

Running head: ASSIGNMENT 6

Assignment 6

Yuqi Zheng

AD 571 Business Analytics Foundations

Boston University

2019 Fall

10/08/2019

Table of Contents

<i>Executive Summary</i>	<i>3</i>
<i>Business Analytics Model and Process.....</i>	<i>4</i>
<i>Business Analytics Model Lifecycle.....</i>	<i>6</i>
<i>Value and Recommendation for Sunnyside</i>	<i>7</i>
<i>Reference.....</i>	<i>8</i>

Executive Summary

This report analyzes Sunnyside's sales performance by establishing different business analytics models. The core business analytics models we used here are the predictive analytics model, descriptive Analytics model, and prescriptive Analytics model.

In the descriptive Analytics model, we used six methods, including five-number summary, standard deviation, the correlation between the sales price and gross square feet, to analyze the sales of Sunnyside. And selected median sale price, standard deviation, and average price per gross square foot as KPIs. K-means clustering was used for comparison with Whitestone with the above three KPIs. The conclusion is that Whitestone is more profitable than Sunnyside.

We used the Predictive Analytics model to predict the sales performance of Sunnyside in the next eight quarters. We found that overall, Sunnyside sales are gradually increasing, and sales data and gross square feet are the two main factors affecting Sunnyside's sales.

Finally, we use the Prescriptive Analytics model to build an optimization problem and use Excel Solver to calculate the most optimal value. Through our analysis, Sunnyside's budget only can hire 1.9372 employees, which is impossible.

According to the above findings, we suggest the Sunnyside manager should change marketing strategies to develop their brand and to maximize the profit in the future.

Business Analytics Model and Process

First, we briefly listed and understood Sunnyside's data through Power BI. We found that Sunnyside's database contains 22 categories of data, including sale price, gross square feet, and square land feet. Therefore, when we analyze a real estate company, we need to analyze the data of these 22 categories to find out the relationship between each group and whether it will affect the performance of Sunnyside. After understanding the structure of the database, we selected Sales Date and Sales Price from 2012 to 2016 as the KPI of Sunnyside for analysis. We used Power BI's line graph to make the relationship between the two more intuitive. We found that except for the period from 2013 to 2014, the sale price increased with time. And, over 2016, Sunnyside's revenue was about \$ 3.39 million.

Next, we import the data into the R language and try to find out the relationship between the average price of 1 square foot of residential real estate in Sunnyside for each year. To maintain the accuracy of the data, we set up a filter to remove fields with 0 for price and square feet. In order to understand the performance of Sunnyside, we compare Sunnyside with Whitestone and Annandale because it is not enough to look at its data. In the end, we found that the overall sales of these three companies are increasing year by year, but Whitestone is more profitable than Sunnyside and Annadale.

Then, we used the descriptive analytics model to analyze Sunnyside. First, we used R to calculate 1604 homes sold since 2009, and the total sales were about \$1.98 billion. The average price per house is about \$1.13 million. The average size of a property is 4125 gross square feet. The five-number summary provides information about the dataset because it consists of five most essential sample percentiles: minimum, maximum, lower quartile, median, and upper quartile. We want to analyze gross square feet and sales prices by using the five-number

summary to help describe the spread of the data and determine whether or not any data points are outliers. Besides, the separate analysis of four types of status in Sunnyside can make our conclusions more objective and accurate. Here, we only need to consider residential status. It is also necessary to find the standard deviation of sale prices for residential properties and the correlation between sale prices and gross square feet for residential properties. Because the standard deviation expresses the variability of the population and finding relationships could help us be clear about the relation between sale prices and gross square feet. Then we choose the median sale price, standard deviation, and the price of one total square foot of residential in Sunnyside as three main KPIs. We use k-means clustering to group them that objects in the same cluster are more similar to each other than to objects in different clusters. This method could let the data be more visualized. Thus we could see the relations between those three KPIs. Finally, we compare Sunnyside and Whitestone by t-test for the average residential property costs and find Sunnyside is costlier than Whitestone.

Lately, we use a predictive analytics model to perform a deep prediction and analysis of Sunnyside. The sales of Sunnyside for the following eight quarters can be forecasted through R by applying time series analysis as well as the multiple regression model. Through time series analysis and multiple regression models to predict, we find Sunnyside's sales changes in the next eight quarters will not change very sharply. Compared with the sharp rise and fall in previous years, the next two years will rise with slight fluctuations. Lots of predictors of sales was found by using a multiple regression model in Sunnyside. By looking for p-value, sale date and gross square feet are two key factors we need to consider in future forecasts. Besides, we also find the most overpriced and the most prominent bargain properties in Sunnyside.

Finally, we apply the prescriptive analytics model to our dataset by using Excel Solver. Because using an optimization model could help us find the best suitable solution for our objective. Thus, we build an optimization model with parameters, variables, objective function, and constraints in Excel and run the GRG non-linear model to analyze Sunnyside's performance. We get the NPV of Sunnyside is approximate \$1.86 million with 4.66% commission and need to hire approximate 2 employees.

Business Analytics Model Lifecycle

Analysis models are rarely static. Their purpose is to model the organization to allow managers to investigate what is happening and evaluate changes in the way they operate. These changes may include looking at the price, how the department is organized, partnerships, or introducing new products/services into new areas. So, the models used to evolve as further information is discovered, affecting what is being analyzed. We defined it as Model Lifecycle. The model lifecycle management stages include eight processes: identify the business goal; manage the necessary data, model development, model validation, model deployment, model monitoring, model improvement and model retirement.

Defining business goals means letting us determine the type of model we need to build and how to apply it. Because once we know why we are building this model, we can understand what kind of data we need to analyze. In this case, our goal is to analyze Sunnyside's performance. Managing the necessary data is vital to narrow the required range, which reduces the possibility of data overfitting. Although the model development step is usually completed using historical data, we can complete the model development through statistical and visual evaluation data. A model validation stage, we need a comprehensive understanding of which variables have an impact on the model and what the effect is. Thus, we usually choose some

KPIs to see the performance of Sunnyside. There often have seven steps in the deployment of the model stage, which includes validation, choice scoring, report integration, application integration, monitoring, and improvement. Model deployment is the use of new data in the model to gather valuable insights for management decisions. Implementation is to extract value from the model after running it with actual data and new data. The purpose of model lifecycle management is to make the organization proactive and have the ability to expand decision-making capabilities.

Value and Recommendation for Sunnyside

By building models, we make valuable data that can't speak. By analyzing this data, we can let Sunnyside managers know how the company is operating and understand the strengths and weaknesses. And by building a predictive analytics model, we not only understand that Sunnyside's sales situation will change with the change of sales date and gross square feet. It also predicts Sunnyside's performance over the next eight quarters. We anticipate that Sunnyside's sales performance in 2018 will be higher than in 2017. By building a Descriptive Analytics model, we learned that even though Sunnyside has higher sales than Whitestone, Whitestone is more profitable. We, therefore, recommend that Sunnyside build and sell more large gross square feet properties. Because through our analysis, the bigger the gross square foot property is, the better it sells, and the more money it makes. Also, assuming that Sunnyside's promotional efforts are the same every quarter, we recommend that Sunnyside can reduce the amount of advertising and promotion in the first quarter of each year and put more advertising in the last quarter. Because according to our findings, sales in the first quarter of each year tend to be less than in the previous quarter of the past year. And in the last quarter of each year is the best time for sales. Sunnyside can use the extra budget to add more publicity in the quarter with good deals.

Reference

How Smart, Connected Products are Transforming Companies, Porter, M., Harvard Business Review, October 2015

Walk the Line in Managing Big Data Analytics Efforts, Stedman, C., Harvard Business Review, May 2015, Vol. 93, Issue 5

Module 6 (Lectures 11 and 12), Boston University, AD 571