Chapter 3

RESEARCH METHODOLOGY

The proposed study is exploratory in nature based on both primary as well as secondary data. The research is divided into quantitative as well as qualitative analysis of data collected from the primary and secondary sources.

Primary data was collected through field survey from cine goers, producers and producers’ associations. For cine goers, 515 data collected administering a questionnaire comprising 21 questions (closed ended and open ended). These responses were collected from six different places within the state of Maharashtra. These places were Mumbai, Pune, Nagpur, Nashik, Satara, and Sangli based on the respective sizes of their population. Mumbai being the largest city with a sizeable population of around 20 million and home of Hindi film industry and Pune being the cultural capital of the state of Maharashtra were the obvious choice for data collection. Places like Nagpur and Nashik being to other important and large cities in sequence were the next choice for data collection. Satara and Sangli comparatively smaller places with 63 percent and 65 percent literacy rate were also chosen to have a proper mix of data in order to avoid sample bias. However the literacy rate of Thane district was 68 percent (Maps of India), more than that of Satara and Sangli but was avoided because of its vicinity with Mumbai. Often people mistake Thane as a part of Mumbai itself. Had Thane been chosen for the purpose, it would not have been possible to collect diversified data. All responses were collected from cinema theatres (multiplexes and single screen theatres during show time) using convenience method. It was observed during field survey that maximum number of respondents, who watch movies fall between the ages of 16 years to 60 years. The population of cities from where data collected was taken as the basis to decide on the number of responses.

<table>
<thead>
<tr>
<th>City</th>
<th>Mumbai</th>
<th>Pune</th>
<th>Nagpur</th>
<th>Nashik</th>
<th>Satara</th>
<th>Sangli</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of respondents</td>
<td>207</td>
<td>91</td>
<td>84</td>
<td>62</td>
<td>38</td>
<td>33</td>
<td>515</td>
</tr>
</tbody>
</table>
3.1. Research Objectives

The mammoth sized India’s media and entertainment industry reaches millions of people across the world through 161 million TV households, 94,067 newspapers in Hindi, English and other regional languages, around 2,000 multiplexes out of total 12,000 movie theatres, 214 million internet users including 130 million mobile internet users (FICCI-KPMG, 2014). Ernst & Young (2014), the phenomenal growth of India’s media and entertainment industry has opened new opportunities for business of entertainment goods and employment generation. The sheer size and reachability of media and entertainment industry’s contribution to national economy which is poised to touch $ 2 trillion could easily be imagined (Mishra, 9 Oct., 2014). According to a report published in International Business Times, dated 03 may, 2013, the size of Indian film industry had to reach $10 billion by the next year. A report published in Live Mint Indian piracy industry suffered a loss of $ 4 billion alone in 2008 (Verjee, 24 Mar. 2008; FICCI-PWC, 2006). The purpose to discuss the vital statistics pertaining to India’s media and entertainment industry in general and film industry in particular was to high light the importance of this study carried out by the researcher (Deloitte-MPA, 2014).

The proposed study was an attempt to identify the factors responsible for movie piracy and to quantify the loss of revenue to Mumbai based Hindi film industry.

1. To identify the factors responsible for movie piracy
2. To propose a conceptual framework to understand how various factors contribute to movie piracy.
3. To quantify the loss of revenue to the film industry at the box office.
4. To quantify the loss of revenue to the Government of Maharashtra at the box office in the form of entertainment tax.

3.2. Research Design

This study was based on both primary as well as secondary data. Primary data was collected into two separate sets for cine goers and cine producers including office bearer of their association.
The questionnaires were designed and hypotheses were formulated. Later the hypotheses were tested. Quantitative and qualitative analyses both were based on primary as well as secondary data. In quantitative analysis hypotheses were formulated and tested using data collected from cine goers. For quantification of loss of revenue at box office collection both primary and secondary data were used. Data collected from cine goers was extrapolated for the purpose. Similarly qualitative analysis was done using both primary as well as secondary data. Primary data was collected from cine producers and office bearer of their associations and interpretive structural model (ISM) and MICMAC analysis were developed to develop conceptual framework of movie piracy. Apart from conceptual framework case study analysis was also done to develop an insight about the role of technology both in encouraging piracy as well as to counter it. In nutshell the research design comprised of:

1. Quantitative analysis:
   a. Demographic factors contributing to movie piracy: hypotheses developed and tested. Primary data from cine goers was used for this analysis
   b. Loss of revenue at box office collection was estimated based on both primary data collected from cine goers and secondary data as well. Secondary sources of data were reports published on newspapers, journals, periodicals, magazines and empirical studies.

2. Qualitative analysis:
   a. ISM and MICMAC analyses were done to develop conceptual framework of movie piracy. Primary data collected from cine producers and office bearers of their associations was used for the very purpose.
   b. Case study analyses were also based on primary as well as secondary data. For primary data the ex-chief technology officer of UFO Digital Cinema was used to understand the role of technology to curb the menace of piracy at theatres during the legitimate screening of movie. Reports published on newspapers were the secondary source of information.

3.3. Data Collection Tools
1. For primary data collection field surveys were conducted administering separate sets of questionnaire for cine goers and cine producers/producers’ associations. Each set of questionnaire contained 21 questions both open and closed ended. Cine goers were targeted at theatres during the show time whereas cine producers were met on appointment at their respective offices located in Mumbai.

2. For secondary data websites of Central Board of Film Certification, National Film Archives, production houses, publications of periodicals, research papers published on reputed journals, news- papers, reports of Federation of Indian Chambers of Commerce and Industries (FICCI), KPMG, a leading risk, internal audit, financial services, business advisory, corporate governance consultancy with its head office located in Netherlands, Ernst & Young, Pricewater House Cooper (PwC), Canada Motion Pictures Distributor’s Association (CMPDA), Motion Pictures Association of America (MPAA), IPSOS, the second largest market research firm based in United Kingdom and many others were searched and accessed.

3.4. Sampling Design

Hindi films produced in Mumbai film industry are widely watched across the world. Their immensely popularity has made a deep impact on the viewership and with each day passing the viewership of Hindi films seems to be increasing. The present study was exploratory in nature based on both primary as well as secondary data and quantitative as well as qualitative analyses. The primary data was collected from cine goers to study their behavioural pattern of watching movies. Since the population of audience for Hindi films across the world was too large in general (within the state of Maharashtra where the field survey was conducted for data collection in particular), to determine on the adequate sample size was of utmost significance in order to keep sample error at its lowest possible level, and ensure stable and authentic results of the research undertaken. Since the population of cine goers was determined, list and contact details of cine goers were also not available it was not possible to select data randomly. Therefore, convenience method of data collection was used by the researcher.

The accuracy of a research depends on sample size. The result of an under taken research obtained on the basis of an adequate sample size is believed to be more stable and authentic. Therefore, to decide and determine on the authenticity of research conclusion an adequate
sample size is very crucial. Among various rules to determine adequate sample size, the one suggests that sample size be determined as a function of the number of variables (Guadagnoli, 1988). By and large, statistical literature in this connection concludes that the sample size needed for stable results is directly proportional to the number of variables under consideration. On the basis of existing literature, Thorndike (1978) proposed two rules connecting the sample size and the number of variables. One informal guide is that there should be at least 10 respondents for each variable. Thorndike observes that an addition of 50 more respondents to the said number would ensure sufficient sample size for variables being considered for analysis. According to this rule the number of respondents (N) required is more than or equal to ten times of number of variables and an addition of 50 respondents. The second rule proposes somewhat more stringent condition. According to the second rule, the sample size should be equal to or greater than the square of the total number of variables (N is greater than or equal to number of variables square) and the addition of 50 more would be good enough. In order to achieve stability of results, the second rule which is more stringent for sample size was followed in this study.

The sample size of this research was based on the following formula of Thorndike:

\[ N \geq 10 \times k + 50 \]

Where N is the sample size, and k is the number of independent variables.

The exploratory factors for analysis of reasons responsible for movie piracy in this study involved nine variables. Those variables were age, occupation and gender of respondents who download movies from internet websites, and/or buy pirated CDs. Two variables regarding awareness about economic/financial implications of movie piracy on the legitimate players, and its legal consequences were also identified. One variable was identified as reasons for watching pirated movies.

Based on the formula the sample size was, \( N \geq 10 \times 9 + 50 = 140 \)

Thus the analysis based on number of variables satisfies both the rules proposed by Thorndike (1978) relating to sample size of 515 for this study.

Another set of questionnaire was prepared to collect data from the cinema producers, and their associations. This set of questionnaire contained 21 questions including both open and closed ended questions. Since the researcher despite his all efforts could collect data only from 12
producers and their associations, a quantitative analysis with such a small size of sample was not possible. Therefore, it was decided to use the data (information) in qualitative analysis.

3.5. Tools for Data Analyses

A questionnaire containing 21 questions was designed to collect data from the cine goers who watch movies. Once the sample size for cine goers was decided, it was the next step to collect data from the field survey. It was observed during the field survey that to target respondents at cinema theatres would be the right approach for two reasons:

a. In cities people were a bit apprehensive to entertain unknowns (researcher) at their homes.

b. It was difficult to identify the audience who watched movies. Even if in a family all members who watched movie it was further expected to segregate those who preferred movies at theatres from those who preferred watching movies at home. Therefore, to overcome these difficulties the researcher visited cinema theatres during the screening of movies to collect data from the respondents (movie goers). The biggest advantage was that there was no need to identify who was audience, who was not. It was implied that a person standing in a queue in front of the box office was no other than audience.

Similarly the data from the cinema producers and association office bearers of producers’ associations were also collected administering the questionnaire with a separate set of 21 questions.