International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 05 | May - 2023 SJIF 2023: 8.176 ISSN: 2582-3930

Advanced-Data Analysis of Air Conditioner Distribution through E-Commerce Platform

Khushit Mehta[1], Dr. Sunil Maggu[2], Bhaskar Kapoor[3]

[1] Student, Department of IT, Maharaja Agrasen Institute of Technology, Delhi INDIA
[2]Assistant Professor, Department of IT, Maharaja Agrasen Institute of Technology, Delhi INDIA
[3]Assistant Professor, Department of IT, Maharaja Agrasen Institute of Technology, Delhi INDIA

Abstract - Due to the rising temperatures and humidity along with a rise in income levels throughout the world as well as the growing acceptance of air conditioners as a utility and necessity rather than a luxury product, the air conditioning (AC) sector is predicted to see enormous development. A deeper Analysis of AC distribution over the E-Commerce platform is the main idea of this study. Platforms like Flipkart, and Amazon have shown tremendous growth. Various high-selling brands like Voltas Hitachi have shown top-level sales in both offline channels as well as online sales. Bluestar and Samsung have been giving fierce competition in online channels. The study started with gathering data from an E-Commerce platform by using web scraping tools. Once the data was collected it was cleaned using data cleaning tools of checking nan values and removing unnecessary data rows. The data visualisation of the case study of the Air Conditioner data on the Flipkart platform till December 2022 was performed and various plots like Bar-graph, Pie-chart, Histogram, Heatmap, etc. were plotted. Many parameters were taken into consideration like the number of reviews, star rating, and number of products of a particular brand, and results were obtained. Some key results are like in offline sales Voltas and Hitachi are market leaders with Bluestar and Samsung close behind. The number of products offered by number is provided by Voltas and Hitachi as well as giving customers a wide range of products to choose from. Overall, the customer seems satisfied because most of them have given 3.8 – 4.8 star ratings to the products. As technology advances, we'll see an increase in the sales of products online grow many folds.

***Key Words*:** AC, star rating, Voltas, Hitachi, E-Commerce, Flipkart

1. INTRODUCTION

The term "air conditioning" is frequently used to refer to the machinery employed in the process of eliminating heat, cold, and moisture from the inside of a populated place. With the aid of a gaseous refrigerant, a conventional air conditioner employs a mechanical device called a "heat pump" to transport heat between the interior and outdoor units. Heat is transferred from the interior exchanger to the outside exchanger via an air conditioner. The outdoor unit's compressor liquefies the refrigerant while cooling to remove heat. Conversely, while heating a room, heated pressure gas produced by the outside unit's heat exchanger and compressor is condensed to heat the air. The system may rotate between these two cycles to provide cooling and heating. Due to the rising temperatures and humidity levels throughout the world as well as the growing acceptance of air conditioners as a utility rather than a luxury product, the air conditioning (AC) sector is predicted to see enormous development. Through the projection period, it is also projected that the advent of technologically sophisticated air conditioners, such as air conditioners with inverter technology and air purification technologies, would have a beneficial impact on the air conditioner market. The industry is also anticipated to be fueled by elements like the tourist and construction sectors' optimistic expansion. Demand for air conditioning systems is predicted to rise as a result of the growing population, which is likely to have a favorable impact on residential and commercial buildings.

2. LITERATURE SURVEY

[1].Soni and Vishal Dinesh Kumar gave insights from the current study on how artificial intelligence affects online business. With its explosive growth over the past few years, the e-commerce age might be regarded. The development of several platforms in tandem with technology advancements might be helpful for keeping up with trends and identifying market demands. Therefore, the focus of this study was on how artificial intelligence is used in e-commerce. [2]. Dilip Kumar Sharma and his co-authors studied the products from Flipkart and Amazon and suggested an e-commerce portal based on user requests, and the user can purchase the product from any site depending entirely on user preferences. In this search paper, data such as product names, prices, reviews, and ratings related to suggested products from Flipkart and Amazon are scraped using a robotic process automation tool. Data is cleaned to remove noisy values.[3]. Manisha Jain along with the co-authors researched star ratings on Appliance labels to tell customers about the energy use of their appliances and required labeling is a key policy instrument for encouraging the adoption of energy-efficient equipment. Since 2006, India has implemented standards and labeling programs for numerous appliances in accordance with international best practices. By evaluating the value consumers in India place on energy labels, this article offers empirical support for this strategy. A discrete choice experiment is used to determine the willingness to pay (WTP) estimates for energy labels and more efficiency as stated on labels in air conditioner purchasing choices. The parameters for selected air conditioner qualities are estimated using a mixed logit model. The parameters for selected air conditioner qualities are estimated using a mixed logit model. The findings indicate that consumers give labels a favorable value. When creating alternative energy efficiency programs in combination with the labeling program, the WTP for higher star ratings is assessed.[4] The author researched the AC distribution in the region of New Delhi, India, and how weather and

development of various zones were important. The main aim was to identify the hotspots of the region so that big companies can focus on areas where the prospects of sales are maximum.[5] E-commerce is gaining momentum as the sale and purchase of goods transition from offline to online platforms. his study focuses on developing a product recommendation system for e-commerce platforms, aiming to cater to user preferences. Collaborative Filtering, specifically the Slope One algorithm, is utilized to generate recommendations based on user input ratings and product relationships represented by a domain ontology.[6] This research examines the impact of e-commerce business practices on tax avoidance through an empirical analysis. The study focuses on European parent companies in the retail trade sector from 22 countries. The findings indicate that e-commerce firms exhibit a significantly higher degree of tax avoidance compared to traditional firms. However, it is noteworthy that traditional firms have also shown an increasing tendency to engage in tax avoidance during the study period, leading to a narrowing of the gap between the two types of firms. The results remain consistent when considering various measures of tax avoidance, timeframes, and sample selection criteria.[7] Kumar J and co-authors worked on the challenges, opportunities, and development, of sustainable development in the electrical sector in India.[8] Anam Malik and co-authors in 2019 research Findings suggest that residential air-conditioning presents a significant demand response opportunity, approaching perhaps 9% of total peak demand in some circumstances

3. METHODOLOGY

Purpose: Initially we have to look at the review and ratings of various ACs as per the Flipkart web scraped data. The star rating of various ACs along with discounts will help in deeper analysis for a better understanding of the data. Large MNCs require this type of careful analysis so that they can target the product which is yielding the best results.

A. Workflow Setup

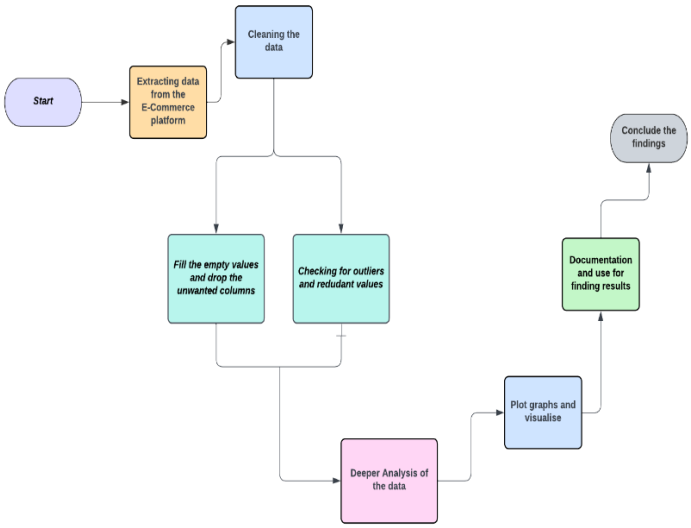

Fig1. The flow of work

B. Tech Stack and Tools

Python is a good language for practical programming and quick learning. A programmer by the name of Guido van Rossum developed a robust high-level programming language.[9] For conducting standard data operations and analysis on such data sets, the Pandas library offers integrated, simple procedures. It seeks to serve as the core of Python's statistical computing system in the future.[10]A Portable Python Plotting Package. matplotlib is a portable 2D plotting and imaging package aimed primarily at the visualization of scientific, engineering, and financial data. matplotlib can be used interactively from the Python shell, called from Python scripts, or embedded in a GUI application (GTK, Wx, Tk, Windows).

C. Data Understanding

Data required for the research was collected from a reliable source (confidential) and different features were analyzed according to the demand and requirement of the research. Preparation: The data in the .csv format was in raw form and had to be prepared for usage, further the data was cleaned, transformed, and certain unnecessary columns were dropped. This data understanding step gives us more perspective on how to move ahead with our analysis.

D. Data Cleaning

You may determine what needs to be cleaned up or processed by looking at feature values until you find the range or distribution of values that are normal for each characteristic. Incomplete or missing data might include the erroneous data form for a column, the wrong measurement units for a certain column, or a lack of examples for a given class.

There are numerous major advantages to this data cleansing procedure, like:

1. By doing this, significant flaws and inconsistencies that are unavoidable when combining data from many sources into one dataset are removed.

2. Having data cleaning software would increase everyone's productivity since they will be able to simply extract the information they require from the data.

3. Happy customers and content employees result from fewer errors.

4. The capability to visualize the many functions, what your data is intended to do, and where it originates.

The core of data science is based on these 5 pillars which are in turn the basis of any machine learning model

Fig2 is the result of data cleaning showing that the column values have no null or irrelevant data. Data cleaning makes data analysis very easy and gives better results.

```
ac.isnull().sum()
```

```
Name              0
Brand             0
Ratings           0
No_of_ratings     0
No_of_reviews     0
Product_features  0
MSP               0
MRP               0
Discount          0
dtype: int64
```

```
ac.nunique()
```

```
Name              239
Brand              28
Ratings            21
No_of_ratings     168
No_of_reviews     114
Product_features  296
MSP               192
MRP               205
Discount           50
dtype: int64
```

Fig2. Look at data

```
UNIQUE BRANDS

ac.Brand.unique()
```

```
array(['CARRIER', 'IFB', 'SAMSUNG', 'Voltas', 'LG', 'Whirlpool', 'Blue',
       'Lloyd', 'MarQ', 'realme', 'ONIDA', 'MOTOROLA', 'Panasonic',
       'Daikin', 'Hitachi', 'Thomson', 'Godrej', 'Haier', 'TCL', 'O',
       'TOSHIBA', 'Nokia', 'Midea', 'LIVPURE', 'Sansui', 'CANDY',
       'iFFALCON', 'LumX'], dtype=object)
```

Fig3 Unique Brands

Image 3 shows the number of all unique brands available on Flipkart showing the wide range of products for the customer to choose from.

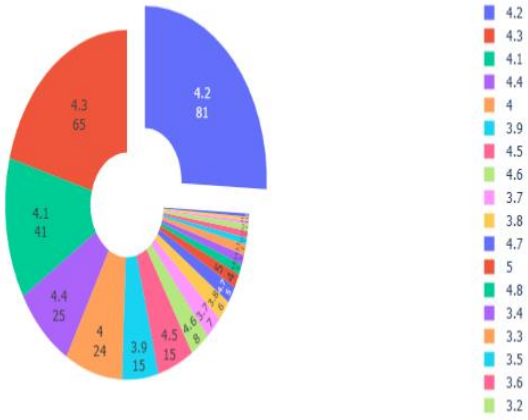

Fig 4. Pie Chart Ratings

Fig.4 is a pie chart of user star ratings for the different AC products they have bought from Flipkart (E-Commerce platform). An average rating of 4 stars on a scale of 5 means the customers are happy with the products they received from the E-Commerce platform. This image is a pie representation of Ratings given VS Count of each unique rating

E. Data Visualization

The cleaned data were checked for redundancy and various graphs like pie charts, line graphs, and bar graphs for plotted for better understanding and assessment.

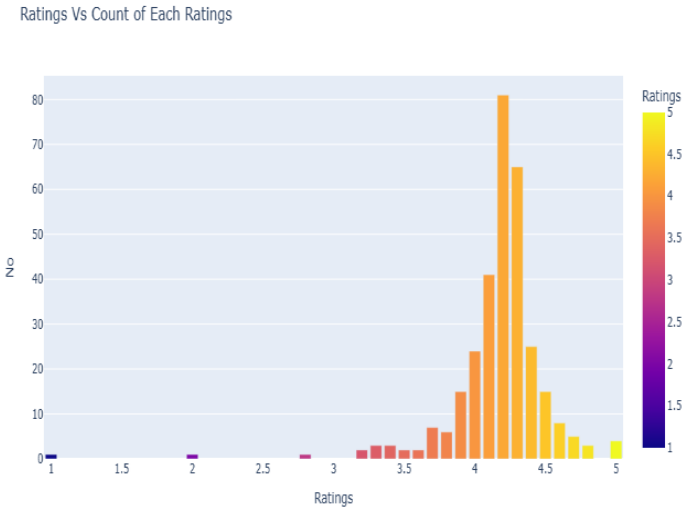

Fig5. Heatmap rating vs count

Fig.5 is a Bar-Heat map of the customer review star ratings which shows a better view of customer satisfaction as the majority rating range is from 3.7 stars – 4.6 stars, indicating good reviews and customer satisfaction.

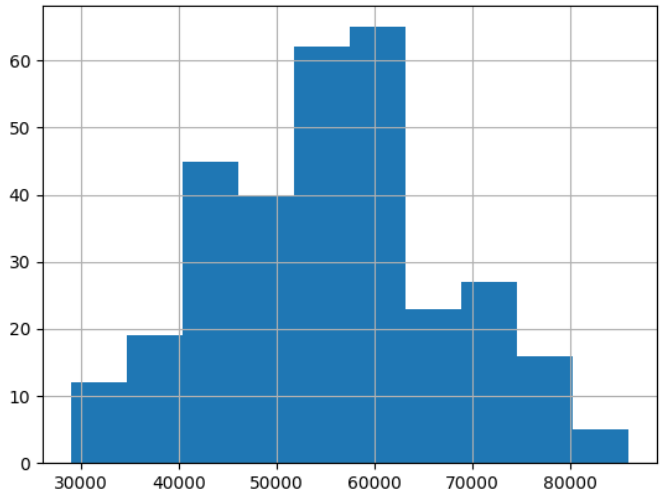

Fig6. Histogram of MRP

The Fig6 is a histogram of MRP of the AC available giving an insight into the price range of products. We can say the average price ranges from 40-60K INR. This average cost also shows that in the Indian market, the best rates to launch an AC should lie in this price range category so as to maximize sales.

4. RESULTS AND DISCUSSION

The AC in today's world is selling like fans because of the rise in the overall temperature of the Earth's surface. People prefer to buy a Three-star rating AC because they have the best price and middle-class people prefer to buy products that suit their budget.

Big Brands like Voltas, Hitachi, Bluestar, Samsung, and Daikin who are the market leaders in offline sales are also the major players in the Ecommerce game. Platforms like Flipkart, and Amazon are a major convenience to everyone

who possesses a smart device as they have basically eradicated the need to visit shops to buy commodities.
They are providing

- Services
- Ease of delivery
- Easy Returns
- Fast payment modes

People in return give reviews based on the satisfaction of the product they buy from the E-commerce platform by giving star ratings.
The most given star rating by customers was 4.2. The number of unique products offered by a brand is 36 (Voltas).

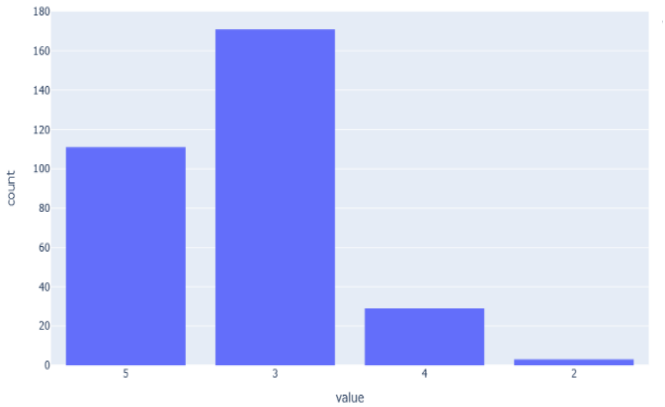

Fig7. Bar Plot Star

Fig.7 is a histogram of AC count under a broader umbrella of star ratings, the above graph shows many people have bought and given reviews to 3star AC because the price of 3star AC is comparatively less as compared to more efficient 4&5star AC.

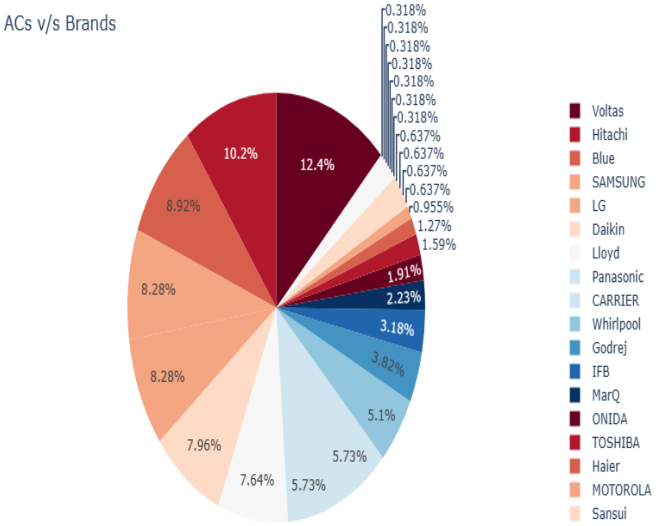

Fig8. Heat Pie chart

Fig.8 is a detailed pie chart of the percentage distribution of AC on the E-Commerce platform with 12.4% of the total market Voltas is the leader even in the online sector. Hitachi is close behind at position two showing a huge online presence in the Indian market. BlueStar, Samsung, LG, and Daikin are the top brands with a large number of products.

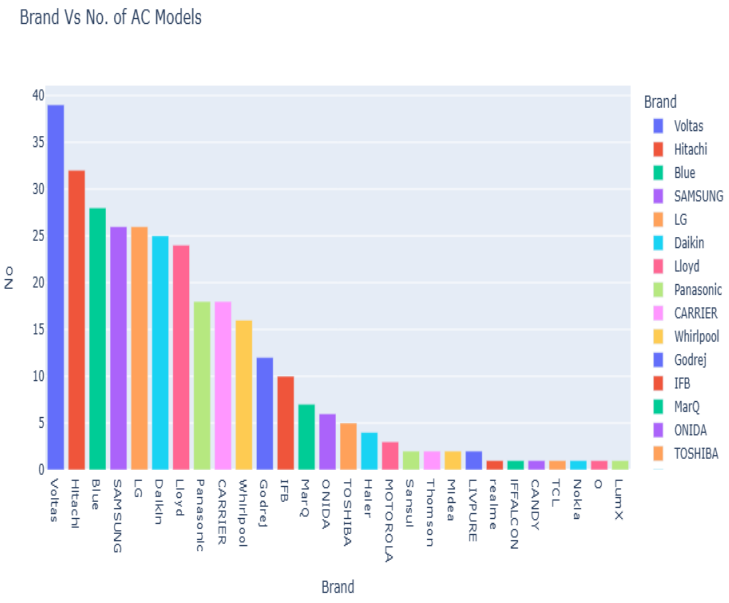

Fig9. Brand vs Number of Models

Fig9. Shows how many different types of models up for display are for sale. This is done to attract customers towards a particular brand so that he or they might have the best catalogue available for a choice of product.

5. CONCLUSION

Voltas, Hitachi leads the way in online sales as well. Many brands have better market penetration in the online channel rather than the offline channels. The customer gives a positive review for most of the products brought. The review rating has a median value of 4.5 showing great satisfaction. The average price range of the product brought is 50-60k INR. Big brands like Voltas Hitachi have shown aggressive marketing by displaying a large number of models on sale giving customers a wide range of products to choose from. The share of Voltas is 12% and of Hitachi is 10% and in the offline sales too they had the best market share.

6. FUTURE SCOPE

Now that the analysis of both online and offline channels has been done, we will try to apply some mathematical tools to further expand the scope of sales. The results of the study can be given to various MNCs so that they can some reference and focus on the product which is high in demand.

REFERENCES

1. Soni, Vishal Dineshkumar, Emerging Roles of Artificial Intelligence in Ecommerce (July 11, 2020). International Journal of Trend in Scientific Research and Development | Volume 4 | Issue 5 | August 2020 | pp.223-225, Available at SSRN: https://ssrn.com/abstract=3648698.
2. Dilip Kumar Sharma, Sarika Lohana, Saurabh Arora, Ashutosh Dixit, Mohit Tiwari, Tripti Tiwari,E-Commerce product comparison portal for classification of customer data based on data mining, Materials Today: Proceedings, Volume 51, Part 1,2022,Pages 166-171,ISSN 2214-7853, https://doi.org/10.1016/j.matpr.2021.05.068.(https://www.sciencedirect.com/science/article/pii/S2214785321036440)

3. Manisha Jain, Anand B. Rao, Anand Patwardhan, Consumer preference for labels in the purchase decisions of air conditioners in India, Energy for Sustainable Development, Volume 42,2018, Pages 24-31, ISSN 0973-0826, https://doi.org/10.1016/j.esd.2017.09.008.
4. Analytical review of AC Distribution in the Delhi Region based on Climate change and Development KhushitMehta https://ijsrem.com/download/analytical-review-of-ac-distribution-in-the-delhi-region-based-on-climate-change-and-development/.
5. N. M. S. Iswari, W. Wella and A. Rusli, "Product Recommendation for e-Commerce System based on Ontology," 2019 1st International Conference on Cybernetics and Intelligent System (ICORIS), Denpasar, Indonesia, 2019, pp. 105-109, doi: 10.1109/ICORIS.2019.8874916.
6. Josep M. Argilés-Bosch, Antonio Somoza, Diego Ravenda, Josep García-Blandón, An empirical examination of the influence of e-commerce on tax avoidance in Europe,Journal of International Accounting, Auditing and Taxation,Volume 41,2020,100339,ISSN 1061-9518, https://doi.org/10.1016/j.intaccaudtax.2020.100339.(https://www.sciencedirect.com/science/article/pii/S1061951820300409)
7. Kumar. J, C.R., Majid, M.A. Renewable energy for sustainable development in India: current status, future prospects, challenges, employment, and investment opportunities. Energ Sustain Soc 10, 2 (2020). https://doi.org/10.1186/s13705-019-0232-1
8. Malik, A.; Haghdadi, N.; MacGill, I.; Ravishankar, J. Appliance level data analysis of summer demand reduction potential from residential air conditioner control. Appl. Energy 2019,235, 776–785. [CrossRef]
9. Mckinney, Wes. (2011). pandas: a Foundational Python Library for Data Analysis and Statistics. Python High-Performance Science Computer.
10. Barrett, Paul & Hunter, J. & Miller, J.T. & Hsu, J.-C & Greenfield, P.. (2005). matplotlib -- A Portable Python Plotting Package.