# Website Usability Evaluation through Sentiment Analysis of the User Population of the IRCTC Website

**Dr. Rinku Sharma Dixit***

*Associate Professor, Department of Business Analytics, New Delhi Institute of Management*
*60 & 50 (B&C), Tughlakabad Institutional Area, Behind Batra Hospital, New Delhi, Delhi 110062*


**Prof. Shailee Choudhary**

*Assistant Professor, Department of Business Analytics, New Delhi Institute of Management*
*60 & 50 (B&C), Tughlakabad Institutional Area, Behind Batra Hospital, New Delhi, Delhi 110062*

## Abstract

IRCTC is the subsidiary of Indian Railways that that handles the catering, tourism and online ticketing operations of the Indian railways. It is one of the biggest e-commerce portals of India and primarily operates through its website. Website Content and Features are designed keeping the user population in mind, which range from all sections and levels of the society. Therefore this paper was designed consequent to a survey to test the acceptability of the Website and its features by the user community. The survey was designed specifically to test the suitability and relevance of the Website Content & its Features, the design of the User- Interface elaborating the Visual Appeal, the Website Capabilities and the Overall Impression of the Website along with specific comments from the users. Analysis of the responses to the various aspects under the above stated questions along with sentiment analysis of the comments has been presented in this paper. The results suggest that the Website has above average capabilities and usability. But certain features of the website such as speed, performance, interface through the mobile, placement of certain useful features on the website, detailed in the paper, need improvement have been discussed in the Conclusions. The results of analysis have been depicted through graphs prepared with Visualization tool. It has been specifically identified that results of the questionnaire provide a mathematical overview of the opinions, but sentiment analysis of the comments provide a more elaborate and detailed review of the public perception about the Website content and its features.

**Keywords:** Sentiment Analysis, Website Features, Data Analytics, Website Usability Evaluation, Visualization, Satisfaction Index

## INTRODUCTION

Indian Railways, launched in 1853 with its first passenger train running between BoriBandar and Thane, with 400 passengers [1] is one of most successful ventures today. Since then, it has come a long way adding a number of features to the basic facility such as online ticketing and now railway tourism. Railways has grown to become the largest public transporter and largest employer of the country by carrying more than 2.3 crore passengers daily and employing almost 13 lakh people, respectively [1]. Railways' total running track stood at 93,902 kms with 7,349 stations and 2,77,987 wagons, as on March 31, 2017, according to Indian Railways official website.

IRCTC (Indian Railway Catering and Tourism Corporation) is a subsidiary of the Indian Railways that handles the catering, tourism and online ticketing operations of the Indian railways. IRCTC has emerged as the biggest e-commerce portal in India, and has to its credit a number of records in terms of number of tickets booked in a day. Within a short span of its going online, it had become the largest and the fastest-growing e-commerce website in the Asia-Pacific region. It provides the following services for Indian Railways:

1. **Online Ticketing:**

   a) *Online Ticketing:* Online ticketing, introduced by IRCTC, changed the face of railway ticketing in India. It pioneered internet-based rail ticket booking through its website, as well as from the mobile phones via WiFi, GPRS or SMS. In addition to e-tickets, Indian Railways Catering and Tourism Corporation also offers I-tickets that are basically like regular tickets except that they are booked online and delivered by post [2].

   b) *PNR Status:* The PNR (Passenger Name Record) Number is also made available for status checking. Status may checked online or through SMS.

   c) *Shubh Yatra:* Shubh Yatra, a loyalty program for frequent travellers has also been launched, that offers discounts to passengers on all the tickets booked round the year after by paying an upfront fee.

   d) *Rolling Deposit Scheme (RDS):* A scheme called Rolling Deposit Scheme (RDS) has been launched, which is a hassle-free e-ticket booking scheme that allows passengers to reserve seats against advance money kept with the corporation. IRCTC has also added flights and hotels booking facilities to their line of online reservation services [3] [4]

2. **Tatkal Scheme**

Under the Tatkal scheme, passengers who plan their journey at short notice can book their tickets through the Indian railways internet portal. The booking starts one day prior to the departure of the train from source station [6].

## 3. Tourism

Indian Railways Catering and Tourism Corporation also organises budget and deluxe package tours for domestic and foreign tourists.

a) ***Bharat Darshan:*** Bharat Darshan is a popular tourism package for budget tourists covering important tourist destinations across India. Luxury tourism packages have also been designed that involve special luxury trains such as Buddhist Circuit Train and Maharajas' Express [7].

b) ***Adventure Tourism Packages:*** Special adventure tourism packages have been offered that include water sports, adventure and wildlife treks, etc. Tours may be customised also as per specific requirements offers an added attraction. Rail Tourism India, is a new venture of Indian Railways that provides direct catering and tourism packages to users [7].

## Objectives

The Objectives of the Study are:

- Evaluate the IRCTC Website Usability
- Understand and analyse the Features and their Effectiveness in the IRCTC Website from the user perspective.
- Suggest measures to improve the Website Effectiveness and Usability

## RESEARCH METHODOLOGY

**Research Study:** The study was based on a Survey Research Method after collecting actual facts and figures. The Survey was conducted in the form of an Online Survey Questionnaire during December 2017 to January 2018. Information about the survey, including the link, was emailed to a vast user population. The responses received were thereafter analysed.

**Type of Data:** The study is descriptive in nature and works on both Primary and Secondary Data. The Primary Data was collected through the online Questionnaire filled by the User Population. While the Secondary Data was collected through various sources such as official reports, websites and the various information repositories available on the Internet.

**Sample Size & Description:** The Link was mailed to vast user population including students, teachers, other professionals and general public. The responses received were thereafter analysed. Total 540 responses were received. The user community belonged to all age groups ranging from 20 years to 60 years and above. The maximum respondents were from the age group of 20 to 30 years closely followed by those in the age group of 30 to 40 years.

## Sentiment Analysis

Sentiment Analysis is the process of computationally identifying, processing and categorizing text data to obtain opinions / opinions and standpoints regarding themes, topics, products, etc. Another definition says Sentiment analysis is a technique used to detect positive and negative opinions about specific products or services using large amounts of sources such as customer feedbacks, forums, blogs and other social media [10]. There are various methods to recognise or mine people's attitudes, ranging from human analysis to supervised and unsupervised machine learning with programming languages such as Python and R ("bag of words"). Sentiment analysis can be based on object/feature, when we determine the expressed sentiment on different objects of entities or based on document level. The latter determines the polarity of, for example, news items or product reviews. There are various ways to show sentiment, from the most common three polarity levels positive, neutral and negative to polarity scales of, for example -10 to +10. In the current study the sentiments have been categorized as per the three polarity levels i.e. positive, neutral and negative. Mathematical responses provide categorical data but Sentiment Analysis adds meaning to it. It has been analyzed in this paper that elaborate requirements of the user can only be obtained by analyzing their detail comments through techniques such as Sentiment Analysis.

## Data Analysis & Interpretation:

The entire Questionnaire was divided into four sections with successive sub sections. Each of the assessment criteria in each section was evaluated on a 5 point scale ranging from 1 to 5 with point 1 as Extremely Dissatisfied, 2 as Somewhat Dissatisfied, 3 as Neutral, 4 as Somewhat Satisfied and 5 as Completely Satisfied. Analysis of the results was also done based on Age Groups (20-25, 25-30, 30-40, 40-50, 50-60 and Above 60) and the reactions of the respondents in respective age groups have been studied and depicted graphically.

***Section 1 focussed on Website Content & Features*** and had the following 9 assessment parameters:

1.1. IRCTC and its Services can be easily understood from the Website

1.2. Home page content is impressive and informative

1.3. It is easy to find information on the website

1.4. Usefulness of the information provided by the website

1.5. Ease of finding how to contact us

1.6. Clicking on links takes me to what I expect.

1.7. The organization of information on the system screens is clear.

1.8. The vocabulary on our website is appropriate for the intended audience.

1.9. The organization of the site is logical and clear.

## ANALYSIS & INTERPRETATION

### 1.1. IRCTC and its Services can be easily understood from the Website
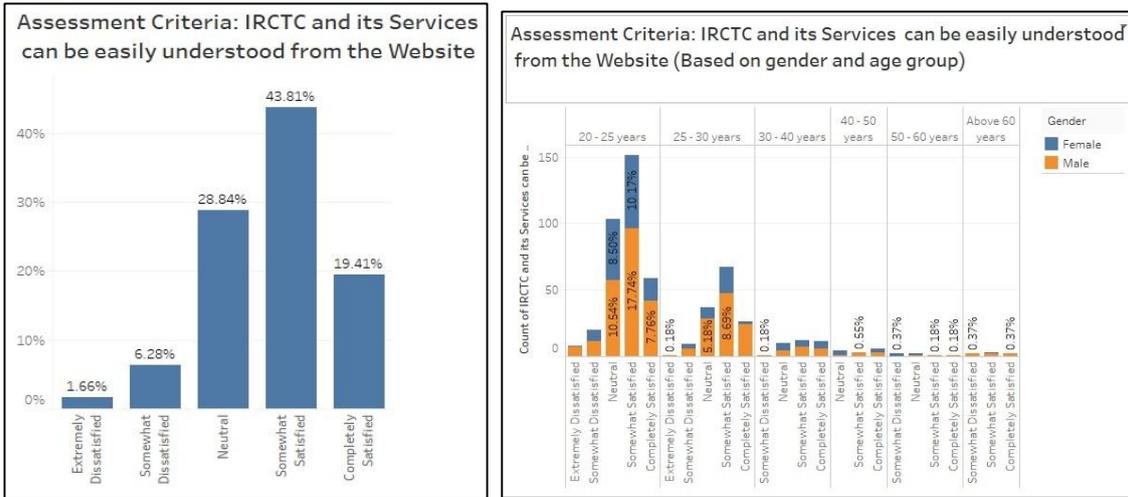


**Figure 1.1 a)** IRCTC and its Services can be easily understood from the Website,
**b)** IRCTC and its Services can be easily understood from the Website (Based on Gender and Age Group)

**Analysis:** As is evident from the graph in Figure 1.1a, the user population seems to be quite satisfied with depiction of the IRCTC and Services through its website. Almost 63% users have given score more than the average while only 8% users have given poor response, which can be attributed to personal grievances faced by specific individuals. Hence the average user population is showing satisfaction.

Figure 1.1b shows a more detailed report. It is clearly evident, that the maximum number of users are from the age groups 20-30, showing that the website is used mostly by young people and there too the number of male users exceeds that of female users.

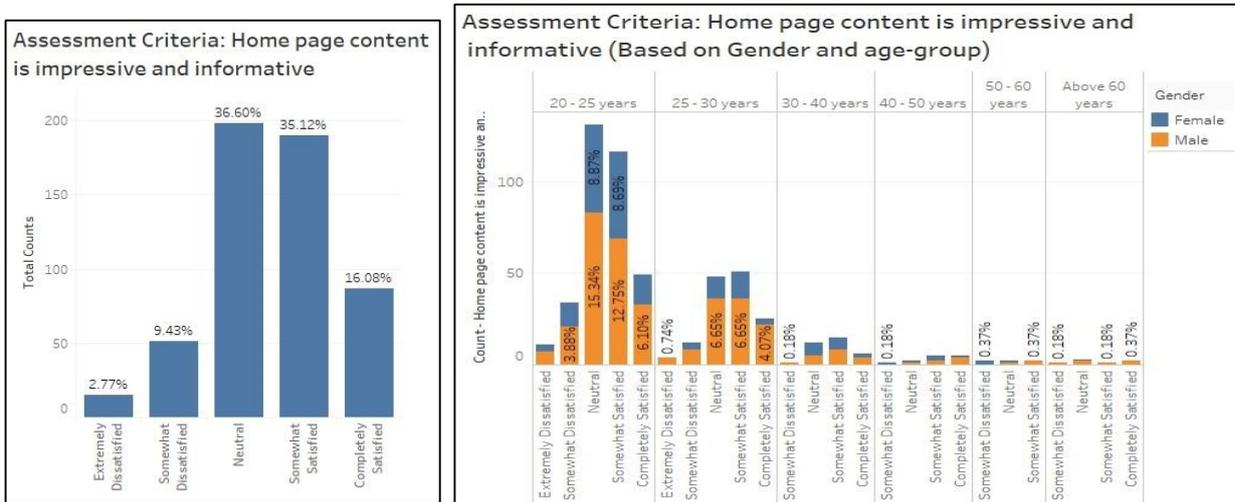### 1.2. Home page content is impressive and informative



**Figure 1.2 a)** Home page content is impressive and informative,  **b)** Home page content is impressive and informative (Based on Gender and Age Group)

**Analysis:** Figure 1.2a shows that the maximum numbers of users have given average marks to this parameter. This suggests that the average people expect that the home page content can be made more impressive and informative. Almost 10% users

seem to be dissatisfied with the home page and its content. Figure 1.2b shows the same trend with regard to gender and age groups also. The young population which is quite internet

friendly expects the website to be more impressive and informative.

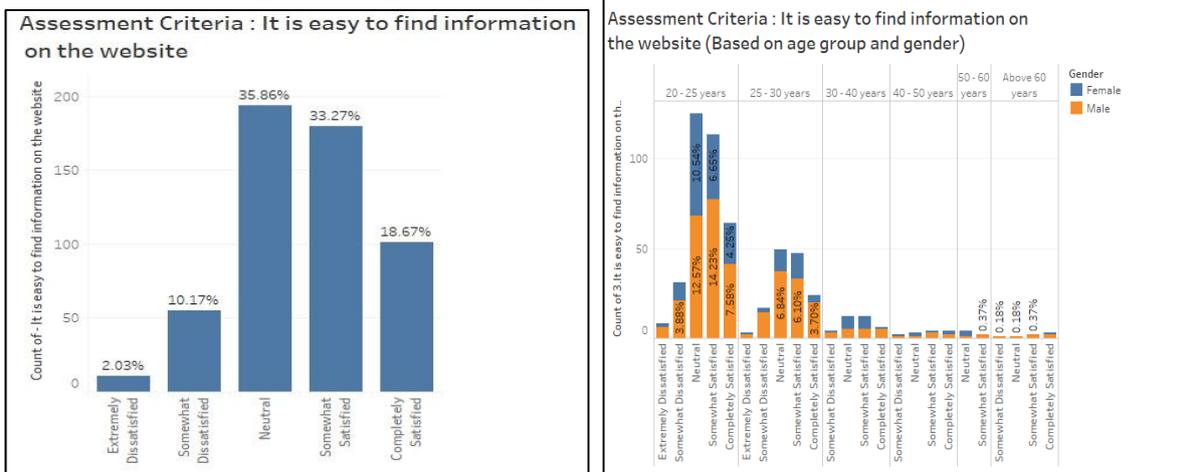### 1.3. It is easy to find information on the website



**Figure 1.3 a)** It is easy to find information on the website,  **b)** It is easy to find information on the website (Based on Gender and Age Group)

**Analysis:** The sub-parameter 3 is related to sub-parameter 2 and depicts the same trend that the young population expects more ease in finding information on the website.
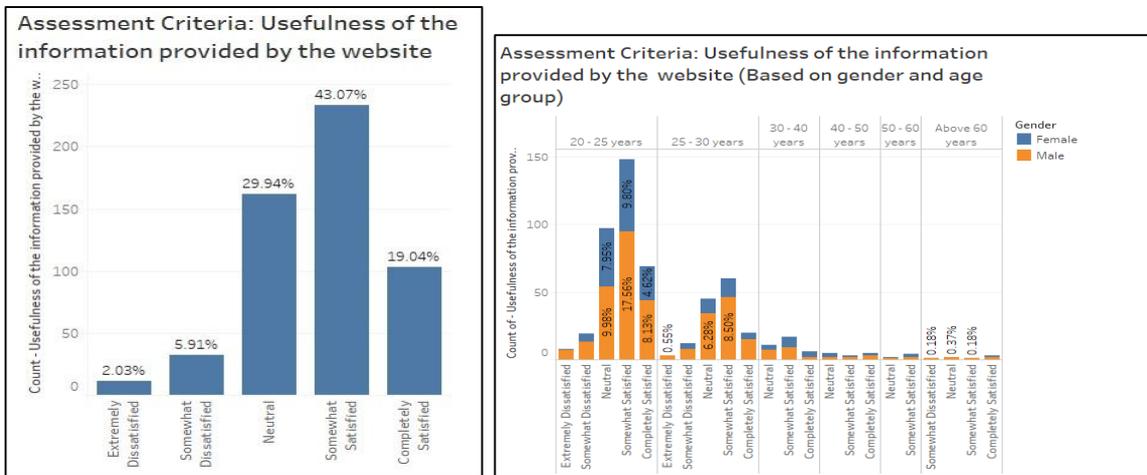
### 1.4. Usefulness of the information provided by the website



**Figure 1.4 a)** Usefulness of the information provided by the website,   **b)** Usefulness of the information provided by the website (Based on Gender and Age Group)

**Analysis:** As per Figure 1.4a, more than 50% users are satisfied that the information provided by the website is useful. This shows that the website is helpful and provides all information required for booking train tickets and for making journey decisions as is also evident from the awards and accolades provided to IRCTC.
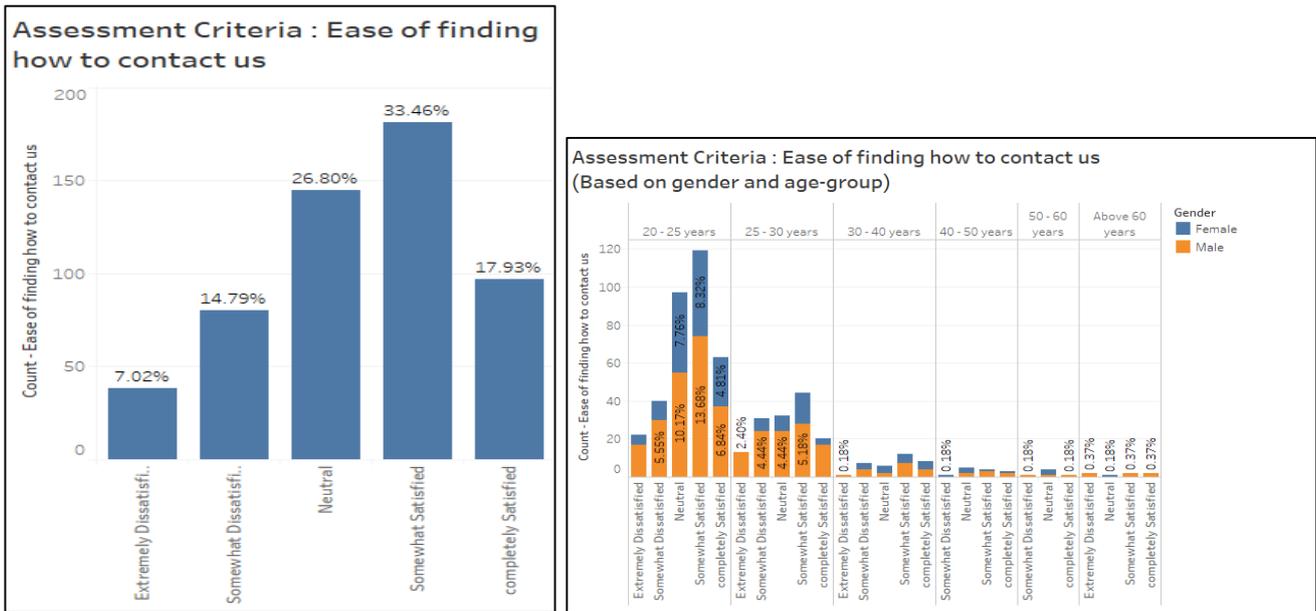
## 1.5.  Ease of finding how to contact us



**Figure 1.5 a)** Ease of finding how to contact us,   **b)** Ease of finding how to contact us (Based on Gender and Age Group)

**Analysis:** As per Figure 1.5a and 1.5b, most of the younger population feels that it is easy to find the contact details but those in upper age groups find it difficult.

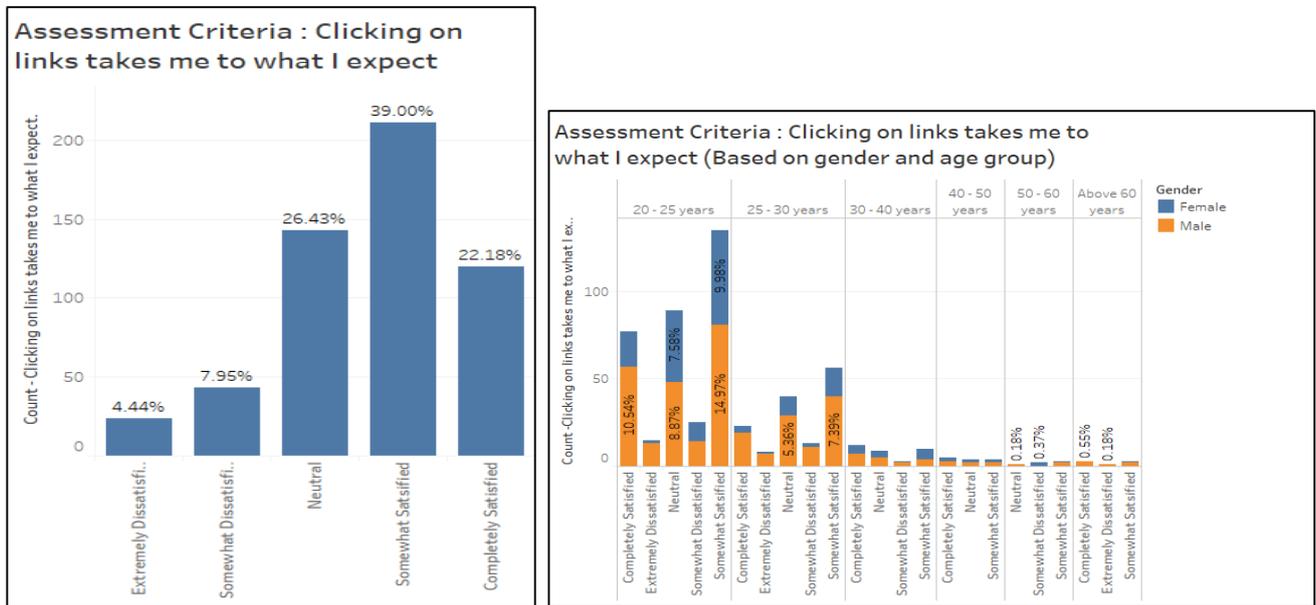## 1.6.  Clicking on links takes me to what I expect



**Figure 1.6 a)** Clicking on links takes me to what I expect, **b)** Clicking on links takes me to what I expect (Based on Gender and Age Group)

**Analysis:** Figure 1.6a and 1.6b show that more than 80% user population is satisfied with the fact that the links lead to the expected links and information meaning thereby that the information on the website is organised as per user expectations and the other general standards.

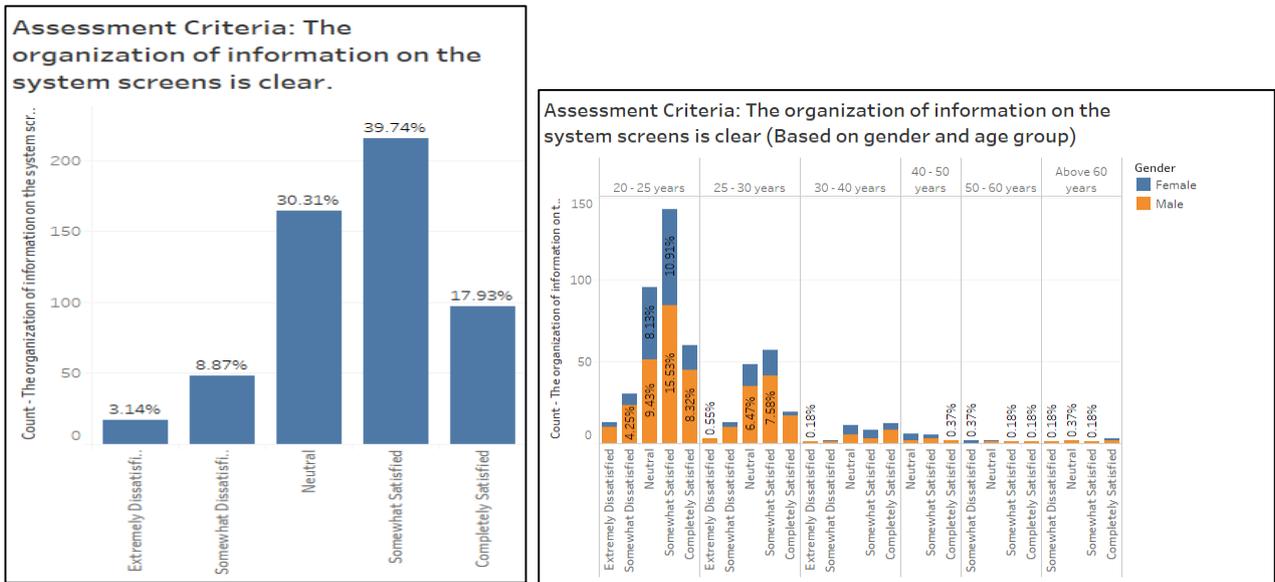*1.7. The organization of information on the system screens is clear.*



**Figure 1.7 a)** The organization of information on the system screens is clear**, b)** The organization of information on the system screens is clear (Based on Gender and Age Group)

**Analysis:** Almost 90% users are satisfied with the organization of the information on the website. This is quite related to sub-parameters discussed above and therefore show the same trend.

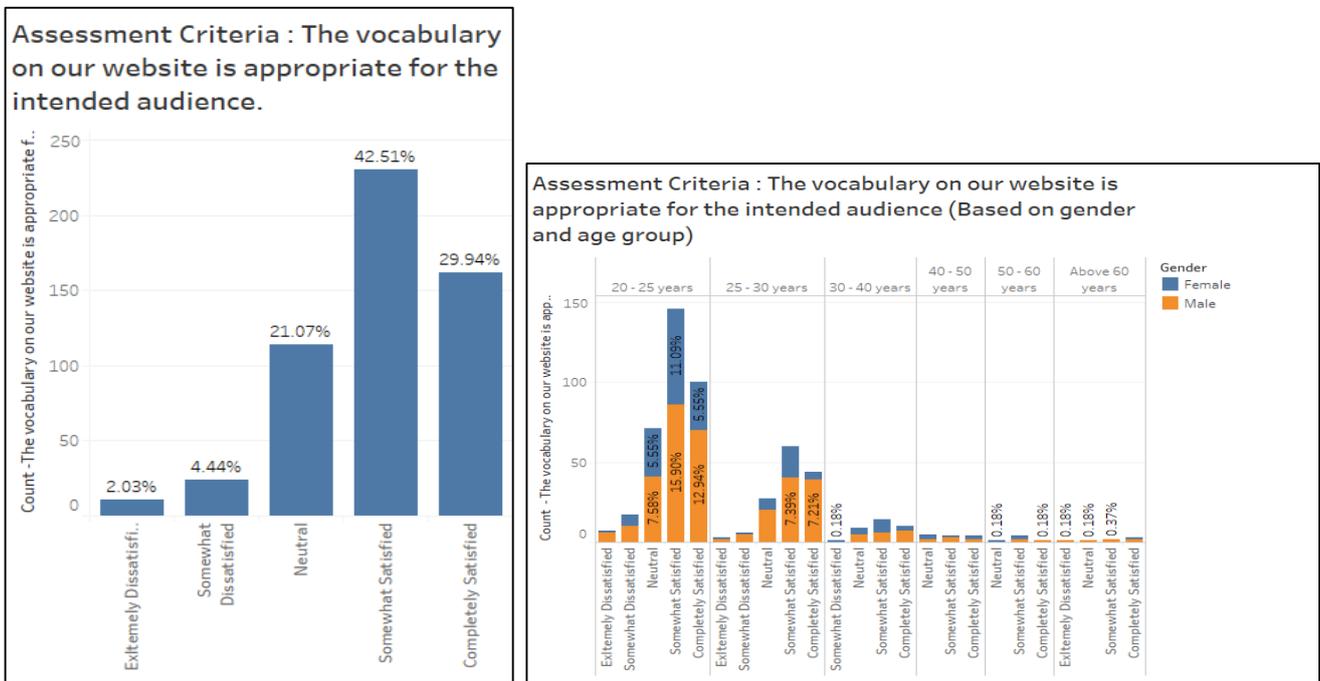*1.8. The vocabulary on our website is appropriate for the intended audience*



**Figure 1.8 a)** The vocabulary on our website is appropriate for the intended audience, **b)** The vocabulary on our website is appropriate for the intended audience (Based on Gender and Age Group)

**Analysis:** This parameter is also in conformity that the vocabulary is appropriate for the intended audience and people are able to understand the terms stated there.

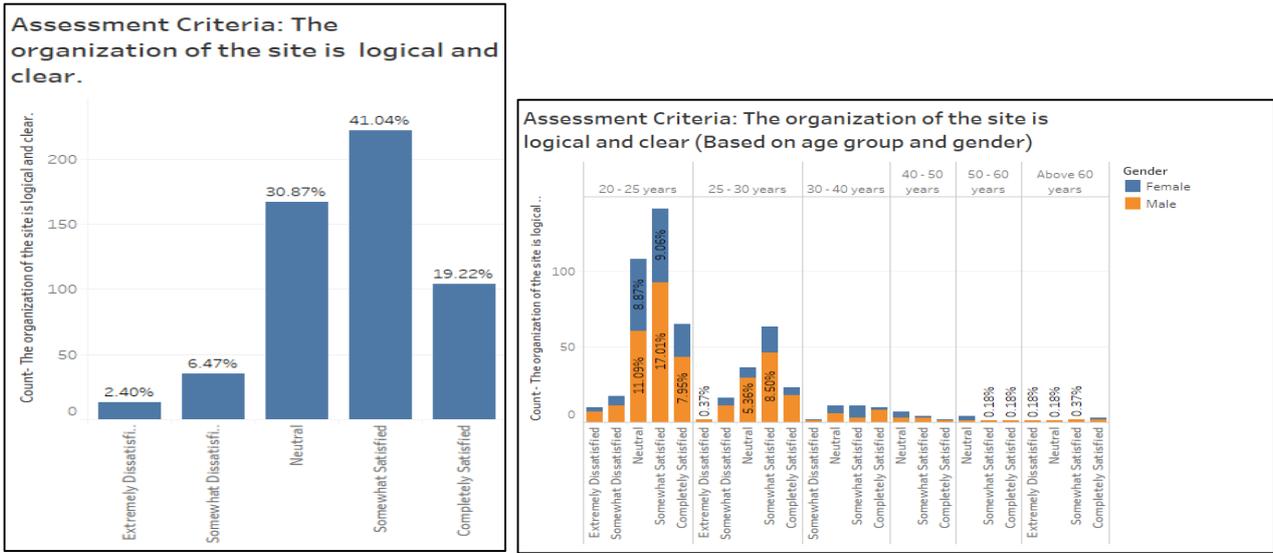### 1.9.  The organization of the site is logical and clear



**Figure 1.9 a)** The organization of the site is logical and clear, **b)** The organization of the site is logical and clear (Based on Gender and Age Group)

**Analysis:** The analysis of the parameter states that the site is logically organized and users find it as per their expectations.

***Section 2 focussed on the User- Interface elaborating the Visual Appeal of the IRCTC website***. It had the following 7 assessment parameters:

2.1. Readability of the website (Consider typeface, font size, color contrast)

2.2. Appropriateness of Terminology related to task

2.3. Website informs about its progress

2.4. Display of Error messages

2.5. Look and feel of the website

2.6. Ease of navigating the website

2.7. It is easy to use this site upon my first visit.

Figure 2 below shows the cumulative impact of the various factors for User Interface and the Visual Appeal of the Website. It is evident from cumulative effect of the various sub-parameters that the User Interface are considered as quite satisfactory by most of the user population. Almost 70% users are satisfied while those dissatisfied or not satisfied are around 20%.
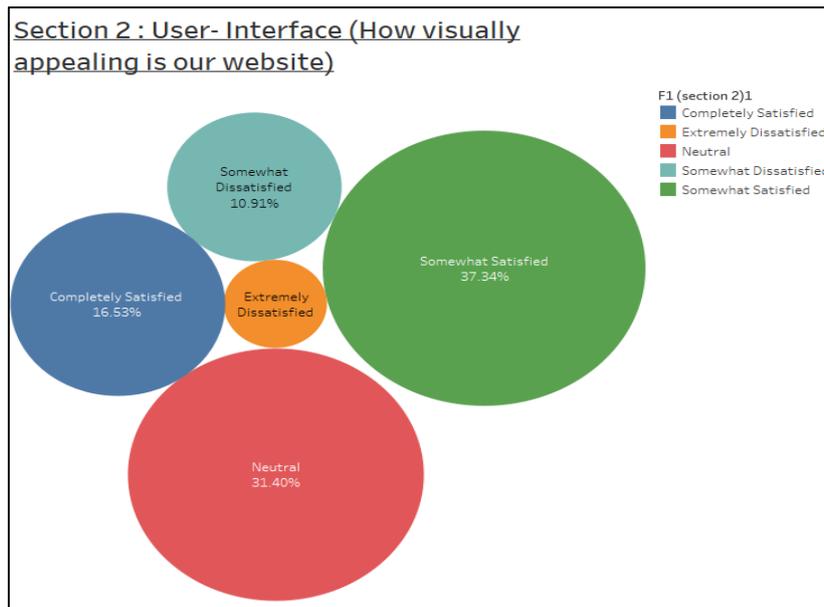


**Figure 2** Overall Satisfaction Index: Section 2-User Interface

The Analysis of the individual parameters is as below:

### 2.1. Readability of the website (Consider typeface, font size, color contrast)

**Analysis:** As per the trend depicted most of the user population agrees that the site is readable and the various aspects of readability i.e. fonts, typeface, size and color-contrast are satisfactory.

### 2.2. Appropriateness of Terminology related to task

**Analysis:** The user population is quite contented with terminology used on the website for the various functions and features.

### 2.3. Website informs about its progress

**Analysis:** The user population has said clearly that the website informs properly of its progress as the various features are used and the various actions are initiated.

### 2.4. Display of Error messages

**Analysis:** The younger population seems to be agreeing that the error messages are properly reflected but the upper age groups seem to be expressing their discontentment over this issue.

### 2.5. Look and feel of the website

**Analysis:** Here again the younger population seems to be happy with the look and feel of the website while the upper age groups are slightly dissatisfied.

### 2.6. Ease of navigating the website

**Analysis:** The trend here is similar to the earlier one that it is easy to navigate the website for the younger population but some of the upper age groups are not comfortable.

### 2.7. It is easy to use this site upon my first visit.

**Analysis:** The trend here is similar to the earlier one that it is easy to use the website for the younger population on their first visit but not so much for the those in the upper age groups.

*Section 3 was designed to check the Website Capabilities.*
The four sub-parameters in this category are as below:

- 3.1. Website speed and performance
- 3.2. Designed for all levels of users
- 3.3. Website load quickly
- 3.4. Website optimized for mobile
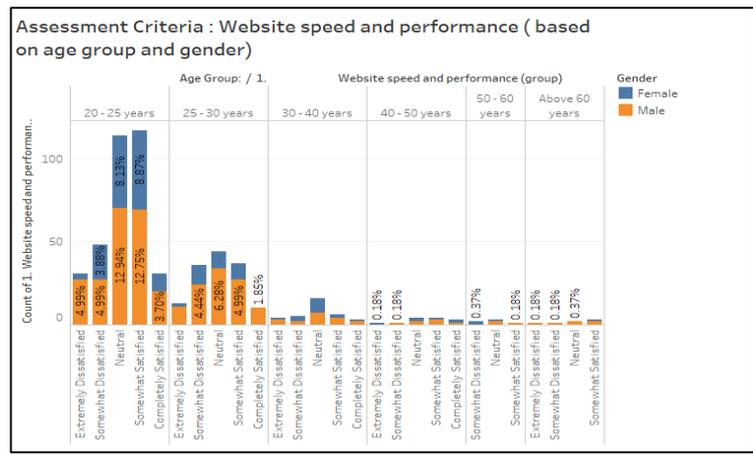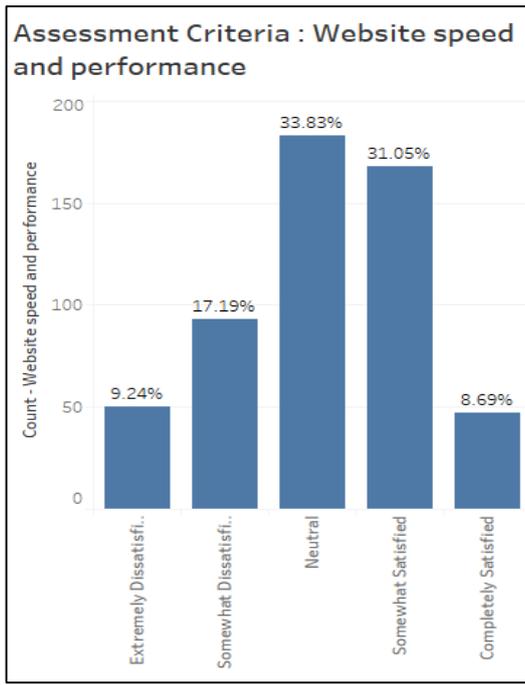
### 3.1. Website speed and performance



**Figure 3.1a** Website speed and performance, **b)** Website speed and performance (Based on Gender and Age Group)

**Analysis:** Website Speed and Performance seems to be an issue with many users as the maximum number of users are neutral while voting for this website capability and almost 25% users seem to be dissatisfied with this factor.
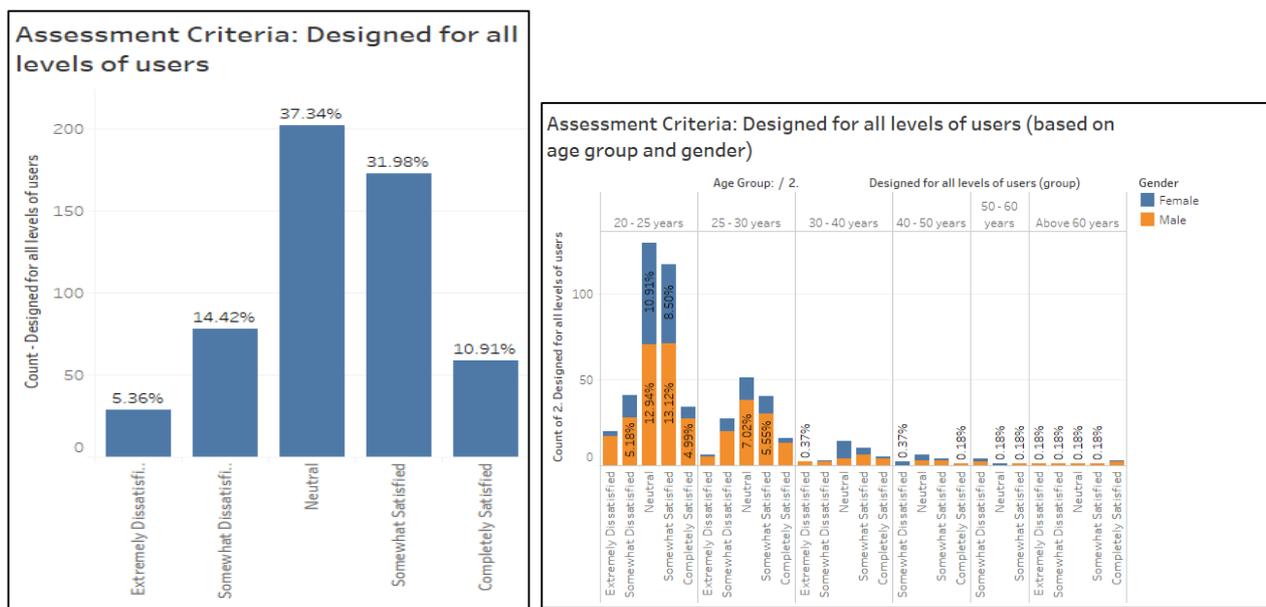
### 3.2. *Designed for all levels of users*



**Figure 3.2a** Designed for all levels of users,  **b)** Designed for all levels of users (Based on Gender and Age Group)

**Analysis:** This factor also seems to be an issue with users as the maximum users are neutral.
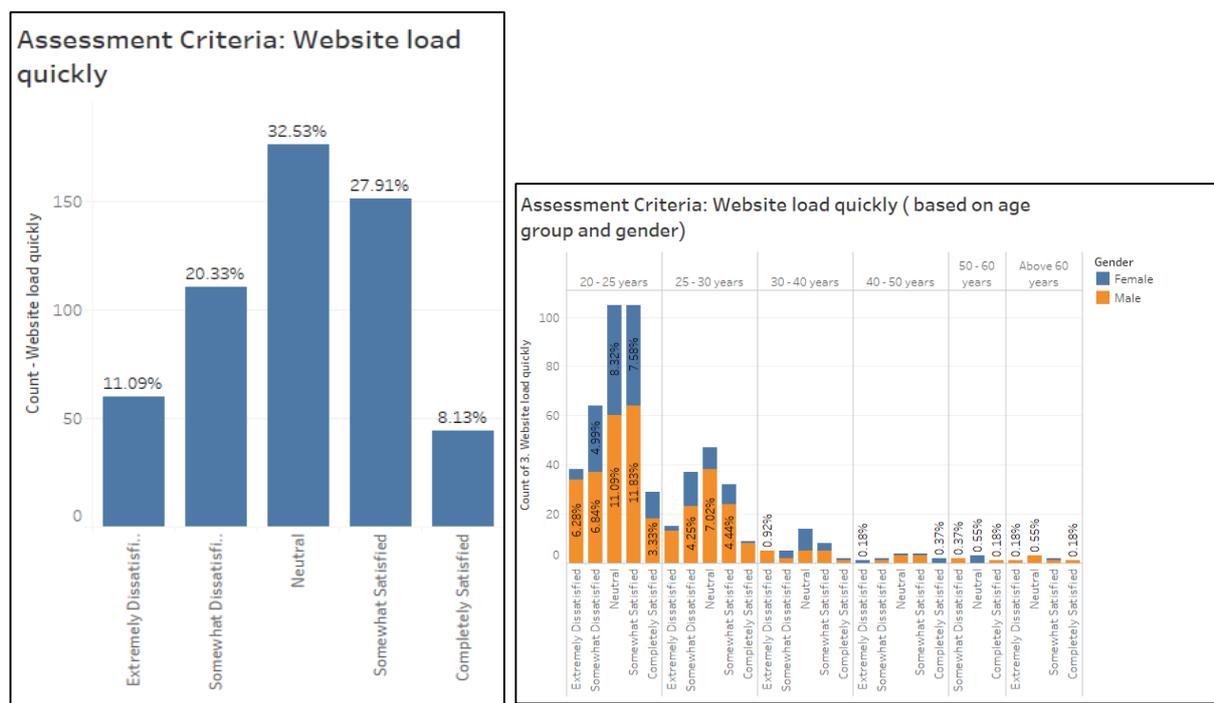
### 3.3. *Website load quickly*



**Figure 3.3a Website load quickly, b) Website load quickly (Based on Gender and Age Group)**

**Analysis:** The speed of loading of the website is a major issue as maximum people are neutral giving it a average rating and almost 33% have given it a rating below average.
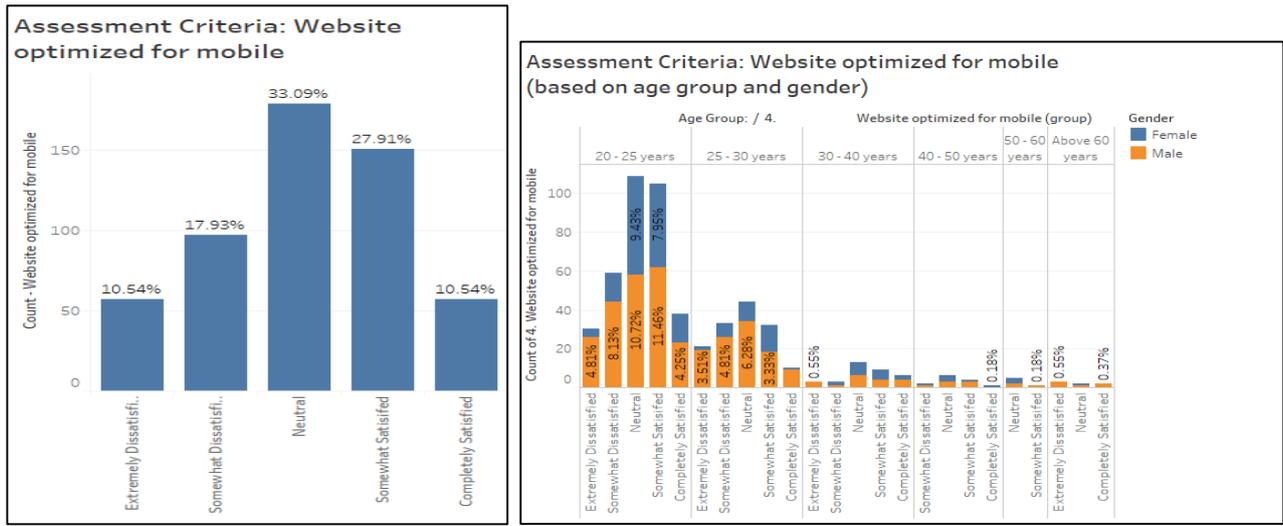
### 3.4.  Website optimized for mobile



**Figure 3.4a** Website optimized for mobile,  **b)** Website optimized for mobile  (Based on Gender and Age Group)

**Analysis:** This factor again has maximum users in neutral category, but he number of votes for above average rating is more than that for below average rating.

It is evident from cumulative effect of the various sub-parameters that the Website Capabilities have been rated as average by the maximum number of users. A large proportion of users seem to be dissatisfied for this factor.

*Section 4 focusses on the Overall Impression of the Website* and has four criteria. The criteria are that whether the website is Informative, Visually Pleasing, Not Impressive, Unfriendly to User. The factor has been analysed against the frequency of usage of the website, categorised age wise, by the users. The graph shows that most of the users from all age groups consider the Website Informative and Visually pleasing. In the same section users have been asked to provide comments for improvements in the website. Sentiment Analysis of the comments so obtained is depicted in Figure 4.2
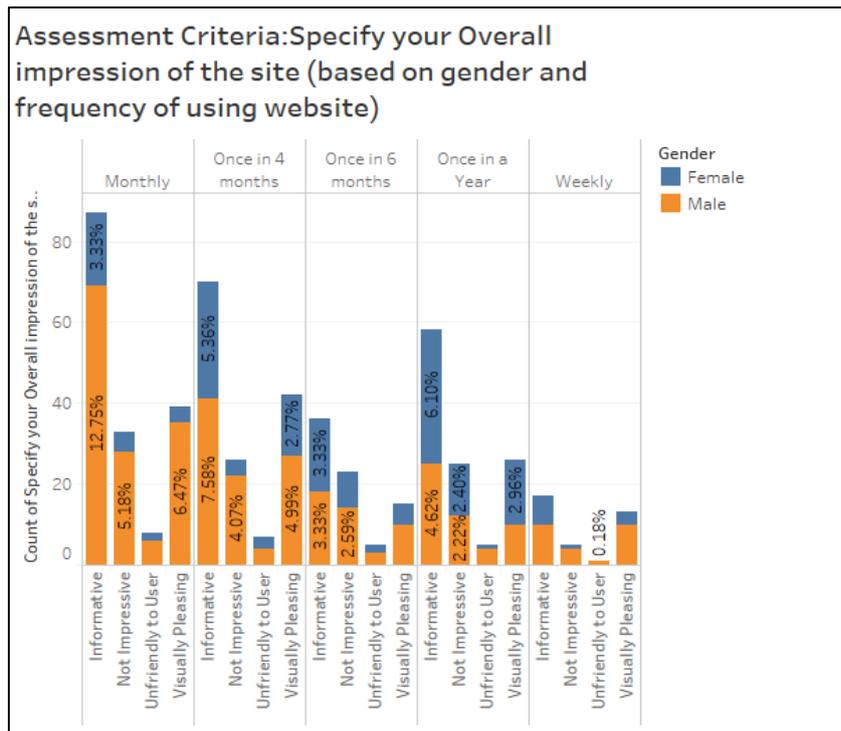


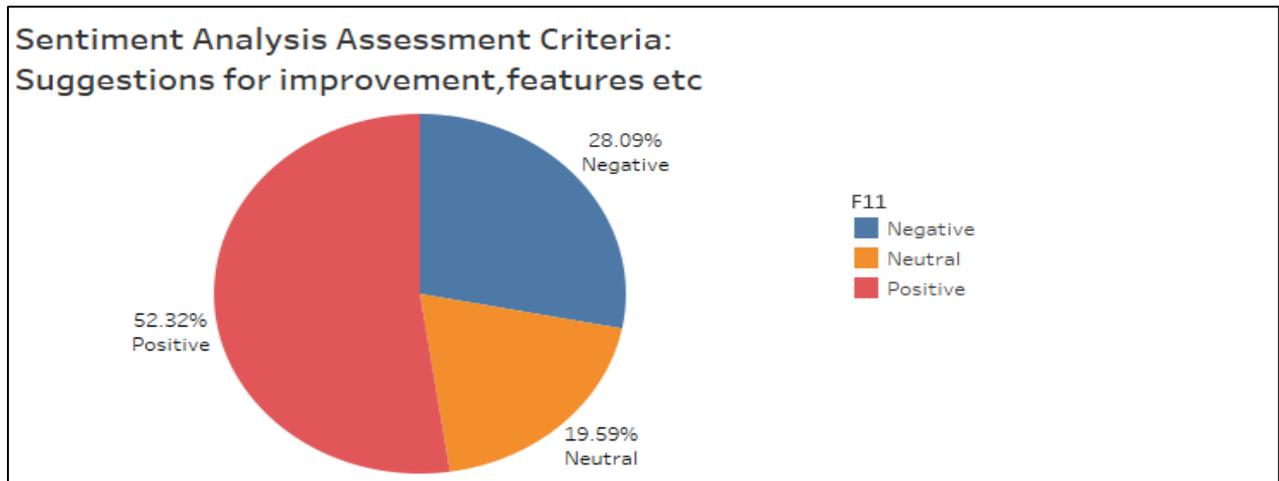**Figure 4.1** Overall Impression of the Website

**Figure 4.2** Sentiment analysis Assessment Criteria: Suggestions for improvement, Features etc

**Analysis:** All the comments received have been analysed using the Data analytics tool "R", an open source software and is further visualised using Tableau. Sentiment Analysis is the process of computationally identifying, processing and categorizing text data in order to obtain opinions and attitudes towards certain themes, topics, products etc. In our analysis we used sentiment R Package and classified them to positive, neutral or negative comments. This R package is based on lexicon, which is a technique that uses a dictionary of words to assess whether bunch of text is positive, neutral or negative. The comments received therefore suggest that almost 52% of users suggest that website is fine in its current form. 20% users are neutral and have some discontentment over certain factors of the website. While 28% users are dissatisfied and have suggested improvements or enhancements to the website.

## CONCLUSIONS

The IRCTC Website Usability Evaluation was done on for features Website Content & Features, the User- Interface and its visual appeal, Website Capabilities and performance and the Overall impression of the Website. The survey was designed to focus on the User Experience while working on the IRCTC Website, either for extracting information or booking railway tickets.

The IRCTC Website Usability Evaluation & Sentiment Analysis of the User Comments, help to draw the following conclusions.

- IRCTC Website has been rated as having above average capabilities and usability.
- The Website is used mostly by the younger population between age groups of 20-30.
- The Website content and features are good but need some enhancements to be appreciable by the user community. Specially the home page content has been marked as average and thus can be enhanced. Finding the desired information on the website has also been reflected as average suggesting therefore the users take time to reach the desired content on the website.

- The user Interface has been appreciated. But the older population has shown some dissatisfaction over issues as display of error messages and the likes.
- The Website speed and performance has been pointed to as a point of concern. The website suitability for the various age groups and working through mobile are specific areas of concern.
- The users' comments were specially directed towards the following issues:
  - Improving the aesthetics of the website, removing useless information from the Home page, organizing the information by creation of tabs etc.
  - Prioritising the railway ticket bookings by separating the railway ticket booking from flight and hotels etc.
  - Editing of Users Profile has been highlighted as an issue. Creation of multiple profiles is possible and checked.
  - Website at times is not able to handle the traffic causing problems to the users.
  - Location of Tatkal tab on the earlier website was better while on the current website the tab is not locatable.
  - The number of payment gateways currently available needs to be enhanced.
  - Request to add Journey Feedback on the website can help users reach out to the department.
  - The website should be compatible with more regional languages.
  - Introduction of an IRCTC wallet.

## REFERENCES

[1] "From First Train In 1853 To Online Bookings Via IRCTC, Indian Railways Covered A Long Distance In Its 165 Years", Economy, NDTV Profit Team, Updated: April 17, 2018 19:06 IST

[2] https://en.wikipedia.org/wiki/Indian_Railway_Catering_and_Tourism_Corporation

[3]   Indian Railway Catering and Tourism Corporation to launch 'Rolling Deposit Scheme' for e-ticket - Economic Times. Articles.economictimes.indiatimes.com, 28 August 2012.

[4]   "Indian Railways to introduce Rolling Deposit Scheme for E-Ticketing".

[5]   "Welcome to Indian Railway Passenger reservation Enquiry", Indianrail.gov.in, 30 June 2006.

[6]   "IRCTC Lite Version For Tatkal Launched". PNRStatusIRCTC.in. 2016-05-28.

[7]   "Rs 1 lakh a night on Maharajas' Express", Gupta, Jayanta (19 March 2010), Times of India.

[8]   "Opinion mining and Sentiment Analysis", Bo Pang and Lillian Lee, Foundations and Trends in Information Retrieval Vol. 2, No 1-2 (2008) 1–135.

[9]   "A Sentiment Analysis as a Tool to Identify The Status Of Universities: The Case of ITU", Mine Işık, Başar Öztayşi, Kübra H. Fenerci, Proceedings of the 2012 International Conference on Industrial Engineering and Operations Management Istanbul, Turkey, July 3 – 6, 2012 1118