

NOVEC Customer Segmentation for Forecasting

Project Proposal

Predicting energy usage in order to meet current
and future demands across different customers

Anita Ahn
Meselework Aytenifsu
Randall Barfield
Daniel Kim

10/06/2016



Contents

- 1.0 Summary 3
- 2.0 Introduction 3
 - 2.1 Background 3
 - 2.2 Problem Statement..... 3
 - 2.3 Project Description..... 4
 - 2.4 Assumptions 4
- 3.0 Scope/ Criteria of Success 4
- 4. Technical Approach..... 5
- 5.0 Project Plan 5
 - 5.1 Work Breakdown Structure..... 5
 - 5.2 Schedule..... 6
- 6. Risk Analysis 8

1.0 Summary

This project seeks to study NOVEC's customer behavior in electricity usage and to develop a method in segmenting the population into different groups that allows the company to accurately predict future energy demand. A sample of 5 years' consumption data was collected by NOVEC and inputted into SQL database to study patterns and run statistical analysis by NOVEC team of GMU students.

2.0 Introduction

2.1 Background

NOVEC stands for Northern Virginia Electric Cooperative and is one of the largest electric distribution cooperatives in the country. It is a locally based and owned electric distribution system located in Manassas, Virginia. Currently, NOVEC services about 651 square miles of area with more than 6,880 miles of power lines and provides electricity to more than 155,000 home and businesses in multiple counties such as Fairfax, Loudoun, Prince William, Stafford, and Fauquier. Some of its bigger and well-known clients include Potomac Mills Outlet Mall, Verizon, and AT&T. Reliable electricity distribution is important for all these businesses to run their daily operations. As NOVEC is in the process of building a new service center in Loudoun County, a model that can predict electric usage will greatly benefit NOVEC as they start expanding their customer size in new areas.

Electricity plays an important part in running the daily lives of Americans. It is used to power schools, office buildings, and small to large corporations. Although it is seemingly an unlimited supply of energy, it actually takes careful planning ahead of time in order for suppliers to purchase electricity in advance for its customers. Many companies like NOVEC that provide electricity have to purchase enough energy a day in advance in order to meet all the demands of customers that will be using it the next day. This requires knowledge of daily and seasonal trends, but also in depth knowledge about customer's behavior patterns. It is especially difficult to predict electric usage when these companies have no information on how its customers will behave in terms of energy usage. Over predicting electric usage and buying too much energy will lead Electric Companies to incur a sunk cost that will reduce the profit it makes and waste resources while under predicting will lead to unhappy customers who will have no access to electricity in their buildings and homes. This presents a great problem for Operations Research Analyst. If there is a way to predict electricity usage for NOVEC's customers, the company will be able to purchase an efficient amount of energy to distribute to its clients.

2.2 Problem Statement

NOVEC has sample customer data from a stratified random sample of all of its customers. NOVEC would like to determine if the stratified sample it has can be used to segment its customers by their contribution towards NOVEC's peak demand and total energy purchases. NOVEC would like to know the recommended number of segments and the

characteristics of those segments. NOVEC would like to know how well those segments represent the overall system behavior (the entire customer base).

2.3 Project Description

NOVEC has the last five years of data, from 2011 to 2015 on daily electricity usage for its customers. The data has “Customer Group” data, which identifies what type of client it is – *Residential*, *Small Commercial*, or *Large Commercial*. It also has a unique “Map Location Number” that tells you the geospatial location of the client, and finally an “Account Number” that can tell you if the client has changed or not. By studying this data and finding a way to segment the customers into groups according to similar consumption patterns, the final goal is to build a model that will be able to predict the overall electricity usage for the entire population of NOVEC’s customers.

With the amount of data available, it is important to scope the problem into more manageable parts. Currently, NOVEC’s peak electricity usage happens in the month of July. When looking at the daily peak of electric usage, it is around 7 p.m. time frame. Since electricity usage changes from month to month depending on the seasonal fluctuations and temperature changes, the initial focus for this project will focus on the month of July. If a method to segment the customers into groups that accurately predicts the total electricity usage for July is found and applicable to all other months, the project will have met its goal. If the same customer segmentation does not serve as a good predictor for total energy usage in other months, the model may have to be run differently for each month, or it may signify a necessity to segment the customers in a different way.

2.4 Assumptions

There are assumptions and limitations that have to be made prior to starting this project. One limitation is the absence of demographic information about the house or building that the electricity is being delivered to; there is no information on whether the house is small or big, old or new, or whether it is only electric or uses gas. For commercial buildings, there is no data on what type of business the building operates, which will limit the analysis in distinguishing certain types of business buildings. Another limitation of the data is that it is a stratified sample and the sample over represents heavy users and under represents light users. This sample has been collected over time and is not possible to recollect data at this point to balance out the heavy user and light user ratio. Although it is unclear whether or not this will have a huge impact in the reliability of our project, this limitation has to be taken into consideration when concluding results from the analysis.

One assumption that we would have to make is that if the Account Number of the customer does not change. The Account Number uniquely identifies a customer so unless the Account Number changes, we will assume that the client is the same client owning the same type of business, which requires similar energy usage.

3.0 Scope/ Criteria of Success

A successful project outcome will be identifying customer groups that accurately depict the total amount of electricity used. With this information on groupings, NOVEC will be able to run a predictive model that can predict the amount of electricity that will be demanded, which

will allow the company to purchase efficient amount of electricity to meet consumer demand and minimize the costs associated with wasted energy. Although the initial focus of the project is to this accurately for the month of July, the ultimate goal will be to find customer segmentation that can predict the daily amount of electricity used in all months.

4. Technical Approach

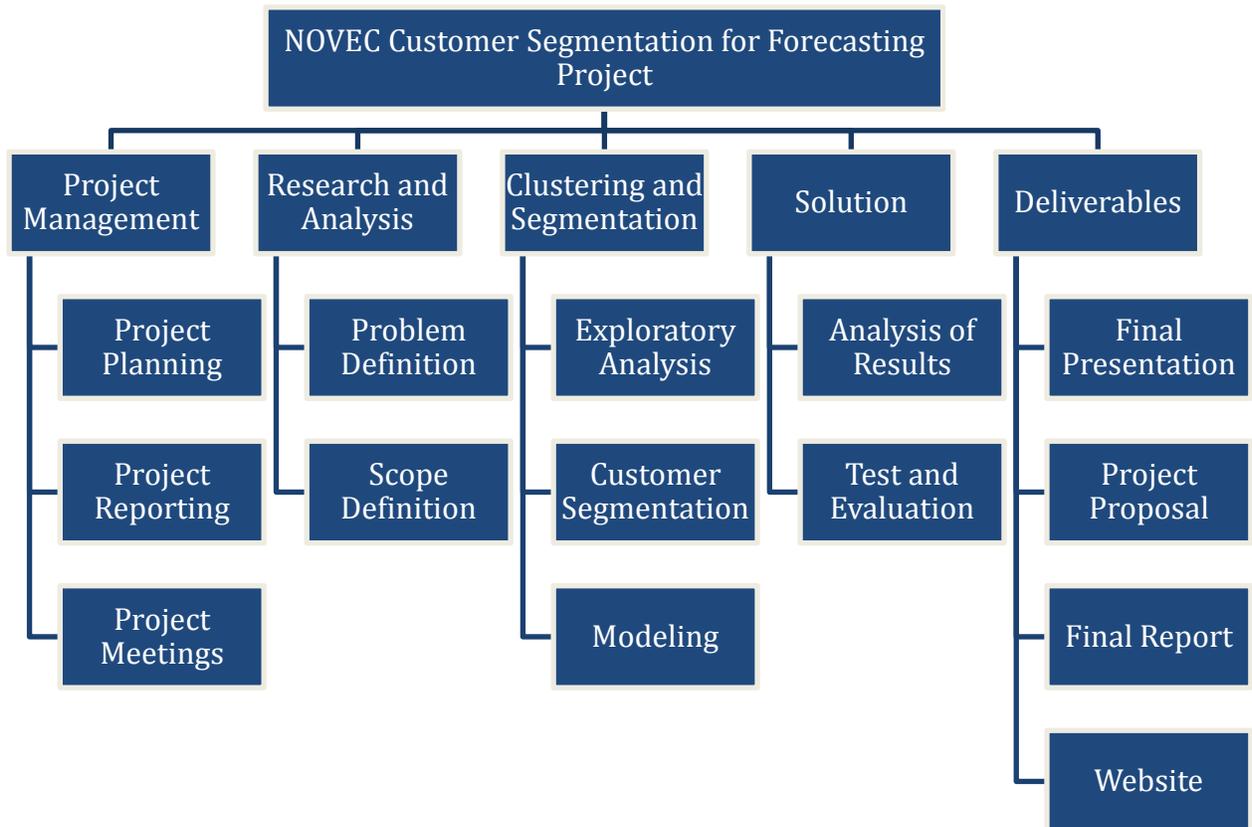
To handle the big amount of data that is available from NOVEC, it will be necessary to use a program that can handle large amounts of data. The group chose to work with SQL to do the initial analysis and R, SAS and Weka to do clustering, generate plots and run analysis on finding patterns of customer data. Using different graphs like line, bar, and histograms, general information about the data can be learned; such as total electricity usage, total number of data points, and total number of unique accounts among different customer groups. Deeper exploratory analysis can be done by using correlation plots, correlation matrix, and clustering dendrograms. Once a similar pattern of behavior among consumers is found, it will be easier to group the customers accordingly and test if the groupings serve as a good predictor for total population's energy usage.

5.0 Project Plan

5.1 Work Breakdown Structure

A Work Breakdown Structure (WBS) was developed to assist in scheduling, evaluating and managing project tasks and deliverables. The WBS has been decomposed into five components: project management, Research and Analysis, Clustering and Segmentation, Solutions and Project Deliverables. Project management consists of project planning, project team meetings, tracking to determine earned value reporting metrics. The purpose of these tasks is to ensure the project team remains focused on sponsor needs, within budget and in time.

Deliverables include final presentation, project proposal, final report and a project website. The Research and analysis consists of problem definition, context, scope, and requirements. It also includes customer segmentation and clustering analysis, which is critical for delivering the solution to be tested and evaluated. The solution also includes analysis of results and group recommendations for the problem



5.2 Schedule

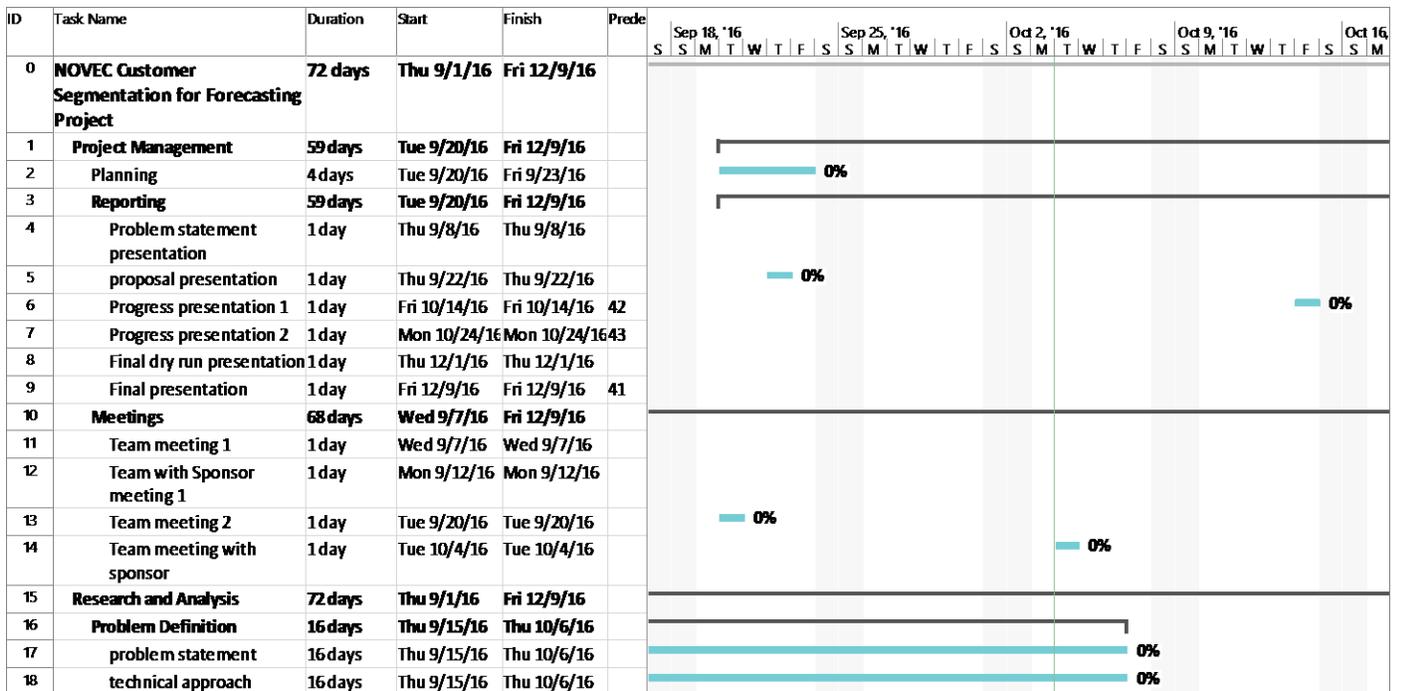
The major milestones planned for the NOVEC Customer Segmentation for Forecasting project are provided in table xxx. These milestones provide a framework for the deliverables and major project briefings.

Milestone	Date
Team Organization and Project Description	Sep 1,2016
Problem Definition Presentation	Sep 9,2016
Project Proposal Presentation	Sep 22,2016
Project Proposal Report	Oct 6, 2016
In Progress Review 1	Oct 13,2016 (20 min)

Professor Working Group Meeting	Nov 3, 2016
In Progress Review 2	Nov 10, 2016
Draft Final Report /Meeting with Professor	Nov 19, 2016
Final Presentation Dry Run	Dec 1, 2016
Final Presentation /Submission Deliverables and Website	Dec 9, 2016 Friday

Table: Project Milestone

The following plan depicts the baseline schedule for the NOVEC Customer Segmentation for Forecasting project.



To mitigate the risks, The Sponsor is aware of the issues with the sample data and recommends a rigorous documentation as we go along using different tools and algorithms for analysis on these data.

biography:

https://www.novec.com/About_NOVEC/index.cfm