



SMU

SINGAPORE MANAGEMENT
UNIVERSITY

School of
Information Systems

ANLY482 Analytics Practicum

Visualization of Consumer Satisfaction

Mid-Term Project Report

1st October 2014

Prepared By:

Mohamed Yousof Bin Shamsul Hameed | Kee Eng Sen

Table of Contents

1 Project Overview	3
1.1 Background.....	3
1.2 Motivation.....	3
1.3 Objective	3
2 Approach	4
2.1 Technology and Tools.....	5
2.2 Project Schedule.....	5
3 Development of Dashboard.....	6
3.1 Prototyping.....	6
3.1.2 Exploration of charts	7
4 Accomplishments and Deliverables	13
4.1 Challenges Faced and Reflections	14
5 Next Steps.....	15
5.1 Exploring Predictive Model.....	15
5.2 Deployment of dashboard	15
5.3 Final Dashboard.....	16
6 Appendix	17
6.1 Dataset.....	17
6.2 Exploratory Data Analysis	18

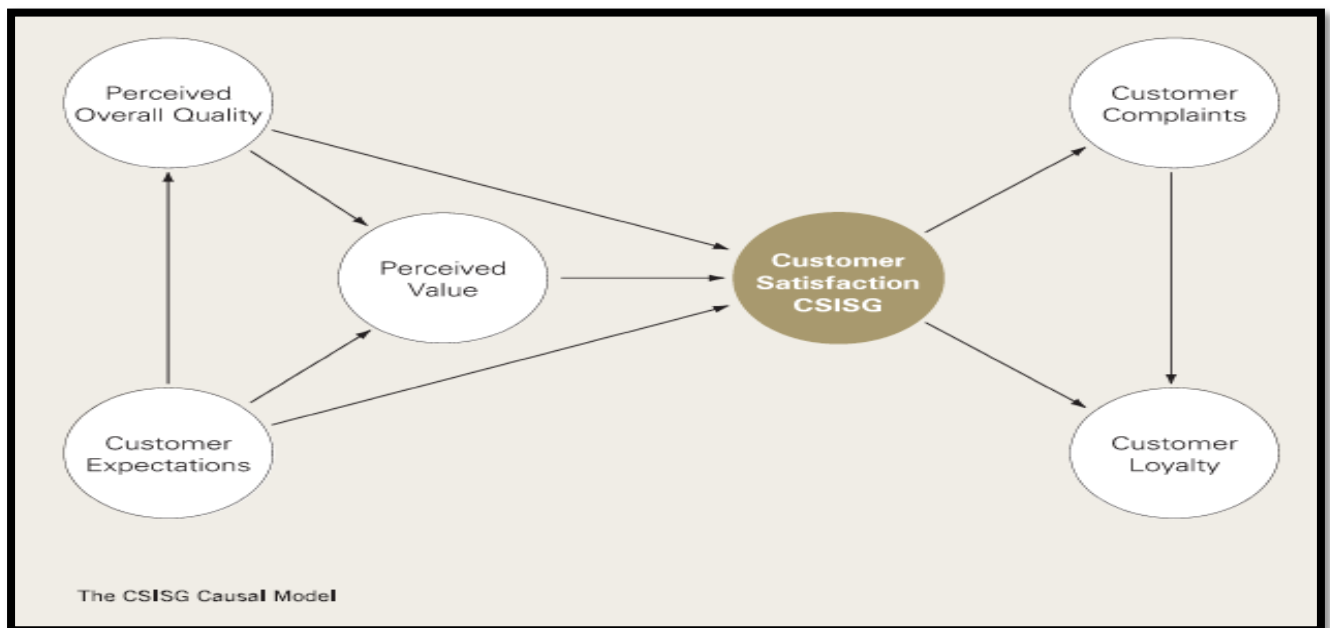
1 Project Overview

1.1 Background

Consumer research has been a hot topic. Businesses and government agencies are interested to know the satisfaction levels of Singaporean consumers and effectively take actions that can create valuable and meaningful impact in the society. This project explores these satisfaction levels. It uses the respondent level data from the Customer Satisfaction Index of Singapore (2008-2013) for the following sub-sectors:

- MRT System
- Public Buses
- Taxi Services
- Supermarkets
- Hotels
- Budget Airlines

The metrics that measure satisfaction includes:



Note: Extracted from the CSISG Brochure

1.2 Motivation

The motivation for doing this project was based on our interest in presenting the data in an interesting way that will help users and business leaders derive fresh insights from the data. Through the visualizations, we hope to illustrate customer satisfaction in various measures as shown in the CSISG Causal Model above to provide the users with a better understanding of the data.

1.3 Objective

The primary scope of this project is to produce an Interactive Dashboard that shows market trends and customer satisfaction visually. We will try to incorporate a 'fun' element into the

dashboard so that users and business leaders will be interested to play with it during events and roadshows organized by the Institute of Service Excellence.

The secondary scope of this project will be to deploy the interactive dashboard onto the client's preferred server and also create the documentation for the processing of the dataset for the interactive dashboard.

The tertiary scope of this project will be to automate the data formatting and processing of the datasets through backend codes for the dashboard. Finally, we will try and develop any predictive analysis models based on the datasets if possible.

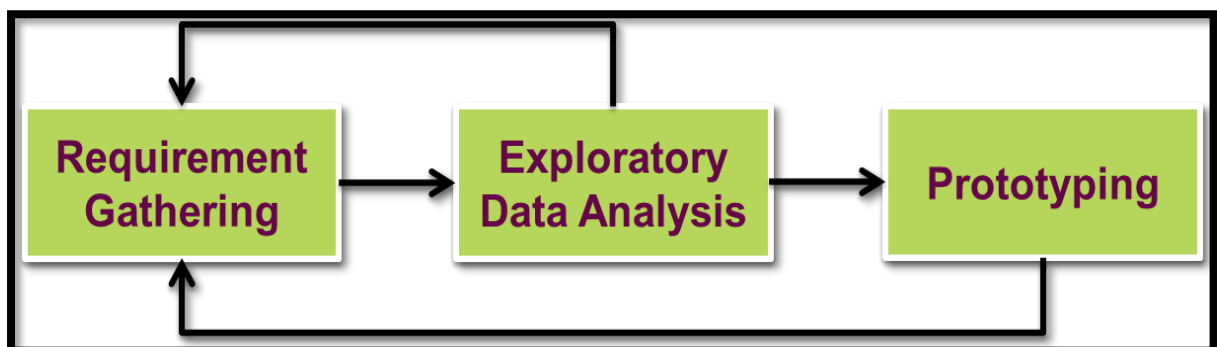
2 Approach

Our primary focus is building the Interactive Dashboard, relying mainly on d3.js as it provides us with the option of customizing our visualizations and charts. We will be taking up the SDLC Software Prototype Model to help us achieve our main objective.

We have conducted the Exploratory Data Analysis on the dataset given to us to explore the characteristics and discover any trends that are present in the data. We then tried to visualize the data using different charts to determine which chart is the best to show meaningful information.

After our initial analysis, we proceeded with the Prototyping of our Interactive Dashboard. The process of Prototyping will be explained more in detail in the report. This is an iterative process as we will continue to gather feedback from our supervisor and sponsor on the proposed dashboard and then come back to the drawing board to enhance our dashboard to suit our sponsor's needs.

Once the Prototyping process is completed, we will then proceed to finalize our Interactive Dashboard.



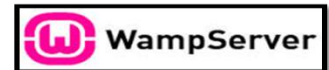
2.1 Technology and Tools

dc.js – dimensional charting

Crossfilter



D3-tip by Caged - LabRatRevenge



2.2 Project Schedule

Weeks 1-2

- Kick-off meeting
- d3.js and javascript

Week 4

- EDA
- Project Proposal Due

Week 7-8

- Mid-Term Presentation
- Mid-Term Report

Week 13-14

- Final Project Presentation
- Final Project Report

Week 3

- Receive project data
- Understanding of data and initial EDA

Week 5-6

- Development of prototypes for dashboard
- Exploration of other tools and technologies

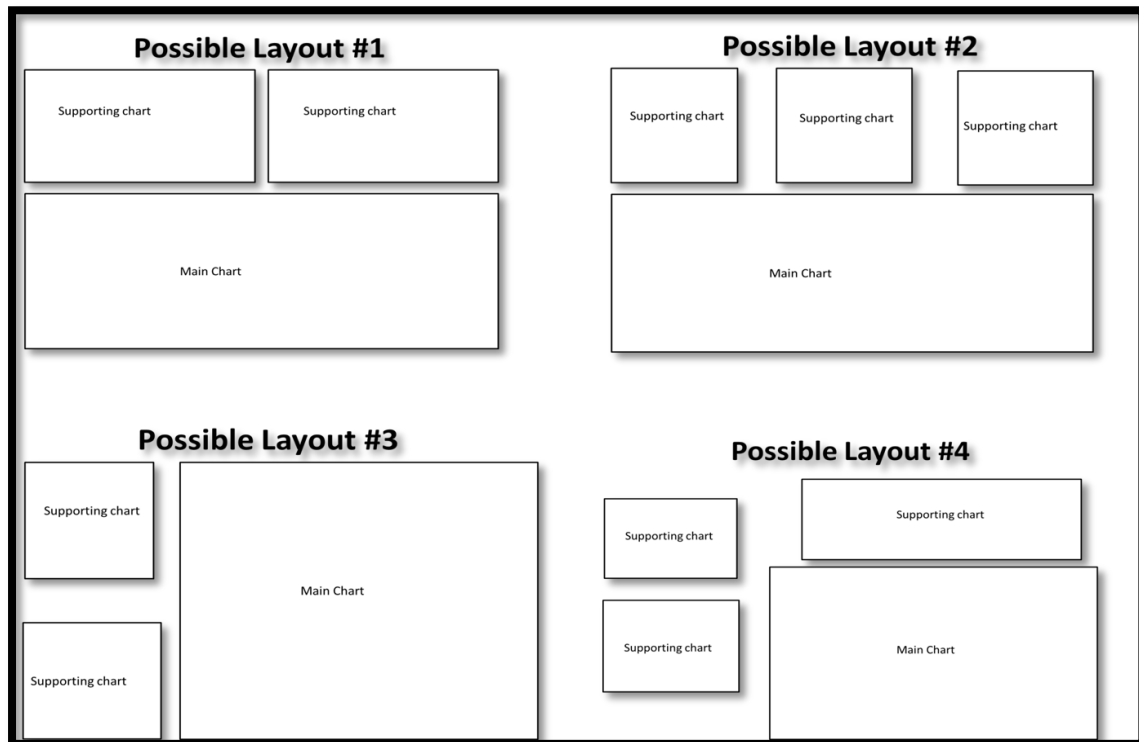
Week 9-12

- Finalize development of dashboard
- Deployment of Dashboard

3 Development of Dashboard

3.1 Prototyping

Possible Layout of Dashboard

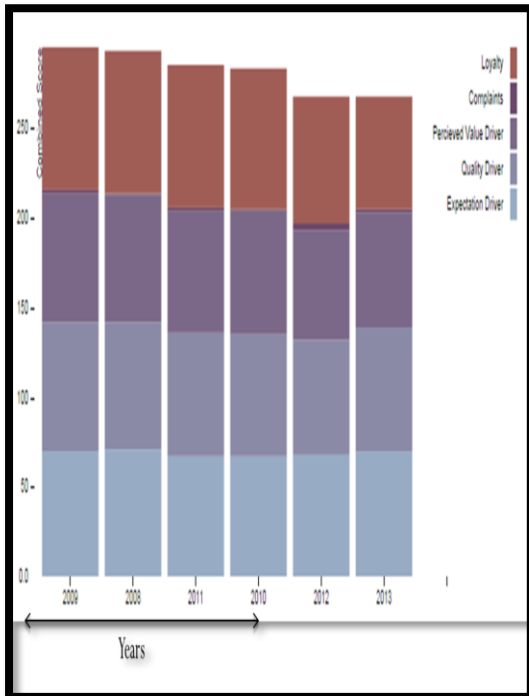


Several possible layouts of the dashboard

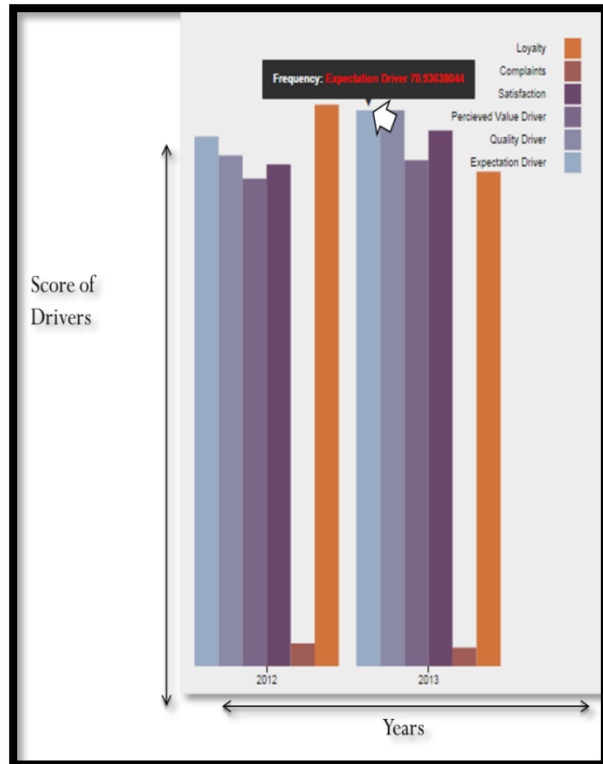
Before we began our development of the dashboard, we began to explore several possible layouts for the final interactive dashboard, as it will help use better visualize which charts might be more appropriate in which area of the dashboard. With the layout of the dashboard in mind, we began to develop various charts using d3.js with the dataset available, and then look into the possibility of combining the various charts into a final interactive dashboard based on the layouts.

3.1.2 Exploration of charts

Development of Bar Charts



Initial Stacked Bar Chart



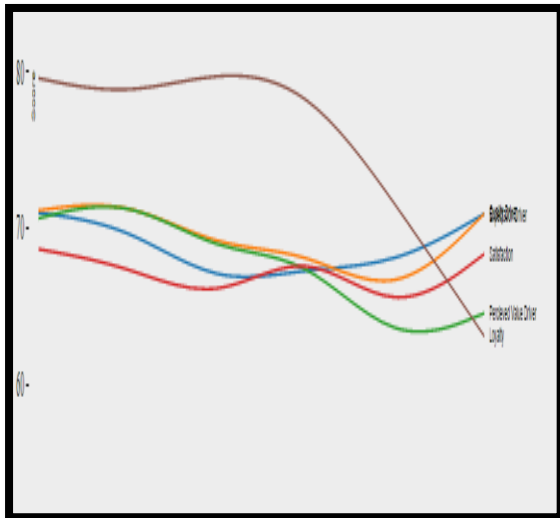
Individual Bar Charts with tooltip

As we were exploring how to develop charts using d3.js, we started out with the most basic chart with the Stacked Bar Chart. Based on the data we received, the Stacked Bar Chart was not a good visualization of the trend of the scores over time.

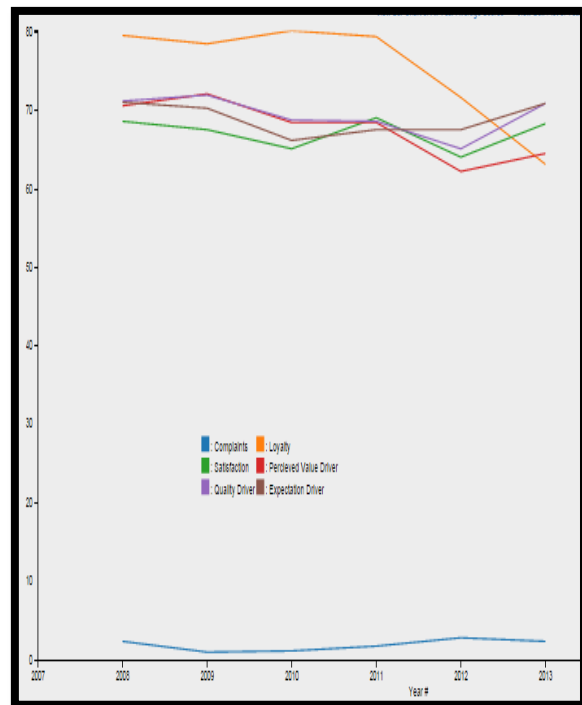
Next, we proceeded to separate each driver and placed them into individual bars, thus showing the average scores for each driver per year. We also added a tooltip feature such that the value of each bar will be shown when a user mouse over on the different bars.

However, after further exploration, we realized that bar charts do not show the trends of the scores effectively, as it is difficult to show how the scores of a specific driver changes year-on-year when they are not side-by-side. Moreover, it is difficult to gauge the height of the bars visually. After further consultation with our sponsor, we decide to remove this bar chart which shows the 6 drivers together.

Development of Line Charts



Smooth Line Chart with individual lines

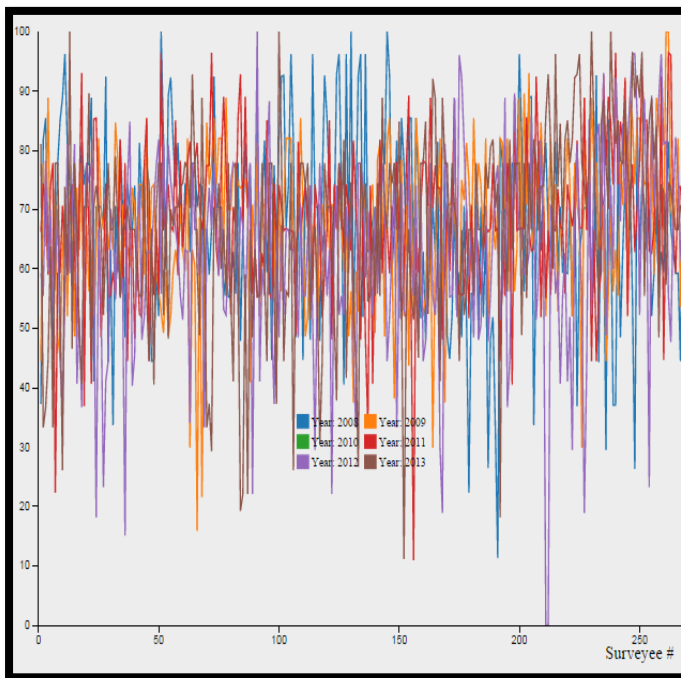


Straighthen Line Chart

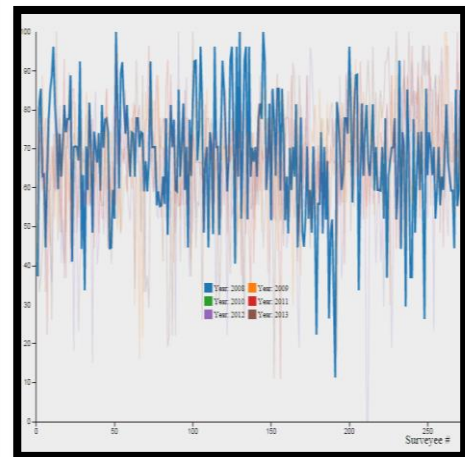
We then developed line chart showing all the drivers over the years. We were only able to develop a line chart with smoothen lines at first, as we were still trying to figure out the correct codes required to straighten the lines. As the smoothened line chart is not very visually appealing, we figured out ways to straighten the lines to better show the trends from each year. We also shifted the labels to the empty space for users to identify the lines better.

Although our sponsor appreciated the line chart, he feels that each driver should have a line chart of its own, instead of plotting all in a chart. He proposed placing the 6 line charts together in the final dashboard, with 3 line charts in each row, for users to visualize the year-on-year change based on the demographic attributes filtered.

Development of Multi-Linear Chart



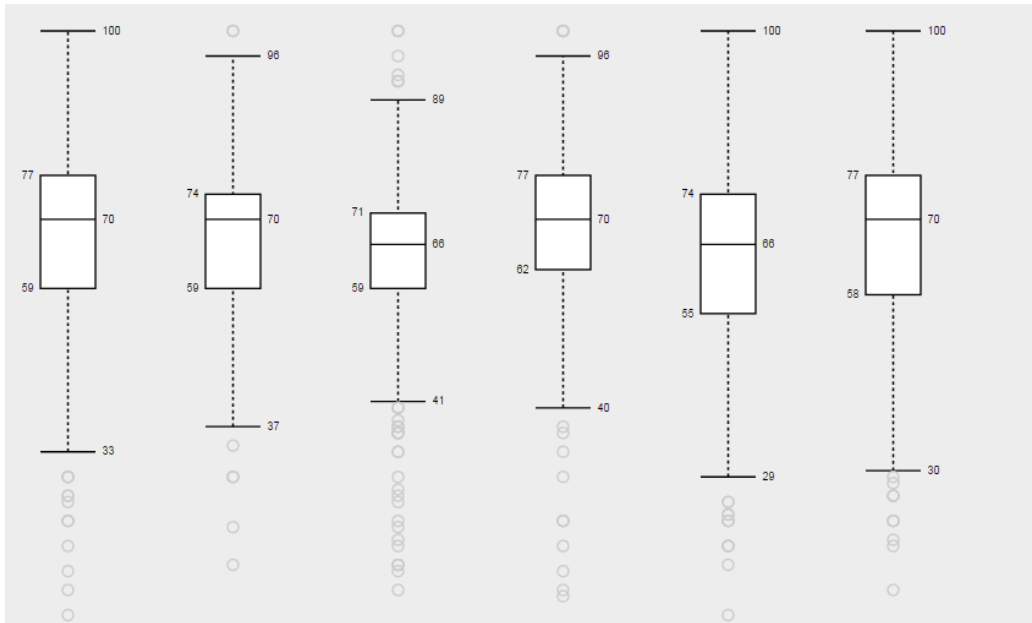
Multi-Linear Chart for different years



Highlighting of a single year of scores

We also developed a multi-linear chart for the Satisfaction driver for all the survey respondents. Our initial idea was to discover any trend or pattern available from plotting this chart; however, we realized that this type of chart is best used for continuous time series data. The x-axis in the chart is actually the sequence of the survey respondents from the data set, which explains the big variations in the scores. This is not very useful as we cannot derive any insights from the chart, even though we separated each year's data with a different colored line and allowed highlighting. We concluded that this chart would not be useful to our sponsor and proceeded to remove it entirely.

Development of Box Plot

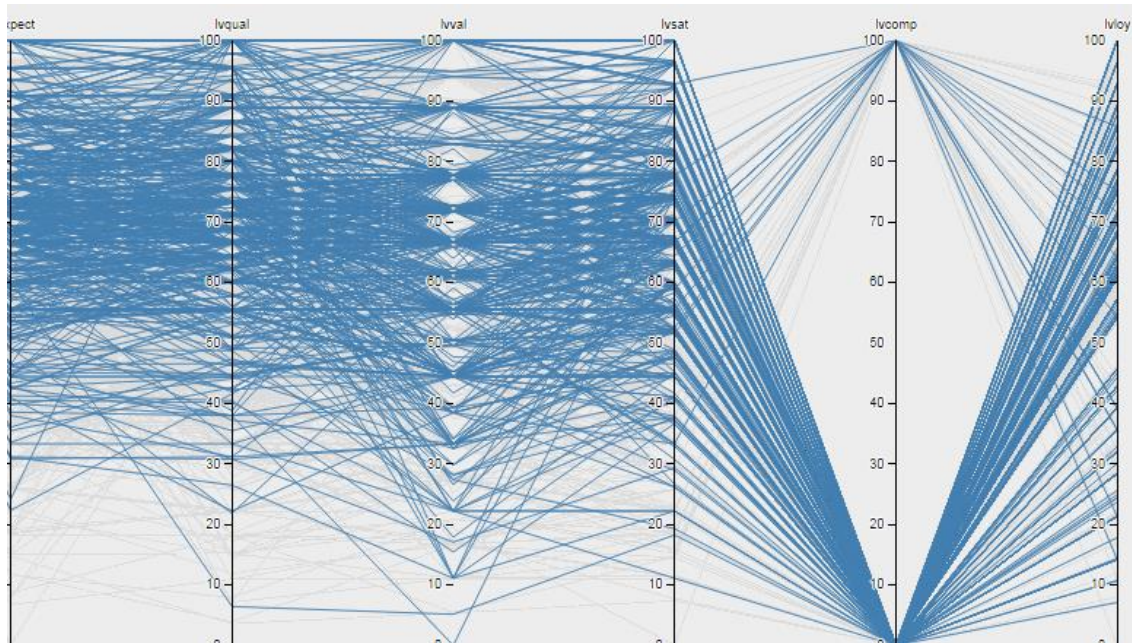


Box plots of Satisfaction Scores by year

During our analysis of the data, we felt that it would be useful to visualize some statistics of the Satisfaction driver using Box Plot, as it shows the min, max, median and inter-quartile range. Then we further broke it down by year, and we were able to visualize the trend of the Satisfaction level of the respondents.

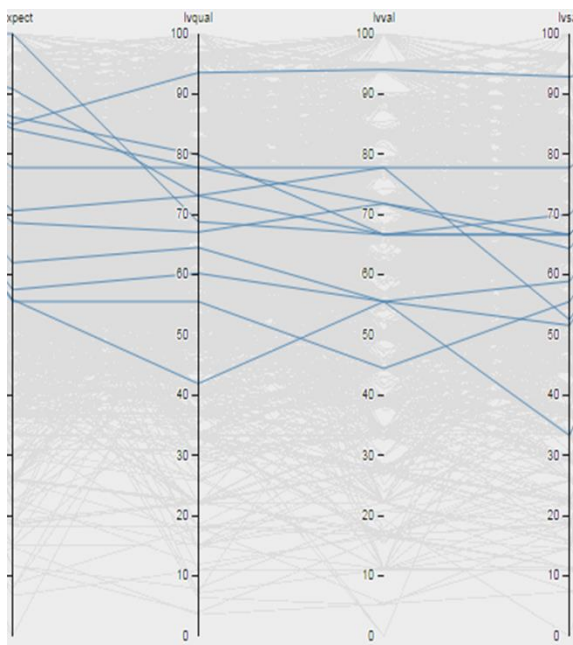
Although we were confident that the Box Plots will be useful, upon further consultation with our client, we realized that Box Plots are not easily understandable by the average person who is not trained in statistics, or not familiar with the terms. As we do not know whether our target audience understands Box Plots, it was advisable not to include it in our final dashboard. Through this, we understood the concept that visualizations should be simple enough for the average user to understand what story it is painting just from the first glance.

Development of Parallel Coordinates Chart



Parallel Coordinates Chart for the different drivers

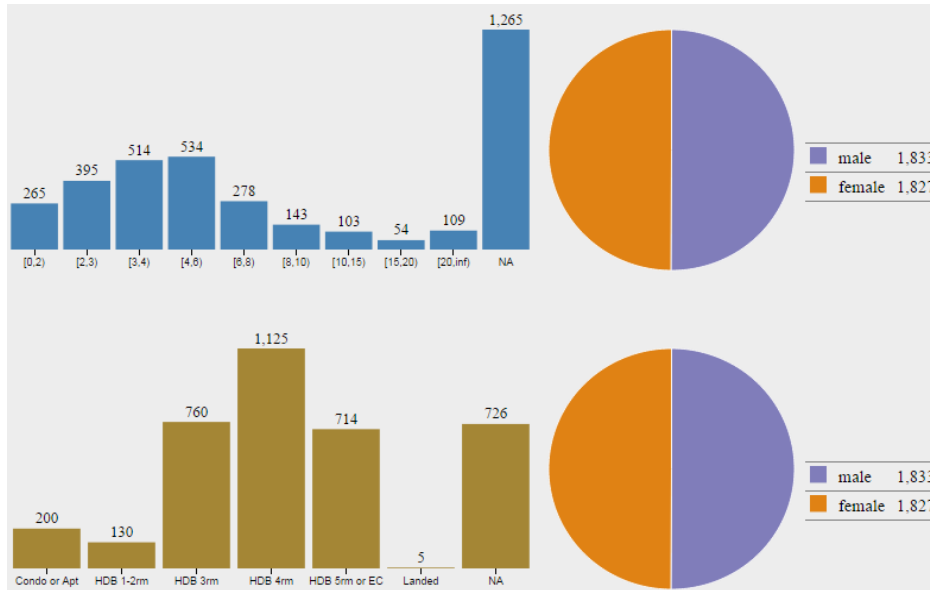
After going through several charts that might be useful to visualize high dimensional data, we decided to incorporate the Parallel Coordinates Chart. At first glance, this chart may be very confusing and cluttered, as each line represents each row of survey respondent's data. To overcome this issue, this chart allows users to filter the lines to be shown based on the value range on each y-axis. Each y-axis consists of the value driver.



Highlighting of specific range of values

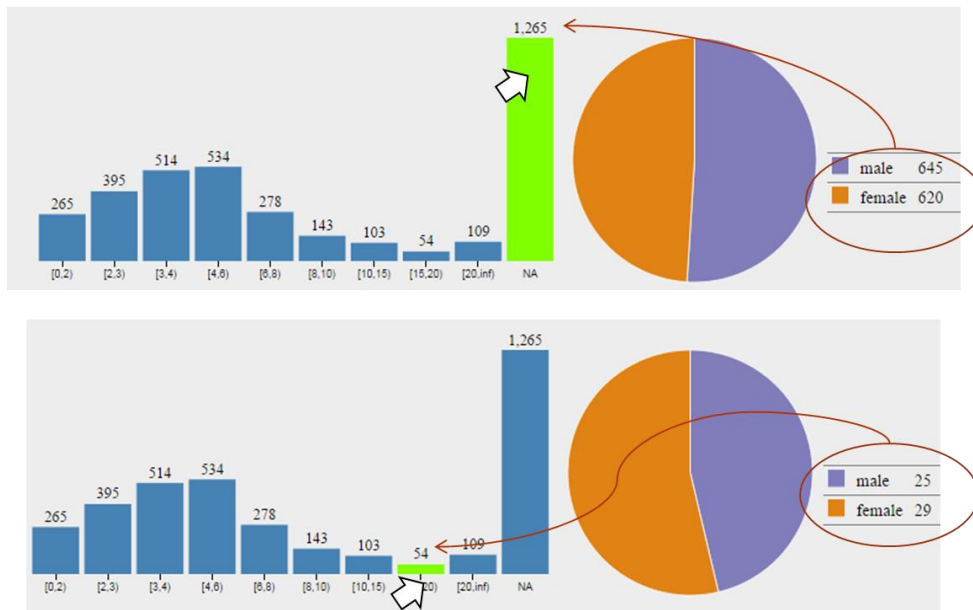
Although we felt that this will be a very useful chart, our supervisor and sponsor did not think likewise, as this chart is not easily understandable. Users may not appreciate the overview of the data. We decided to remove this chart from our final dashboard.

Initial Development of Interactive Dashboard



Interactive dashboard for demographic

After several attempts at developing simple charts with d3.js, we proceed to develop interactive charts for the demographic variables, as a stepping stone for our final dashboard. The initial idea of this simple interactive chart consists of a Histogram and a Pie Chart. Both charts can be interchangeable with the demographic variables in the dataset.



How the pie chart changes with the highlighting of the histogram

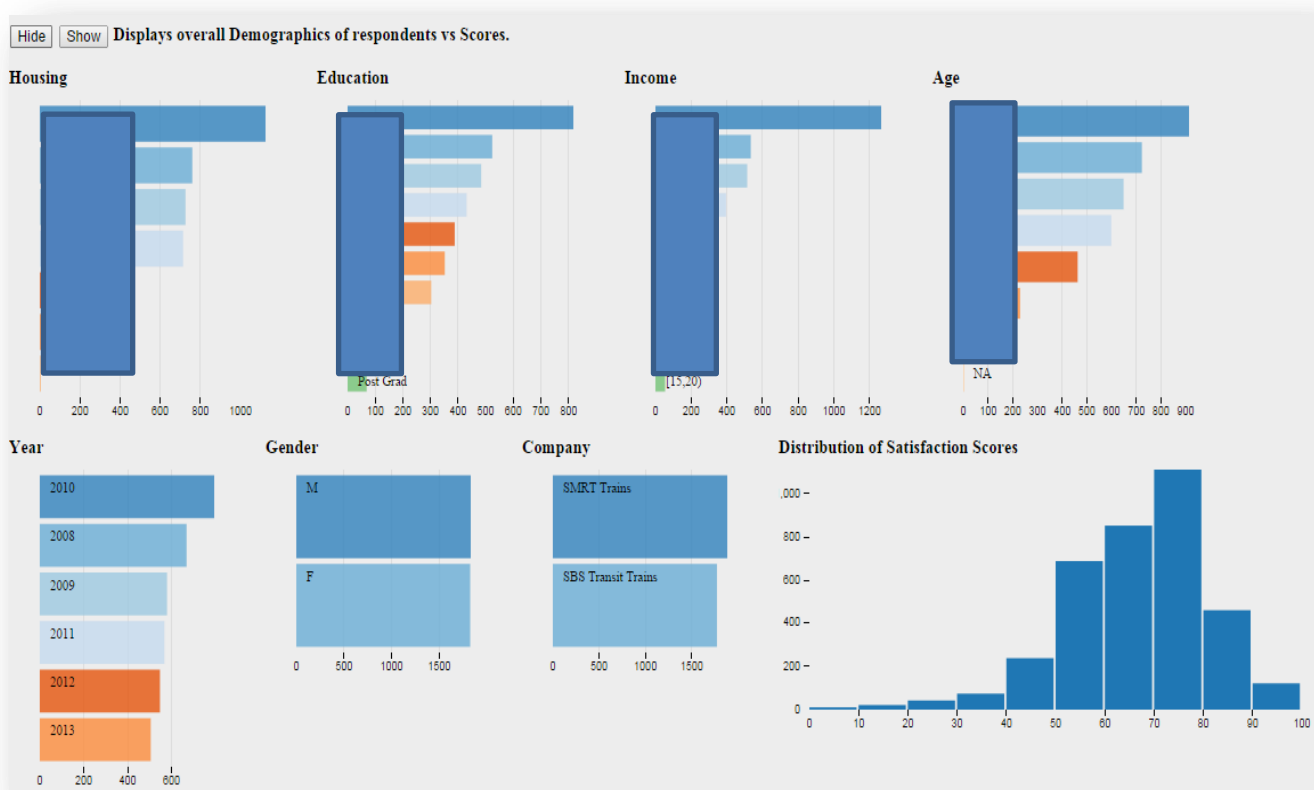
By moving the mouse over to a bar or the pie chart, the chart figures and size will change accordingly. The highlighted bar or pie acts as a filter for the entire chart, and the chart reacts to the filter interactively.

Through this initial development of an interactive chart, we got a better understanding of developing the final interactive dashboard consisting of different variables, as our main focus of the dashboard is in the various value drivers.

4 Accomplishments and Deliverables

1) Prototype

- 2) After iterating through our development process, we finally came up with a working interactive dashboard prototype using different libraries and technologies. Initially, we split our visualizations into 3 different categories: Quick Glance, Parallel Coordinates and the final interactive Demographic Dashboard. However, as mentioned in the development process in [previous sections](#), we will be removing the Quick Glance and Parallel Coordinates section as it was not very useful to our sponsor. We will focus our effort mainly in perfecting the interactive demographic dashboard as it is the



closest to our final product.

Our sponsor mentioned the following changes to be made to the dashboard to make it more useful:

- Remove the Donut and Pie Charts and make them horizontal bar charts like other attributes
- Change the values of all the variables to percentages instead of aggregated values

Besides suggesting the following changes, our sponsor also requested for an entire new dashboard with the following features in mind:

- Develop line charts for the each of the 6 value drivers and place them in a 2 by 3 format
- Add all the demographic variables as filters on top of the line charts

The proposed draft of the final dashboard will be described further in the [Final Dashboard](#) section below.

3) Project Wiki Update, Mid Term Report and Presentation

From the start of the project until the Mid-Term milestone, our team has successfully updated our Wiki on a regular basis. We have also completed and submitted this Mid-Term Report for grading. Our Mid-Term Presentation is scheduled on 8th October 2014, 1500 at SIS Seminar Room 2-4.

4.1 Challenges Faced and Reflections

Challenges Faced

Our team faced numerous challenges from the start of the project, as we were not very sure how to begin with our project and which direction should we head towards. Once we chose our project, we began to pick up javascript and d3.js, as we understood that those were the technologies widely used for visualizations. Along the way, we realized that we had to pick up several other libraries and study new frameworks in order to create the interactive dashboard that we require.

We also felt that the scope of this project were big at times, as we not only have to take care of developing the dashboard, we had to juggle between managing the stakeholders, updating the documentations and maintaining our wiki page. We often wished that we had more team members to share our workload as we cope with other modules.

Reflections

Through the first half of this project, we definitely realized that communication between various stakeholders is very important in a real-world project. We learnt this through the hard way as there were minor hiccups along the way. We learnt from this and hope to apply this in our work in other areas in future as well.

We also learnt to seek clarification and help from our supervisor and instructor whenever we are in doubt, especially in areas which we are not familiar with, in terms of visualizations and analysis. Their valuable inputs definitely guided us along the way as we went through the process of analysing the dataset and choosing the main points to visualize on the dashboard.

Finally, we picked up several visualization design principles along the way as we are working with sponsors who really hope to use our final product and present it to industry managers in the real world. We understood that the interactive dashboard must be simple and intuitive enough for the average person to understand and use it.

5 Next Steps

5.1 Exploring Predictive Model

Besides developing our main interactive dashboard, we have tried to come up with a Predictive Model that calculates the satisfaction scores of a respondent based on the input scores for the different questions available. This is just a concept and our sponsor may not require us to do it as they have their own regression models and coefficients readily available. Our sponsor also stated that if we are interested in developing the model, they can provide us with the regression coefficients which they have calculated and we will use them in our models

Fill up the form below and see what the predicted satisfaction score is !
Expectation Driver + Quality Driver + Perceived Value Driver =

Customer Satisfaction Driver : 0

Question	Score	picture
The overall quality you were expecting to experience, where 1 means 'you were expecting very low quality' and 10 means 'you were expecting very high quality' ; ?	<input type="text"/>	
How well were you expecting them to meet your personal requirements where 1 is not very well and 10 is very well?	<input type="text"/>	
How often were you expecting things to go wrong things where 1 now means very often and 10 means not very often?	<input type="text"/>	
The overall QUALITY you experienced, where 1 means very low and 10 means very high?	<input type="text"/>	
How well your personal requirements were met where 1 is not very well and 10 is very well?	<input type="text"/>	
How often things went wrong where 1 now means very often and 10 means not very often?	<input type="text"/>	
The price (INSERT NAME) charges, given the quality of their services?	<input type="text"/>	
The QUALITY of (INSERT NAME) products and services, given the fares they charge?	<input type="text"/>	

Snapshot of the predictive model

5.2 Deployment of dashboard

We are also looking into getting more information from our project sponsor on the deployment of the dashboard as we proceed with finalizing our dashboard.

The following are some of the information that we will need before we proceed with the deployment:

- System Requirement to display dashboard
 - Software Requirements
 - Server Configurations
- Data Requirement

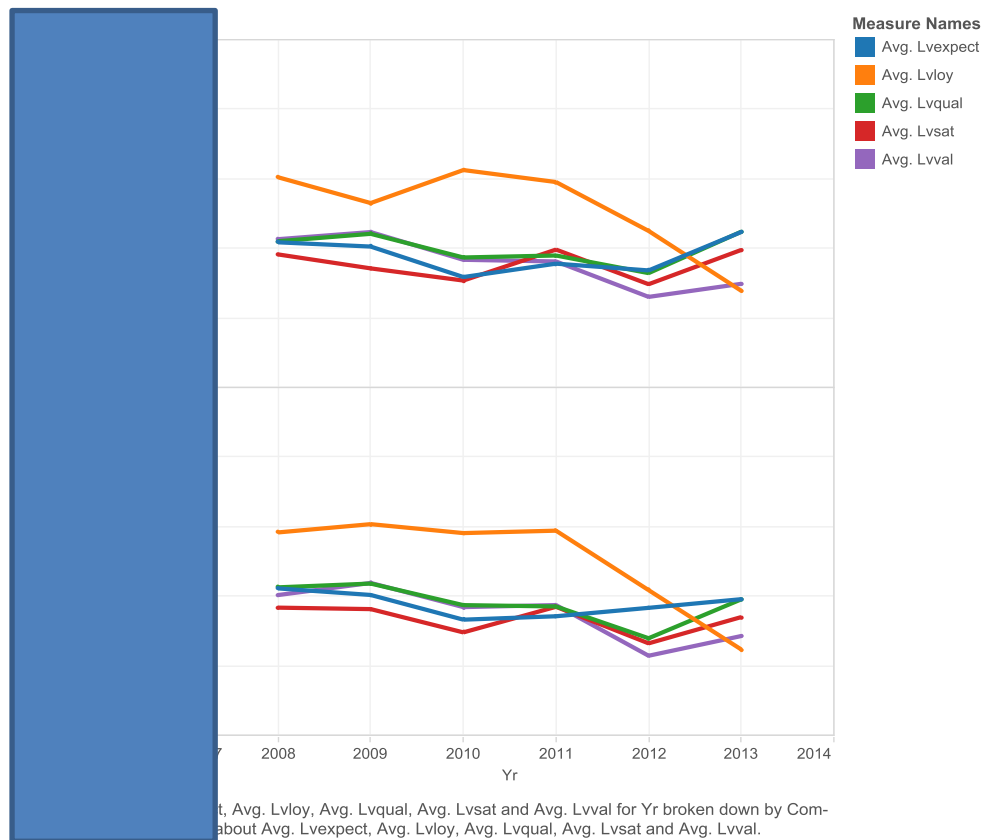
5.3 Final Dashboard

Our sponsor highlighted some changes to be made to the interactive dashboard during our most recent meeting. Instead of all the bar charts showing the distribution of the demographic variables, they preferred using the demographic variables as a filtering option for the 6 drivers showing the year-on-year changes in scores. Below is a draft of the Final Interactive Dashboard made using Tableau Software.

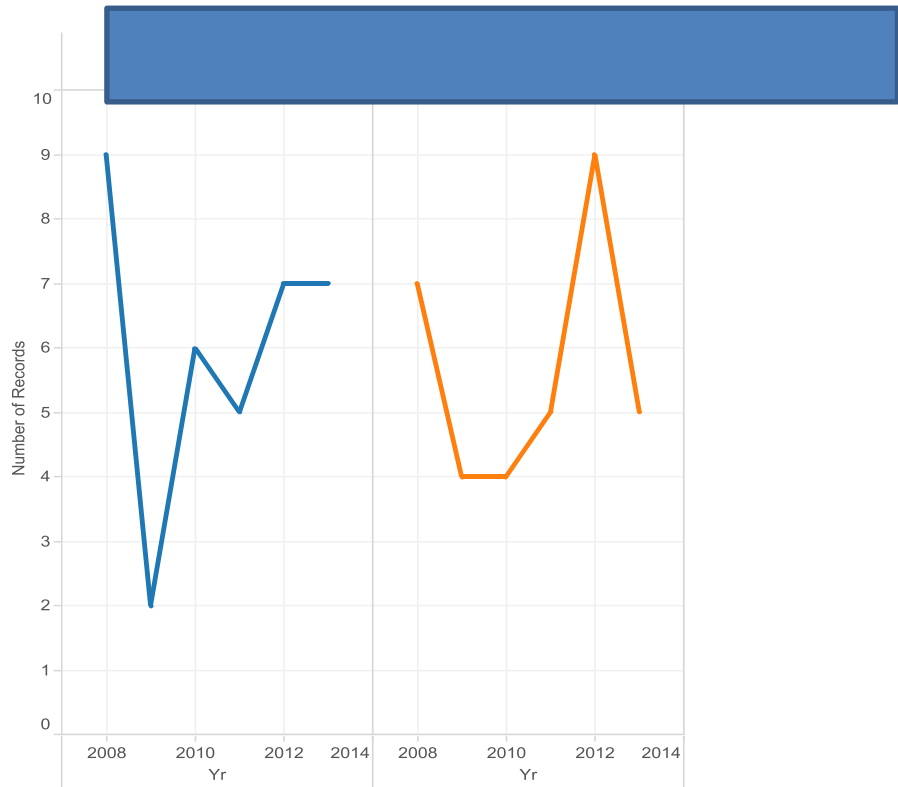


6.2 Exploratory Data Analysis

Trends



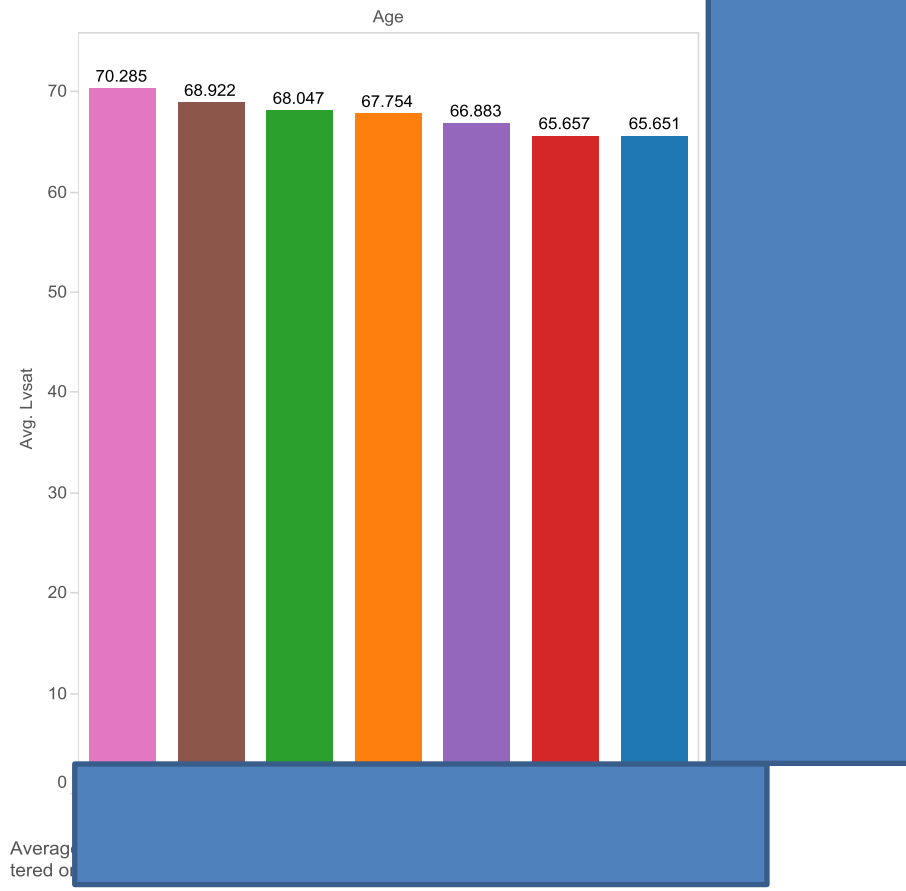
Number of complains



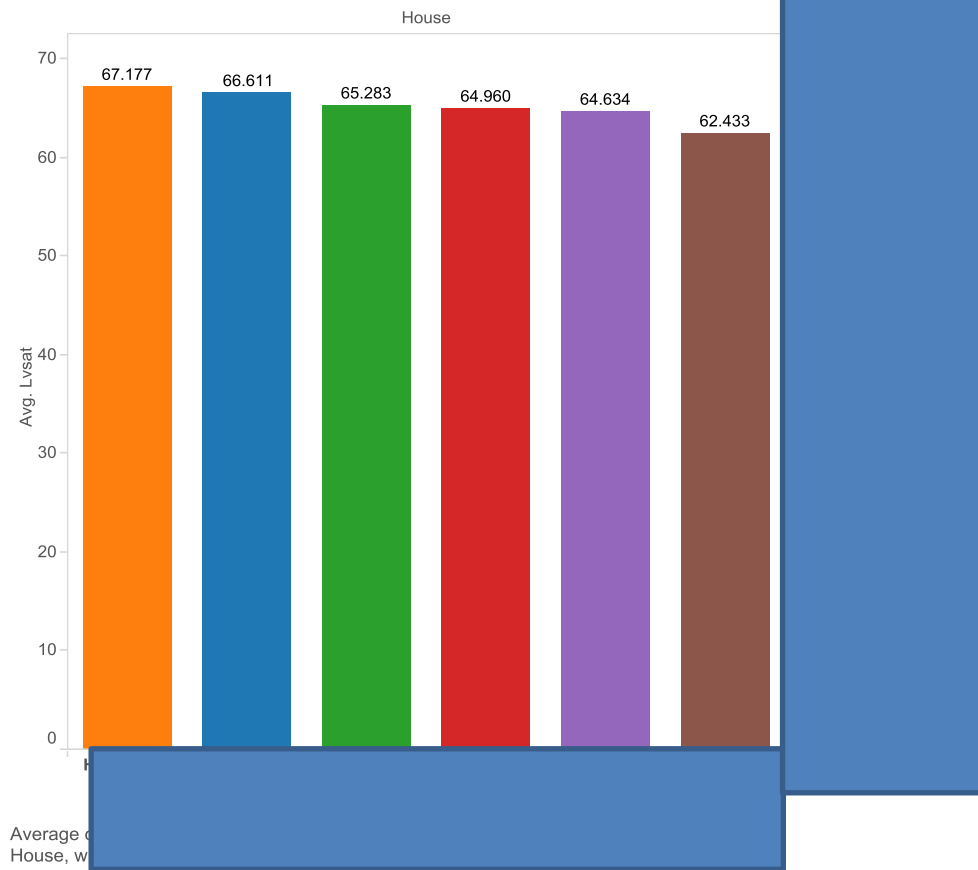
The trend of sum of Number of Records for Yr broken down by Comp and Company. Color shows details about Company. The view is filtered on Comp, which keeps 1.



age



housing



[Redacted text block]