. This data is suitably merged to transform it to annual data after taking care of the missing data by substituting it with the previous quarter data for few operators. But such instances are found in very rare cases only.

It should be mentioned that the variable CDP is the call drop reported by the Service Provider to TRAI which is also collected as part of the TRAI data. The maximum permissible CDP is 2% as mandated by the service standard prescribed by TRAI. A cursory glance thru the data shows that for all the periods, the CDP exceeding the ceiling is almost negligible. It is therefore concluded that the SACFA filing cannot be correlated to this variable CDP. Therefore, this variable is not given any consideration in the analysis.

The list of variables:

(i) The Response variable:

1. The number of SACFA applications filed(SACFA_REF_NO)- Quantitative (Integer)

(ii) The explanatory variables:

1. TSP –Categorical
2. LSA-Categorical
3. YEAR-Categorical
4. Subscribers – Quantitative (Integer)
5. AGR-Quantitative (Float in Rs. Crore)
6. REVENUE(REV)-Quantitative (Float in Rs. Crore)
7. Call Drop Percentage(CDP)- percentage(Float)
8. Spectrum –Quantitative (Float in MHz units)

7. Literature review: -

(i) SCAFA Clearance:- The Standing Advisory Committee on Radio Frequency Allocations (SACFA) has been constituted as per the Department of Telecommunications order dated 05.02.1966, with the Secretary (Telecom) being the chairman and the Wireless Advisor to the Government of India as one of the Member. The Joint Wireless Advisor to GoI is the Member Secretary of this committee. The membership of this Committee is open to all major wireless user Ministries and Administrative Departments of the Government of India. The Committee may co-opt other advisers/members as necessary when specialised questions are discussed. The function of the Committee will be to make recommendations on major frequency allocation problems requiring coordination amongst the various wireless users in the country and as such, referred to the committee by the administrative Ministries/Departments of the Government. The Committee is advisory in nature and the final responsibility for decisions and related implementation will remain, as hitherto, with the Department of Telecommunications.

The main functions of the committee, during its inception, were to make recommendations on:-

1. To sort out the problems referred to the committee by various wireless users, siting clearance of all wireless installations in the country, etc.
2. Issue of SACFA clearances for fixed wireless sites after coordinating with the members.

To meet the above mentioned objectives, the SACFA secretariat issues a clearance for the installation of any new fixed wireless site (FWS) within the country. The SACFA clearances are issued after also getting a ‘no objection' from the various SACFA members who carry out detailed technical evaluations, including field surveys. The technical evaluation carried out by the field units of the various member ministries is done primarily by taking into considerations the following parameters-

a) The FWS shall not cause any Aviation hazard.

b) The FWS shall not cause an obstruction to line of site of existing/planned networks.

c) The FWS shall not cause Interference (Electro Magnetic Interference (EMI)/Electro Magnetic Compatibility (EMC)) to any existing or proposed networks.

There are mainly three types of clearances for the towers:

a) Full Siting Category

b) Mast Height Category

c) Additional Antenna Category

The first two categories are distinguished by the power of the transmitters employed for the antenna and height of the towers. The third category is the additional antenna proposed in an existing site which has been cleared with reference to the height of the mast.

(ii) SACFA for telecom companies: - The first stage in the network planning is acquiring the spectrum. The company do network planning and as part of their plan, they have to establish Base station (BTS) within particular geographical area. Once the sites are determined, then next step is to file SACFA applications. As this is a time taking process, generally, the TSP s applies for SCAFA in a bulk in the SACFA online site. Once the clearances are issued by SACFA secretariat, they can deploy those towers and complete the network.

(iii) Roll out liability clause:-As part of the license contract, the telecom companies have to be compliant to the roll out obligations. Rollout obligations to cover a particular area in an LSA as per the conditions attached with acquiring spectrum through auction: (The following conditions are extracted from the Notice Inviting Applications, precursor to the Auction of spectrum in March 2021)

"For 700 MHz, 800 MHz, 900 MHz & 1800MHz bands (a) In LSAs other than Metro LSAs. The following rollout obligations for non- metro LSAs have to be fulfilled by the bidders acquiring spectrum through this auction: -
Phase 1: Coverage of 10% District Headquarters (DHQ)s/ Towns by the end of first (1)year from effective date of Notice Inviting Applications or date of assignment of spectrum won in this auction process whichever is later.
Phase 2: Coverage of 50% DHQs/ Towns by the end of three (3) years from effective date of license or date of assignment of spectrum won in this auction process whichever is later.
Phase 3: Coverage of 10% Block Headquarters (BHQ)s by the end of third (3) year from effective date of license or date of assignment of spectrum won in this auction process whichever is later.
Phase 4: Coverage of additional 10% BHQs (Cumulative 20% BHQs) by the end of fourth (4) year from effective date of license or date of assignment of spectrum won in this auction process whichever is later.
Phase 5: Coverage of additional 10% BHQs (Cumulative 30% BHQs) by the end of fifth (5) year from effective date of license or date of assignment of spectrum won in this auction process whichever is later. "

The following is mentioned in the roll out of obligations.

"At present, minimum sixty (60) calendar days duration is prescribed for grant of SACFA clearance. While examining the compliance of rollout obligations, the maximum delay in grant of the SACFA clearance beyond the above mentioned prescribed duration shall be excluded from the duration set for rollout obligations. For the purpose of calculating the delay in grant of SACFA clearance of individual Base Station site, the date of 'WPC acceptance number' for SACFA clearance application would be treated as the 'start date' and the date on which the SACFA clearance is granted in the online system of WPC would be considered as the 'end date'. The difference between the 'start date' and 'end date' after excluding the above mentioned prescribed period would be considered as the delay in grant of SACFA clearance for that Base Station site for the limited purpose of calculating the delay in compliance of rollout obligations. If the SACFA clearance for a particular Base Station site has been granted within prescribed duration, then for that particular site, delay would be considered as 'zero' days. "

All TSPs acquiring spectrum has to fulfil those obligatory requirement in the different period of time spell. A proper planning to deploy the BTS is required to avoid the penal clauses. Therefore there will always be urgency for the clearance of SACFA sites for telecom companies.

(iv) Spectrum allocation in auction:- The spectrum will be allotted and the specified frequencies won by the TSP will allocated by the WPC Wing after completing the formalities for payment of spectrum fee etc. This may generally take 30 days after the auction process. With the spectrum rights are acquired by the TSP, the network planning and SACFA process can commence.

The author could not find any work in this particular field of research, involving SACFA previously done for any period, though there are technical papers related to the spectrum allocation and the mobile services in India.

8. Explorative Data Analysis: It is an exercise to unveil the trends and patterns from the data using statistical means often with the help of different plots. The following explorative analysis is done on the SACFA data.

1. First, the statistics of the number of SACFA filing for each year as per each TSP:

SACFA filing	2014	2015	2016	2017	2018	2019	2020
Aircel	12755	25804	19120	4422	187	Nil	Nil
Bharti	71206	132197	258216	246032	231738	159058	66488
Idea	46447	92121	99987	118465	64125	Nil	Nil
Jio	73183	183785	153514	340379	312499	243332	39271
Reliance	3928	3543	5152	842	Nil	Nil	Nil
Tata	10844	15063	10522	1479	317	Nil	Nil
Vodafone	66845	75764	87429	117256	55326	26	Nil
Voda Idea	Nil	Nil	Nil	Nil	29454	183938	47864

Please note that the data shows the total number of applications filed in a year across all LSA s.
It can be seen that the Bharti and JIO has the maximum number of filing.

2. The AGR data by taking mean across all LSA for each year is shown here. The unit is in Rs. Crores. The AGR of JIO is almost Nil in the first three years and goes on to become the topper from 2018 onwards.

AGR	2014	2015	2016	2017	2018	2019	2020
Aircel	648	1119	3617	775	1475	Nil	Nil
Bharti	2445	2463	2832	2265	1590	1684	1580
Idea	1607	1806	2183	1540	591	Nil	Nil
Jio	Nil	1	20	675	1754	2672	2561
Reliance	376	433	271	139	Nil	Nil	Nil
Tata	837	864	983	441	183	Nil	Nil
Vodafone	1762	1925	2206	1598	704	666	Nil
Voda Idea	Nil	Nil	Nil	Nil	1079	1827	2053

3. The data showing the average subscribers by the TSP s across all LSA is shown below:

Subscribers	2014	2015	2016	2017	2018	2019	2020
Aircel	6840689	9849173	17270528	10452934	16297148	Nil	Nil
Bharti	11559928	12149125	14083374	17183448	19090268	18429800	17897248
Idea	9806850	10640520	12578463	13489576	47123425	Nil	Nil
Jio	1	Nil	3991454	8979908	15701597	21936465	21151981
Reliance	4932976	6287120	4103158	2670782	Nil	Nil	Nil
Tata	5203751	4750459	5070241	2554348	1539855	Nil	Nil
Vodafone	10065336	9954204	11477522	12198820	44093722	64998735	Nil
Voda Idea	Nil	Nil	Nil	Nil	29873387	21438137	22696714

JIO's customer base is also increasing from 2016 onwards which is in line with its trends in AGR

4. The data about the Spectrum acquired by the TSP in each of the years is tabulated. Please note that the spectrum represents the quantum acquired by the TSP in that particular year, not to be

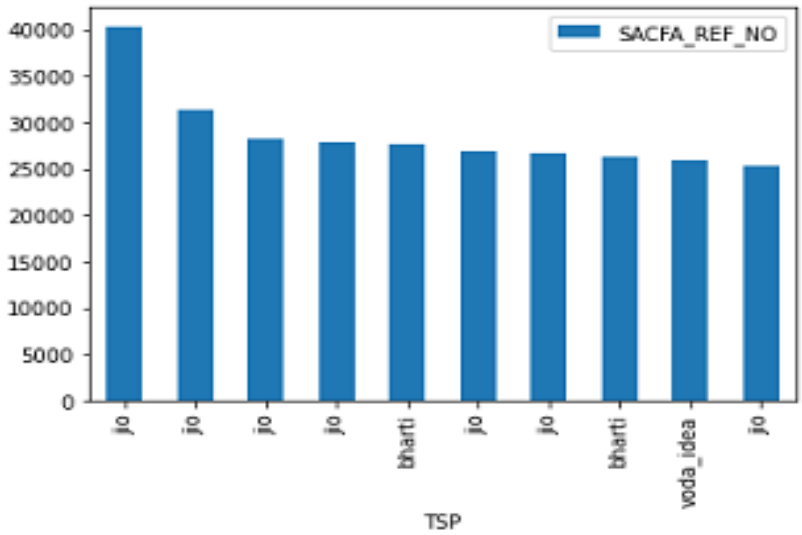
confused by the total spectrum holding of the company.

Spectrum	2014	2015	2016	2017	2018	2019	2020
Aircel	0	0.37	7.3	0.31	Nil	Nil	Nil
Bharti	0.03	5.96	5.05	8.53	8.52	Nil	Nil
Idea	0.66	4.1	5.12	7.87	7.41	Nil	Nil
Jio	Nil	4.3	2.44	6.13	6.11	Nil	Nil
Reliance	Nil	0.09	2.66	Nil	Nil	Nil	Nil
Tata	Nil	Nil	1.04	0.43	1.25	Nil	Nil
Vodafone	0.84	5.53	3.95	6.52	7.92	Nil	Nil
Voda Idea	Nil	Nil	Nil	Nil	Nil	Nil	Nil

5. When the data is further grouped by LSA wise, the total size of the data becomes (838, 9) which means there are 838 rows (records) and 8 columns in the data. An extract of the data in terms of the numbers and LSA wise can be seen in the following tables:-

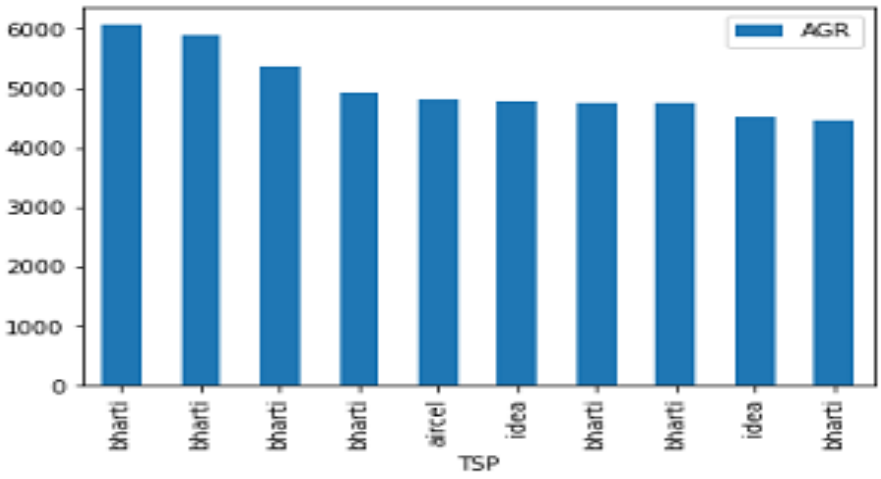
From the data, the first ten records with highest SACFA filing in any LSA by any TSP for any year along with the bar plot can be seen here:-

	TSP	YEAR	LSA	SACFA_REF_NO (In Numbers)	SUBSCRIBERS(In Numbers)	REV (In Rs Crores)	AGR(In Rs crores)	Spectrum(In MHz)
1	JIO	2018	Madhya Pradesh	40357	210448920	2826.17	2364.64	5.000
2	JIO	2017	Andhra Pradesh	31443	132768180	1267.43	982.80	6.875
3	JIO	2017	Tamil Nadu	28221	105892050	1090.74	850.19	6.875
4	JIO	2017	Maharashtra	27925	115947050	1164.79	902.78	6.875
5	Bharti	2016	Delhi	27631	113664850	5908.98	4015.44	0.000
6	JIO	2018	Karnataka	26792	152580540	2061.72	1481.50	6.875
7	JIO	2017	UP(East)	26635	101337273	974.13	731.94	2.840
8	Bharti	2018	UP(East)	26308	327244490	3488.32	1754.01	2.500
9	Voda idea	2019	Maharashtra	25867	396765420	5979.25	3891.38	0.000
10	JIO	2017	Rajasthan	25243	87761290	842.01	646.44	3.710



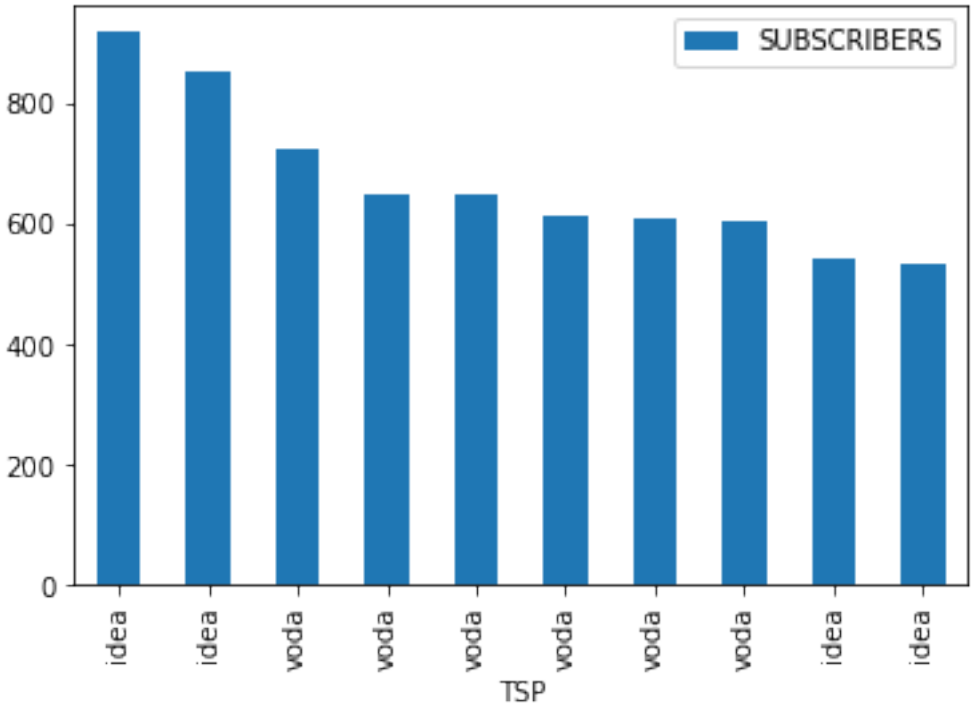
(6)From the data, the first ten records with highest AGR in any LSA for any year and the respective plot:

	TSP	YEAR	LSA	SACFA_REF_NO (In Numbers)	SUBSCRIBERS(In Numbers)	REV (In Rs Crores)	AGR(In Rs crores)	Spectrum(In MHz)
1	Bharti	2016	Karnataka	22413	220015440	7368.35	6067.07	8.8
2	Bharti	2015	Karnataka	11174	208407390	7091.53	5894.06	8.6
3	Bharti	2016	Andhra Pradesh	17130	251126120	6262.41	5365.33	10.2
4	Bharti	2015	Andhra Pradesh	6767	232270590	5771.26	4931.66	8.8
5	Aircel	2016	Tamil Nadu	13951	217053860	5252.71	4799.00	10.0
6	Idea	2016	Maharashtra	14256	254882000	5416.25	4783.72	9.0
7	Bharti	2016	Tamil Nadu	17455	183176950	5619.41	4752.67	5.0
8	Bharti	2014	Karnataka	6717	177726406	6077.38	4749.38	0.0
9	Idea	2015	Maharashtra	10124	230366280	5165.04	4507.46	9.0
10	Bharti	2017	Karnataka	25162	242513360	6268.54	4446.82	5.0



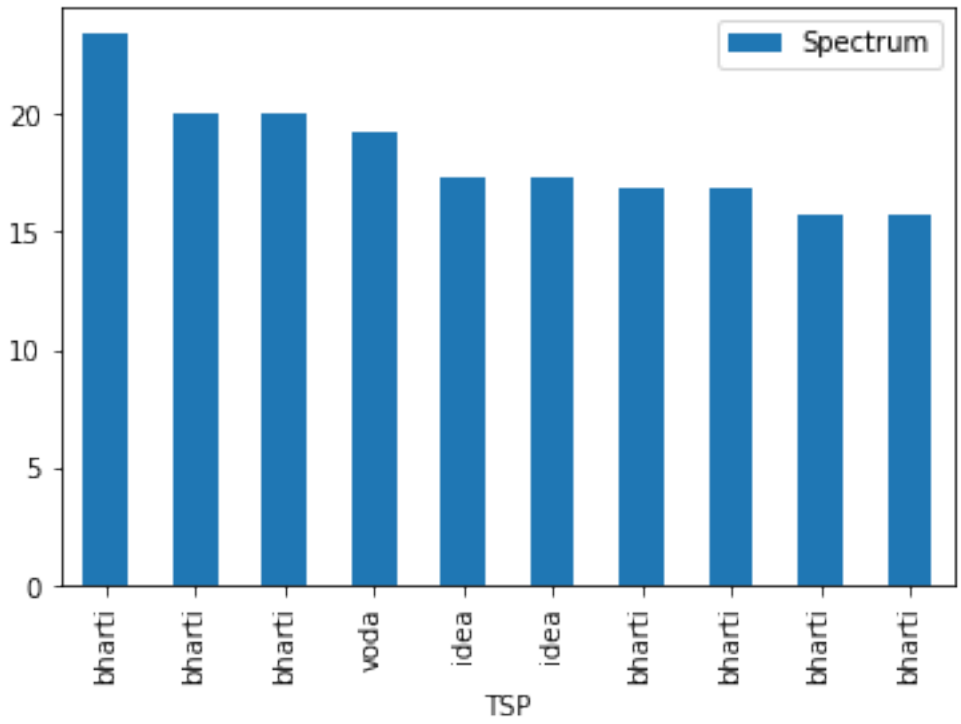
(7)The first ten records with highest Subscribers in any LSA for any year followed by the bar plot:

	TSP	YEAR	LSA	SACFA_REF_NO (In Numbers)	SUBSCRIBERS(In Numbers)	REV (In Rs Crores)	AGR(In Rs crores)	Spectrum (In MHz)
1	Idea	2018	Maharashtra	5206	91933704	1992.53	1453.53	11.0
2	Idea	2018	Madhya Pradesh	7965	85273154	1493.68	1045.84	17.3
3	Voda	2018	UP(East)	1888	72576087	1581.62	814.72	8.5
4	Voda	2019	West Bengal	26	64998735	1068.81	666.28	0.0
5	Voda	2018	West Bengal	4475	64998735	1068.81	666.28	8.8
6	Voda	2018	Tamil Nadu	7605	61366820	1505.95	1141.18	5.0
7	Voda	2018	Gujarat	8113	61174387	1598.38	1043.37	11.5
8	Voda	2018	Maharashtra	3212	60671388	1668.10	1038.90	12.5
9	Idea	2018	Andhra Pradesh	6240	54295308	1220.06	936.07	5.0
10	Idea	2018	UP(West)	9948	53566843	991.54	553.06	8.5



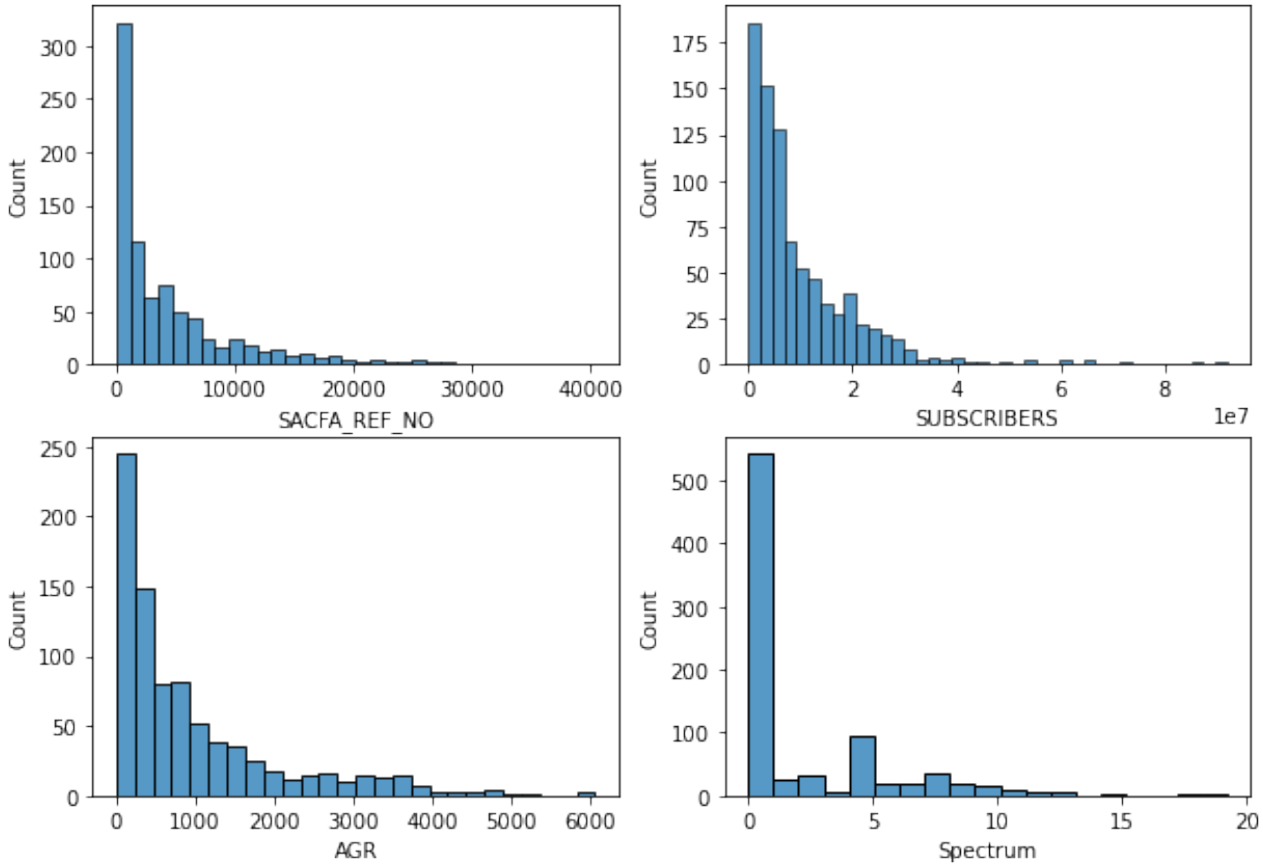
(8)From the data, the first ten records with highest Spectrum acquired in any LSA for any year followed by the bar plot:

	TSP	YEAR	LSA	SACFA_REF_NO (In Numbers)	SUBSCRIBERS (In Numbers)	REV (In Rs Crores)	AGR(In Rs crores)	Spectrum(In MHz)
1	Bharti	2018	Rajasthan	8588	23357048	3042.56	1578.08	23.4
2	Bharti	2017	Bihar	19895	31181437	4550.48	3478.26	20.0
3	Bharti	2018	Bihar	9225	39346424	4124.33	1901.50	20.0
4	Voda	2015	Mumbai	8568	8612645	4430.34	3100.67	19.2
5	Idea	2018	Madhya Pradesh	7965	85273154	1493.68	1045.84	17.3
6	Idea	2017	Madhya Pradesh	8134	26420511	3369.89	2849.59	17.3
7	Bharti	2017	Assam	4521	6766210	1114.99	816.73	16.9
8	Bharti	2018	Assam	6581	8382180	1059.64	630.58	16.9
9	Bharti	2017	North East	3993	4426986	864.19	545.47	15.7
10	Bharti	2018	North East	4859	5335815	720.32	368.40	15.7



[Note:-*More number of plots on the SACFA distribution, LSA wise can be seen in the Annexure at the end of doc.* ***The EDA and OLS Regression were performed using various Python libraries]***

(9)The histograms of the 3 explanatory variables and the one response variable can be seen here:



The SACFA_REF_NO is the number of SACFA applications filed in absolute numbers

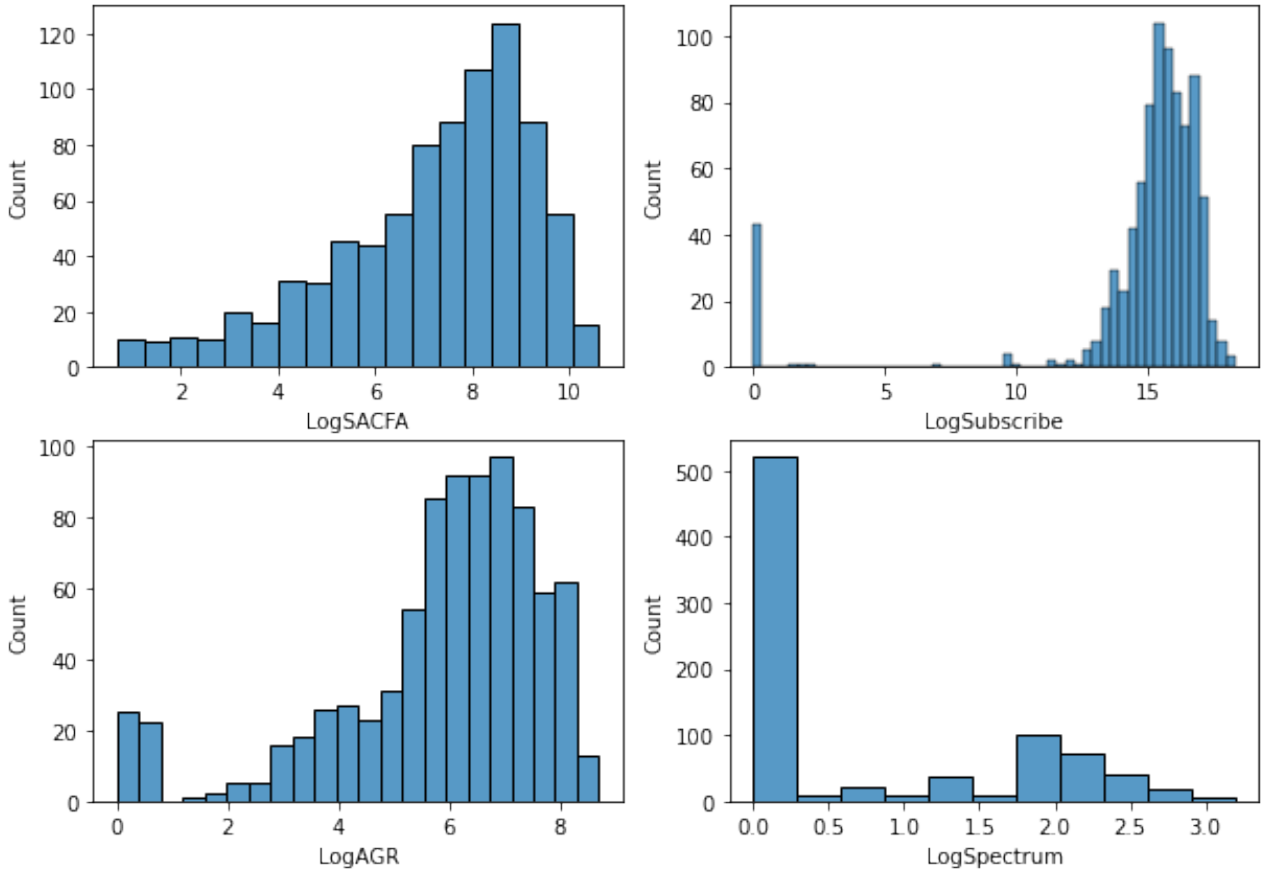
The SUBSCRIBERS is the number of subscribers in Crores.

The AGR is the Aggregate Gross Revenue in Rs Crores.

The Spectrum is the spectrum holding in MHz

The plots are all following a right skewed distribution with unimodal, with the peak occurring at zero value. This is due to the fact that there are about 300 SACFA records where the numbers of filing are less than 1000 in some LSA s. Similarly there are about 50 records where both the Subscribers and AGR values are shown as zero. This data pertains to JIO during the years 2014 and 2015. In addition to this, there are about 130 records where the spectrum purchased as zero in a particular year. This is due to the fact there are no spectrum allocation in the years 2019 and 2020 to the three TSP operating during this time-Jio, Bharti and Vodafone Idea. The above mentioned records are pulling the plots towards right skewed position.

(10)One of the solutions to remove the skewedness is conversion into logarithmic scale. After doing this exercise, we get the following plots:



It can be seen from the above plots that the logarithmic transformation has helped in changing the shape into a normal distribution plots. However the plots are still showing the prominence of the zero values which need to be curtailed before the regression can be done.

Impact of removing zero values: - Now let's calculate the impact of removing the zero values in the AGR, Subscribers and Spectrum mentioned above. If all the zeros mentioned earlier, there will be a net reduction of records from 838 to 658, which comes to about 21 percentages reduction in data points. It may be noted that it is about removing the independent variables with zero values that adversely affect the normality condition of the overall data. If we remove these records, the change that is going to be reflected in the SACFA filings can be assessed as follows:

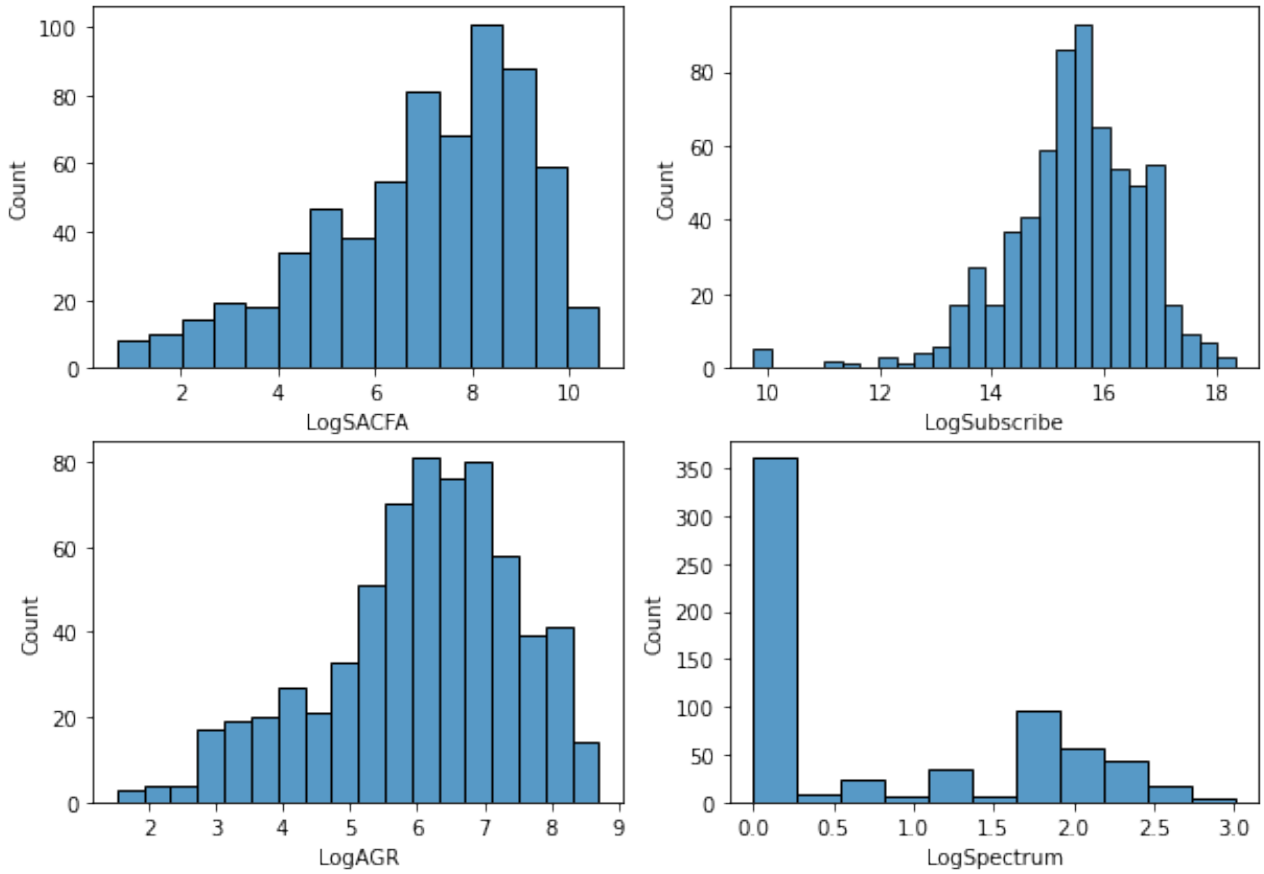
Total SACFA filing before removal of zero data points= 3709923

Total SACFA filing after removal of data points= 2712347

SACFA filing lost in the exercise=997576

Percentage deduction=997576/3709923*100=26 percentage.

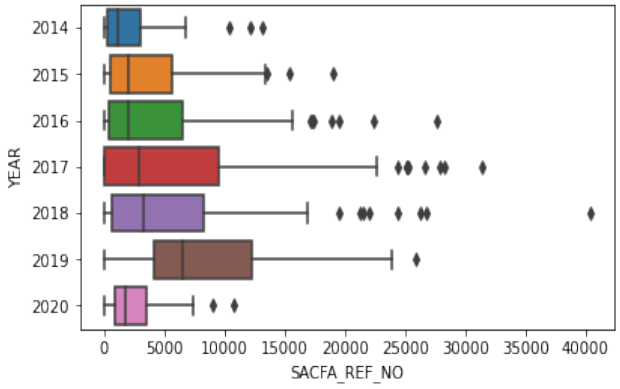
(11) After removing these zero value outliers, we get the following histograms:

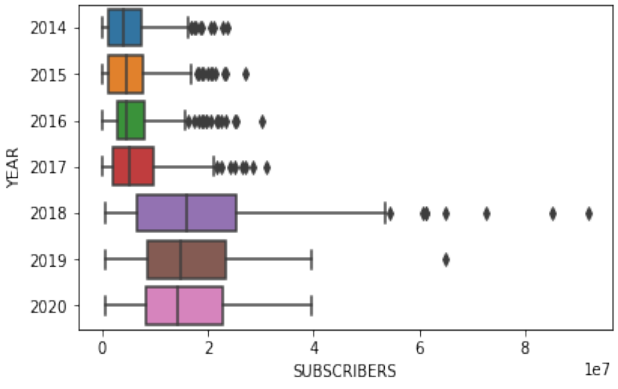


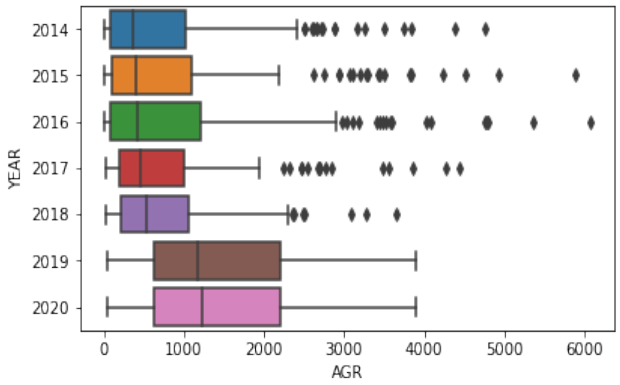
This shows some improvement in the data distribution and it augurs well for performing the regression. The spectrum plot still has about 350 numbers of zero values. Since spectrum is expected to be an important variable that determine the filing of SACFA, it is not possible to remove all the zero values, with the understanding that spectrum allocation is nil, one of the possible value in those cases. However, in this pre-processing exercise, the entire data in the year 2019 and 2020 have been taken down and virtually limiting the study period to 2014-2018. As the operators are also only 3 in these two years (2019 and 2020) with all the records having their spectrum holding value to zero, it is felt that removing these 130 data points would reduce the influence of the zero points in the developed model but retaining the importance of Spectrum in the remaining data points.

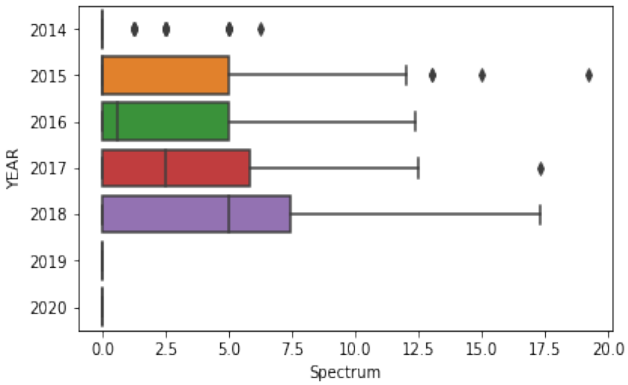
(12) Box plots of the input and output variables:-

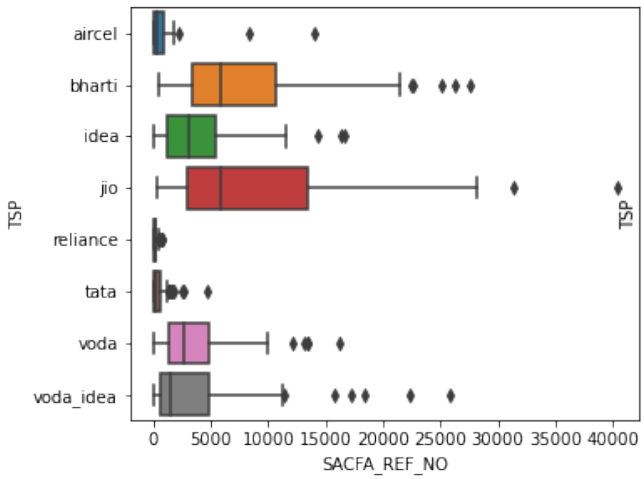
Two box plots are displayed here that shows the spread of the data and median of each variable. It also shows the outliers of each variable and clearly shows the data is skewed due to these outliers. This plot is before the logarithmic conversion is performed. The first plot shows the distribution of the input and output variables against the respective year. The maximum SACFA filing took place in 2019 and maximum subscribers in 2018. From the second plot, it can be seen that JIO has filed the maximum SACFA filing overall and Bharti recorded the highest AGR in this period.

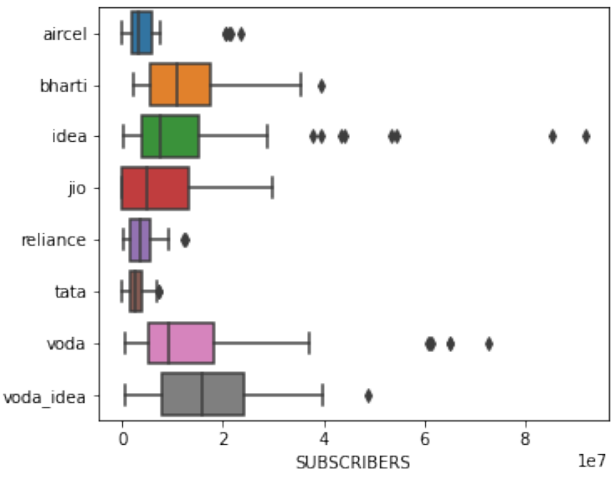

YEAR
2014
2015
2016
2017
2018
2019
2020
0
5000
10000
15000
20000
25000
30000
35000
40000
SACFA_REF_NO

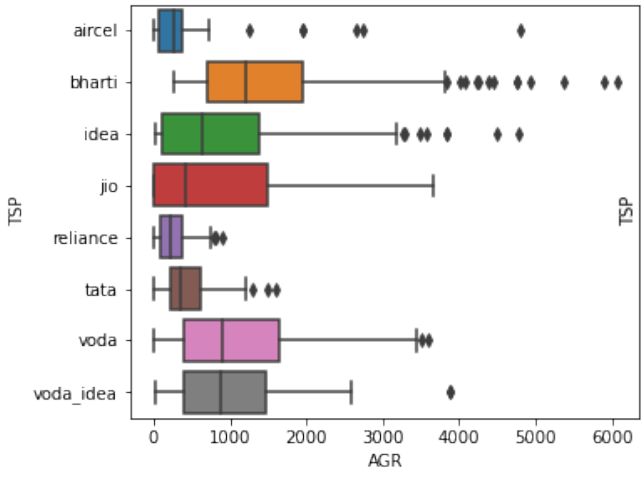

YEAR
2014
2015
2016
2017
2018
2019
2020
0
2
4
6
8
1e7
SUBSCRIBERS

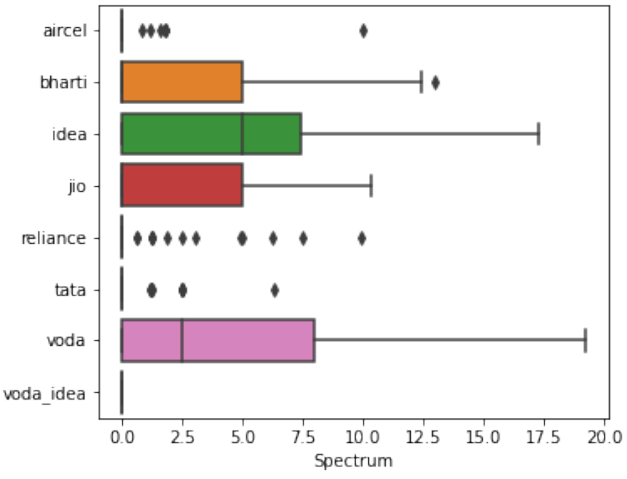

YEAR
2014
2015
2016
2017
2018
2019
2020
0
1000
2000
3000
4000
5000
6000
AGR


YEAR
2014
2015
2016
2017
2018
2019
2020
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
Spectrum

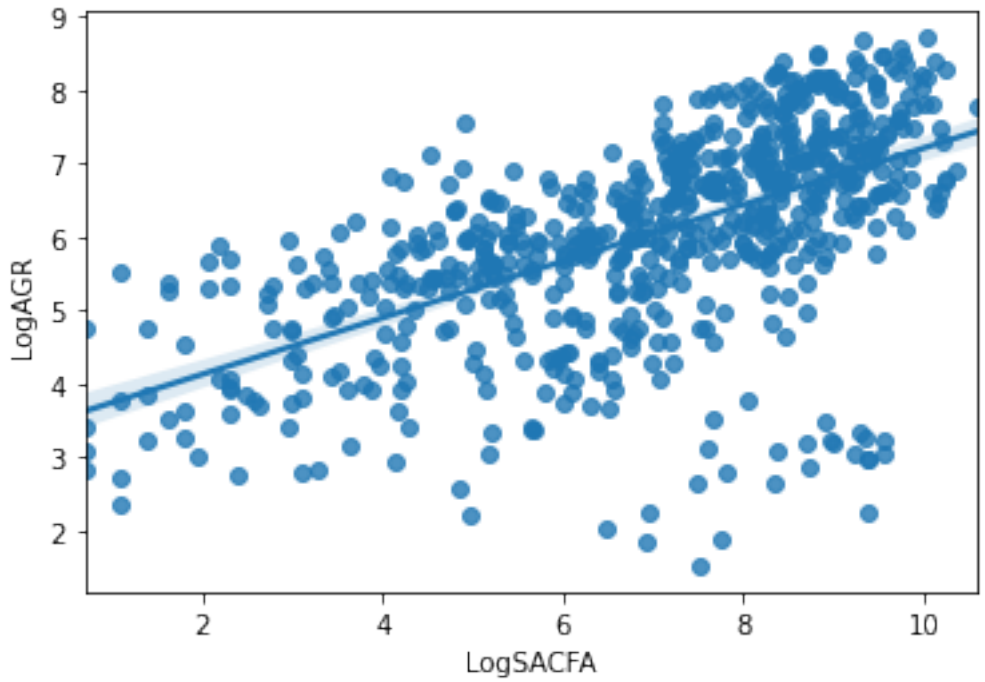

TSP
aircel
bharti
idea
jio
reliance
tata
voda
voda_idea
0
5000
10000
15000
20000
25000
30000
35000
40000
SACFA_REF_NO


TSP
aircel
bharti
idea
jio
reliance
tata
voda
voda_idea
0
2
4
6
8
1e7
SUBSCRIBERS

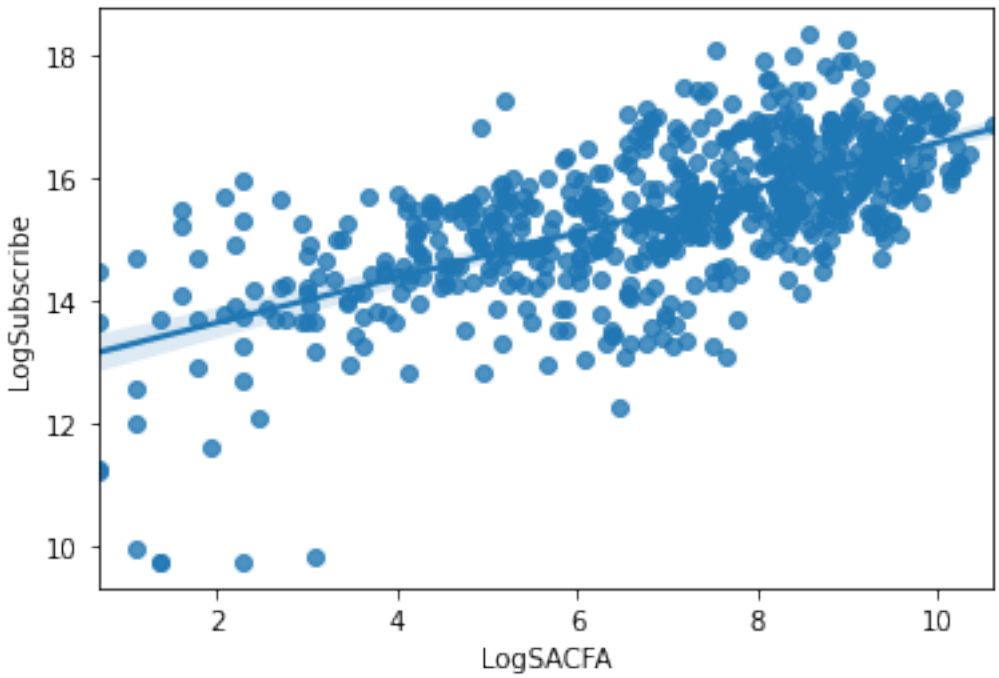

TSP
aircel
bharti
idea
jio
reliance
tata
voda
voda_idea
0
1000
2000
3000
4000
5000
6000
AGR


TSP
aircel
bharti
idea
jio
reliance
tata
voda
voda_idea
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
Spectrum

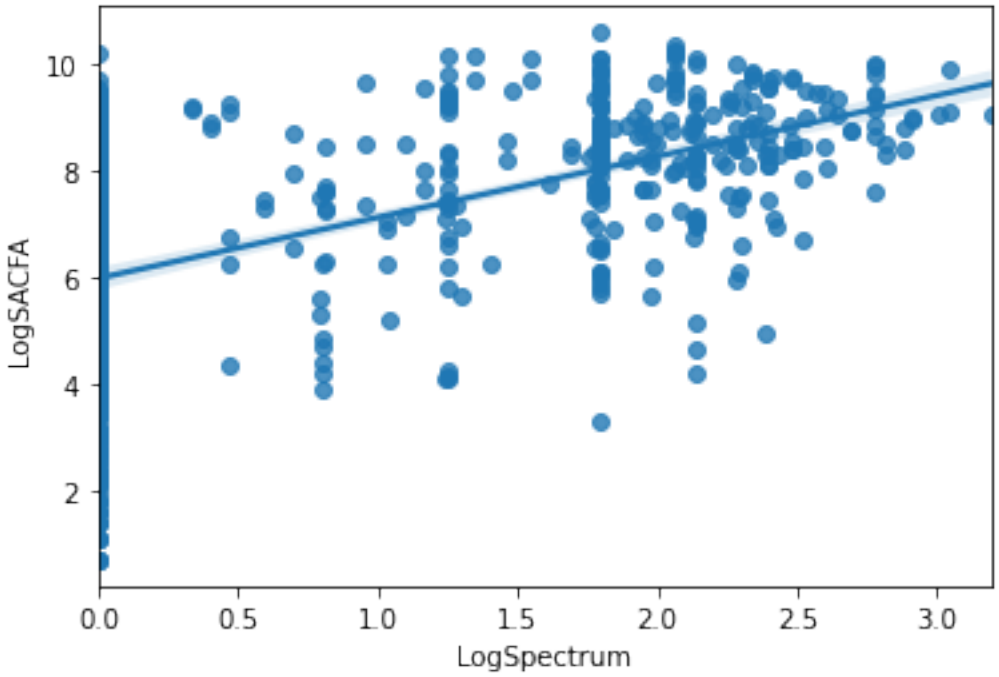
(13) Scatter plots: - What is next is to test the hypothesis. This is done through the OLS regression. The regression is performed only for the quantitative variables. The scatterplots can be seen here:



There are few outliers at high value of SACFA with small AGR value, which is not influential in drawing the line of fit. The Pearson correlation coefficient in the above case: 0.582



The Pearson correlation coefficient in the above case=0.651



The Pearson correlation coefficient in the above case=0.535

In general the outliers may be pushing down the Pearson coefficient, which is a measure to show the linearity of the data.

9. Linear Regression and fitting of the model to the data:-

The OLS regression has produced the following outcome:-

OLS Regression Results

Dep. Variable:	LogSACFA	R-squared:	0.559
Model:	OLS	Adj. R-squared:	0.557
Method:	Least Squares	F-statistic:	276.4
Date:	Fri, 30 Apr 2021	Prob (F-statistic):	8.12e-116
Time:	20:28:13	Log-Likelihood:	-1169.0
No. Observations:	658	AIC:	2346.
Df Residuals:	654	BIC:	2364.
Df Model:	3		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	-5.7584	0.855	-6.736	0.000	-7.437	-4.080
LogAGR	0.3371	0.058	5.792	0.000	0.223	0.451
LogSubscribe	0.6490	0.070	9.263	0.000	0.511	0.787
LogSpectrum	0.7630	0.059	12.931	0.000	0.647	0.879

Omnibus:	7.543	Durbin-Watson:	1.032
Prob(Omnibus):	0.023	Jarque-Bera (JB):	10.262
Skew:	-0.085	Prob(JB):	0.00591
Kurtosis:	3.587	Cond. No.	257.

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

The p-values are all statistically significant pointing to the stable regression coefficients. It may be mentioned that in this exercise, the whole population has been used for analysis rather than multiple samples. Sampling error is therefore not relevant here. However we have limited the period of analysis to 2014 to 2020 even though the SACFA process for the service providers started after the 2G technology was introduced in 1995 or so. This transforms to a 20 year period of SACFA data has not been considered in this study. As the target population is 2014-2020, we say that the entire data that generated out of an independent organic process (SACFA clearance) has been chosen without any bias in selection. The data is independent and identically distributed and this fact can be verified from the plots generated from the model.

Another point that can be observed from the regression model is that the R squared value, which is known as goodness of fit, is about 56 percentages. This means that the OLS equation is able to explain only 56 percentage of the variability of the response variables with reference to the explanatory variables. This points to the existence of other variables that has not been considered and important in influencing the goodness of fit. It means the filing of SACFA may depend on many other variables apart from what have been considered in this study. Whatsoever may be those unknown variables, but within the selected variables, the SACFA filing numbers can be correlated with the following equation:

```
log(SACFA_REF_NO+1)=-5.7584+0.3371X(log(AGR+1))
                    +0.649X(log(SUBSCRIBERS+1))
                    +0.763*(log(Spectrum+1))
```

This regression outcome shows that the number of SACFA filing is linearly associated with all the three variables -namely its financial revenue (AGR), the number of consumers it is serving (SUBSCRIBERS) and its Spectrum holding. The regressed model can be interpreted as one unit change in AGR results 0.3371 times the log(AGR+1) increase in the number of SACFA filing in any LSA when other independent variables number of Subscribers and Spectrum are held at zero value. Similarly one unit change in the number of Subscribers causes 0.649 times log (Subscribers+1) increase in the number of SACFA filing in that LSA when other independent variables AGR and Spectrum are held at zero value. And one unit change in the Spectrum in MHz causes the increase in the number of SACFA filing by 0.763 times log (Spectrum+1) when other variables like AGR and Subscribers are kept at zero value.

It was also noticed that when the SA (LSA) variable was included in the regression, the respective p-values were not significant indicating that there is no difference among the 4 classes of LSA -metro, Class-A, Class-B and Class-C- in specifying the response variable. That is,

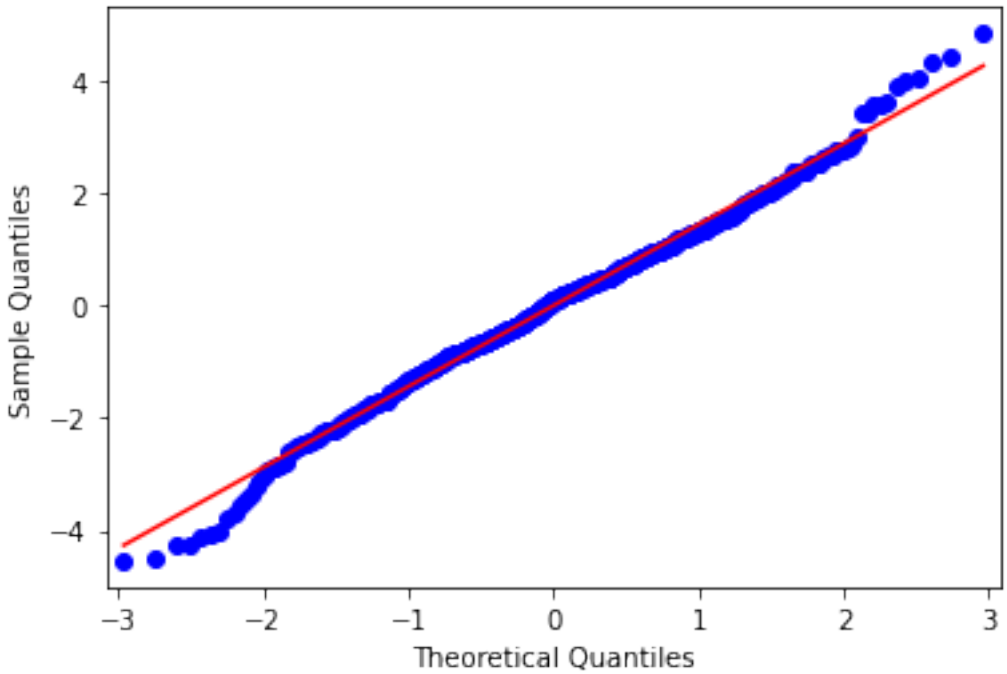
the equation above holds good irrespective of the LSA class wherever the SACFA is filed. Therefore the basic predictors that determine the SACFA applications filed are AGR, Number of Subscribers and the Spectrum availability with 56% specification accuracy.

10. Regression plots:-

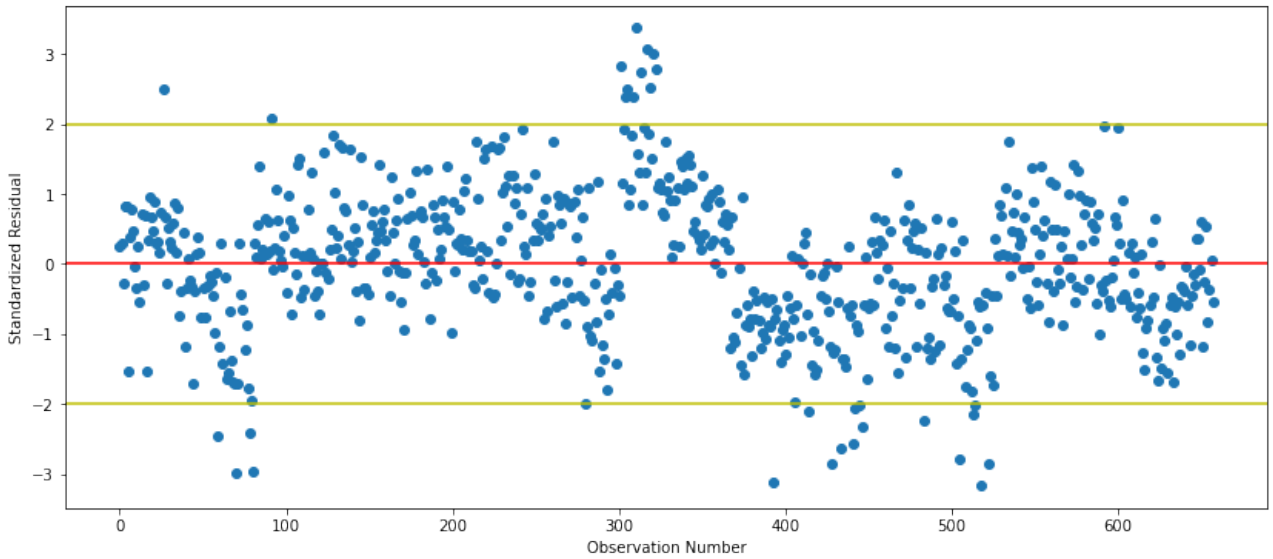
It is mandatory that the following conditions are to be met by the data for achieving goodness of fit in regression.

Normality, Linearity, Independent and Homoscedasticity: - The last two terms together is referred to as IID (independent and identically distributed). If the data do not meet the conditions, then the model will said be poorly fit. The normality and linearity has been satisfied to a larger extend as seen from the plots earlier. The IID can be seen from the standardised residual plots with zero mean and constant variance. The QQ plot is probability distribution showing how much normal is the distribution of the residual. The plots almost follow the normal distribution of the residuals and can be seen the data is almost normally distributed with few outliers.

(1)QQ Plot:-



(2) The standardised residual plot:-



The residuals can be seen lying within 2 standard deviations. As a thumb rule, maximum of 5% residuals can have an absolute value of more than 2 beyond which the error in the model is not acceptable. As the residuals in this case are less than 5%, the regression model can be accepted.

11. Forecasting based on the regression: - By substituting the known values AGR, Subscriber numbers and the Spectrum acquired in the regression equation, it would be possible to predict the number of SACFA filing in a year. The spectrum auction was done in 2021 and the three operators have acquired spectrum in different bands. The AGR and Subscriber data are available for the year 2020. Based on these data, the number of SACFA applications that are filed is calculated and depicted in the table below. It should be understood that this is based on the mean values of the response variable arrived at using the regression. The actual number may vary on either side of this mean value of the SACFA filing that is calculated for each of the Service Areas. The change in AGR and Subscribers' data for the year 2021, which is not available presently, would also influence the final SACFA figures. The expected SACFA filing numbers for each LSA are presented with these limitations.

	Number of SACFA applications likely to be filed in next year based on the available predictors for this year.		
LSA	Bharti	Voda Idea	Jio
Andhra Pradesh	1432	695	15123
Assam	3279	134	3591
Bihar	11622	377	15489
Delhi	721	506	9525
Gujarat	3916	1052	11670
Haryana	2653	299	5862
Himachal Prades	3517	27	757
Jammu & Kashr	3122	20	2148
Karnataka	9593	1406	10796
Kerala	3150	832	5608
Kolkata	448	292	4379
Madhya Prades	7120	800	12047
Maharashtra	6131	1423	19949
Mumbai	450	430	6939
North East	2194	63	1558
Orissa	3377	90	6712
Punjab	5865	370	8682
Rajasthan	9460	475	14197
Tamilnadu	12695	855	14581
UP(E)	19561	1568	20909
UP(W)	7620	1065	11631
West Bengal	6621	1007	9977
Total	124547	13786	212130

12. Conclusion: -

What the evidence has shown is that the response variable is associated with all the three explanatory variables chosen for the study. Therefore, the Null Hypothesis is rejected and the alternate hypothesis that there is a linear association between the dependent variable and the independent variables stands proved.

It was shown that the goodness of fit of the statistical model is about 56 % indicating that the variability of the SACFA numbers could be explained only 56 percentages by the predictor variables. That shows there are other factors that could influence the SACFA filing. The availability of land and building for BTS installation, the ROW (Right of way), the local regulations, the financial condition of the TSP, the availability of technology and electronics, TRAI regulations and many more could be explored to reduce the specification error in the response variable. The presence of these factors could point to the possibility of taking further studies in this area.

Every SACFA site is uniquely identified by the latitude and longitude of the site. Further the SACFA numbers is a direct indicator of the network deployment by a TSP in any LSA since it is linked to a transmitting station (BTS) serving a particular area in that LSA. It has been found from the study that the SACFA filing is linearly associated with the financial capacity of the

TSP as well as the number of consumers it is serving in an LSA. Therefore the quantity of SACFA filing is a key parameter explaining the reach or the spread of the telecom network in a particular geographic area and can be counted as a proxy to the presence of a vibrant telecom network.

13. Reference:

1. Telecom Regulatory Authority of India (https://www.trai.gov.in/)
   (a) Quarterly Financial Data Pertaining to Gross Revenue, Adjusted Gross Revenue, License Fee and Spectrum Charges in respect of all the Telecom Service Providers for the years 2014 to 2020.
   (b) The Indian Telecom Services Performance Indicator Report for various periods.
2. Department of Telecommunications (Spectrum Management | Department of Telecommunications | Ministry of Communication | Government of India (dot.gov.in)
3. SACFA statistics from WPC.
4. Basic Econometrics – Damodar N Gujarati /Sangeetha.
5. www.coursera.com.- OLS regression fundamentals.

14. Abbreviations:-

2G-Second Generation telecom network.

3G-Third Generation telecom networks.

5G-5th Generation telecom network.

AGR-Aggregate Gross Revenue

BTS-Base Transmitting Station.

DoT-Department of Telecom.

LSA-Licensed Service Area

LTE- Long Term Evolution.

OLS-Ordinary Least Square.

SACFA-Standing Advisory Committee for Frequency Allocation

TSP-Telecom Service Provider.

TRAI-Telecom regulatory Authority.

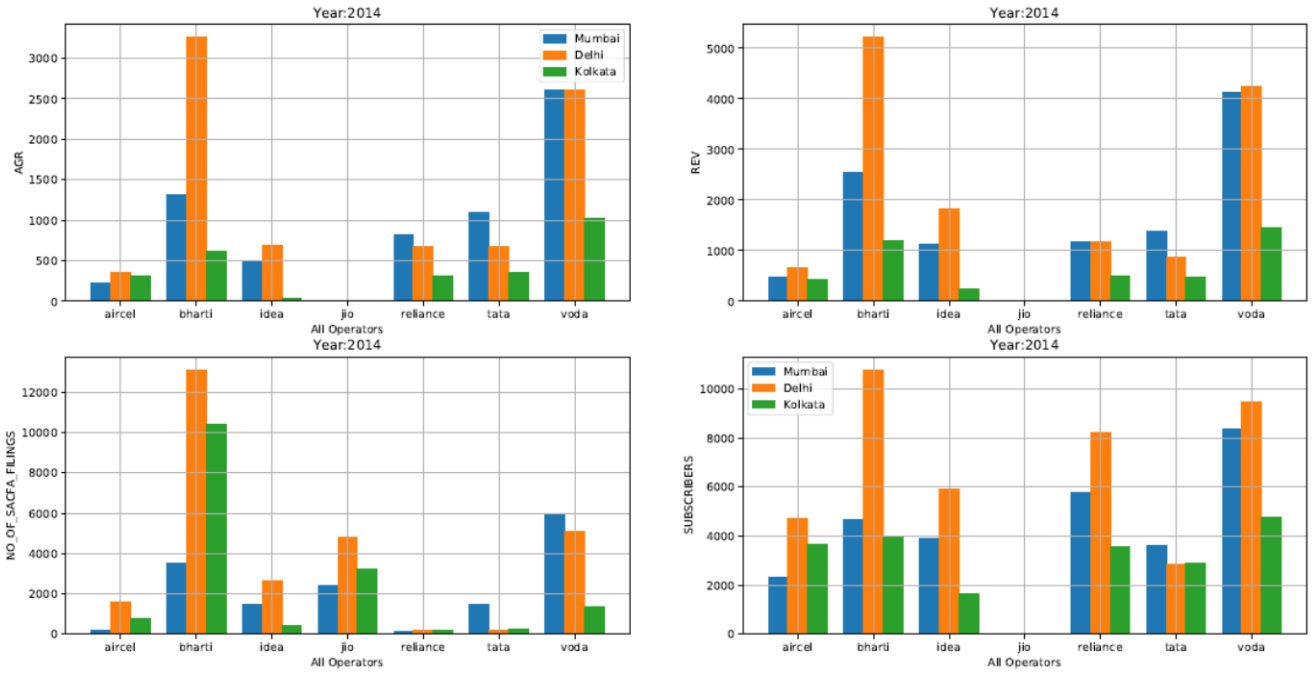
VOLTE- Voice Over LTE network.

WPC-Wireless Planning and Coordination.

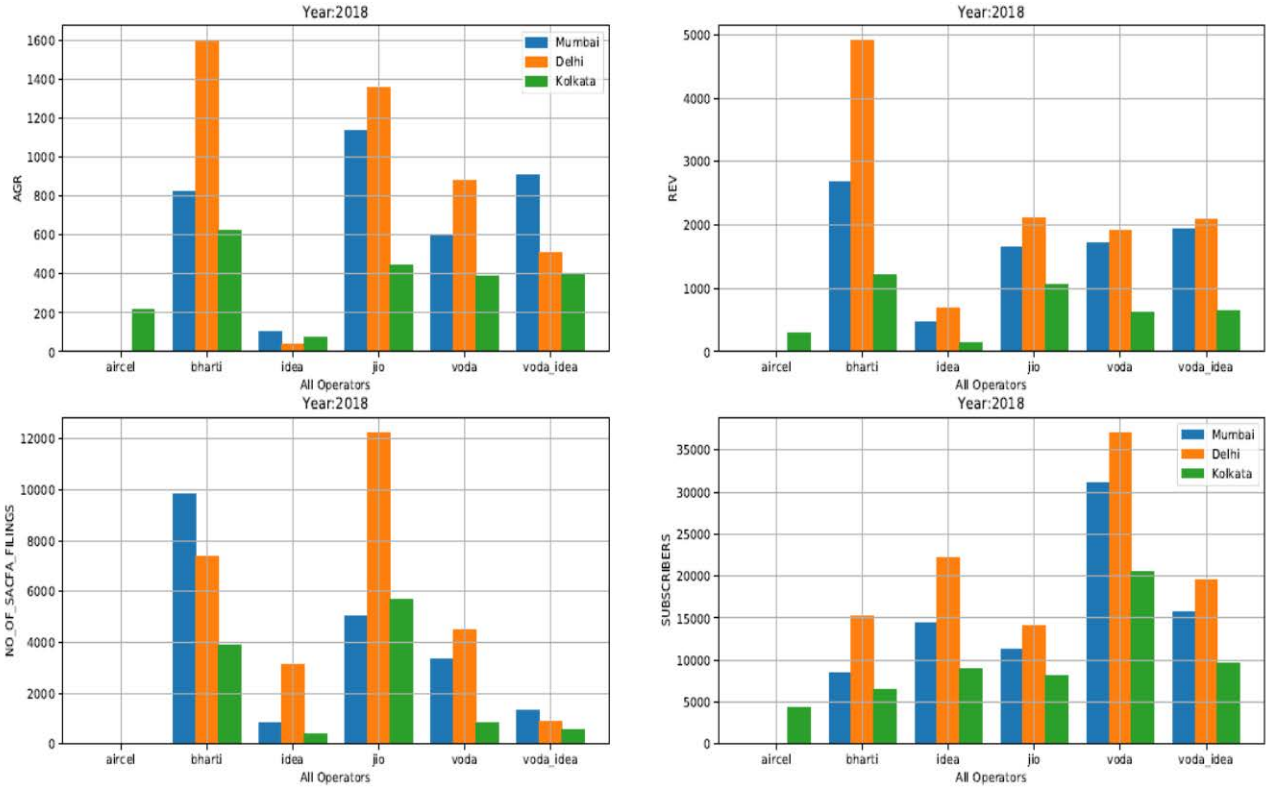
Annexure-I

Bar charts showing the comparative distribution of the dependent and independent variables in different LSA:

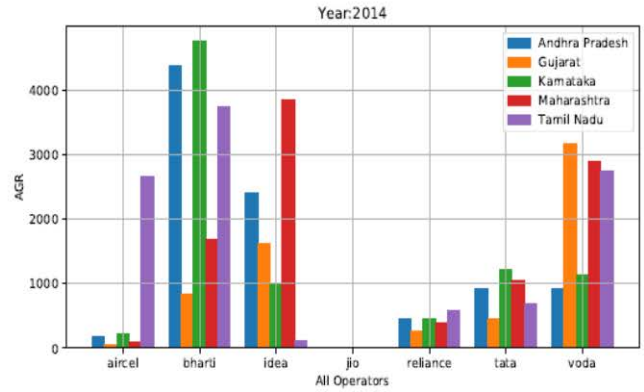
1. For the LSA s in the Metros in the year 2014

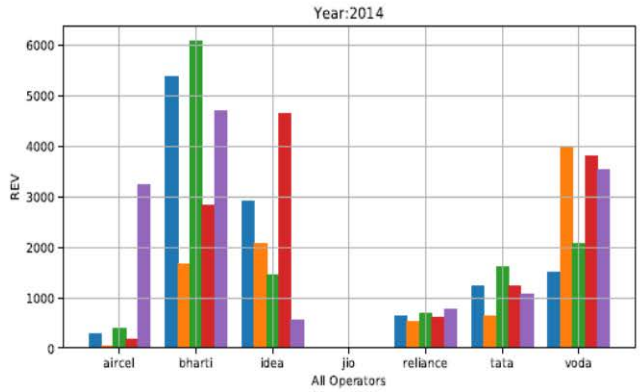


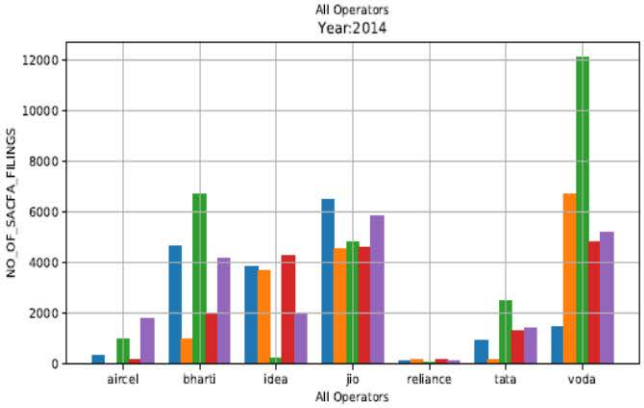
2. For the LSA s in the Metros in the Year 2018:

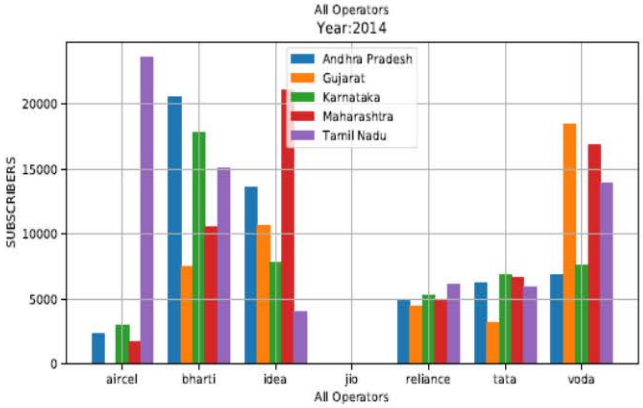


3. For the Class A LSAs in the Year 2014:-

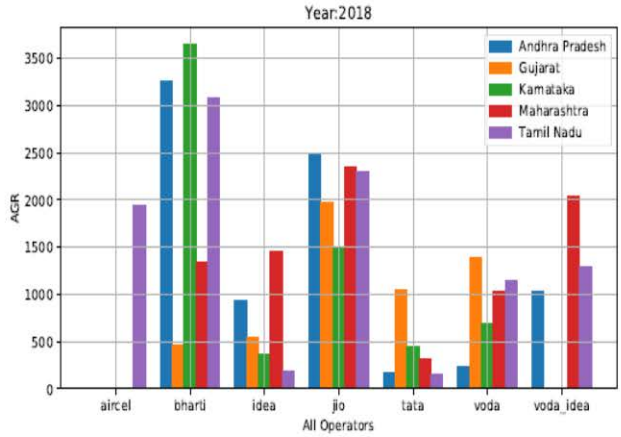


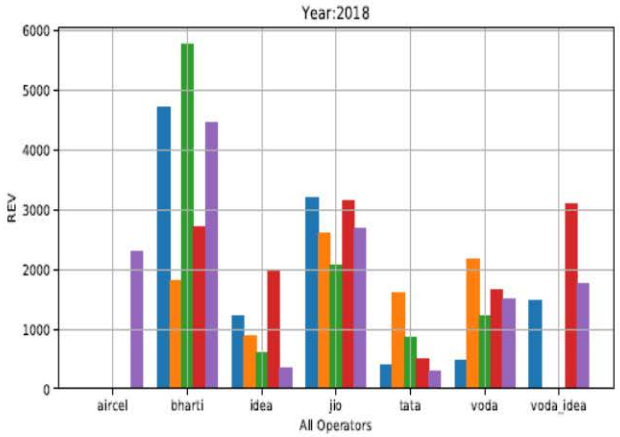


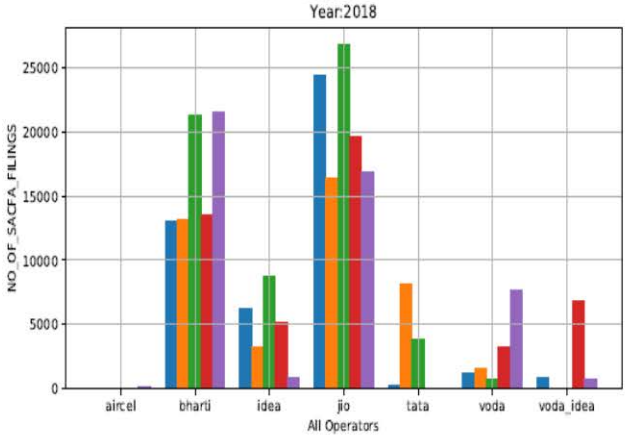


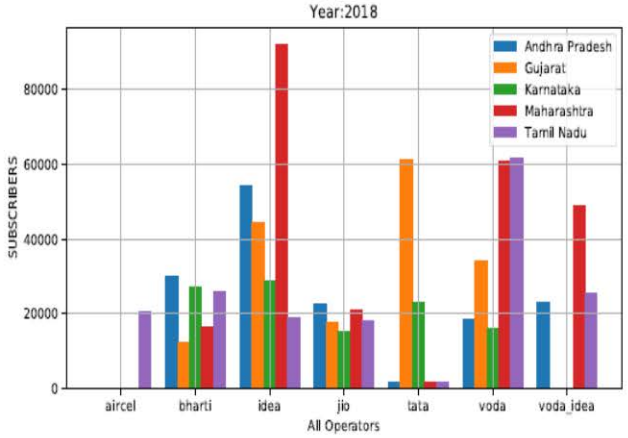


4. For the Class A LSAs in the Year 2018:-

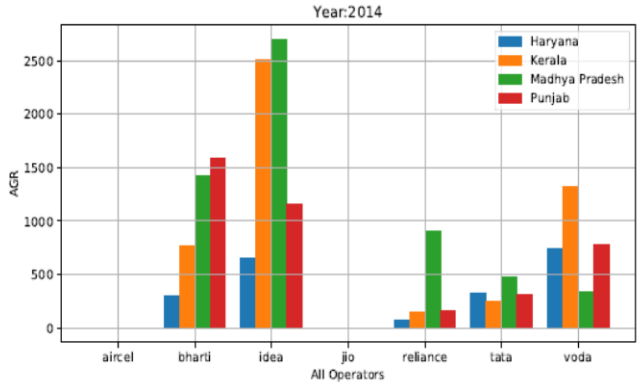


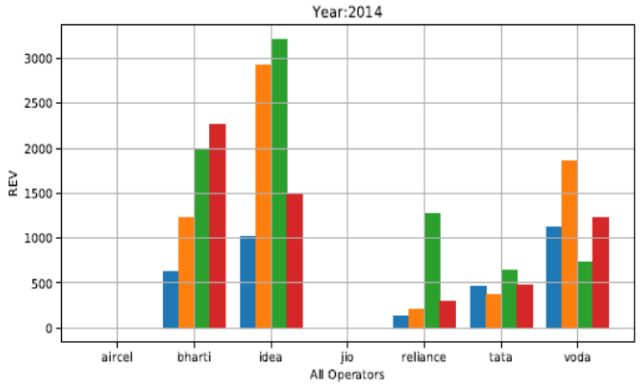


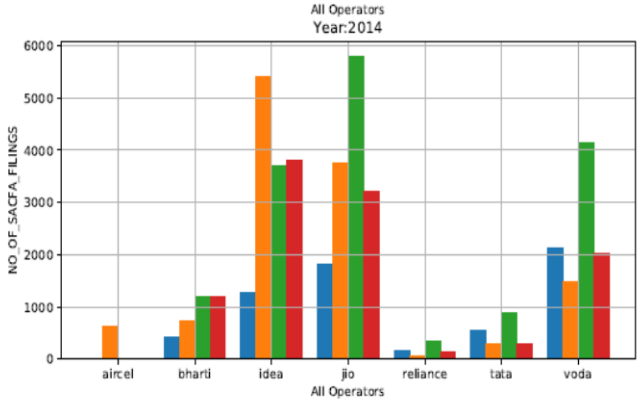


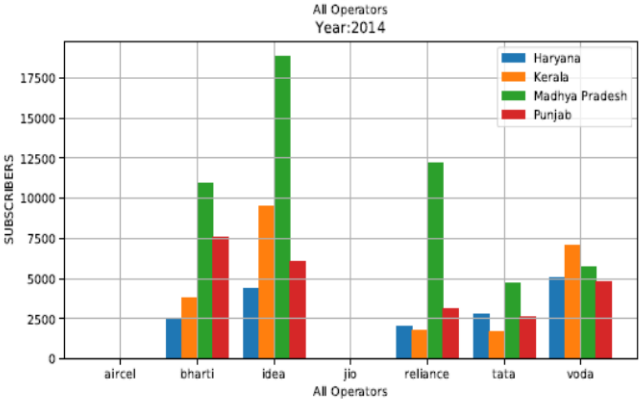


5. For the Class B1 LSAs in the Year 2014:-

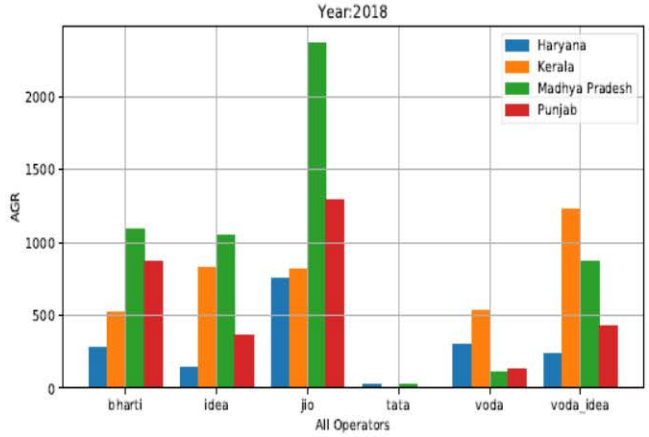


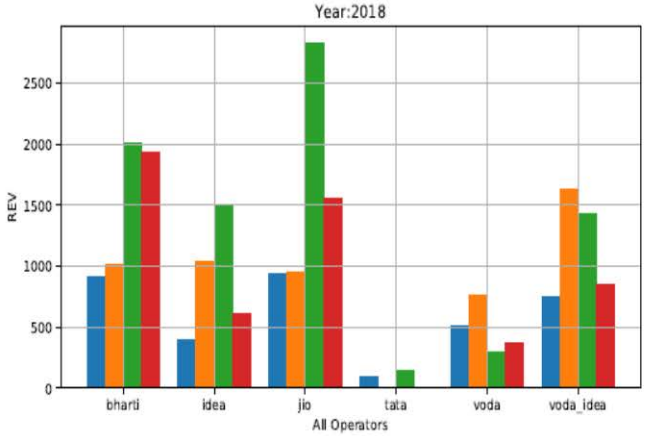


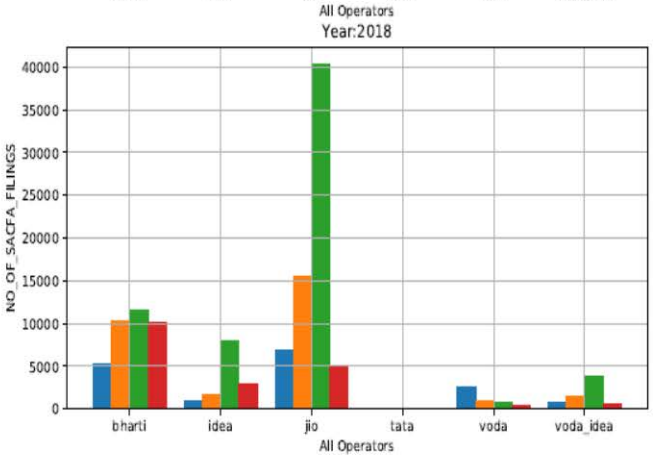


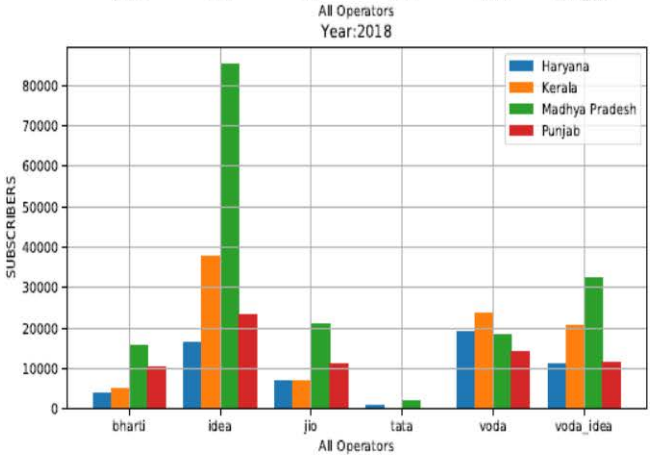


6. For the Class B1 LSAs in the Year 2018:-

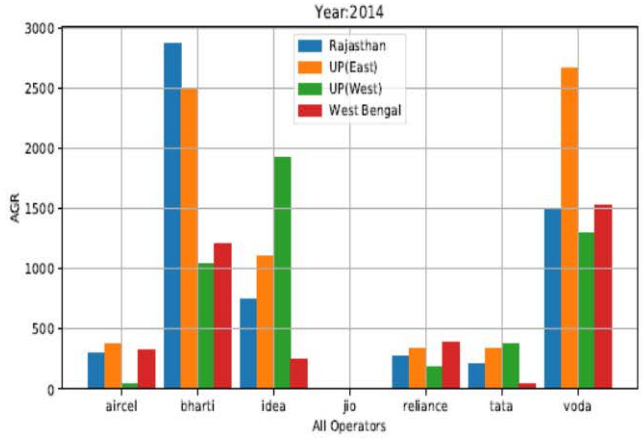


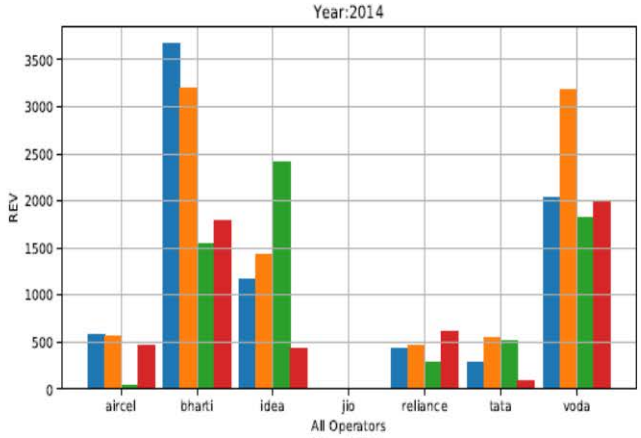


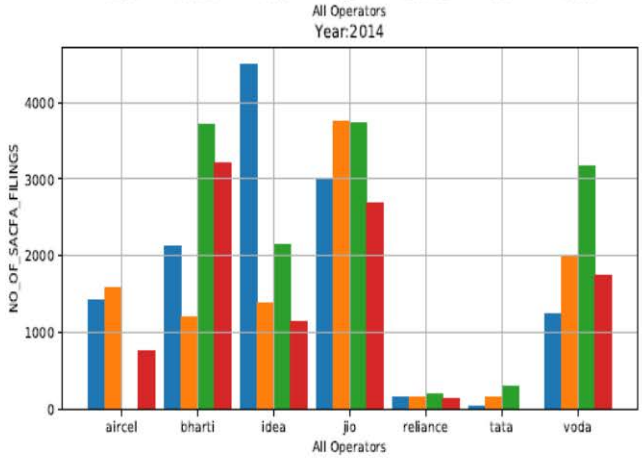


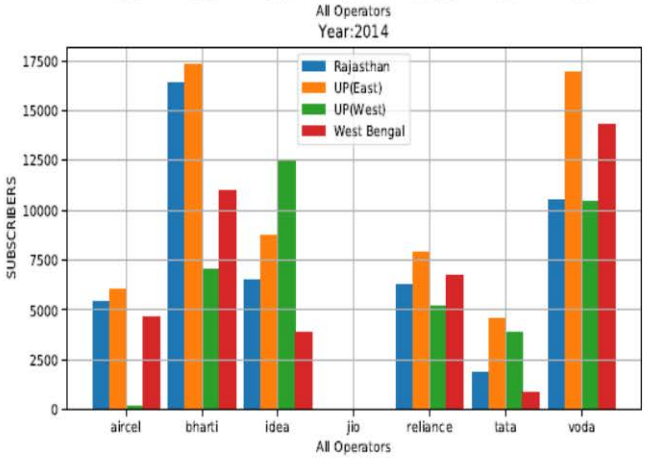


7. For the Class B2 LSAs in the Year 2014:-

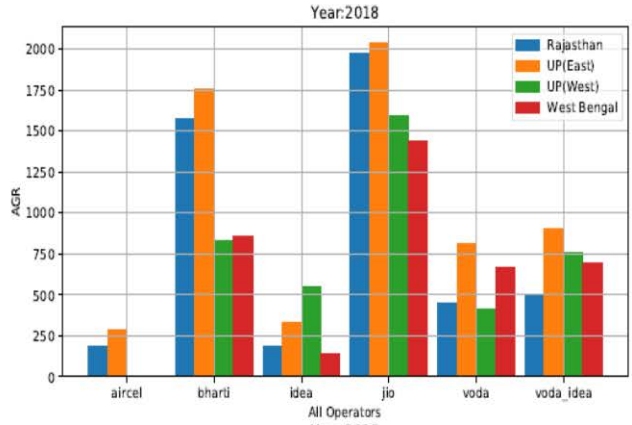


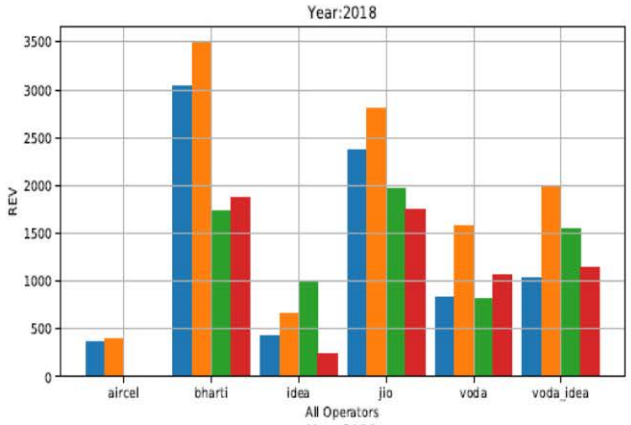


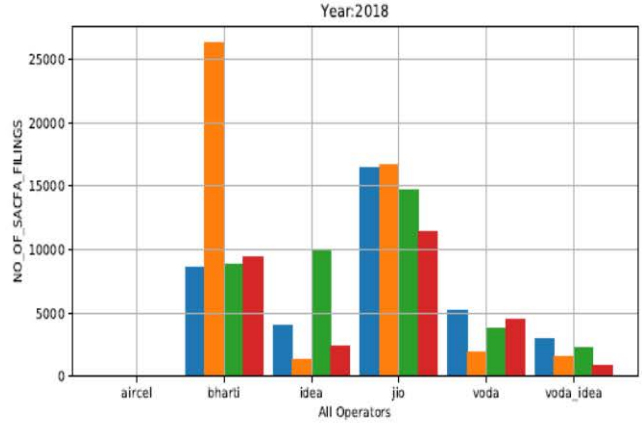


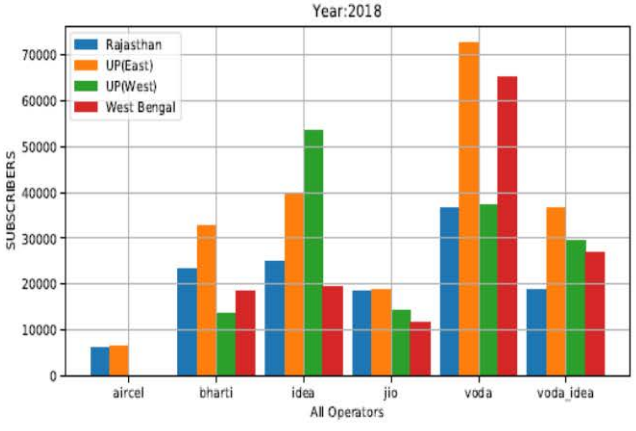


8. For the Class B2 LSAs in the Year 2018:-

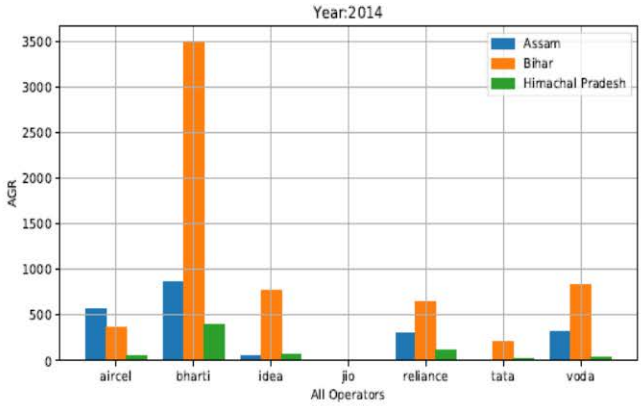


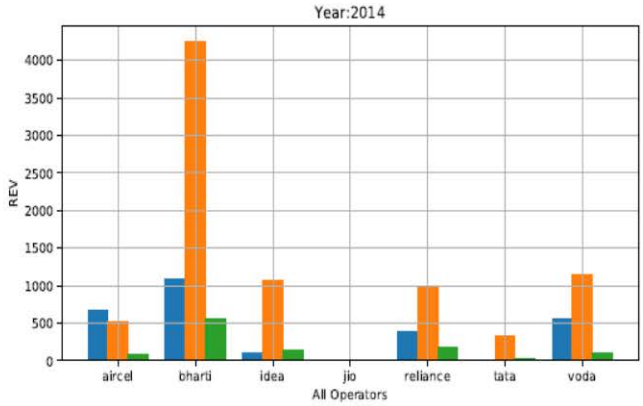


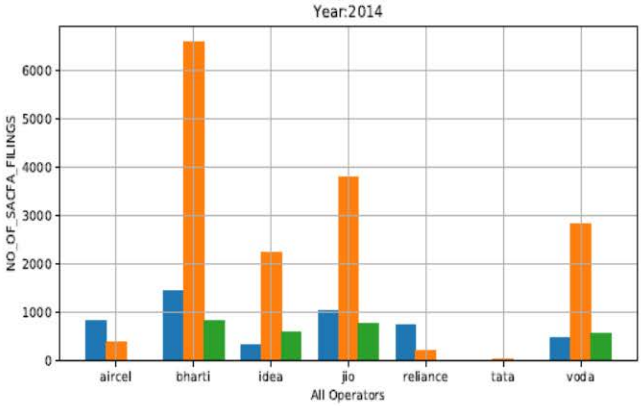


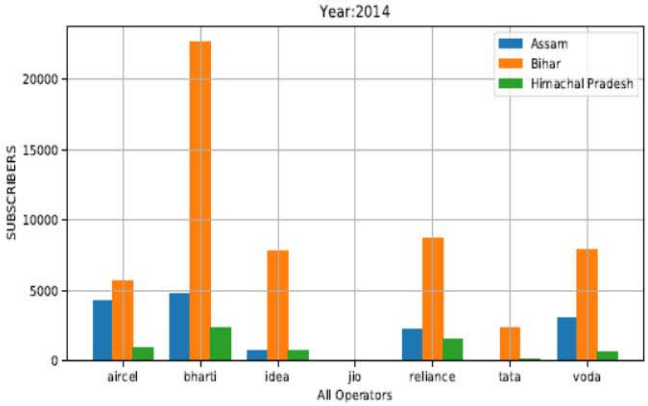


9. For the Class C1 LSAs in the Year 2014:-

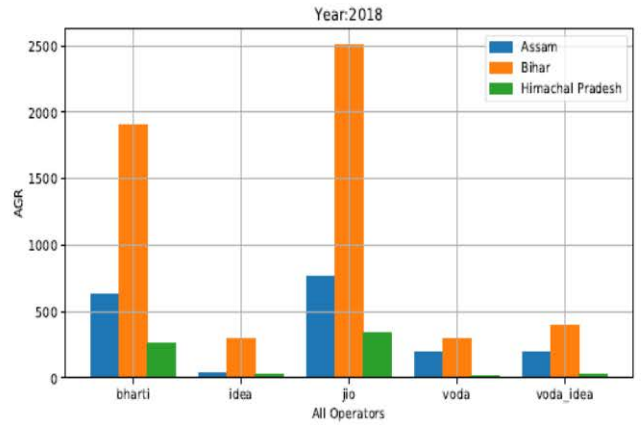


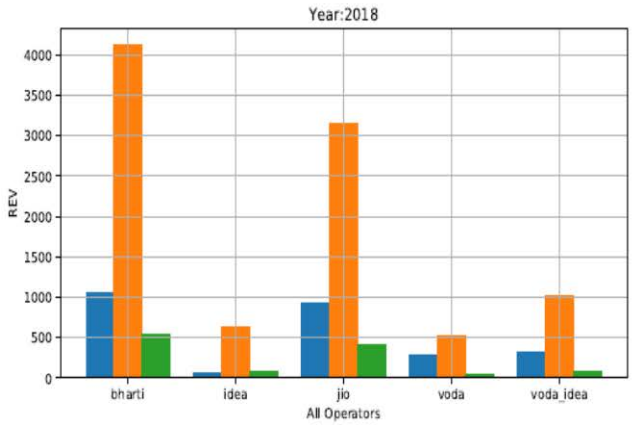


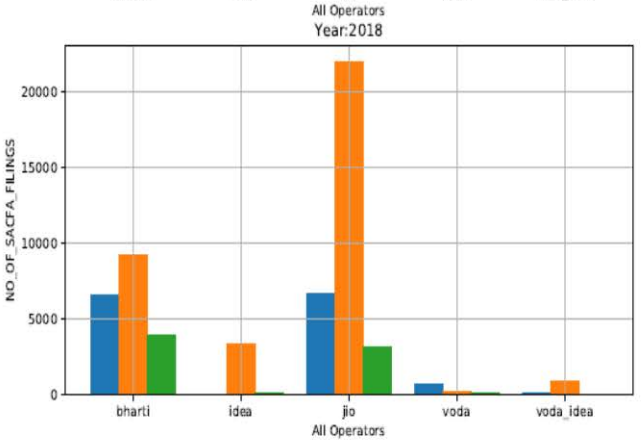


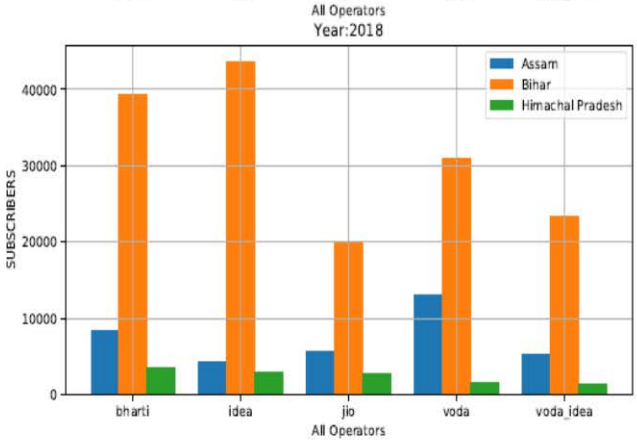


10. For the Class C1 LSAs in the Year 2018:-

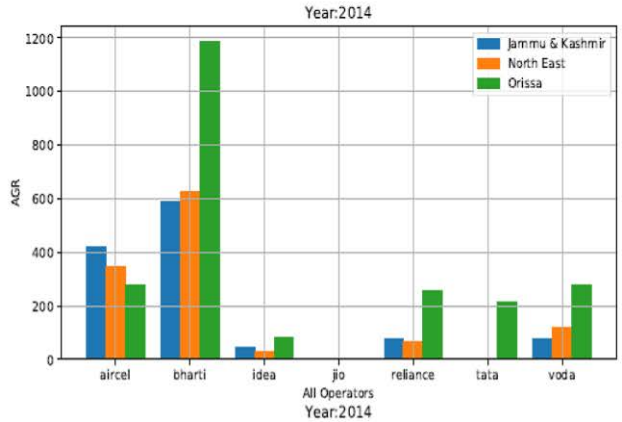


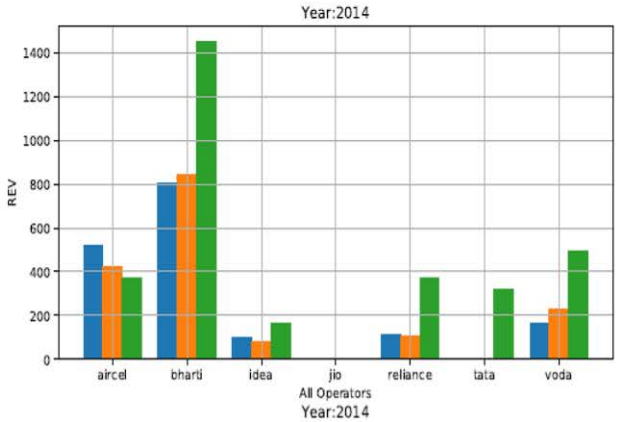


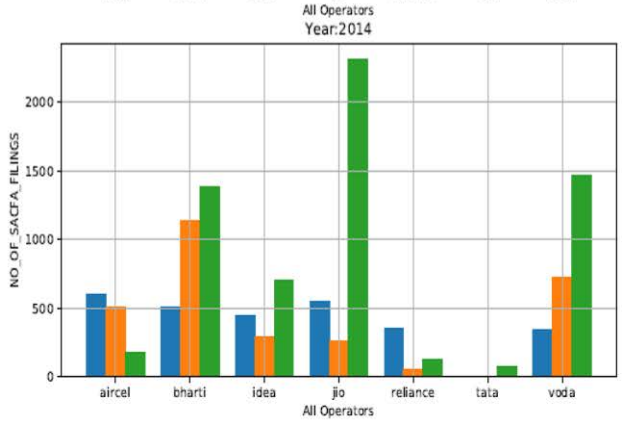


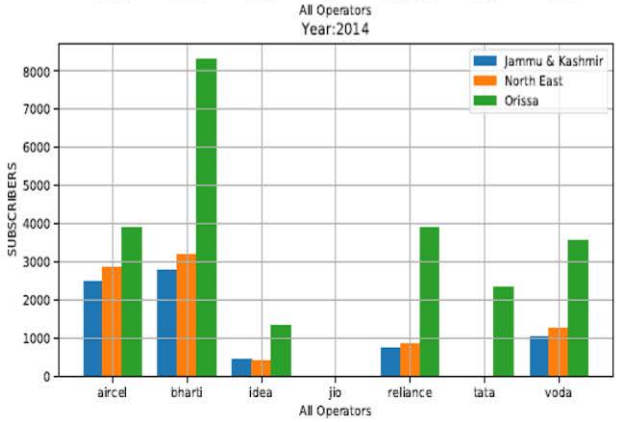


11. For the Class C2 LSAs in the Year 2014:-









12. For the Class C2 LSAs in the Year 2018:-