Investigating Web Size Metrics for Early Web Cost Estimation

Sajjad Asif
This thesis is submitted to the Faculty of Computing at Blekinge Institute of Technology in partial fulfillment of the requirements for the degree of Master of Science in Software Engineering. The thesis is equivalent to 20 weeks of full time studies.

Contact Information:
Author(s):
Sajjad Asif
E-mail: saao10@student.bth.se

University advisor:
Prof. Emilia Mendes
School of Computing, BTH

Faculty of Computing
Blekinge Institute of Technology
SE-371 79 Karlskrona, Sweden

Internet : www.bth.se
Phone : +46 455 38 50 00
Fax : +46 455 38 50 57
ABSTRACT

Context Web engineering is a new research field which utilizes engineering principles to produce quality web applications. Web applications have become more complex with the passage of time and it’s quite difficult to analyze the web metrics for the estimation due to a wide range of web applications. Correct estimates for web development effort play a very important role in the success of large-scale web development projects.

Objectives In this study I investigated size metrics and cost drivers used by web companies for early web cost estimation. I also aim to get validation through industrial interviews and web quote form. This form is designed based on most frequently occurring metrics after analyzing different companies. Secondly, this research aims to revisit previous work done by Mendes (a senior researcher and contributor in this research area) to validate whether early web cost estimation trends are same or changed? The ultimate goal is to help companies in web cost estimation.

Methods First research question is answered by conducting an online survey through 212 web companies and finding their web predictor forms (quote forms). All companies included in the survey used Web forms to give quotes on Web development projects based on gathered size and cost measures. The second research question is answered by finding most occurring size metrics from the results of Survey 1. List of size metrics are validated by two methods: (i) Industrial interviews are conducted with 15 web companies to validate results of the first survey (ii) a quote form is designed using validated results from industrial interviews and quote form sent to web companies around the world to seek data on real Web projects. Data gathered from Web projects are analyzed using CBR tool and results are validated with Industrial interview results along with Survey 1. Final results are compared with old research to justify answer of third research question whether size metrics have been changed. All research findings are contributed to Tukutuku research benchmark project.

Results “Number of pages/features” and “responsive implementation” are top web size metrics for early Web cost estimation.

Conclusions This research investigated metrics which can be used for early Web cost estimation at the early stage of Web application development. This is the stage where the application is not built yet but just requirements are being collected and an expected cost estimation is being evaluated. List of new metrics variable is concluded which can be added in Tukutuku project.

Key Words: Early Web Size Metrics, Effort Estimation, Tukutuku, Web Quote forms
# CONTENTS

ABSTRACT ................................................................................................................................. I

CONTENTS ................................................................................................................................. 1

1 INTRODUCTION ..................................................................................................................... 6
  1.1 BACKGROUND .................................................................................................................. 6
  1.2 AIMS AND OBJECTIVES ............................................................................................... 7
  1.3 PROBLEM DEFINITION ............................................................................................... 8
  1.4 THESIS STRUCTURE ..................................................................................................... 9

2 RELATED WORK ................................................................................................................... 10

3 METHODOLOGY .................................................................................................................. 14
  3.1 RESEARCH QUESTIONS ............................................................................................... 14
  3.2 RESEARCH DESIGN ..................................................................................................... 15
    3.2.1 Survey 1 (S1) ......................................................................................................... 15
    3.2.2 Industrial Interviews ........................................................................................... 15
    3.2.3 Survey 2 (Quote form) ......................................................................................... 15
    3.2.4 Data Synthesis and Comparison ......................................................................... 16

4 SURVEY .................................................................................................................................. 17
  4.1 INTRODUCTION ............................................................................................................ 17
  4.2 KEYWORDS AND SEARCH STRING ........................................................................... 17
  4.3 INCLUSION/EXCLUSION ............................................................................................ 19
  4.4 SAMPLE QUOTE FORM FROM THE WEB ................................................................ 19
  4.5 RESULTS OF SURVEY (S1) ....................................................................................... 21
  4.6 FINDINGS OF SURVEY (S1) ....................................................................................... 23

5 INDUSTRIAL INTERVIEW ..................................................................................................... 24
  5.1 PURPOSE OF INTERVIEW ........................................................................................... 24
  5.2 SELECTION OF INTERVIEW SUBJECTS .................................................................. 24
  5.3 STUDY INSTRUMENTS ............................................................................................... 25
  5.4 INTERVIEW STRUCTURE ............................................................................................ 25
  5.5 DESIGNING .................................................................................................................. 25
  5.6 INTERVIEWEES ........................................................................................................... 25
  5.7 TRANSCRIBING ........................................................................................................... 26
  5.8 ANALYZING AND VERIFYING ................................................................................... 26
  5.9 ANALYSIS AND RESULTS OF INTERVIEWS ......................................................... 27
  5.10 FINAL SELECTED METRICS .................................................................................. 29
  5.11 DISCUSSION AND SUMMARY ............................................................................... 29

6 SURVEY 2 (QUOTE FORM) .................................................................................................. 31

7 ANALYSIS AND DISCUSSION ............................................................................................ 33

8 VALIDITY THREATS ............................................................................................................ 43
  8.1 INTERNAL VALIDITY ..................................................................................................... 43
    8.1.1 Selection of quote form ......................................................................................... 43
    8.1.2 Interviewee Selection ........................................................................................... 43
    8.1.3 Missing Information during Interview .................................................................. 44
  8.2 EXTERNAL VALIDITY .................................................................................................. 44
    8.2.1 Validity of Interview Participants ........................................................................ 44
List of Figures

Figure 1 Thesis Structure .......................................................................................................... 9
Figure 2 Research Methodology ............................................................................................. 16
Figure 3 Total selected URLs ............................................................................................... 19
Figure 4 QDA Model .............................................................................................................. 27
Figure 5 CBR Model .............................................................................................................. 34
## List of Abbreviations & Acronyms

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>RQ</td>
<td>Research Question</td>
</tr>
<tr>
<td>CSS</td>
<td>Cascading Style Sheet</td>
</tr>
<tr>
<td>HTML</td>
<td>Hyper Text Markup Language</td>
</tr>
<tr>
<td>WQM</td>
<td>Web Quality Model</td>
</tr>
<tr>
<td>QDA</td>
<td>Qualitative data analysis</td>
</tr>
</tbody>
</table>
1 **INTRODUCTION**

The chapter contains an insight into the background of the selected research domain, problem definition, aim and objectives of this thesis.

1.1 **Background**

Web engineering is a new research field which utilizes engineering principles to produce quality web applications [6]. Web applications have become more complex with the passage of time and it's quite difficult to analyze the web metrics for the estimation due to a wide range of web applications [1]. Correct estimates for web development effort play a very important role in the success of large-scale web development projects. Feedback can be obtained in order to control, understand and improve processes and products by applying measurement principles to measure the qualities of the web application and their development processes [2]. A number of techniques for estimating development efforts have been proposed which are derived from different software project measures. This is also necessary to calculate estimates throughout the whole software development lifecycle [3].

Web engineering is the use of scientific, engineering, and management principles and systematic approaches with the aim of successfully developing, deploying and maintaining high quality Web-based systems and applications [35]. Web effort estimation is the process by which effort is forecasted and used as a basis to predict costs and allocate resources effectively, so enabling projects to be delivered on time and within budget [36].

Planning in Web applications development represents one of the complex tasks especially when effort, time and cost estimation is needed. A number of methods have been developed in order to measure effort estimation for classic software development. Web applications development is different as compare to classical software development and there is lack of models or methods which provides more reliable web development effort estimation [37]. Web applications are being a core need for businesses.

Web applications development companies have been increased in recent years in order to meet the needs of the market. There is growing number of companies willing to have the web presence and as a result number of companies are increased who provides bidding to develop Web applications. Different technological solutions are being used to develop Web applications. In order to win the bids, companies provide unrealistic schedules which become very difficult to develop within estimated time and budget. A realistic effort estimate at the early stage is the key to the success of the project and it leads project managers to manage resources effectively. It also helps the Web companies to benchmark its work efficiency across its own and competitor's projects.

Estimation is one of the most critical components of a successful software project. Its a classical challenge for all project managers to complete the project on time and within budget [40]. Software projects need an average 30-40% more effort than estimated effort [41]. There are many estimation models but no model is able to adequately measure effort of web applications [3, 29, 38, 39]. Many practitioners see the measurement of the web application is critical now when the importance of web applications is growing. Quality of the software has become a very important driver for competitiveness.

Web companies need to have a better effort and cost estimate approaches in order to survive in the competitive web development market. If any web company provides an inaccurate estimate then its difficult for that company to survive in the market. Inaccurate early web cost estimation can leads towards the failure of the project as well which can affect the Web...
companies business. It has become very critical for Web companies to have an accurate early web cost estimation. More accurate effort and cost estimate in Web applications development leads towards the success of the project and long-term customer’s relationship. The web application is a great way to build distributed applications. Application structure is that the front-end, back-end, and connection between them. Hyper Text Markup Language (HTML) remains approachable to a large number of people and very easy and flexible to build pages and designs. Cascading Style Sheet (CSS) makes it easy to create designs and formatting.

Web applications provide multi-platforms support. Industries such as banking, education, travel and hospitality and Government utilized Web-based applications to increase and improve their operations [50]. Web allows building intranet corporate Web applications which can run only within the boundaries of individual organizations. Web applications for e-commerce and communication sector make it one of the most important branches of software industry [51]. Web application development is new and growing industry. There is a variety of web applications which are fully functional systems that provide business to business and business to customer solutions.

The difference between Web application development and generic software development is their development processes, technologies, quality factors, and measures. Some factors which exist in Web applications do not exist in other software development i.e. Hosting, Responsiveness, Communication protocols and security measures etc. Development of generic software is being dominated by IT engineers where a good knowledge of programming, databases etc. are very necessary while in Web development there is a wider variety of developers, graphics designers, testers, writers where little programming skills can still work. Now web pages are being created even by automated software or similar solutions where no programming skills needed.

Early web cost estimation is being done on very early stage when requirements are being collected. This estimate is a starting point towards a final agreement between customer and Web Company. The customer can add more requirements or decrease requirements but this estimate gives an overall idea to the customer in the start about the expected cost of the web project. Web companies can also provide hours estimate on the base of calculated effort and become easy for the customer to convert pricing on the base of estimated effort hours.

1.2 Aims and objectives

The main purpose of this research is to identify cost drivers and size metrics for early web cost estimation based on current practices of many web companies worldwide.

This research aims to revisit previous work done by Mendes [4]. Purpose of doing this research again is to check whether still early web cost estimation trends are same or changed? One source to check this change is through Case-based Reasoning [20]. What are most important factors being considered now? Is there any new size metrics or factors available? How are web companies measuring web effort estimation for web projects? This research will answer all these questions and will help to identify new trends, existing or new size metrics and approaches that companies are following nowadays. This research will provide more up-to-date and improved results about web size metrics for early cost estimation as previous research was done 10 years before.

The aim of my study is to investigate early web cost estimators and validate the results of any change before and after 2002. Following are objectives of this research:
To investigate size metrics and cost drivers used by web companies for early web cost estimation
- To find out the most commonly used predictors of the companies
- To design a data entry form comprising most commonly used web size metrics for web projects in order to get early web cost estimation. Perform analysis on the results by comparing current and old research work.

Two surveys and one industrial interview is conducted in this research. Feedback obtained from research survey one S1 and Industrial interviews will be used to prepare web quote forms to gather data on web projects worldwide. All these findings will be contributed to Tukutuku benchmarking project. The Tukutuku Benchmarking project [7] was initiated by Mendes in 2002. The aim of the Tukutuku project is to collect data on web worldwide web projects to be used to develop web cost estimation models which are based on the early effort estimation and productivity benchmark within the web companies.

1.3 Problem Definition

The problem domain of this research is to identify effort estimation for early web cost estimation based on current practices through different web development companies worldwide.

Quality web applications require appropriate Web application development and project management processes. For project managers, cost estimation is one of the crucial steps at the beginning of any new web project. Accurate cost estimates are an essential element of being able to provide competitive bids and remaining successful in the market. Over- and underestimates of the expected costs have a significant influence on a company’s reputation and competitiveness. The industry still reports difficulties with determining accurate cost estimates. Different surveys indicate that only very few companies use cost estimation models. There are a number of practically relevant reasons for not using cost estimation models: Many cost estimation models are not very accurate [34].

Experience shows that model-based estimates do not perform considerably better than estimates solely based on expert judgment. Another reason for the rare use of cost estimation models in practice is the lack of sufficient, explicit past project data to systematically build a cost estimation model. Many organizations do not have enough resources for the required measurement methods and tools. Data collection is a time-consuming process that needs to be carefully planned because consistency and completeness are two key drivers for the usefulness of the data for effort estimation. Only data on recent projects allows for cost estimation research that is meaningful for current projects [34].
1.4 Thesis Structure

Figure 1 Thesis Structure
2 Related Work

Rollo was the researcher to investigate the issues of measuring the size metrics of Web applications for early cost estimation. While Cleary proposes size metrics for Web cost estimation dividing them into two types: web size metrics for hypermedia and software applications [4].

Early web-based development and functional size measurement are considered the economical and time-saving approach in these days. By using these measurement principles to evaluate the quality and development of existing Web applications, we can obtain feedback that will help us understand, control, improve, and make predictions about these products and their development processes [3] [15].

Effort estimation accuracy requires knowledge about the size and nature of the impact of effort estimates on the software work. The effort depends upon development team skill, experience, market popularity, and milestones. Projects with incomplete requirement specifications and cost were prone to effort estimate while to optimistic requirements and high priority on quality had led to effort overruns. To estimate effort to develop a project is good because in this case, the cost is measured for client and Development Company before the project is broken down into activities and corresponding project plan is implemented. This effort estimate is also called pre-planning estimate [3].

As there is not a standardized development technique for web development projects, so web prediction for early development is challenging. Although a variety of technological solutions are available for web developers to deliver the product in time yet big web projects need consideration. Secondly, web engineering literature is scattered when comparing effort prediction approaches. Case-based reasoning (CBR) provides estimates by comparing the current problem to be estimated against a library of historical projects [20].

In this technological era, almost every organization has a corporate website or any business related websites which can support their businesses. Due to this increasing trend of websites in the businesses, more complex web applications are being designed. It has become a challenge for web application managers and developers to create correct development time, effort and size estimates in order to meet customer requirements.

There are many techniques and methodologies available to estimate size and effort in classical software application development but not in web application development because web application development is different than traditional software development [53] [54].

Size is considered a primary input to software cost models. Mendes state that a web software application represents conventional software applications that depend on the Web or use the Web's infrastructure for execution. Typical applications include legacy information systems such as databases, booking systems, knowledge bases etc. Many e-commerce applications fall into this category. They also postulate that for Web development, cost can be very difficult to estimate because: (i) There is no standard way of sizing Web applications (ii) Each can be created using diverse technologies such as varieties of Java (Java, servlets, Enterprise java Beans, applets, and Java Server Pages), HTML, JavaScript, XML, XSL, and so on (iii) Web project’s primary goal is to bring quality applications to market as quickly as possible, varying from a few weeks to 6 months (iv) People involved in Web development are represented by less experienced programmers, users as developers, graphic designers and new hires straight from university (v) Typical project team size is small, between 3 to 7 members (vi) Processes employed are in general heuristic, although some organizations are
starting to look into the use of agile methods. They also believe that company-specific datasets provide the better basis for accurate estimates. However, the time required to accumulate enough data on past projects from a single company may be prohibitive [4].

Due to increased demand for web projects today, lack of quality is one of the main problems due to their short life-cycle. A good mechanism to control such problem is the optimal use of metrics. In the literature, hundreds of metrics are available but no guidance for their use. Hence a 3-D web quality model (WQM) was refined and presented for efficient web metrics classification [21] [22].

Other software cost estimation models and techniques include parametric models (SLIM, cocomo, checkpoint, SEER), expertise-based techniques (Delphi, rule-based), learning-oriented techniques (neural, case-based), dynamics-based models (abdel-hamid-madnick), regression-based models (OLS, Robust), and composite-Bayesian techniques (cocomoll) for integrating expertise-based and regression-based models. It is also found that neural and dynamics-based techniques are less mature but no one method or model should be preferred over all others [23].

Research [42] presented the use of analogy and two algorithmic models, stepwise multiple regression and linear regression to estimate the authoring effort of Web applications. Results showed that estimation by analogy is a better technique [42]. Research [31] presented two prediction models using statistical techniques named as linear regression and stepwise multiple regression. The results of both models were compared and stepwise regression technique was considered as the best technique as compared to linear regression.

Research [21] compared the prediction accuracy of CBR techniques to estimate the effort to develop Web hypermedia applications. Reifer [43] proposed a new size metric named as web objects. Web objects were implemented using model WebMo, an algorithmic model to estimate effort for the web application. Research [3] shows that "Weighted Euclidean Similarity Measure" gave the correct prediction accuracy among three different CBR techniques to predict the effort estimation of web hypermedia applications. Research [44] proposed to estimate the development effort using CWADEE method which is about creation estimate from 24 to 72 hours.

The research [45] investigated the relationship between size and effort of Web applications. Results showed that Web Objects size measure gave results comparable to Function Points in case of large datasets. The research [46] showed the use of simple models like median effort and complex models like Bayesian Networks for estimation. Manual Stepwise Regression technique was considered are more prominent. The research paper [47] showed some existing analogy-based software estimation tools were used by project managers. A system prototype named as EffortEst was implemented and evaluated. EffortEst provides a better estimation of software project effort.

According to research [4], two most common size metrics used for Web cost estimation were “the total number of Web pages” (70%) and “which features/functionality to be provided by the application” (66%).

LOC (Lines of code) is also being used to size web applications because LOC depends upon the development language but it can be done after the project is finished. FSM (Functional Size Measurement) can be applied on the early stage of web application development. FSM (Functional Size Measurement) is a concept that how to measure the size of software considering functional requirements required by the customer. Functional Points Analysis (FPA) was first developed by Albrecht to support this concept. FSM technique can be used
for early web effort estimation and it’s not dependent on any technology or programming language and it can be used during the whole development life cycle [56].

Marco [57] analyzed FSM and COSMIC (Common Software Measurement International Consortium) methods for software sizing and explained that FSM highlighted some issues while using FSM for web size estimation. On the other hand, COSMIC method was useful to get web application effort estimation. It is easy to understand, simple and cost-effective implementation. It was possible to have some extent accurate estimation for large projects using COSMIC. COSMIC-FFP [58] is a method to measure a functional size of the software to estimate the size of the software applications.

Web application development is becoming very popular on the base of conceptual models and these models can be corrected and reviewed any time during the application development process. It can be helpful for estimation managers to do a detailed analysis of the application before starting the development [37].

The fact is that none of the approaches suggested any technique for applying estimation scenario in certain cases. Industry datasets were used in most of the studies but not academic ones. There are relatively small numbers of research papers which investigated early web cost and effort estimation on the early stage of Web application development [55] [37].

Abrahão [59] provided the comparison between procedures of web applications and FSM methods. There was no use of components and standard notations for effort estimation models creation.

Gathering information on real industry projects is very difficult. Most of the companies don’t share their project’s data. There are no any data repositories available which can be used to gather information or decision making or effort estimation. Technologies are being changed so rapidly that it has become difficult to make a standard model for estimation. Different companies use different approaches and projects nature and size are different. According to research [37], the creation of estimation models on the basis of real industry web projects is the only solution for early web cost and effort estimation. A method was presented for web application effort estimation on the base of the conceptual model and functional size measurement based on COSMIC method [37].

Nagar [60] suggested multi-stage effort estimation for software applications. Every stage of software development required an effort estimate and it means the company has to do effort estimation more than once to make project successful within planned time, effort and money. Each effort estimation model has different limitations according to nature of the projects.

Research [61] proposed a method to calculate early web effort and cost estimation for Web applications. The author suggested to calculate estimate functionality required on each web page. The effort is estimated in hours and each functionality is multiplied by estimated hours. Complexity factor is multiplied and a rough estimate is created for one page and so on. But the author couldn’t really validate the model correction and functional.

Predicting software development effort with high accuracy is still a challenge for estimation managers. Reza [62] suggested an approach to develop correctness of software effort and cost estimation using the structure of data set of a web application. Author has calculated effort estimation by doing clustering based on functional measurement, sub-functional measurement and complex calculations of web applications. There are different cost estimation methods available and can be classified into three different categories: Expert Judgement, Algorithmic models, and machine learning. In the expert judgement, estimation is done on the base of data and related projects. Algorithmic models, different effort, and estimation techniques are part of it like COCOMO. Machine learning, this technique is the
replacement of algorithmic models and it includes like fuzzy logic, regression trees [62]. In the proposed algorithm, effort and cost estimation are based on the conditions of dataset i.e. how many sub functional measurement type. COCOMO required a lot of data in order to do estimation but it provides clear results while on the other hand Expert Judgement is fast predicted estimation approach. Analogy works well if it has lot of projects data [62].

Research [63] presented new size metric, two models WEBMO+ (Web model) and VPM+ (Vector Prediction Model) are suitable to fast estimate the development effort for web applications. Popular size metrics, function points and SLOC (source line of code) are not suitable for all web estimation as it doesn’t consider all Web objects (template, buttons etc). Vector-based Approach is depended upon two concepts i.e. Vector Size Measure (VSM) to measure the size of the system and Vector Prediction Model (VPM) to estimate the effort [63]. Many size metrics are proposed for Web applications like Application points, Object points, and Multimedia points but Web objects (WO) is more appropriate [63]. WO calculation sheet include function point predictors (internal logical files, external interface files, external inputs, external outputs, external inquiries), number of HTML, XML and query lines, Number of multimedia files, Number of scripts, Number of web building blocks. Usage of each Web Object predictors is used in a scale like Low, Average, and High. Effort estimation can be calculated from past projects data using both models, WEBMO+ and VPM+.

Effort estimation methods for early development stages of the web application is compared but did not recommend any specific method [47]. Kaur [64] compared Case Base reasoning and Ordinary Least Square Regression techniques using OO-HFP size measure and concluded that Case Base Reasoning (CBR) is better estimation technique for early Web cost estimation. Rosmina & Suharjito [65] proposed a combined model FHSWebEE which consists of OOmFPWeb, Web metrics for web application size measurement and functional size measurement by web applications. Kaur & Sood [66] compared 3 different analogy software (ANGEL, ACE, ESTOR) and proposed an enhanced estimation model EffortEST. Results of EffortEST are similar to ESTOR. Estimation is more accurate if there is a large number of analogues exist.

Research [55] compared different web resource estimation techniques using systematic literature review. Results show that Case-base reasoning (CBR) is the best approach used for Web resource estimation.
3 METHODOLOGY

According to Creswell, research is a study which is beyond individual experience and ideas [7]. Qualitative, Quantitative and Mix research are three different types of methods used for research. Mixed methods include the collection and mixing or integration of both quantitative and qualitative data [5]. Mixed methods research has increased in popularity in recent years. Its emphasis on collecting, analyzing, and mixing both quantitative and qualitative data. The reason of using Mix methodology in this research is to get more in-depth results as one data source is not enough and initial results need to have further investigation. Surveys and industrial interviews were carried out to get optimal results. Research methodology and research design are presented in this chapter to answer the research questions.

3.1 Research Questions

Following are research questions for this research:

RQ1- What are size metrics and factors used for early web effort estimation by Web companies?

*Motivation:* The motivation behind this research question is to know the size metrics and factors for early web effort estimation being used by Web companies. The answer to this question provides the main set of metrics which are important for any Web application estimate. Answer to this question provides the base for RQ2.

RQ2- Which size metrics are most commonly used in early web effort estimation by Web companies?

*Motivation:* The motivation behind this research questions is to find the most commonly used metrics in early web cost estimation. There are many metrics which are being used for early estimates but the purpose of this question to list down top few metrics which are very critical. Answer to this question provides the base for RQ3.

RQ3 - Have the trends changed between the original work and this revisited research with regard to the Web predictors that are used the most in practice?

*Motivation:* The motivation for this research question is to compare these revisited research results with old research which was conducted in 2002. Motivation is to identify any differences in web predictors if trends have been changed or still the same during early web cost estimation.

The optimal results will be achieved by conducting two surveys and one Interview method. The first survey will be conducted to gather data about size metrics, cost drivers, profit metrics and contingency. List of most commonly used metrics will be created from the survey results. Results of survey one will be validated by industrial interviews. Results of both validations will be used to create web project data entry form which will be used to gather data on web projects. List of most influential effort predictors will be created after gathering data on many real projects worldwide.
3.2 Research Design

This study has been conducted in following phases:

1. Survey (S1)
2. Industrial Interviews
3. Survey 2 (quote form)
4. Data synthesis & Comparison

<table>
<thead>
<tr>
<th>Research Question</th>
<th>Methodology</th>
</tr>
</thead>
<tbody>
<tr>
<td>RQ1</td>
<td>Survey (S1)</td>
</tr>
<tr>
<td>RQ2</td>
<td>Industrial Interviews and Survey (S2)</td>
</tr>
<tr>
<td>RQ3</td>
<td>Data synthesis and Comparison</td>
</tr>
</tbody>
</table>

Table 1 – Research Questions & Methodologies

The research methodology was designed in order to get desired results. Surveys and Industrial Interviews were used in this research. This thesis research was divided into 5 different phases which are explained below:

3.2.1 Survey 1 (S1)

A survey was conducted as the first step in this research. Survey (S1) was based on an online search. The aim of the survey was to identify factors and size metrics used by web companies in order to get early web cost estimation. Maximum web companies were targeted and web quote forms were downloaded. Google was used as a search engine in order find and download web quote forms. Google is high ranking search [13] [14] and that’s why it was used for survey 1. After downloading quote forms, each form was studied and transcribe into excel sheet with metrics list. Different search queries were performed in order to get an optimal list of metrics. The motivation of using Survey for RQ1 was to revisit old research and Survey was used in old research as well. Another reason was to find maximum metrics from web quote forms and there was no any other way to find maximum quote forms around the world except a survey using an online search engine.

3.2.2 Industrial Interviews

Interviews were conducted in order to validate results of Survey 1. There are different kinds of interviews research methodology exists but I used semi-structured interviews technique. The motivation for the selection of semi-structured interviews is explained in section 4. The validation was done based on the ranking of metrics selected from survey 1. Participants for these interviews were having experience in effort and cost estimation in real web projects. All participants were from software engineering domain and have many years working experience in software estimation and development.

3.2.3 Survey 2 (Quote form)

A quote form was designed using results gathered from interviews. Purpose of quote form was to get feedback from global companies about early web cost estimation used on real projects. Quote form was sent to companies worldwide to get data on projects and did a follow-up to get feedback on quote form. Detail about this section is explained in section 5.
3.2.4 Data Synthesis and Comparison

Data was collected on real projects from worldwide companies and an analysis was performed. CBR (Case base reasoning) data analysis technique was used to analyze the project’s data. Tukutuku implementation was done using collected data.

A detailed comparison was done between old research [4] and new research. Comparison of results showed trends and changes in early web cost estimation used by web companies. It highlighted the practices followed by the Web companies for early web cost estimation and also showed new techniques and methods which did not exist in old research. Detail comparison is discussed in section 6.

![Research Methodology Diagram]

Figure 2 Research Methods
4 Survey

4.1 Introduction

A survey was conducted in order to answer the first research question. Purpose of this survey was to gather web size metrics and factors used by web companies for early web cost estimation. The target audience in this survey was web companies which provide web quote form online on their websites. There was no need to contact them personally; all forms were downloaded from their websites.

In order to target and select appropriate survey population, it was important to select best web search engine. There were different options available about search engines and I selected top rated search engine. A web search engine (www.google.com) was used to carry out search and to acquire survey population of for this research. The reason for selection of this particular search engine was that currently, it stands as number one free search engine [13] [14]. Main objectives of this search were to find the list of size metrics and factors which are being used by web companies to get initial web cost estimation using online web quote forms.

Search queries were finalized in this research is an extended version of old research conducted by Mendes [4]. Same search query was used to find web quote forms which are mentioned by Mendes [4]. Two more search queries were added to get correct and extended results, so total three search queries were executed on the search engine in order to find maximum web quote forms. The motivation for using more search queries was to get maximum data which can help to find maximum number of size metrics being used by the web companies.

Research [4] [15] conducted by Mendes was also considered to find more relevant keywords in order to create more search queries. Some keywords were defined along with their synonym and based on those keywords and synonyms initial search was conducted to gather relevant data. Different sets of keywords combination were used to execute the search on the search engine. Following are steps which were used to create search queries for the search engine in order to conduct this survey.

4.2 Keywords and Search String

The following steps were performed in order to develop the search strings:

- Search queries from old research
- Major search terms were extracted from research questions and title
- Synonyms of search keywords terms were used
- Boolean operator AND was used to combine major search terms and synonym
- Guidelines and assistance was taken from the librarian about search tips

Following are set of keywords which were used to develop search strings:
Queries were designed with this aim to get online quote forms from the web companies which are being used to get early web cost estimation. All queries were executed separately on Google search engine. List of URLs was visible in the Google search results and then each URL link was opened and verified for quote form. Selection of URLs was done after applying inclusion and exclusion criteria mentioned in section 4.3. Each selected URL was opened and quote form was downloaded. All the set of metrics and factors were documented in an excel sheet. Any metric or quote form which was duplicate was excluded from the final list. Following were queries which were executed on the search engine in order to get metrics list from quote forms:

1. “How to accurately estimate a web design project”

This query was taken from the research conducted by Mendes [4]. Total 507 records were found when it was executed on the search engine. After applying inclusion and exclusion criteria then only 30 URLs were selected which had relevant detail about quote form. The motivation for the selection of this query was to evaluate the difference between old and current research.

2. “Web development project estimator”

This query was executed on the search engine to get more relevant records about this research. This query provided total 447 records and 45 were selected after applying inclusion and exclusion criteria. Selected URLs were providing quote forms for early web cost estimation.

3. “Quote form web development project”

This query provided total 620 URLs and 137 were selected after applying inclusion and exclusion criteria. Duplicate URLs were not selected.
4.3 Inclusion/Exclusion

Inclusion and exclusion criteria were defined in order to get accurate and quality results. Following were considered in inclusion criteria:

- Select those URLs where quote form is in English
- Select those URLs where enough useful information provided in quote forms
- Select those URLs where quote is about Web application

An exclusion criterion was defined to filter quality results and it includes following factors:

- Do not consider those URLs where URLs are duplicate or repeated
- Do not consider those URLs where there is very limited information available on quote forms
- Do not consider those URLs where follow-up links are opening and do not direct any relevant page
- Do not consider those quote forms where metrics are being duplicated

Firstly, Query 1 was executed and results were collected then Query 2 was executed and results were collected. Duplicate URLs were excluded from the results of Query 2. Lastly, Query 3 was executed and results were collected and duplicate results were excluded by following the exclusion criteria.

4.4 Sample quote form from the web

Following is a sample quote form where the user can answer different questions and check the estimated quote about the web application.
<table>
<thead>
<tr>
<th>Feature</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of pages</td>
<td>$2,000</td>
<td>$3,000</td>
</tr>
<tr>
<td>Style of design</td>
<td>$3,000</td>
<td>$5,000</td>
</tr>
<tr>
<td>Copywriting # of pages</td>
<td>$1,000</td>
<td>$1,500</td>
</tr>
<tr>
<td>SEO w/ Placement Guarantee</td>
<td>$4,000</td>
<td>$6,000</td>
</tr>
<tr>
<td>Responsive Design</td>
<td>$3,000</td>
<td>$3,000</td>
</tr>
<tr>
<td>Database Integration</td>
<td>$4,000</td>
<td>$10,000</td>
</tr>
<tr>
<td>e-Commerce Functionality</td>
<td>$2,000</td>
<td>$4,000</td>
</tr>
<tr>
<td>CMS</td>
<td>$4,000</td>
<td>$10,000</td>
</tr>
<tr>
<td><strong>Total Estimated Quote</strong></td>
<td><strong>$23,000</strong></td>
<td><strong>$42,500</strong></td>
</tr>
</tbody>
</table>
4.5 Results of Survey (S1)

There were total 212 URLs where quote form was available for customers. These quote forms were being used for early web cost estimation. The different set of metrics and factors were used on those quote forms. Each quote form was varying from company to company. Different Web companies were asking different questions in the quote forms in order to get the better understanding of the customer’s needs and get an overall idea about the estimate. Companies were using customer’s provided feedback to estimate initial cost of the projects. Some companies implemented cost calculator on the quote forms so that visitor can see estimated price about the Web application on the base of quote form input.

All the quote forms were downloaded and gathered all metrics list from these quote forms to make a final document comprising questions strength and their repetitions.

Results from the survey were collected in a simple way into an excel sheet. Different Web companies asked a variety of questions in their quote forms. Some companies asked long and detailed questions while some other companies asked short and few questions depending on their needs. Majority of the questions were to estimate the Cost, Time and amount of effort (majorly technical) that are likely to be employed during a Web project development and management.

Total 212 web companies were selected to collect web quote forms. The below mention table contains a list of categories. Those categories further hold set of metrics and a percentage was provided about asked metrics. The percentage is the weightage that describes how many times the metrics were used by the Web companies. Results of Survey 1 gave a list of metrics and then mostly used metrics were shortlisted on the base of occurrences in the quote form. The selection was made on the basis of repetition and percentage.

Following are the result of Survey (S1):

<table>
<thead>
<tr>
<th>Category</th>
<th>Metrics</th>
<th>Repetitions</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Static Metrics</td>
<td>Static graphics contents (logo, photo, flyer, brochure)</td>
<td>66</td>
<td>31.32</td>
</tr>
<tr>
<td></td>
<td>Dynamic graphics contents (Include multimedia, flash, interactive dynamic forms)</td>
<td>62</td>
<td>29.48</td>
</tr>
<tr>
<td></td>
<td>Number of Pages</td>
<td>107</td>
<td>50.47</td>
</tr>
<tr>
<td>Dynamic Metrics</td>
<td>Public features (membership, login, forums, RSS, Map, File Upload, Site Search, multilingual)</td>
<td>21</td>
<td>9.90</td>
</tr>
<tr>
<td></td>
<td>Layouts and functionality</td>
<td>24</td>
<td>11.32</td>
</tr>
<tr>
<td></td>
<td>Responsive support (mobile devices)</td>
<td>57</td>
<td>26.89</td>
</tr>
<tr>
<td></td>
<td>Social Media package (Twitter, Facebook, Google+, YouTube Channel)</td>
<td>53</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>Search Engine Optimization support (meta-tag keyword optimization, SE Submission, link building, Press release)</td>
<td>98</td>
<td>46.23</td>
</tr>
<tr>
<td></td>
<td>Metrics for e-commerce (Cart, credit card processing, online payment or PayPal, e-business products)</td>
<td>102</td>
<td>48.11</td>
</tr>
<tr>
<td></td>
<td>Additional features</td>
<td>106</td>
<td>50</td>
</tr>
<tr>
<td>Cost Drivers &amp; Budget Metrics</td>
<td>New application or update existing application</td>
<td>83</td>
<td>39.15</td>
</tr>
<tr>
<td></td>
<td>Available project budget</td>
<td>79</td>
<td>37.26</td>
</tr>
<tr>
<td></td>
<td>Start date for the project</td>
<td>28</td>
<td>13.21</td>
</tr>
<tr>
<td></td>
<td>End date for the project</td>
<td>45</td>
<td>21.23</td>
</tr>
<tr>
<td>Web company metrics</td>
<td>Contact information</td>
<td>63</td>
<td>29.71</td>
</tr>
<tr>
<td></td>
<td>Company description</td>
<td>52</td>
<td>24.53</td>
</tr>
<tr>
<td></td>
<td>Target audiences</td>
<td>12</td>
<td>5.66</td>
</tr>
<tr>
<td></td>
<td>Extra comments</td>
<td>39</td>
<td>18.40</td>
</tr>
</tbody>
</table>
Table 3 – List of Metrics after S1

Data is gathered from 212 web companies. All data gathered from different quote forms was organized into five different categories and all identified metrics were explained according to repetition and percentage. To calculate the percentage, I used following formula:

\[
\text{Percentage} = \frac{\text{No of size metrics repetitions in quote forms} \times 100}{\text{Total no of companies}}
\]

Collected data from survey 1 is divided into five different categories which are following:

**Static metrics**

Formally web application static metrics are available in initial stages of the SDLC (software development life cycle) and it deals with structural features of the web application. These metrics do not deal with responsive features of the web application. So, there is no dynamic functionality involved in these kinds of metrics.

**Dynamic metrics**

Web application dynamic metrics are available on customer requirements mostly at the stage of the SDLC. These metrics belong to dynamic behavior of the web application and it supports object-oriented features [15].

**Cost Drivers & Budget Metrics**

Cost drivers are the major factors which have the influence on the effort and time to develop web applications. Examples include Size and Complexity of the application, the participation of user organization etc. It is the hard target to project the actual amount to be invested in the project [24].

Cost driver includes the factor which affects cost. Changes in cost drivers are also one of the causes in the overall cost of the project [24].

**Web Company related Metrics**

Web Company related metrics are set of metrics about the company profile, user accounts, login, analytics, and feedbacks.

**Web Management**

Web project metrics normally use strategic and managerial purposes by project leadership and a software team collaborates with leadership to fulfill their requirements. It includes CMS, promotional and marketing plan for the website, search engine optimization and social media features.
4.6 Findings of Survey (S1)

Survey 1 presented number of metrics. Occurrences and percentage were calculated on the base of gathered data. There is a number of other factors which are not size metrics and those metrics are excluded from the final list. For example Start Date of the project, End date of the project, domain-specific feature like e-commerce, project management support etc. Following 5 top metrics were identified in the Survey 1 which most of the companies asked in their quote forms. This selection was made on the basis of repetition of these metrics in quote forms by Web companies:

- Number of web pages/features
- Number of static/dynamic graphics contents
- Responsive implementation
- Search Engine Optimization support
- Contents management support
5 INDUSTRIAL INTERVIEW

Interviews are explained as a discussion which involves a minimum of two people [10]. Interviews are efficient method to extract and investigate information regarding relevant research by interrogating professionals from industry [9]. In order to get information and understanding from industry, various professionals were interviewed [8]. Face-to-face and telephonic were two different methods of conducting interviews. There were three different methods to conduct an interview that includes structured, unstructured and semi-structured. When the questions include Yes & No then it’s called structured interviews. Unstructured interviews are those in which interviewee is a source of questions with their answers. Semi-structured interviews are the combination of structured and unstructured interviews.

According to the nature of research and in order to gather rich data from industry, semi-structured interviews approach was adopted [16]. The semi-structured approach is one which allows two-way communications and in order to get detailed information from the interviewee. During interview, session’s researcher asked a prepared list of questions and conducted a session of follow up questions [9]. Interviewees were also asked few addition questions which were raised during first few interview sessions. Interviews were conducted from the different region of the world and Skype/telephonic interviews were conducted. Few face-to-face interviews were also conducted.

5.1 Purpose of Interview
The main reason to conduct interviews was to validate results of survey 1. Another reason to conduct interviews was to validate research from the industrial point of view. The results of interviews provided partially answer to research question 2. To conduct industrial interviews researcher design questions in such a way to grab insight knowledge from the interviewee. The advantages of conducting these interviews were to get a deep knowledge about web cost estimation from industrial practitioners.

5.2 Selection of Interview Subjects
To gather and collect relevant information from industrial practitioners it was necessary to select most appropriate candidates. A selection criterion was established and only those participants were selected who fulfills those criteria. Following criteria was set to select candidates for interviews in order to get appropriate results:

- Those participants who were currently working in the software industry
- All of the selected interviewees must have a minimum of two years of experience
- Participants must have experience in estimation of different web projects
- Participants should be part of the reputed organization and holds a solid position
- In order to get better and accurate results selection of participants were made from different countries
To make research trustworthy and authenticated the targeted participants were the software engineer, consultant, software manager, software developer and requirement engineers.

5.3 Study Instruments
In order to fetch appropriate data from industrial participants, researcher defined and designed a study instrument. The study instrument used for industrial interviews was results of survey 1 and the table is provided in section 3.4. This helped in findings answer of research question 2 partially. Open-ended questions were used as a base for the designed instrument to get in-depth feedback from participants. To make research more authenticated a follow-up questions discussion session was also held with each participant to get as much as information possible. The follow-up discussion session assisted a lot in getting the latest information about the estimation size metrics. The discussion and interviews session with each selected candidate was about 20-35 minutes maximum. Industrial interviews results were gathered and analyzed.

5.4 Interview Structure
To carry the interview more structured and accurate few steps were followed, planning, analysis, during interview feedback and post-interview analysis [11].

5.5 Designing
The intention of designing interview was to define a strategy to gather required information. Interview structure and design of industrial interview were designed properly in order to get quality results. Special effort was put during designing and structure of interview so that better industrial aspects can be achieved. Both open-end and explicit questions were designed in order to conduct this research. The interview was designed from results of survey 1. In this research, a single study instrument was designed. All questions in study instrument were achieved from survey 1 result. This was designed to validate survey 1 results and to get the industrial opinion on identified metrics.

5.6 Interviewees
Total 15 industrial interviews were conducted where 6 were face to face interview and 9 were on Skype interviews. These interviews were conducted with mutual understanding and availability of interviewee. All these interviewees had experience in estimating web application projects and this was the core reason to select these interviewees. A small discussion was conducted with interviewees before having a formal interview. Short description of the research work was explained so that interviewee can understand questions properly and answer correctly. The goal of interviews was shared with interviewees. As a next step, set of designed questions were asked along with following up questions. A general discussion session was held after each interview to grab more in-depth information. All interviews were recorded with the permission of interviewee. Keynotes were also written along with the recording. Each interview took around 20-35 min.
5.7 Transcribing
Interviews were transcribed and results were gathered. Purpose of doing this step was to gather all relevant information and discard unnecessary information. Interviews were recorded and notes were also taken. Recorded interview was considered during re-writing details if needed. After each interview session interviews were transcribed using recording and written notes. Once the interviews were transcribed then it was sent to interviewee to validate. To make result more efficient and error free, all interviews were transcribed into XMind.

5.8 Analyzing and Verifying
Qualitative data analysis method was used to analyze the transcribed data which was gathered from interviews [12]. The investigated output was reserved for verification. Qualitative data analysis (QDA) has three phases which include notice, collect and think about things [12].
Qualitative data analysis was used to analyze all gathered data. It is non-linear analysis model which consists of different features which are following:

1. It is spiral in nature and assists in repeating steps
2. Each step is considered as a complete process
3. Due to recursive nature of model researcher needs to collect more data for the same problem

Qualitative data analysis method has three phases from the analysis as shown in the figure. The phase one allows the researcher to observe and code relevant data which was provided by all interviewees. All the data which was recorded and note created during each interview session was used for this purpose. The next step was to make an analysis on the bases of the recorded interview so that researcher can fetch the related information which was considered as metrics. Later these metrics were further sorted according to the number of replies received from interviewees. The final phase was critical from all three phases, a careful analysis was done on interviews results.

5.9 Analysis and Results of Interviews

This table contains all results that were achieved after interviewing 15 practitioners from industry.

<table>
<thead>
<tr>
<th>Metrics</th>
<th>No of times</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do you consider static graphics contents (logo, photo, flyer, and brochure) as an important size metrics?</td>
<td>[Int-1], [Int-2], [Int-3], [Int-4], [Int-5], [Int-7], [Int-8], [Int-9], [Int-10], [Int-11], [Int-12], [Int-13], [Int-15]</td>
</tr>
<tr>
<td>Do you consider dynamic graphics contents (Include multimedia, flash, interactive dynamic forms) as an important metrics? (10)</td>
<td>[Int-1], [Int-2], [Int-3], [Int-4], [Int-5], [Int-6], [Int-9], [Int-10], [Int-11], [Int-15]</td>
</tr>
<tr>
<td>Are number of pages a necessary metrics which calculating size?</td>
<td>[Int-1], [Int-2], [Int-3], [Int-4], [Int-5], [Int-6], [Int-7], [Int-8], [Int-10], [Int-11], [Int-12], [Int-13], [Int-14], [Int-15]</td>
</tr>
<tr>
<td>- If template available, then no need for pages count (14)</td>
<td>[Int-1], [Int-2], [Int-3], [Int-4], [Int-5], [Int-6], [Int-7], [Int-8], [Int-9], [Int-12], [Int-14], [Int-15]</td>
</tr>
<tr>
<td>Public features (membership, login, forums, RSS, map, file upload, and site search multilingual) are also considered during size metrics</td>
<td>[Int-1], [Int-2], [Int-3], [Int-4], [Int-5], [Int-6], [Int-7], [Int-8], [Int-9], [Int-12], [Int-14], [Int-15]</td>
</tr>
</tbody>
</table>
estimation? (12)

What about layouts and functionality? (10)  [Int-1], [Int-4], [Int-5], [Int-6], [Int-7], [Int-8], [Int-9], [Int-10], [Int-11], [Int-14]

New application or update existing application does this matter during estimation? (12)  [Int-1], [Int-3], [Int-4], [Int-5], [Int-6], [Int-8], [Int-9], [Int-10], [Int-11], [Int-12], [Int-13], [Int-14]

What about additional features? (7)  [Int-2], [Int-5], [Int-7], [Int-9], [Int-11], [Int-14], [Int-15]

Responsive support (iOS and android) (14)  [Int-1], [Int-2], [Int-3], [Int-4], [Int-5], [Int-6], [Int-7], [Int-8], [Int-9], [Int-10], [Int-11], [Int-12], [Int-13], [Int-15]

Social Media package (Twitter, Facebook, Google+, YouTube Channel) (9)  [Int-1], [Int-3], [Int-5], [Int-7], [Int-9], [Int-10], [Int-12], [Int-13], [Int-15]

Search Engine Optimization support (meta-tag keyword optimization, SE Submission, link building, Press release) (11)  [Int-3], [Int-4], [Int-5], [Int-6], [Int-7], [Int-8], [Int-9], [Int-11], [Int-12], [Int-13], [Int-15]

Web Hosting / Domain Name/ Cloud hosting (9)  [Int-2], [Int-3], [Int-4], [Int-5], [Int-7], [Int-9], [Int-11], [Int-12], [Int-13]

Metrics for ecommerce (Cart, credit card processing, online payment or PayPal, e-business products) (3)  [Int-6], [Int-7], [Int-9]

Maintenance and support services
- Not in initial cost estimate (13)  [Int-1], [Int-2], [Int-3], [Int-4], [Int-5], [Int-6],[Int-9], [Int-10], [Int-11], [Int-12], [Int-13], [Int-14], [Int-15]

Marketing and Promotional strategies (banner and campaign, custom blog, site survey, news post/newsletter, AdSense, classified or Ads, email marketing or mailing list) (2)  [Int-6], [Int-9]

Contents management support (CMS Training, Content development, writing, copyrighting, cookies audit, xml sitemaps, content transfer) (9)  [Int-1], [Int-2], [Int-3], [Int-4], [Int-5], [Int-7], [Int-9], [Int-11], [Int-14],

<table>
<thead>
<tr>
<th>Table 5 – Interview Results</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Metrics</strong></td>
</tr>
<tr>
<td>Documentation</td>
</tr>
</tbody>
</table>

The table below contains an extra metrics which were provided by interviewee additionally.

<table>
<thead>
<tr>
<th>Metrics</th>
<th>No of Interviewees</th>
</tr>
</thead>
<tbody>
<tr>
<td>Documentation</td>
<td>2</td>
</tr>
</tbody>
</table>
5.10 Final Selected Metrics

Following table holds the list of web size metrics and cost drivers which were finalized after conducting industrial interviews. Aim to carry out this research is to identify size metrics and cost drivers for early web cost estimation on the base of current practices of several web companies worldwide. "Total Number of pages/features" is a most commonly used size metric to measure the early web cost. Similarly, many new size and cost factors identified in this revisited research.

<table>
<thead>
<tr>
<th>#</th>
<th>Metrics</th>
<th>No of times</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Number of pages/features</td>
<td>14</td>
</tr>
<tr>
<td>2</td>
<td>Responsive implementation</td>
<td>14</td>
</tr>
<tr>
<td>3</td>
<td>Static graphics contents (logo, photo, flyer, and brochure)</td>
<td>13</td>
</tr>
<tr>
<td>4</td>
<td>Maintenance and support</td>
<td>13</td>
</tr>
<tr>
<td>5</td>
<td>General Public features (membership, login, forums, RSS, map, file upload, and site search, multilingual)</td>
<td>12</td>
</tr>
<tr>
<td>6</td>
<td>New application or update existing application</td>
<td>12</td>
</tr>
<tr>
<td>7</td>
<td>Search Engine Optimization support (on and off page optimization)</td>
<td>11</td>
</tr>
<tr>
<td>8</td>
<td>Dynamic graphics contents (Include multimedia, flash, interactive dynamic forms)</td>
<td>10</td>
</tr>
<tr>
<td>9</td>
<td>Application layout and core features functionality</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>Social Media package (Twitter, Facebook, Google+, YouTube Channel)</td>
<td>9</td>
</tr>
<tr>
<td>11</td>
<td>Web Hosting / Domain Name/ Cloud hosting</td>
<td>9</td>
</tr>
<tr>
<td>12</td>
<td>Contents management support (CMS Training, Content development, writing, copyrighting, cookies audit, xml sitemaps, content transfer)</td>
<td>9</td>
</tr>
<tr>
<td>13</td>
<td>Additional features (security, performance)</td>
<td>7</td>
</tr>
<tr>
<td>14</td>
<td>Domain specific application features i.e. ecommerce, healthcare</td>
<td>3</td>
</tr>
<tr>
<td>15</td>
<td>Marketing and Promotional strategies (banner and campaign, custom blog, site survey, news post/newsletter, AdSense, classified or Ads, email marketing or mailing list)</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 7 – List of Metrics after Interviews

5.11 Discussion and Summary

The major contribution of this chapter was to present results which were gathered after execution of various interviews. The feedback from industry practitioners and relevant information helped in getting the desired data. After by applying qualitative data analysis method on gathered data researcher managed to achieve final list of metrics. All these interviews were based on the Survey 1 results. Moreover, gathered and identified results of the interview were considered as an input for quote form. Survey 2 (quote form) was designed from the results of interviews. The reason of this survey study was to gather data on real projects and validate results of interviews.
There are few metrics which were answered with the entity of “May be” means that practitioners consider those metrics depending on customer demand and type of project. To avoid biasness in research, researcher considers those as excluded metrics which has results like “May be”. Only those matrices were considered which were answered as a yes by industry.
6 SURVEY 2 (QUOTE FORM)

A quote form was designed using results gathered from interviews as final metrics in section 5.10. Purpose of quote form was to get feedback from global companies about early web cost estimation used for real projects. Quote form was built using an online service http://www.survio.com/. Quote form consists of total 21 questions which were divided into different sections. It includes questions about basic info about the company, teams as well along with size and cost metrics questions. Quote form is available in Appendix.

The contact information of different companies was gathered from different websites, friends, and references. Companies were contacted via telephonic calls and emails. Basic information was provided on phone calls about the quote form and project. Quote form was sent via email and shared on different social media channels. It was told to the companies that final research results will be shared with them as a compensation for filling quote forms using real web projects data. These research results can help companies to improve their cost estimation process.

Follow up phone calls were made to the companies and asked them to fill the quote forms. Finally, 29 valid projects data were collected.

Final stats show that quote form was opened using the direct link which was the main visiting source.

Most of the participants took only 5-10 min in order to fill quote form and provided real projects data. This graph shows average time of completion for quote form:
Total 13% companies shared projects data while other 87% companies didn’t provide projects data. It shows that not every company wanted to share their project’s data. Total 29 projects complete data was collected while other 157 started but then didn’t share projects data. I tried to follow up with different companies and could manage to get data on 29 projects.

In the next step, data was transformed into excel sheet. Each project data was taken care one by one and all 29 projects data was merged into one excel sheet which is available in the Appendix (Projects data collected via Quote Form). It was a simple process to copy results from online Survio service and then paste it into excel sheet under the relevant project. Complete survey (Quote Form) can be found at this link: https://www.survio.com/survey/d/G0E4D8Q9N4N6Q2K7Q
7 ANALYSIS AND DISCUSSION

Tukutuku research project was initiated in 2002 by Mendes [4]. Data gathering initiative was called Tukutuku benchmarking project. The purpose of Tukutuku project is to gather data on web projects to be used to develop web cost estimation models based on early effort predictors and to benchmark productivity in the web companies [4].

Survey 1 was conducted in order to get a list of cost factors and size metrics used by web companies. Data was collected from 212 quote forms. 15 Interviews were conducted in order to validate results of Survey 1. A quote form was designed on the base of most commonly used and verified cost and size metrics.

Data from 29 real projects were collected from web companies around the world. All projects data was stored online on cloud-based service. I have manually fetched all 29 projects data and transformed into excel sheet. A set of Tukutuku database variables are created on the base of current new projects data and new approaches being used in web projects development. The table contains variable names, different scaling, and description.

Variables for Tukutuku database:

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Scale</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country</td>
<td>Nominal</td>
<td>Company country name</td>
</tr>
<tr>
<td>TeamExp</td>
<td>Ratio</td>
<td>Web application development experience of teams</td>
</tr>
<tr>
<td>EffortEstExp</td>
<td>Ratio</td>
<td>Average web effort estimation experience</td>
</tr>
<tr>
<td>DevTeam</td>
<td>Ratio</td>
<td>Average number of developers in development team</td>
</tr>
<tr>
<td>WebStaticPag</td>
<td>Ratio</td>
<td>Number of static web pages implemented</td>
</tr>
<tr>
<td>WebDynamicPag</td>
<td>Ratio</td>
<td>Number of dynamic web pages implemented</td>
</tr>
<tr>
<td>DevWRD</td>
<td>Ratio</td>
<td>Number of devices support for web responsive design</td>
</tr>
<tr>
<td>Support</td>
<td>Ratio</td>
<td>Maintenance and support services</td>
</tr>
<tr>
<td>StaticImg</td>
<td>Ratio</td>
<td>Use of static images</td>
</tr>
<tr>
<td>StadModule</td>
<td>Ratio</td>
<td>Number of standard modules provided by client or purchased</td>
</tr>
<tr>
<td>DevType</td>
<td>Nominal</td>
<td>Type of development either new or existing</td>
</tr>
<tr>
<td>SEO</td>
<td>Ratio</td>
<td>Number of SEO features</td>
</tr>
<tr>
<td>DynamicContent</td>
<td>Ratio</td>
<td>Number of dynamic content added</td>
</tr>
<tr>
<td>CustomMod</td>
<td>Ratio</td>
<td>Number of pre-developed custom module given by customer</td>
</tr>
<tr>
<td>SocialMed</td>
<td>Ratio</td>
<td>Number of social media packages integrated</td>
</tr>
<tr>
<td>CMS</td>
<td>Ratio</td>
<td>Number of CMS integrated</td>
</tr>
<tr>
<td>AddFeature</td>
<td>Ratio</td>
<td>Number of addition feature implement in web application</td>
</tr>
<tr>
<td>Devteam</td>
<td>Ratio</td>
<td>Size of development team</td>
</tr>
</tbody>
</table>

Table 8 – Tukutuku Database Variables

A list of variables created which can be used to build a cost model. Some variables are excluded from the list and exclusion was done on the base of following criteria:
- If instances of the variable are zero
- If the variable is of categorical type
- If the variable is about basic info
- If the variable is related to some other variable where both variables cannot be excluded
The technique used to build cost model is estimation by analogy (Case-based reasoning). In order to build estimation model for currently collected data from 29 projects, various options were available and I selected case-based reasoning (CBR) technique. Case-base reasoning is a model which solve the new problem by reusing past information of similar kind of problems [25] [26].

In this approach, all information regarding past solved problems is maintained which in future is used to solved new problems. All the experiences are stored in databases as different cases. In order to solve a new problem then help is taken from stored cases in the database. There are many advantages of case base reasoning [27] [28]:

- The approach allows taking the decision in the same way as human does in real time.
- Quantitative and qualitative data is handled by CBR.
- Every time a new problem is considered as a new case
- It has the ability to learning from failure and success of past solved problem
- CBR use stored information and solution in order to solve a new problem.
- It is simple, flexible and easy to implement as compare to other approaches
- All steps for software cost estimation are supported effectively

Aamodt and Plaza described the case-base reasoning in four different step which is following [28]:
1- Retrieve – Any previous case which is relevant to the current problem
2- Re-use - Use of current or previous case in various ways
3- Revise - Reusing the previous case to provide the solution for new one
4- Retain - Addition of new solution in the present case base database which helps CBR system to get larger and making its valuable resource for future

Figure 5 CBR Model

Variables are designed on the base of numerical and ordinal data collected from real web projects. Numerical data shows results as numeric values while ordinal data shows results as a symbolic or specific range.

Term “case-based reasoning” consists of three words. Case is an experience of the solved problem, the base is a collection of such cases and reasoning term means that approach intended to draw conclusions using cases. [30].

CBR gives estimates after doing the comparison between current problems against a collection of historical projects/cases. Similarity features are compared with completed
projects. An estimate is taken from completed projects and the new estimate is calculated. There are different techniques to do similarity assessment but nearest neighbor algorithms using a weighted Euclidean distance metric is being used in software engineering and web engineering [31]. Watson compares different CBR techniques and found that Weighted Euclidean distance was the most accurate method for predicting development effort for web hypermedia software [49].

Case-based Reasoning consists of [29]:

- Describing the project where estimation is required i.e. finding projects attributed which can have the impact on effort.
- Use of dataset to find similar projects where the effort was calculated
- Use of adjusted effort values to create predicted effort value

There is the number of parameters for decision making while using CBR [29]:

- Scaling – conversion of attribute values according to a defined rule such as all attributes have the same level of influence.
- Analogy adaptation – way to create estimation for new project
- No. of analogies – Number of cases which are similar to each other and used to generate estimation.
- Selection of feature subset – the best subset of features which gives the more correct estimation
- Similarity measures – this is about the level of similarities between the cases

There can be more parameters and each parameter can be divided into more in-depth detail in order to adjust values in CBR tool. CBR tool generates the output according to given configurations.

In this research, statistical data gathered from real 29 company web projects which consist of numerical and ordinal types. There are different data statistical techniques available in order to analyze projects data. Case-based reasoning is the estimation technique which is very suitable for this kind of data where there is numerical, ordinal or any other type of data.

There are different open source CBR tools available. Some of the tools are added in this list:

- CBR Shell
- FreeCBR
- jCOLIBRI
- myCBR
- eXiTCBR

It was a challenge to select one CBR tool in order to do analysis on the project data in order to find most commonly used cost and size metrics by web companies for web application development. FreeCBR tool was selected to perform this analysis and following were reasons to choose this tool [28]:

- It’s an open sources java implementation
- Easy to use as compare to other tools
- Better accuracy
- Flexibility to add different of variables i.e. numeric, bool, string, multi-string and float
- Simple GUI and easy to configure
Selected variables were added in FreeCBR tool and relevant type was selected for each variable according to tool’s suggested type. Data of all 29 projects were added using FreeCBR tool.

<table>
<thead>
<tr>
<th>Country</th>
<th>Start (years)</th>
<th>End (years)</th>
<th>Years (years)</th>
<th>Dev (days)</th>
<th>Staff (days)</th>
<th>DevType (days)</th>
<th>StaffType (days)</th>
<th>DevRMD (days)</th>
<th>StaffRMD (days)</th>
<th>DevTypeRMD (days)</th>
<th>StaffTypeRMD (days)</th>
<th>BDD (days)</th>
<th>CommentContent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pakistan</td>
<td>0-1 years</td>
<td>0-2 years</td>
<td>1-3 years</td>
<td>10</td>
<td>5</td>
<td>2</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>USA</td>
<td>0-1 years</td>
<td>0-2 years</td>
<td>1-3 years</td>
<td>10</td>
<td>5</td>
<td>2</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>India</td>
<td>0-1 years</td>
<td>0-2 years</td>
<td>1-3 years</td>
<td>10</td>
<td>5</td>
<td>2</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>10</td>
<td>3</td>
</tr>
</tbody>
</table>

...
As a validation part, the specific case was added to the Search fields and a default scaling type is selected in order to find the relevant case. The Fuzzy linear scale is useful in my case. If fuzzy linear is "=" then it means if exact match then distance = 0 else needs to use the NormalDistance algorithm [48] in FreeCBR. A project-specific data is added input for “Search Value” and then click on Ok button.
Once Fuzzy linear Scale is executed with given Search Values then it returned project data with maximum Hit %. Results collected from the case on the base of existing cases show in percentages. These percentages show most relevant matched case of existing project data with newly search case using FreeCBR tool.

Different cases were run through using FreeCBR tool and compared it with the existing projects cases. Each executed case returned most commonly relevant project data in percentage and then relevant variables collected from the project data. Multiple cases were run through and final list of variables found on the base of occurrences and validity.

Projects data were compared with new search cases and it validated the mostly used size and cost metrics. These were most commonly used metrics which were validated using FreeCBR tool:
- Total number of pages or features
- Responsive implementation

“Number of pages or features” is the one of most important metric for size and cost estimation. Adding new features and pages will increase the cost and time of the web application.
“Responsive implementation” is the 2nd critical factor for size and cost estimation. If a web company develop a web application which is responsive (works on mobile, iPad, tablet, desktop and different screen sizes) will increase the cost and time estimate in order to complete the web project. Responsive web design is a term derived by Marcotte in his article in 2010 (Marcotte) and next year he wrote his book “Responsive Web Design” [52]. In new book Marcotte proposed new techniques and some design patterns. .Net magazine announced "responsive design" as the second development trend of the year 2012 [52]. The main objective of RWD (Responsive Web Design) is to deliver high-quality user experience no matter about the size of the display. RWD provides a user-friendly easy reading, navigation and viewing contents of web applications on different devices. RWD techniques provide the solution of different design, layout, menu, navigation patterns. In order to implement RWD, developers need to write CSS media queries or to use different available libraries to make application as responsive. It takes extra effort from developers to build the application with responsive support and takes more time for testers to verify Web applications on different screen sizes and browsers. Some mostly common CMS like WordPress, Drupal has built-in support for responsive but not all solutions. Developers need to use different libraries or need to write design specific media queries in style sheets which are the extra line of codes.

Results of Survey 1 highlighted that “Number of pages or features” was topmost size and cost metrics for early web cost estimation and it’s validated in interviews and also in Survey 2. Results of Interviews highlighted that “Number of pages or features” was the top critical size and cost metric for early web cost estimation and it’s validated with results of Survey 2.

A similar research was conducted in 2002 by Mendes [4] and results indicated that two most commonly used for web cost estimation were:
- Total number of web pages
- Which features/functionality to be provided by the application

<table>
<thead>
<tr>
<th>Research 2002</th>
<th>Research 2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Total number of web pages</td>
<td>- Number of pages or features</td>
</tr>
<tr>
<td>- Which features/functionality to be provided by the application</td>
<td>- Responsive implementation</td>
</tr>
</tbody>
</table>

Table 9 – Research Comparison

1. Research results about top factor in 2017 are slightly different from the research results of 2002 about early web cost estimation. In 2002, there were “Total number of web pages” was the one of the top factor used for estimation while now in 2017 it is “Number of pages or features”. Most companies talk about features now instead of pages. Some companies ask about pages but most of the companies ask about the number of features.

2. In 2002, second highest factor was “which features/functionality to be provided by the application” while in 2017 it’s totally different and it is “Responsive implementation”. Responsive implementation is a development technique where the web application is designed in a way that can work on all devices, screen sizes, and all browsers. This approach does not exist in 2002 while now most of the clients ask
for responsive web application and it increases cost and effort for the web application.

3. There are the new set of metrics which are found in 2017 while those did not exist in 2002.

- **Search Engine Optimization (SEO)**
  SEO has become a high demand feature for web applications in this competitive era. SEO is a technique which optimized website contents and brings it in search results. There are different techniques within SEO i.e. On-page optimization and off-page optimization. Every company wants their product, website on the top of search engines in order to get more popularity and sales. Most of the companies use the standard set of SEO modules. Adding SEO to the website affects the size and cost of web applications.

- **Maintenance and support**
  Maintenance and support have become an integral part of web application development. Most of the companies add it as a primary part of the web application development cost estimate.

- **Standard pre-developed set of modules**
  Use of pre-developed standard modules is very common practice in web application development by web companies. Most of the common modules are available in the market to use as a purchased component or open source component for the web application. It reduced the cost and time to complete web application projects.

- **Social Media packages**
  Social Media packages integration is very popular now in web applications. Companies want their presence on social media. Social media can be Blogs, Facebook, Twitter, Instagram etc. Different integration packages are available to integrate it on web application and it needs extra effort to build. These packages have also become part of web application cost estimates.

- **Content Management Support (CMS)**
  Copywriting and managing contents along with modules handling are modern techniques in web applications. Some companies add these as part of web application development.
Recent research shows some new methods and models to measure web cost estimation at an early stage. Following is comparison of new methods with my research results:

1. **Functional Size Measurement (FSM) [56]**
   This method can be used to measure early web effort and cost estimation. FSM is a concept that how to measure the size of software considering functional requirements required by the customer. Line of codes (LOC) can be calculated and an estimate can be provided. It’s not suitable for large-scale applications. In my research results, there are metrics which is completely based on the functional size i.e., “number of pages/features”, “responsive implementation”, “general features”, “dynamic features implementation”.

2. **Common Software Measurement International Consortium (COSMIC) [57]**
   FSM is not perfect model as it showed issues during a case study while measuring effort for web development. COSMIC is useful to get a better estimate for large-scale web applications but very specific to nature of the projects. COSMIC-FFP is an enhanced method to measure a functional size of the software to estimate the size of the software applications.

3. **Multi-Stage Effort Estimation [60]**
   Estimation can be done on each stage of software development phase but it will take resources. This approach is not suitable for early web cost estimation.

4. **Page-wise effort estimation [61]**
   An approach is suggested that calculate an estimate for each web page. First, check that how many pages needed for the website and then measure the estimated hours for each page. Functions can be calculated and a rough estimate can be created by multiplying a static complexity factor. Summing up all web pages estimate will give a final rough estimate for early web cost estimation. This approach can only work for basic web applications. In my research results, there is one core metric is “total number of pages and features” which is being used to measure early web cost estimation along many other size metrics.

5. **Functional Measurement Type [62]**
   The effort is being calculated by doing clustering based on functional requirements, sub-functional measurement and complex calculations of the web applications. My research results also show more focus on functional size metrics where developers need to write functions and sub-functions in order to make a functional web application.

6. **WebMO+ & VPM+ [63]**
   Both models WebMO+ and VPM+ are suitable for web application effort estimates. Web Objects (template, buttons etc) is a good way to list down all the objects of the web application and then calculate effort of the objects. Many size metrics are proposed for Web applications like Application points, Object points, and Multimedia points but Web objects (WO) is more appropriate. Web object calculation sheet consists of following:
   - internal logical files
   - external interface files
   - external inputs
   - external outputs and external inquiries
   - number of HTML, XML and query lines
   - number of multimedia files
   - number of scripts
- number of web building blocks

Usage of each Web Object predictors is used in a scale like Low, Average and High and effort is calculated. Effort estimation can be calculated of past projects using both models.

If we look into online quote forms on company websites then there are only few objects from above list are being asked while quote forms ask many other metrics which are concluded in my research. Logical files can be related to dynamic functions implementation in my research. The number of HTML, XML and query lines is equal to the number of pages and features in my research. My research also shows the number of multimedia files (static, dynamic graphics, flash etc) which is equal to above web object “number of multimedia files”. Number of Web building blocks in this research is related to my research i.e. “number of public features”, “number of additional features”. Refer to table 7.0 for further detail.

7. Case-based reasoning and Expert Judgement [64][65][66]

Studies show that case-based reasoning is still the best approach for early Web cost estimation when there is a lot of projects data available. Expert judgment approach is quick to give effort estimate. Another model FHSWebEE is also suggested which consists of OOMFPWeb, web metrics for web application and functional size measurement. Another analogy approach EffortEST is suggested after comparing three analogy software (ANGEL, ACE and ESTOR). All these software use projects data to provide estimate using old projects data. I also use case-based reasoning approach using real web projects data in my research. I used FreeCBR tool which is good and easy to use.

There are also algorithmic models and machine learning methods available for web effort estimation [62]. Most of the model and methods depends upon the large set of web projects data as I mentioned above. There is no any repository available where data of web projects can be collected and that’s why Tukutuku project was initiated to store project specific metrics details. Many newly created methods have limitations and difficult to get a precise estimate for early web cost estimation. Most of the approaches which have been suggested for web size estimation using functional points have limitations, it has dependencies on real implementation decisions, not according to standards or not according to latest technological trends [59].

Prediction Technique
Data collected on projects had fixed data and also some open-ended responses as well. Dataset included both collinearity and outliers. Case-based reasoning is considered as the best approach to use for this type of dataset. Stepwise Regression technique can be used to benchmark the compared results [29]. Stepwise regression technique is a statistical approach where a prediction model is built and characterizes the relationship among independent (i.e. the number of web pages) and dependent variables (i.e. total effort) [33].

This approach builds the model by adding at each stage, the independent variables with the association of dependent variables. The main purpose is to find set of independent variables (predictors) which explains the variation independent variables [32]. Stepwise regression technique is widely used as the benchmark and considered as good prediction technique [29]. I added all 29 projects data into SPSS and executed Step-Wise regression analysis. Data as dependent variable needs to be in numeric type while my projects data has the mixed type like ordinal, ratio and numerical. CBR was best technique which was suitable in my case.
8 \textbf{Validity Threats}

Predictor variables were based on my survey results which were collected via survey form to gather data on projects worldwide. I used the predictors which are more useful or meaningful for the web companies. Data was collected on given input and it was based on company’s own methods. There was no any other solution used to gather data on companies web projects. Predictors are documented which are received from companies. There is a risk that some company may provide incomplete data.

Different companies have different approaches to their web projects. There can be different size and teams for the projects. Estimate depends upon the nature of the project and some internal and external factors. Some companies had their estimates added in planning phase while some just added on the base of their meetings or weekly plans. Data on web projects were gathered from worldwide and there was no quality check was performed on projects provided data. There were different domains of the applications and it was difficult to generalize the results.

There can be different kinds of validity threats:
   a) Internal Validity: “Internal validity is estimated reality about inferences concerning casual relationships” [17]

   b) External Validity: “It is related to amount to which conclusion in our studies would hold for another person in other person places and at other times” [17]

   c) Construct Validity: “Construct Validity is a generalization of our measurement to concept four measures” [17]

   d) Conclusion Validity: “The statistically significant relationship between treatment and outcome is related to conclusion validity” [17]

8.1 \textbf{Internal Validity}

8.1.1 Selection of quote form
The internal validity threat in Survey S1 was missing of any important web quote form online. Various strategies were considered in order to overcome this threat. The first step was the selection of the search engine, the selected search engines are ranked number one among all. Next was query formulation. Initially, various combinations were executed and finally with the help of supervisor final query was designed. There were thousands of URLs available on the base of the search query and it was a risk to miss any important URL. To overcome this risk all forms were downloaded which were retrieved after query execution.

8.1.2 Interviewee Selection
Selection of interviewee was a threat to this study. Few of interviewees were selected from personal network and rest were identified online. Prior to interview, all the selected participant’s profiles were studies in order to check if they have knowledge and experience in particular domain in which research is being conducted. This pre-work approach reduced this risk.
8.1.3 Missing Information during Interview
There were many chances of missing some important information while conducting interviews. In order to overcome this, a strategy was decided before conducting each interview. Permission to record each interview was taken from each participant. All the interviews were transcribed immediately after each interview and were send to the same person for validation.

8.2 External Validity
8.2.1 Validity of Interview Participants
15 interviews were conducted from a participant which were from different countries. The limited time was a reason to conduct interviews with fewer participants. There were chances that generalization of data cannot be big enough which could be a possible risk. The survey was conducted to overcome this threat.

8.2.2 Validity of Survey Participants
An online survey was designed and was posted on various networking website, groups, and social media. There was a positive response which was received around the world. On the first page of survey, a briefed introduction was provided which help in targeting the accurate participant.

8.3 Construct Validity
8.3.1 Survey 1
During survey 1 there was a chance of missing an important factor. To overcome this risk an excel sheet was created in which data from each and every single quote form was written.

8.3.2 Interview Data Analysis
This might be a threat to analysis of data as there can be issues in collection and analysis of data. To overcome this transcribed interview was analyzed using qualitative data analysis method. This approach helped in reducing risk.

8.4 Conclusion Validity
Conclusion validity is related to results obtained from the study which ensures results reliability [18] [19]. This threat was kept in noticed at the beginning of this research. A survey was conducted at the initial level of study. To validate the result of survey an interview was conducted. The results which were achieved after interviews were used to create a quote form and validate interview results.
9 CONCLUSION AND FUTURE WORK

This research investigated metrics which can be used for early Web cost estimation at the early stage of Web application development. This is the stage where the application is not built yet but just requirements are being collected and an expected cost estimation is being evaluated.

Chapter 4 explains that Survey 1 was conducted and list of size metrics was gathered. This was a huge survey and 212 Web project price quote form was gathered in order to get the list of size and cost metrics from the Web companies around the world. The outcome of this section is the final list of size metrics and cost factors for Web cost estimation. The sample population was the Web companies that use Web forms to give price quotes for Web application projects. Occurrences of each metrics were also measured and identified top rated metrics and it was “Total number of Pages/features”.

Chapter 5 explains that fifteen Industrial interviews were conducted on the base of the list of metrics generated by Survey 1. Different Web companies were selected to have interviews in order to validate results of Survey 1. Interviews results were transcribed into the more readable format. Results of industrial interviews were used to validate results of Survey 1. Results of Industrial interviews also leads toward making a final set of size metrics which can be part of Tukutuku project. The most used metric in Interviews results according to ranking was “Total number of Pages/features”.

In Chapter 6, a quote form was designed on the base of the results of Interviews. An online service was used to gather data on real Web projects. 29 real projects data were collected from different Web companies around the world. Data were analyzed by Case Base Reasoning (CBR) method. FreeCBR tool was used to do the measurements and data analysis. “Total number of pages/features” and “Responsive implementation” are the top metrics. If there are more number of pages or features, then cost and time estimate will increase. Similarly, there will be the difference in size and cost if a company built a responsive web application and it will increase cost and time estimates.

Direction for future work

- Upgrade Tukutuku database with new metrics. It should have all latest set of metrics which are being used these days by Web development companies.
- Make Tukutuku database available online for researchers so that anyone can access it and use it for research.
- Get more real web projects data from Web companies to build the large dataset of web projects that can help Web companies to estimate cost. More and more projects data can be helpful for cost estimation and effort comparison.
- Find the early size and cost metrics for desktop and native mobile apps. Do the comparison of these metrics along with Web metrics and find out the differences. Can there be a unique set of metrics which can be common in Web, Desktop, and native mobile apps? How cost estimation can vary from Web to Desktop and native mobile apps.
REFERENCES


[38] T.C. Hooi, Y. Yusoff, Z. Hassan, Comparative study on applicability of WEBMO in Web application cost estimation within Klang Valley in Malaysia, Computer and Information Technology Workshops, CIT Workshops 2008, IEEE 8th International Conference on, pp.116-121, July 2008


# APPENDIX

## Interviews Results Table

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Int1</th>
<th>Int2</th>
<th>Int3</th>
<th>Int4</th>
<th>Int5</th>
<th>Int6</th>
<th>Int7</th>
<th>Int8</th>
<th>Int9</th>
<th>Int10</th>
<th>Int11</th>
<th>Int12</th>
<th>Int13</th>
<th>Int14</th>
<th>Int15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Static graphics contents</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Dynamic graphics contents</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>No. of pages</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Public features</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Layout and functionality</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>New or existing application</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Additional features</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Responsive support</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Social media package</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Search engine optimization</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Web/Domain/Cloud hosting</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Metrics for ecommerce</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Maintenance and support</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Marketing &amp; promotion</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Web content management</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

1 = Yes  
0 = No  
2 = May be/both/depends

Table 10 – Interview Results
Industrial Interviews transcribing using Xmind
Investigating Web Size Metrics for Early Web Cost Estimation

The aim of this research is to investigate size metrics and cost drivers used by Web companies for early Web cost estimation.

This survey will take only 5-10 min and you will get detailed research results regarding Web size metrics for early web cost estimation after analysis.

Note: This is an academic research project and all of your feedback will be used in research. The analysis of this research will be presented as thesis research in the Blekinge Institute of Technology (www.bth.se)

If you have any questions then you can reach me at saao10@student.bth.se

What is your company name?

What is your email address?
What is the country name where your company is established?

Type a sentence

250 characters remaining

What kind of Web project where you have worked?

Type a paragraph

1900 characters remaining

How many years of experience in Web application development?

- 0-3 years
- 3-6 years
- 6-9 years
- 9+ years
How many years of experience in Web Effort Estimation?

- 0-3 years
- 3-6 years
- 6-9 years
- 9+ years

How many developers worked on the Web application?

- 1-3 developers
- 4-7 developers
- 7+ developers

How many Web pages (static pages only) implemented in the Web application?

Type a sentence

How many Web features (dynamic pages only) implemented in the Web application?

Type a sentence

What devices are supported for Web Responsive Design (WRD)? RWD is a technique where Web design are created in order to provide optimal viewing and interaction experience on different devices.

- Mobile
- Pad
- Tablet
- Desktop PC
- Other

Type your answer
Is this Web application needed to have maintenance and support?
- Yes
- No

Are the static graphics contents (logo, photo, flyer, and brochure) added in the Web application?
- Yes
- No

How many pre-developed standard set of modules are used during Web application development?
- 0
- 1-5
- 6-10
- 10+

Is it the new Web application development or maintenance of existing Web application?
- New Web application
- Enhancement of existing Web application

What kind of Search Engine Optimization features added in the Web application?
- On page optimization
- Off page optimization
- Both
- None
16. How many dynamic contents (include multimedia, flash, interactive dynamic forms, widgets) developed for Web application?
   - 0
   - 1-5
   - 6-10
   - 10+

17. What artifacts are pre-developed from third party company and hand over to Web company (You)?
   - Graphical User Interface
   - Database Structure
   - Flow Diagrams
   - Documentation
   - Other
     Type your answer

18. What Social Media packages integration done in the Web application?
   - Facebook
   - Twitter
   - Google+
   - YouTube Channel
   - Instagrams
   - SnapChat
   - Blog
   - Other
     Type your answer
What kind of contents management support provided in the Web application?

- CMS Training
- Contents Development
- Technical Writing
- Contents proof reading
- Graphical contents
- Products portfolio
- Other
  - Type your answer

What additional features are implemented in the Web application?

- Security
- Performance
- User Interface
- Other
  - Type your answer

If you have used any other size and cost metric during the Web application development then you can mention here. Any comment or suggestion?

- Type a paragraph
Projects data collected via Quote Form

Q1 (Company name) and Q2 (email address) are confidential and not shared in this report but rest of the data is available.

<table>
<thead>
<tr>
<th>Q8</th>
<th>Q9</th>
<th>Q10</th>
<th>Q11</th>
<th>Q12</th>
<th>Q13</th>
<th>Q14</th>
<th>Q15</th>
<th>Q16</th>
<th>Q17</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sweden</td>
<td>0.3 years</td>
<td>0.3 years</td>
<td>3.3 Developers</td>
<td>1</td>
<td>5</td>
<td>0.001</td>
<td>Yes</td>
<td>Yes</td>
<td>1 to 5</td>
</tr>
<tr>
<td>Pakistan</td>
<td>0.3 years</td>
<td>0.3 years</td>
<td>3.3 Developers</td>
<td>12</td>
<td>2</td>
<td>111</td>
<td>Yes</td>
<td>Yes</td>
<td>1 to 5</td>
</tr>
<tr>
<td>USA</td>
<td>6.9 years</td>
<td>6.9 years</td>
<td>7 Developers</td>
<td>Depends on requirements, depends on</td>
<td>4001</td>
<td>Yes</td>
<td>Yes</td>
<td>6 to 10</td>
<td></td>
</tr>
<tr>
<td>India</td>
<td>0.3 years</td>
<td>0.3 years</td>
<td>4 Developers</td>
<td>Depends on requirements, depends on</td>
<td>111</td>
<td>Yes</td>
<td>Yes</td>
<td>1 to 5</td>
<td></td>
</tr>
<tr>
<td>USA</td>
<td>0.3 years</td>
<td>0.3 years</td>
<td>4.7 Developers</td>
<td>9</td>
<td>7</td>
<td>111</td>
<td>Yes</td>
<td>Yes</td>
<td>1 to 5</td>
</tr>
<tr>
<td>UAE</td>
<td>0.3 years</td>
<td>0.3 years</td>
<td>4.7 Developers</td>
<td>May be around 50</td>
<td>111</td>
<td>Yes</td>
<td>Yes</td>
<td>6 to 10</td>
<td></td>
</tr>
<tr>
<td>Sweden</td>
<td>6.9 years</td>
<td>6.9 years</td>
<td>3.3 Developers</td>
<td>5+</td>
<td>100+</td>
<td>Yes</td>
<td>Yes</td>
<td>Zero</td>
<td></td>
</tr>
<tr>
<td>Pakistan</td>
<td>3.6 years</td>
<td>3.6 years</td>
<td>3.3 Developers</td>
<td>3 to 6</td>
<td>100+</td>
<td>Yes</td>
<td>Yes</td>
<td>1 to 5</td>
<td></td>
</tr>
<tr>
<td>Pakistan</td>
<td>3.6 years</td>
<td>3.6 years</td>
<td>3.3 Developers</td>
<td>Many</td>
<td>Many</td>
<td>Yes</td>
<td>Yes</td>
<td>1 to 5</td>
<td></td>
</tr>
<tr>
<td>Pakistan</td>
<td>3.6 years</td>
<td>3.6 years</td>
<td>3.3 Developers</td>
<td>Zero</td>
<td>Depends</td>
<td>Yes</td>
<td>Yes</td>
<td>1 to 5</td>
<td></td>
</tr>
<tr>
<td>UAE</td>
<td>6.9 years</td>
<td>6.9 years</td>
<td>3.3 Developers</td>
<td>New</td>
<td>2 to 3</td>
<td>111</td>
<td>Yes</td>
<td>Yes</td>
<td>1 to 5</td>
</tr>
<tr>
<td>Sweden</td>
<td>6.9 years</td>
<td>6.9 years</td>
<td>3.3 Developers</td>
<td>20 Greater th</td>
<td>111</td>
<td>Yes</td>
<td>Yes</td>
<td>10+</td>
<td></td>
</tr>
<tr>
<td>Pakistan</td>
<td>3.6 years</td>
<td>3.6 years</td>
<td>3.3 Developers</td>
<td>It varies requirements normally 2- Variar as y</td>
<td>111</td>
<td>Yes</td>
<td>Yes</td>
<td>1 to 5</td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td>6.9 years</td>
<td>3.6 years</td>
<td>4.7 Developers</td>
<td>5+</td>
<td>11+</td>
<td>111, website w</td>
<td>Yes</td>
<td>Yes</td>
<td>6 to 10</td>
</tr>
<tr>
<td>Sweden</td>
<td>6.9 years</td>
<td>3.6 years</td>
<td>4.7 Developers</td>
<td>Single page applications</td>
<td>Multiple</td>
<td>111</td>
<td>Yes</td>
<td>Yes</td>
<td>1 to 5</td>
</tr>
<tr>
<td>Pakistan</td>
<td>3.6 years</td>
<td>3.6 years</td>
<td>3.3 Developers</td>
<td>10</td>
<td>200</td>
<td>111</td>
<td>Yes</td>
<td>Yes</td>
<td>6 to 10</td>
</tr>
<tr>
<td>Pakistan</td>
<td>0.3 years</td>
<td>0.3 years</td>
<td>4 Developers</td>
<td>100+</td>
<td>100+</td>
<td>Yes</td>
<td>Yes</td>
<td>1 to 5</td>
<td></td>
</tr>
<tr>
<td>Pakistan</td>
<td>6.9 years</td>
<td>3.4 years</td>
<td>4.7 Developers</td>
<td>120</td>
<td>10a to 50</td>
<td>111</td>
<td>Yes</td>
<td>Yes</td>
<td>1 to 5</td>
</tr>
<tr>
<td>Sweden</td>
<td>3.6 years</td>
<td>3.6 years</td>
<td>3.7 Developers</td>
<td>30 to 50</td>
<td>10a to 50</td>
<td>111</td>
<td>Yes</td>
<td>Yes</td>
<td>1 to 5</td>
</tr>
<tr>
<td>Pakistan</td>
<td>9+</td>
<td>6.9 years</td>
<td>4.7 Developers</td>
<td>We prefer not to use static pages, Database</td>
<td>111</td>
<td>Yes</td>
<td>Yes</td>
<td>1 to 5</td>
<td></td>
</tr>
<tr>
<td>Pakistan</td>
<td>9+</td>
<td>6.9 years</td>
<td>3.3 Developers</td>
<td>1 to 5</td>
<td>3 to 10</td>
<td>111</td>
<td>Yes</td>
<td>Yes</td>
<td>6 to 10</td>
</tr>
<tr>
<td>Pakistan</td>
<td>0.3 years</td>
<td>0.3 years</td>
<td>3.3 Developers</td>
<td>1 to 5</td>
<td>1 to 5</td>
<td>0.001</td>
<td>Yes</td>
<td>Yes</td>
<td>0</td>
</tr>
<tr>
<td>Pakistan</td>
<td>9+ years</td>
<td>3.6 years</td>
<td>3.7 Developers</td>
<td>1 to 5</td>
<td>10 to 20</td>
<td>0.001</td>
<td>Yes</td>
<td>Yes</td>
<td>6 to 10</td>
</tr>
<tr>
<td>Pakistan</td>
<td>3.6 years</td>
<td>3.6 years</td>
<td>3.3 Developers</td>
<td>1 to 5</td>
<td>5 to 15</td>
<td>0.001</td>
<td>Yes</td>
<td>Yes</td>
<td>6 to 10</td>
</tr>
<tr>
<td>Sweden</td>
<td>9+ years</td>
<td>6.9 years</td>
<td>3.3 Developers</td>
<td>5 to 10</td>
<td>10 to 100</td>
<td>111</td>
<td>Yes</td>
<td>Yes</td>
<td>1 to 5</td>
</tr>
<tr>
<td>Sweden</td>
<td>6.9 years</td>
<td>3.6 years</td>
<td>4.7 Developers</td>
<td>Average 10-15</td>
<td>25-30</td>
<td>111</td>
<td>Yes</td>
<td>Yes</td>
<td>10+</td>
</tr>
<tr>
<td>Australia</td>
<td>9+ years</td>
<td>9+ years</td>
<td>3.3 Developers</td>
<td>10+</td>
<td>up to 80</td>
<td>111</td>
<td>Yes</td>
<td>Yes</td>
<td>0</td>
</tr>
<tr>
<td>Australia</td>
<td>9+ years</td>
<td>3.6 years</td>
<td>3.3 Developers</td>
<td>5</td>
<td>10</td>
<td>111</td>
<td>Yes</td>
<td>Yes</td>
<td>1 to 5</td>
</tr>
<tr>
<td>Australia</td>
<td>9+ years</td>
<td>9+ years</td>
<td>3.3 Developers</td>
<td>11 to 2</td>
<td>111</td>
<td>Yes</td>
<td>Yes</td>
<td>10+</td>
<td></td>
</tr>
<tr>
<td>Q18</td>
<td>Q19</td>
<td>Q20</td>
<td>Q21</td>
<td>Q24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000</td>
<td>100000</td>
<td>111</td>
<td>None</td>
<td>Web application for laboratory inventory management</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000</td>
<td>100000</td>
<td>100</td>
<td>None</td>
<td>Web Application</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10000000</td>
<td>100000</td>
<td>1001</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11135201</td>
<td>1000101</td>
<td>111</td>
<td>Web</td>
<td>We used third party API’s. We have worked on CMS - Wordpress, Drupal, Magento and Framework - CodeIgniter, Symphony 2 based projects for clients across the globe.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11135200</td>
<td>100101</td>
<td>111</td>
<td>A good understanding of hollywood entertainment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11135100</td>
<td>111111</td>
<td>111</td>
<td>No suggestions</td>
<td>Mobile applications and mobile-apps</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000</td>
<td>100000</td>
<td>1100</td>
<td>Yes</td>
<td>it was based on the CBS Payment solutions, license handling, web security, content management etc</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000</td>
<td>100000</td>
<td>1101</td>
<td>Nothing special</td>
<td>SharePoint and Bit differs integration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11130000</td>
<td>100010</td>
<td>110</td>
<td>No</td>
<td>Online Shopping Cart</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000</td>
<td>100010</td>
<td>1100</td>
<td>CMM Level</td>
<td>B2B Cloud, Cloud and Integration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10000000</td>
<td>100101</td>
<td>111</td>
<td>No</td>
<td>Business intelligence web projects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11131001</td>
<td>100101</td>
<td>111</td>
<td>UI design process, such as e-commerce, ux design and building a multi-language website for 4 markets.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11130001</td>
<td>111101</td>
<td>111</td>
<td>NA</td>
<td>B2B Web sites, Dynamic web pages and static sites, Enterprise web applications</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000</td>
<td>000000</td>
<td>1100</td>
<td>Story point estimations, M Different ones in my current project. Web platform for pricing analytics and web platform for near real-time communication of events in a financial market.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11101111</td>
<td>100101</td>
<td>111</td>
<td>I have not reused any other Mobile Applications</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10000000</td>
<td>000000</td>
<td>1101</td>
<td>Survey questions were am Selling applications</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000</td>
<td>000000</td>
<td>1101</td>
<td>Responsive website design Electronic Health Records</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10000000</td>
<td>100010</td>
<td>110</td>
<td>Medical</td>
<td>Research employer branding</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11101001</td>
<td>110101</td>
<td>111</td>
<td>Each module is estimated We work on almost all web technologies including PHP, ASP.NET, Java, html, xml, css, responsive design, databases including mysql and SQL server, frameworks I</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11101001</td>
<td>110101</td>
<td>111</td>
<td>Web application about social media</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11101001</td>
<td>111101</td>
<td>111</td>
<td>Web application about education</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000</td>
<td>111101</td>
<td>111</td>
<td>Web application about health care nanotechnology</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000</td>
<td>000000</td>
<td>111</td>
<td>Web application for parental details</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11111001</td>
<td>111111</td>
<td>111</td>
<td>Web application related to ecommerce</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11101001</td>
<td>111001</td>
<td>111</td>
<td>CRM, Raiser’s, Employer Branding, Job portal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11100100</td>
<td>000000</td>
<td>110</td>
<td>UX Design</td>
<td>Client’s website design and development - Web Application</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10100100</td>
<td>001010</td>
<td>110</td>
<td>Penny Auction - Wordpress - Blogs classified mobile apps</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00000000</td>
<td>000000</td>
<td>110</td>
<td>Story point</td>
<td>Online games, online payment systems</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Tukutuku Benchmark website
I developed this website and deployed on my own hosting domain. Contents used in the websites were taken from old website which does not exist anymore.

http://glaxtech.com/tukutuku