Study of Income for Companies that Provide Mobile Internet Service in Colombia

Luis Fernando Pedraza, Cesar Augusto Hernandez, Ernesto Cadena Muñoz
Technological Faculty
Universidad Distrital Francisco José de Caldas
Bogotá, Colombia
e-mails: lfpedrazam@udistrital.edu.co, cahernandezs@udistrital.edu.co, ecadenam@udistrital.edu.co

***Abstract*— The telecommunications services market has been growing to satisfy customer needs. The above has prompted mobile phone operators to enhance their bandwidth and coverage for services like mobile Internet. This has had an impact on the operators' income. Therefore, this paper analyzes the income in the mobile Internet service of the three companies with the most significant penetration in the Colombian market from 2012 to 2022. For this, statistical measures are used. The results show that, in general, the companies' income has been increasing; however, the Claro operator's income is much higher than that of the rest of the operators during this period, despite the competitive dynamics in the market of this sector.**

Keywords-analysis of variance; coefficient of correlation; mobile Internet; statical measures; telecommunications market.

I. INTRODUCTION

The telecommunications services market is moving in increasingly flexible and adaptable environments to customers' needs. In addition, an increase in bandwidth is recognized due to the emergence of various services and applications on the Internet, which increases the volume of data sales, becoming a solid source of income for Internet Service Providers (ISP). Intense competition exists between these ISPs to maintain and attract new customers to their businesses, and they are forced to analyze the quality of the service and experience to improve interaction with their customers. Providers are interested in maintaining these clients and increasing the volume of consumption of their services [1][2].

Analyzing the relevant aspects of telecommunications inevitably involves using statistical or data analysis. We can better understand this by looking at the various aspects directly influenced by the sector, including users, infrastructure, technology, and services. Each factor plays a significant role and can be analyzed or measured on a large scale. This requires establishing trends, using metrics, and making comparisons. The various tools offered by statistical and data analysis facilitate these possibilities.

The importance of statistical analysis is verifiable if various studies carried out in the area are reviewed. For example, a variable of great interest within telecommunications is the clients (users) since the massification of services, bandwidth consumption, and the use of infrastructure will always be a function of these; therefore, for service providers, it is essential to understand the consumption trends and preferences that their customers have according to a given region, as done in [3][4]. Analyzing the industry's economic growth and the factors that influence its fluctuations is common. Researchers in a study cited in [5] analyzed the relationship between telecommunications technology and economic growth. Other studies also analyze the impact of Big Data analytics on telecommunications, specifically on governance, architecture, and other use cases, as shown in [6]. Studies looked at the telecommunications market in Latin American countries and found the need for improved spectrum management that follows established policies. Addressing competition in these markets is crucial to improve competitiveness, reduce the digital divide, and promote regional development. Potential investments are also identified through these studies. Countries that allocate more bandwidth and have competitive markets receive more social benefits [7].

Also, it is important to analyze that rapid expansions of market demands have gradually become saturated with the development of the telecommunications industry. Operators convert their attention from users to income. They try their best to increase telecom income via various measures, one of which being to set an amount of income as a goal. Many factors affect income, such as market demand, the policy of subsidies, the invention of new products, and so on. There are many issues to be determined if we analyze the income from the relationship between factors and the variation of income [8].

This paper aims to analyze the income for the three companies that obtained more than 90% of the total market revenue corresponding to mobile Internet service from 2012 to 2022 to determine how income is distributed in companies in this sector. The data is obtained from the Communications Regulation Commission of Colombia. For this, the following statistical measures are used: central tendency and variability, coefficient of correlation, and the analysis of variance (ANOVA). The results show growth in the income of the three operators; however, the Claro operator is the one that receives the most income from this service.

The structure of this paper is as follows: Section II explains the statistical variables used and the data from the telecommunications operators. Section III shows and describes the statistical results. Finally, Section IV presents the conclusions obtained.

II. STATISTICS FOR TELECOMMUNICATIONS

The telecommunications sector is one of the service-based sectors where both the team and the customer relationships are the backbone of the organization. The success of such an

 ISBN: 978-1-68558-145-9

organization will be based on the success of managing an excellent relationship with the clients. The above is possible with the use of statistics. "Statistics is a science that helps us make decisions and draw conclusions in the presence of variability" [9]. Different types of statistics could be applied depending on the case study.

Statistics allows us to understand customer behavior and define the significant variables essential in analyzing customer relationships. This is done through analytical and statistical tools and techniques such as Explanatory and Confirmatory Factor Analysis, Regression and ANOVA, Clustering Analysis, and Classification Analysis. In addition to collecting and reading data correctly, statistics supports its interpretation by making forecasts [10].

Our statistical analysis presents the basic quantitative measures offered by descriptive statistics (mean, standard deviation, and so on), the relationship between the variables analyzed using correlation, and the verification of a hypothesis through ANOVA for the characterization of income in the mobile Internet service of the three companies with the most significant penetration in the Colombian market. Below, we explain these statistical variables.

A. Descriptive measures

An average is a value typical or representative of a data set. Since such typical values tend to lie centrally within a set of data arranged according to magnitude, averages are also called measures of central tendency.

Several types of averages can be defined, the most common being the arithmetic mean, the median, the mode, the geometric mean, and the harmonic mean. Each has advantages and disadvantages, depending on the data and the intended purpose. In this study, we are interested in the arithmetic mean [11]:

1) The arithmetic mean is a measure of a continuous variable's location or central value. For a sample of observations *x1; x2; . . . ; xn* the measure is calculated as

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \quad (1)$$

2) Quartile deviation: If a data set is arranged in order of magnitude, the median is the middle value (or arithmetic mean of the two middle values) that divides the set into two equal parts. By extending this idea, we can think of those values which divide the set into four equal parts. These values, denoted by Q_1, Q_2, and Q_3, are called the first, second, and third quartiles, respectively, the value Q_2 being equal to the median.

3) Standard Deviation: The standard deviation of a set of *N* numbers $X_1, X_2, \ldots, X_N$ is denoted by *std* and is defined by

$$std = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(x_j - \bar{x}\right)^2} \quad (2)$$

where *x* represents the deviations of each of the numbers X_j from the mean $\bar{x}$.

B. Karl Pearson's Coefficient of Correlation

A British biometrician developed a formula called the correlation coefficient to measure the intensity or degree of linear relationship between two variables.

The correlation coefficient between two random variables *X* and *Y*, usually denoted by *r(X, Y)* or simply r_{xy}, is a numerical measure of the linear relationship between them and is defined as:

$$r(X,Y) = \frac{Cov\,(X,Y)}{\sigma_x \sigma_y} \quad (3)$$

where *Cov (X, Y)* is the covariance between *X* and *Y*, which is a value that indicates the degree of joint variation of *X* and *Y* concerning their means. σ_x and σ_y are the deviations of *X* and *Y*, respectively, which is a measure used to quantify the dispersion of a set of numerical data [11].

C. Analysis of Variance

ANOVA is a statistical tool for analyzing how the mean value of a quantitative response (or dependent) variable is affected by one or more categorical variables, known as treatment variables or factors. While ANOVA allows us to compare the means of more than two populations, it can only tell us whether differences appear to exist, not specifically which population means are different. Consequently, the appropriate hypotheses for ANOVA are H_0: $\mu_1 = \mu_2 = \ldots = \mu_k$ (that the population mean responses are equal, where k is the number of populations) and H_a: at least one of the population mean responses, μ_i, is different [12].

The F-statistic is used to accept or reject the hypothesis. The F-statistic is simply a ratio of two variances. Variance is the square of the standard deviation. Therefore, the F-statistic can be summarized as follows:

A sample variance is calculated as SS/df, where SS is the sum of squared deviations from the mean, and df is the degree of freedom [13].

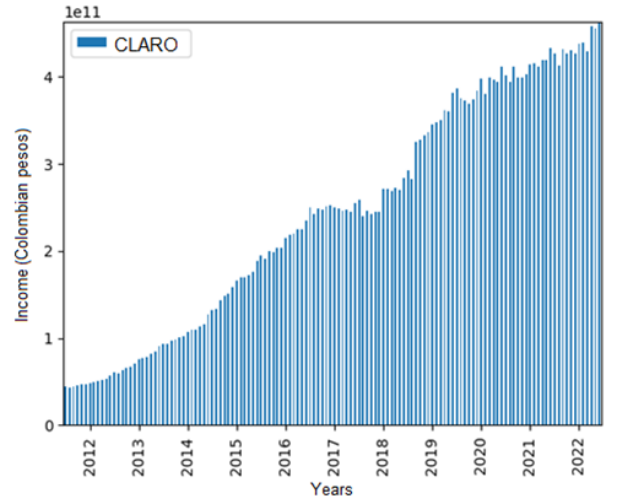
D. Data

The data used for this study corresponds to the sum of postpaid (fixed charge) and prepaid data (on demand). The collection of data from mobile Internet operators in Colombia is carried out on the website of the Communications Regulation Commission (CRC) [14]; then, this data is organized, treated, and debugged to filter the information necessary for the subsequent statistical analysis. With this information, the income of the companies "Claro", "Tigo" and "Movistar" are graphed, as shown in Figure 1.

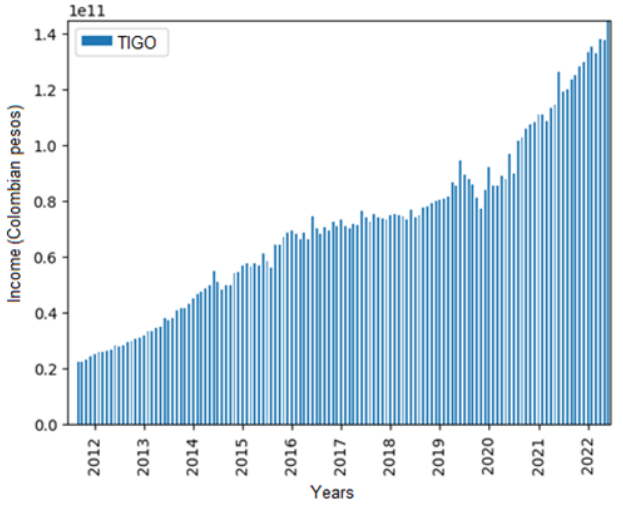
III. RESULTS

The statistical results of Figure 1 are presented below in Table I and Figure 2, which include data count, mean, standard deviation, minimum and maximum sample, and the distribution's first, second, and third quartiles. Table II shows whether there is an increasing trend in income for operators based on the data correlation. At an exploratory level,

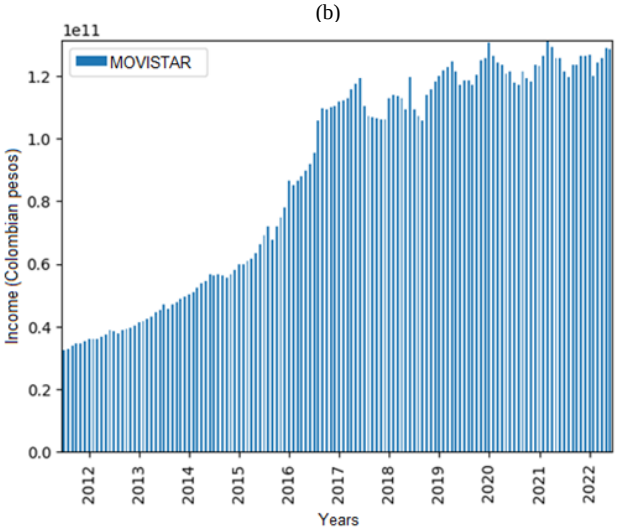
ANOVA is presented in Table III for the operators Tigo and Movistar, since their data have a degree of similarity.



(a)



(b)



(c)

Figure 1. Income in Colombian pesos for mobile Internet companies: (a) Claro, (b) Tigo, (c) Movistar.

TABLE I. STATISTICAL DATA OF MOBILE INTERNET INCOME FOR THE COMPANIES CLARO, TIGO AND MOVISTAR

Index	Claro	Tigo	Movistar
count	132	130	132
mean	242,972,245,121	71,155,903,596	88,576,078,279
std	132,142,308,577	30,241,219,067	34,615,171,364
min	42,717,665,502	22,412,389,160	32,386,111,291
Q_1(25%)	112,687,051,935	49,008,466,545	53,418,243,260
Q_2(50%)	245,620,245,469	71,566,749,718	106,526,211,467
Q_3(75%)	376,555,779,779	85,722,526,660	119,730,605,411
max	462,660,494,175	144,650,203,048	131,331,643,922

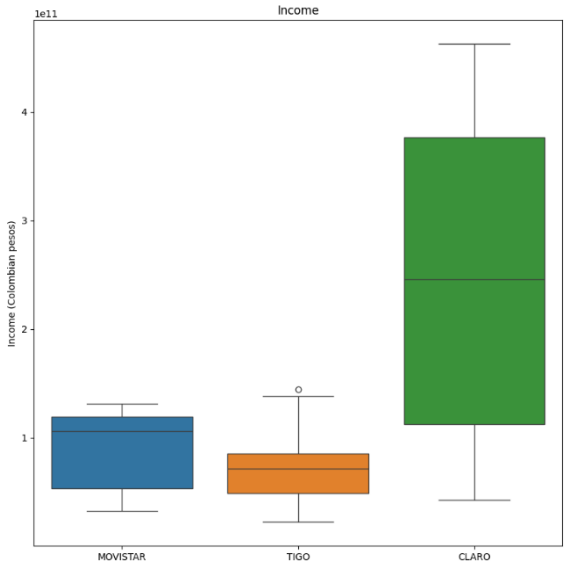


Figure 2. Boxplot for company income from mobile Internet.

TABLE II. COEFFICIENT OF CORRELATION IN MOBILE INTERNET INCOME FOR THE COMPANIES CLARO, TIGO AND MOVISTAR

	Claro income	Tigo income	Movistar income
Claro income	1	0.9586	0.9519
Tigo income	0.9586	1	0.8872
Movistar income	0.9519	0.8872	1

TABLE III. ANALYSIS OF VARIANCE IN MOBILE INTERNET INCOME FOR THE COMPANIES TIGO AND MOVISTAR

	Claro-Movistar income	Tigo-Movistar income
Mean squares	1.57E24	1.98E22
F	168.6319	18.795
Probability	4.197E-30	0.0000208
The critical value of F	3.87719	3.87747

 ISBN: 978-1-68558-145-9

The data exhibited in Figure 1, Table I, and Figure 2 indicate the following:

- Income growth for Claro and Tigo have a linear trend. For Movistar, growth is exponential until 2017, and then it grows oscillatory because of competition in the prices of other operators' plans [15].
- Claro recorded the highest mean income, $242.972.245.121, with a standard deviation of $132.142.308.577. Tigo presented the lowest mean income, $71.155.903.596, with a standard deviation of $30.241.219.067.

In the results of Table II, a high correlation coefficient is observed between the income of the companies evaluated.

With the values in Table III, the relationship of the variances can be compared using the F value to accept or reject the hypothesis. Here, it is observed that in all cases, the value of F is greater than the critical value for F and that the probability value is less than 0,05, which means that the hypothesis is rejected. Therefore, the compared mean values are statistically different.

IV. CONCLUSIONS

The free competition market for the telecommunications sector established in Colombia shows that, for mobile Internet, the operator Claro received, on average, more than twice the income of the operators Tigo and Movistar during the ten years covered by this study.

The high correlation between the incomes of the mobile Internet companies analyzed indicates that operators have increased their income from this service over time.

The ANOVA test between the companies' income confirms the information obtained in Table I regarding the significant difference in the mean income.

Although a high degree of income concentration is observed in the operator Claro, in Colombia, the dynamics of the mobile operator market are one of competition at a technological level and a decrease in rates.

This study shows the income gap between the largest operator and its closest competitors.

It is worth highlighting the work of telecommunications regulatory entities, such as the CRC, which require the frequent sending of information by operators to optimize the processes of collection, processing, and use of information as a tool for regulatory support.

This study does not analyze the consequences of changes in income, such as rates and number portability.

ACKNOWLEDGEMENT

The authors would like to thank the Universidad Distrital Francisco Jose de Caldas, Ministerio de Ciencia, Tecnología e Innovación and Comision de Regulacion de las Comunicaciones for supporting and funding this project.

REFERENCES

[1] M. Fiedler, K. De Moor, H. Ravuri, P. Tanneedi, and M. Chandiri, "Users on the move: On relationships between QoE ratings, data volumes and intentions to churn", in IEEE 42nd Conference on Local Computer Networks Workshops (LCN Workshops), 2017, pp. 97-102.

[2] C. Hernandez, E. Cadena, and L. Pedraza, "Measurement of Competitiveness in the Colombian Mobile Telecommunications Market based on the Linda Index," The Eighteenth International Conference on Systems and Networks Communications (ICSNC 2023) IARIA, Nov. 2023, pp. 10-15, ISSN: 2163-9027, ISBN: 978-1-68558-099-5.

[3] J. Calzada, B. García-Mariñoso and D. Suárez, "Do telecommunications prices depend on consumer engagement?" Information Economics and Policy, vol. 62, Mar. 2023, doi: 101018. 10.1016/j.infoecopol.2023.101018.

[4] A. Shahzad, R. M. S. Yaqub, A. Di Vaio, and R. Hassan, "Antecedents of customer loyalty and performance improvement: Evidence from Pakistan's telecommunications sector," Utilities Policy, vol. 70, Jun. 2021, doi: 10.1016/j.jup.2021.101208.

[5] K. Okeleke, H. F. Aquije and J. Joiner, *The Mobile Economy 2023*. [Online]. Available from: https://www.gsma.com/mobileeconomy/wp-content/uploads/2023/03/270223-The-Mobile-Economy-2023.pdf, [retrieved: 03, 2024].

[6] Comision de Regulacion de Comunicaciones, *Batería de indicadores para el análisis de la competencia en los mercados de comunicaciones*. [in English: Battery of indicators for the analysis of competition in communications markets], [Online]. Available from: https://postdata.gov.co/story/bateria-de-indicadores-para-el-analisis-de-competencia, [retrieved: 03, 2024].

[7] O. S. de Miera Berglind, "Spectrum concentration and market competition. Implications for the use of caps in Mexico," Conference of Telecommunication, Media and Internet Techno-Economics (CTTE), Dec. 2015, pp. 1-8, doi: 10.1109/CTTE.2015.7347228.

[8] M. Wang, Y. Wang, X. Wang and Z. Wei, "Forecast and Analyze the Telecom Income based on ARIMA Model," The Open Cybernetics & Systemics Journal, vol. 9, pp. 2559-2564, Oct. 2015, doi: 10.2174/1874110X01509012559.

[9] D. C. Montgomery and G. C. Runger, Applied statistics and probability for engineers, 7th ed., Singapore: Wiley & Sons, pp. 1, 2014.

[10] M. Abiad, "Increasing telecommunication service performance through statistical analysis", PhD Thesis, Doctoral School of Entrepreneurship, Engineering and Business, University Politehnica of Bucharest, pp. 11, 2021.

[11] M. Spiegel and L. Stephens, Schaum's Outline of Statistics, 6th ed., New York: Mc Graw Hill, 2018.

[12] B. Gunderson, *Statistics 250 Lab Workbook*, Department of Statistics, University of Michigan, 2015. [Online]. Available from: https://open.umich.edu/find/open-educational-resources/statistics/statistics-250-introduction-statistics-data-analysis#handouts, [retrieved: 03, 2024].

[13] H. J. Seltman, *Experimental Design and Analysis*, Carnegie Mellon University, 2018. [Online]. Available from: https://www.stat.cmu.edu/~hseltman/309/Book/Book.pdf

[14] Comisión de Regulación de Comunicaciones, *Postdata.* [Online]. Available from: https://postdata.gov.co/search/type/dataset, [retrieved: 03, 2024].

[15] Comisión de Regulación de Comunicaciones, "Revisión del mercado relevante servicios móviles" [in English: Review of relevant market mobile services], pp.74, 2023.