THE PRICING & CALIBRATION

Revnomix
Hotel Data Analytics

Sameer Kulkarni
REVNOMIX SOLUTION

CONTENTS

INTRODUCTION

In today's competitive world the pricing is the most important decision that any organization has to take on regular basis. Companies increasingly need to make pricing decisions more and more rapidly in order to account for the changing inventory or to respond to the competitive action and market dynamics.

The hotel industry is privy to the dynamic inventory changes. To handle the inventory position of a hotel dynamically the Revnomix has developed a set of new algorithms. One of these algorithms is called "Remaining Capacity Based Pricing". The capacity available to sell after deducting the current business on books and any Out of Order (OOO) rooms is considered as Remaining Capacity. This remaining capacity is the base for pricing. This pricing algorithm is based on a philosophy "Price Low when there are more room to sell and Price More when there are fewer rooms to sell". The algorithm pushes for building a base before the prices are increased and the hotel starts yielding on the available situation.

Since the hotels operate in a dynamic market, there is a need to react to the competitive world. The Revnomix has developed three algorithms aimed at providing the hotels with the Market Appropriate Pricing (MAP). In this MAP approach, the algorithms predominantly use Rates floated by the Hotel's Competitors on the Online Travel Portals or Online Travel Agents (OTA). Along with this Market Pricing information, the algorithms use other sentimental inputs like Ranking on each of the OTAs, etc. The sentimental inputs are considered as they influence the buying decisions of the hotel guests.

REMAINING CAPACITY BASED PRICING

Most of the pricing engines available in the market use the projected demand for pricing. The major flaw of such approach is that, the engine is susceptible to dropping rates closer to the arrival date. This drop is majorly caused by the demand projections by such systems. In most of the times these systems over forecast the demand and rely on the demand materialization patterns observed in the recent times. The demand dynamics in the hotel industry is very volatile in nature. The swings in the demand materialization can cause huge surge or tumble in the rates.

These systems also fail to address the very basic business philosophy that the hotels work with "Yield on the inventory position". Most of the hotels would always wish to sell at higher rates when they have fewer rooms. Conversely, the hotels would like to build a base before hiking up the prices. Since most of these systems work on projected demand and projected demand materialization, there is a possibility that these systems either assume additional demand when there is a quick pick-up ahead of expected demand materialization pace or assume no demand when there is a flat movement in the demand materialization. Both of these assumptions have adverse effect on the hotel pricing.

In the case of additional demand assumption, these systems increase the rates in anticipation of very high demand and then drop the rates closer to arrival when such high demand does not materialize. This causes a very daunting task of explaining the hotel guests or even cancelling and rebooking. This also affects adversely to the hotels' social media sentimental index on the OTA platforms. In todays' digital era, one cannot turn blind eye to such negative impacts.

On the contrary, the assumption of no demand when there is sluggish or flat demand materialization will trigger drop in prices by these systems. And, when the demand actually picks up, these systems try to increase the rates but by then there is every possibility that the hotel would have lost a genuine revenue making opportunity.

In order to arrest such issues and to provide with stable pricing, the Revnomixians have developed a new algorithm - The Remaining Capacity Based Pricing (RCBP). This algorithm derives a price point at which the hotel can sell entire remaining hotel rooms inventory. The RCBP revises the price point at each inventory availability dynamically. The focus of the algorithm is to derive at the price point that will consume maximum remaining inventory of the hotel while generating an optimal revenue. As per the philosophy “Price Low when there are more room to sell and Price More when there are fewer rooms to sell”, the algorithm keeps revising the price upwards every time when a room is sold. This algorithm tries to establishes a linear relationship between available capacity and rates to offer.

In order to draw a linear line, one needs a “***Slope***” and an “***Intercept***”. This being a pricing line with a linear relationship between capacity and rates, it is a downward sloping line. Therefore, the pricing line has a “***Negative***” slope. The equation for linear line is –

$$y = mx + C \tag{1}$$

where

y — Remaining Capacity

x — Rate to Offer

m — Slope of the pricing line

C — Intercept of the pricing line

Since the pricing line is a downward sloping line it can be expressed as –

$$y = C - mx \tag{2}$$

We are interested in obtaining the price for the remaining capacity, we have to solve the above formula (2) to obtain the value of "x" .

Therefore, the formula (2) can be expressed as –

$$\mathrm{x} = \frac{(\mathrm{y} - \mathrm{C})}{-\mathrm{m}} \tag{3}$$

This formula needs “Slope” and “Intercept”. The slope and intercepts are calibrated and stored. The inputs to Slope and Intercept calibration are discussed in the next section.

Inputs to the calibration

The RCBP uses the following inputs to derived the price point for the given capacity –

1. Historical data
2. Total Hotel Inventory or Rooms Capacity

HISTORICAL DATA

The algorithm is designed to consume any amount of historical data to calibrate the pricing line. However, the default requirement is last 24 months of historical data. The algorithm is designed to work even on a minimum requirement of 90days of historical data. The Historical data is the one of the major inputs to the calculation of Slope and Intercept.

The data should contain following details –

1. Transaction ID
2. Arrival Date
3. Number of Rooms Sold
4. Number of days occupied or Length of Stay
5. Room Revenue generated by the transaction

Following other items are calculated based on the "Arrival date" Day of the Week

1. Weekday and Weekend
2. Week Number
3. Month
4. Season

Day of the Week

The "Day of the Week" (DOW) is the calendar day of the arrival date.

Weekday and Weekend

The week is further divided into Weekday (WD) and Weekend (WE). Currently, this parameter is hard set as Weekdays for Monday to Thursday and Friday to Sunday considered as Weekends. During the phase two this is going to be a configurable parameter.

Week Number

The "Week Number" is calculated in such a way that the FIRST week of the year always starts from the first Monday of the Year. This is immaterial of the 1st Jan falling on any other day of the week (e.g. the 1st Jan 2013 was on Tuesday and first Monday of the year was on 7th Jan 2013. As per the algorithm, the week number ONE will be from 7th Jan and the period from 1st to 6th will be part of 52nd week of previous year). This implementation ensures that the years are correctly aligned with each other by the days and week numbers even when one of the years is a leap year.

Month

The "Month" is calculated in such a way that all the months have equal number of days. Every month has 28days and 4 weeks and there will be 13 months in any given year.

Season

Currently the seasons are fixed and are 6 in nature. Each season will have either 8 weeks or 9 weeks. The Season 1 starts from the 49th Week and ends on 4th week and this will have 8 weeks. Similarly, Season 6 will also have only 8 weeks starting from 41st week and ending on 48th week. Rest all the seasons would have 9 weeks. Apart from these 6 seasons, there will one more season "Season 0". This 0th Season will be the superset of all the six seasons. In other words, the entire year is considered as the 0th Season.

TOTAL HOTEL INVENTORY OR ROOMS CAPACITY

The actual rooms Inventory of the hotel is used as one of the major inputs to the calibration of Slope and Intercept.

Calibration of Slope and Intercept

These two are the core elements of this algorithm. The calculation of the Slope and the Intercept are stored in as calibration parameters and to be used for deriving the pricing for the remaining capacity of rooms left to sell for any given date in future.

These two items are calculated for each Season, for each DOW and for all days of the week as well. Below given is a sample table of these two parameters.

Season	All	WD	WE	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
0	-0.030	-0.029	-0.029	-0.031	-0.031	-0.027	-0.028	-0.029	-0.028	-0.029
1	-0.028	-0.029	-0.027	-0.028	-0.036	-0.026	-0.028	-0.025	-0.030	-0.031
2	-0.032	-0.032	-0.029	-0.058	-0.033	-0.030	-0.035	-0.041	-0.030	-0.031
3	-0.033	-0.037	-0.032	-0.033	-0.032	-0.053	-0.037	-0.032	-0.030	-0.030
4	-0.028	-0.030	-0.028	-0.030	-0.032	-0.032	-0.029	-0.029	-0.029	-0.028
5	-0.032	-0.032	-0.029	-0.029	-0.048	-0.036	-0.033	-0.027	-0.028	-0.031
6	-0.033	-0.031	-0.034	-0.042	-0.036	-0.053	-0.030	-0.032	-0.038	-0.031

Table 1. Sample table of Slope

Season	All	WD	WE	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
0	98.88	96.94	99.48	97.64	101.11	91.98	98.36	99.08	98.74	98.60
1	97.66	98.92	96.50	100.16	100.36	93.85	98.89	93.72	103.26	102.71
2	101.68	96.58	95.98	118.78	99.61	97.51	106.49	116.36	101.52	98.35
3	106.96	109.06	107.55	100.76	101.90	133.57	113.02	107.43	104.08	109.20
4	99.53	100.23	100.56	98.15	111.15	102.65	92.88	102.64	102.41	96.20
5	103.34	100.31	98.04	93.97	120.26	110.54	100.12	95.32	99.45	101.26
6	104.81	99.19	108.67	116.81	103.84	129.81	101.58	104.65	116.04	101.01

Table 2. Sample table of Intercepts

The Calibration of Slope and Intercepts is well equipped to handle the data insufficiency. During the analysis, it was observed that there were many instances where the data was insufficient to calculate these two calibration items. The algorithm needs at least 30 transactions for each DOW within a season. Each season, from 1 to 6 will have either 56days or 63days. Therefore, each DOW in a season would have either 8 or 9 occurrences. As per the minimum requirement of 30 transactions, it comes to slightly less than 4 transactions per day. And it was observed that some of the hotels were short of even such a miniscule requirement.

Fallback Mechanism

In case, for any DOW in any of the Seasons, the algorithm detects data insufficiency, the Fallback mechanism is triggered. If the Algorithm detects the insufficient data for a DOW that is a Weekday then the values calculated for WD in the table for the specific season is used and in case of Weekend, the value from WE is used for that season. Similarly, if there is a data insufficiency at WD and WE level for a season then the values from the 0^{th} Season are used.

For example, if the algorithm finds that the data is insufficient to calibrate slope and intercept for Monday of 2^{nd} Season, then the algorithm will use the values calculated under the column WD for calibration. Similarly, if the WD and/or WE values are not calculable for a season, say the 3^{rd} Season, due to data insufficiency, then the algorithm borrows these values from the 0^{th} Season.

Data Arrangement

In order to carry out the calibration the data obtained from the hotel needs to be arranged in a right format. The data is arranged to subset by Season and by DOW and by WD/WE. The steps followed are –

1. Fetch the data from the database. A sample SQL Query
   SELECT ArrivalDate as CheckInDate, DAYNAME(ArrivalDate) AS dow, WEEK(ArrivalDate,7) as WkNum, RoomSold as Nights, COALESCE(ROUND((RoomRevenue/(RoomSold*LOS)),2),0) AS ADR FROM channelproduction WHERE HotelID = T0001;
2. Sort the Historical data by Arrival Date in ascending order
3. Identify the Week number of the arrival date
   The Week number has to be identified in such a way that the first week of the year starts from first Monday of the year.
   a. To identify the week number in MySQL -
      SELECT CheckInDate, WEEK(CheckInDate,7) AS WkNum FROM allTransactions;
      The value 7 in the optional parameter in Week() function of MySQL. This parameter ensures that the first week of the year always starts from first Monday of the year

b. To identify the week number in MS Excel -
IF(CELL<DATE(YEAR(CELL),1,8)-WEEKDAY(DATE(YEAR(CELL),1,6)), 52, WEEKNUM(CELL,2)-1)
This excel formula ensures that the first week of the year always starts from first Monday of the year.

4. Identify the Month by week number. A month will always have 4 weeks only. The week 1 to week 4 will be 1^{st} month and 5 to 8 will be 2^{nd} month and similarly week 49 to week 52 will be 13^{th} month. Thereby, each week will have 13 months (52 ÷ 4 = 13).
5. Identify Season by Week Number.
    a. The 1^{st} season will always start from 49^{th} week and end on 4^{th} Week.
    b. The 2^{nd} season will always start from 5^{th} week and end on 13^{th} Week
    c. The 3^{rd} season will always start from 14^{th} week and end on 22^{nd} Week
    d. The 4^{th} season will always start from 23^{rd} week and end on 31^{st} Week
    e. The 5^{th} season will always start from 32^{nd} week and end on 40^{th} Week
    f. The 6^{th} season will always start from 41^{st} week and end on 48^{th} Week
    g. The 0^{th} Season will have all the weeks from 1^{st} to the 52^{nd}
6. Calculate Average Daily Rate (ADR) for each of the transaction

Once the data is arranged it would look like the one shown in table 3

Date	DOW	WeekNum	SeasonID	Month	Sold	Capacity	ADR
03-Apr-14	Thu	13	2	4	1	35	1600
05-Apr-14	Sat	13	2	4	1	35	2400
04-Apr-14	Fri	13	2	4	2	35	1600
05-Apr-14	Sat	13	2	4	1	35	1600
26-Apr-14	Sat	16	3	4	1	35	2000
19-Apr-14	Sat	15	3	4	1	35	4000
06-Apr-14	Sun	13	2	4	1	35	1600
14-Jun-14	Sat	23	4	6	1	35	2800
30-May-14	Fri	21	3	5	2	35	2000
13-Apr-14	Sun	14	3	4	1	35	3200
13-Apr-14	Sun	14	3	4	2	35	1600

Table 3. Sample table after the data is arranged

Outlier Detection and Removal

The algorithm has built in two levels outlier detection mechanism.

1. The outliers are removed at each season level
2. The above refined data again goes through another round of outlier removal at each DOW level

This is to ensure that the data is not skewed due to –

1. the Special Event Pricing
2. Hotel's aggressive pricing strategy for any particular day or a period of days
3. Wrong data for any particular day or a period of days

The Algorithm uses two methods for detecting the outlier –

1. Interquartile Range Method[1]
2. Percentile Method[2]

INTERQUARTILE RANGE METHOD

The interquartile range (IQR) is a measure of variability, based on dividing a data set into quartiles. Quartiles divide a rank-ordered data set into four equal parts. The values that divide each part are called the first, second, and third quartiles; and they are denoted by Q1, Q2, and Q3, respectively.

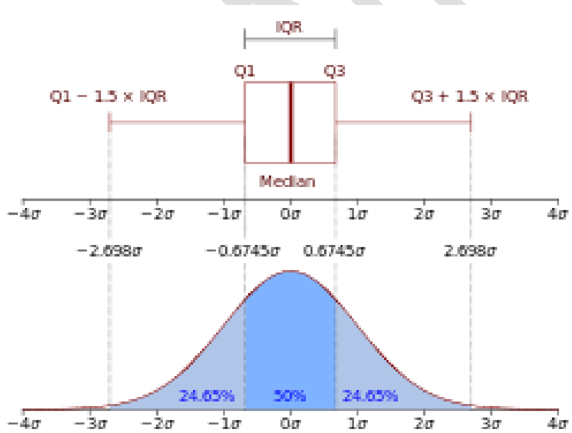


Figure 1. The distribution with IQR calculation

The difference between Q3 and Q1 is IQR

$$IQR = Q3 - Q1 \tag{4}$$

The interquartile range is popularly used to find outliers in the data. Outliers are observations that fall below Q1 - 1.5(IQR) or above Q3 + 1.5(IQR).

[1] https://en.wikipedia.org/wiki/Interquartile_range
[2] https://en.wikipedia.org/wiki/Percentile

Therefore, the data used for the calibration should be

$$Q1 - 1.5(IQR) \leq ADR_n \leq Q3 + 1.5(IQR) \tag{5}$$

Where:

ADR_n – Distribution of Hotel ADR

n – Season Number

The IQR method is slightly modified to suite the hotel data conditions. The Lower Bound and Upper bounds are calculated as

$$Lower\ Bound = Q2 - 2.5(IQR) \tag{6}$$

$$Upper\ Bound = Q2 + 2.5(IQR) \tag{8}$$

There is a possibility that the Lower Bound goes sub ZERO. Hence, the Lower Bound is required to be restricted from breaching into a subzero value.

Now, the formula no. 5 can be expressed as

$$Lower\ Bound \leq ADR_n \leq Upper\ Bound \tag{9}$$

The refined data is used to further subset by each DOW. This set of data is then passed through another round of outlier detection and removal.

PERCENTILE METHOD

The data obtained after removing the outliers using IQR method is then broken into 10 subsets.

1. One each for each DOW (i.e. 7 datasets)
2. One each for Weekday and Weekend
3. One superset containing all DOW data

A percentile indicates how many observations fall below a value. For example, the 25th percentile is a value below which a quarter of observations from the given distribution are found.

In the algorithm, 5th and 95th percentiles are used. All the values below 5th percentile and above 95th percentile are removed as outliers.

$$P05 \leq ADR_{nw} \leq P95 \tag{10}$$

Where:

ADR_{nw} – Distribution of Hotel ADR

n – Season Number

w – Day of the week number

P05 – the value at 5th Percentile

P95 – the value at 95th Percentile

This $'w'$ has maximum limit of 10.

1. ZERO (0) for all days of the week
2. ONE (1) for weekdays
3. TWO (2) for weekends
4. THREE (3) for Monday
5. FOUR (4) for Tuesday
6. FIVE (5) for Wednesday
7. SIX (6) for Thursday
8. SEVEN (7) for Friday
9. EIGHT (8) for Saturday
10. NINE (9) for Sunday

Linear line Fitting

Once the data is processed and clear of all the outliers, the algorithm forces the linear relationship between rooms sold and rates achieved. In order to establish linearity, ‘***z***’ Scores[3] are used.

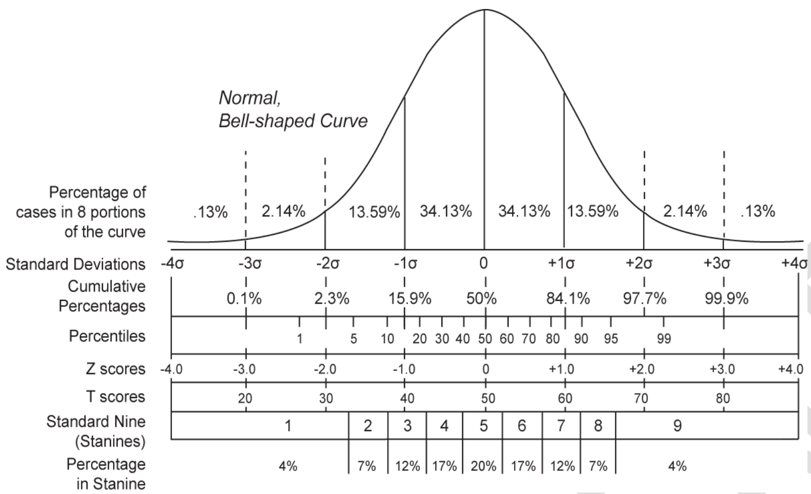


Figure 2. The Standard Normal Distribution and its attributes

The picture given here is to help in understanding the concept of ‘z’ score or Standard Score. It is assumed that a distribution is perfectly linear when it follows the Normal Distribution or Bell Curve. As the linearity is forced on the ADR distribution, the ‘z’ scored are used in this algorithm to fit the distribution.

In the algorithm ‘z’ scores are restricted between -3 and +3. Since this range covers 99.5% of all the observations, the algorithm restricts the ‘z’ scores to this range (-3 to +3).

MU ADJUSTMENT

In order to carry out the linear fitting, MEAN (Mu or μ) and STANDARD DEVIATION (Sigma or σ) are calculated for each of the distributions obtained from formula 10. Using μ and σ, the value at the -3 ‘z’ score is calculated.

$$\tau = (-3 \text{ X } \sigma) + \mu \tag{11}$$

Where ‘τ’ is used to capture the output from the above formula.

[3] https://en.wikipedia.org/wiki/Standard_score

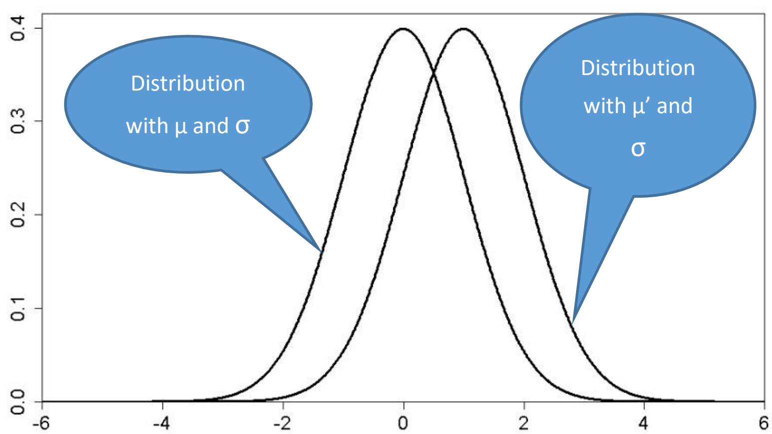
If τ < 0 then adjust μ such the value of τ ≥ 0. In order to adjust the μ, such that the τ ≥ 0, a delta (δ) value is calculated dynamically by passing various numerical values (higher than -3 while incrementing the value by 0.01 in each iteration) in loop till the condition is realized. The formula 11 can be expressed as –

$$\tau = (-\delta \, X \, \sigma) + \mu \qquad \text{such that } \tau \geq 0 \tag{12}$$

Once the δ that provides τ ≥ 0 is obtained, then the μ is adjusted while keeping the σ at the same level while assuming that the new τ is at -3 z-score. The idea is to retain the spread of the distribution while changing the central tendency of it. Also, this ensures that the lower price points do not have much of an impact on the overall pricing. The adjusted Mu (μ') is given by –

$$\mu' = \tau - (-3 \, X \, \sigma) \tag{13}$$

where τ ≥ 0



Once the μ' is calculated and plotted along with the μ and σ the distribution would somewhat similar to the one depicted here.

Figure 3. The shift in the distribution due to the adjusted μ'

VOLUME VS VALUE

The values at each 'z' score is considered as the price point at which the demand would arrive. Using the z score and the volume a table will be created along with the reverse cumulative volume and capacity adjusted volume.

z Score	Rates (τ)	Volume	Cumulative	y Value
-3	4.04	0	61	15
-2.5	777.28	0	61	15
-2	1,550.52	0	61	15
-1.5	2,323.76	0	61	15
-1	3,097.00	32	61	15
-0.5	3,870.24	4	29	7
0	4,643.48	1	25	6
0.5	5,416.72	16	24	6
1	6,189.95	1	8	2
1.5	6,963.19	4	7	2
2	7,736.43	3	3	1
2.5	8,509.67	0	0	0
3	9,282.91	0	0	0

Table 4. Sample table with volume calculation at each price point for hotel with 15 rooms

The volume is calculated using the formula given below

$$\sum_{i=0}^{n} S_i \begin{cases} if\ z = 13\ then\ R_i \geq \tau_z \\ else\ \tau_z \leq R_i < \tau_{z+1} \end{cases} \tag{14}$$

Where

n = number of observations in the dataset

s_i = Rooms sold

R_i – Rate at which the rooms were sold

τ_z – Price point derived by using z-score

REVERSE CUMULATIVE VOLUME

Reverse cumulative volume is calculated (refer column 'Cumulative' in the table 4) at each price point. This with an assumption that, if the lower rate is kept open then entire demand would book at that rate and the demand that books at lower rates would not book at higher rates.

Y VALUE

Then the 'y value' is calculated. The y value is nothing but the expression of cumulative volume with regards to the capacity. Extending the assumption of lower rates attract maximum demand, the rate with maximum cumulative volume is considered to be that rate at which the hotel can sell all its capacity. The y value is calculated as –

$$y = \frac{C_i \times H}{\sum v} \qquad (15)$$

Where:

i Each price point from 1 to 13

C_i Cumulative value at each price point

H Hotel Capacity

v Total volume

ADJUSTED TABLE

Once the y value is calculated, all the rows having ZERO values under the "Volume" column are removed. The assumption here is that the hotel would not be selling any rooms below a certain rate level and above a certain rate level. Hence, all the price points having zero values are removed.

Rates	Volume	Cumulative	y Value
3,097.00	32	61	15
3,870.24	4	29	7
4,643.48	1	25	6
5,416.72	16	24	6
6,189.95	1	8	2
6,963.19	4	7	2
7,736.43	3	3	1

Table 5. Sample table after the adjustment for hotel with 15 rooms

Linear Regression[4]

Now the data is ready to pass through the linear regression to obtain the SLOPE and INTERCEPT. The Rate is passed as x value in the linear regression. Though the algorithm uses the Python inbuilt linear regression model (using the package scipy.stats and importing linregress), here is a very simple way to calculate the slope and intercept.

$$\text{Slope} = {(Min(y) - Max(y))}\Big/{(\text{Max(Rate)} - \text{Min(rate)})} \tag{16}$$

$$\text{Intercept} = Max(y) - (Slope * Min(Rate)) \tag{17}$$

Implementation of RCP

The Slope and Intercept are calculated for each season (along with the whole year depicted with 0th Season) and for each DOW (along with Weekday, Weekend and all DOW) and remaining capacity for any given date are used to derive the pricing using the formula 3.

$$\text{x} = \frac{(\text{y} - \text{C}_{sw})}{\text{m}_{sw}} \tag{18}$$

where

x – Rate to charge

y – Remaining Capacity

C – Intercept

m – Slope

s – Season

w – Day of Week

[4] https://en.wikipedia.org/wiki/Simple_linear_regression

REAMING CAPACITY

The reaming capacity is calculated considering

1. Current Rooms Sold (Business on Books or BoB) for each of the days into the future
2. Room on OOO status

Using the above details, the Remaining Capacity is calculated (Total Capacity – BoB – OOO). In case the hotel is not able to provide with the OOO rooms separately, the remaining capacity is calculated as Total Capacity – BoB. Passing this remaining capacity as an input into the formula 3 along with the Slope and Intercept pertaining to the season and day of week of the date for which the pricing is derived.

Illustration of RCP

Here is an example of the RCP based on various capacity position for a hotel of 75 room capacity for one season.

Occ%	Remaining Capacity	Mon	Tue	Wed	Thu	Fri	Sat	Sun	WD	WE	Season
100	4	7949	7949	7949	7949	8899	8899	8899	7949	8899	9349
90	8	7249	7249	7249	7249	8099	8099	8099	7249	8099	8599
80	15	6599	6599	6599	6599	7299	7299	7299	6599	7299	7799
70	23	5899	5899	5899	5899	6499	6499	6499	5899	6499	7049
60	30	5199	5199	5199	5199	5699	5699	5699	5199	5699	6299
50	38	4499	4499	4499	4499	4899	4899	4899	4499	4899	5499
40	45	3799	3799	3799	3799	4099	4099	4099	3799	4099	4749
30	53	3099	3099	3099	3099	3349	3349	3349	3099	3349	3999
20	60	2399	2399	2399	2399	2549	2549	2549	2399	2549	3199
10	68	1699	1699	1699	1699	1749	1749	1749	1699	1749	2449

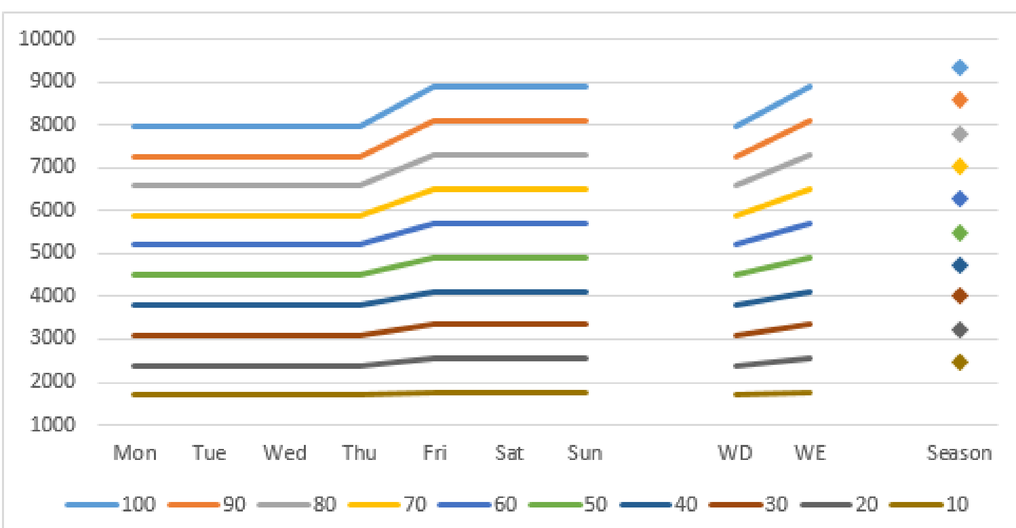


Table 6. Sample table and graph with the Prices based on the different capacity positions

MARKET APPROPRIATE PRICING

Most of the pricing engines available in the market use the median or mean rates based on the competition set definition. Lately many OTAs are also helping the hotels with the pricing based on the market intelligence. Most of these OTAs use the rates being floated on their own portal and do not use rates being floated on others. In most cases these pricing decisions are going to be based on simple market conditions and do not fold in the hotel pricing tactical like adopt lower pricing levels during weekends as the hotel needs more support. And, adopt higher pricing levels during weekdays as the hotel needs to push the rates. Or, adjust the pricing based on the social sentiment indicators. The Market Appropriate Pricing (MAP), that Revnomix Solutions is going to use will have three different options. All the options are designed in such a way that the hotel pricing is neither too low nor too high.

1. MPI Pricing
2. ARI Pricing
3. PQI Pricing

All these algorithms use the competitor rates along with the RCP decision. These also use price distance or price position to arrive at the appropriate pricing. However, the ARI and MPI pricing methods calculate the price ranking as well and use it as an input to the pricing algorithm.

PRICE RANKING

Price ranking is the process of ranking each rates floated by the identified competition in the market. For the purpose of MPI Pricing the rates are ranked in descending order and for ARI Pricing the rates are ranked in ascending order. These ranks are used as an input for calculating the weights while arriving at the hotel pricing. Since the very meaning of ARI is “Average Rate Index”, this algorithm tries to push the rates towards the higher set of rates floated in the market. Therefore, in order to assign higher weights to the higher rates, these rates are ranked in Ascending Order. Conversely, MPI means “Market Penetration Index” where the hotel is interested in gaining more occupancy rather than achieving higher rates. Hence, the rates are ranked in Descending Order in order to assign higher weightage to the lower set of rates.

Here is an illustration that describes the Ranking Methodology

Occupancy Dates	Hotel Aurora Towers	Crowne Plaza Pune City Centre	Shantai Hotel	Hotel Sagar Plaza	The Central Park Hotel
5-Dec-16	2,728.60	4,050.00	3,199.00	3,199.00	3,250.00
6-Dec-16	2,679.09	4,050.00	3,199.00	3,199.00	3,250.00
7-Dec-16	3,988.55	4,050.00	2,799.00	3,199.00	3,250.00
8-Dec-16	4,077.47	6,480.00	2,799.00	4,199.00	10,320.00
12-Dec-16	2,646.09	4,500.00	2,799.00	3,199.00	3,250.00
15-Dec-16	2,618.71	4,995.00	2,799.00	3,199.00	3,250.00
16-Dec-16	2,637.21	3,780.00	2,639.00	3,199.00	2,950.00

Table 7. Sample table with the rates floated by different hotels

Ranking for ARI					
Occupancy Dates	Hotel Aurora Towers	Crowne Plaza Pune City Centre	Shantai Hotel	Hotel Sagar Plaza	The Central Park Hotel
5-Dec-16	1	5	3	3	4
6-Dec-16	1	5	3	3	4
7-Dec-16	4	5	1	2	3
8-Dec-16	2	4	1	3	5
12-Dec-16	1	5	2	3	4
15-Dec-16	1	5	2	3	4
16-Dec-16	1	5	2	4	3

Table 7a. Sample table with the ranks calculated for each occupancy date for each hotel

Ranking MPI					
Occupany Dates	Hotel Aurora Towers	Crowne Plaza Pune City Centre	Shantai Hotel	Hotel Sagar Plaza	The Central Park Hotel
5-Dec-16	5	1	4	4	2
6-Dec-16	5	1	4	4	2
7-Dec-16	2	1	5	4	3
8-Dec-16	4	2	5	3	1
12-Dec-16	5	1	4	3	2
15-Dec-16	5	1	4	3	2
16-Dec-16	5	1	4	2	3

Table 7b. Sample table with the ranks calculated for each occupancy date for each hotel

PRICE DISTANCE

Price distance can be called as price position of all the identified competition hotels vis-à-vis the Client Hotel rates. The Price Position is calculated as –

$$\rho = \frac{(\mathrm{h} - c_i)}{\mathrm{h}} \tag{19}$$

Where

h – Client Hotel Rate

c_i- Competitor Rate

The Client hotel will get ZERO value as the price positioning is in relation to it. After applying this to all the competitors for each day, the output looks something like –

Pricing Position					
Occupancy Dates	Hotel Aurora Towers	Crowne Plaza Pune City Centre	Shantai Hotel	Hotel Sagar Plaza	The Central Park Hotel
5-Dec-16	0	0.48	0.17	0.17	0.19
6-Dec-16	0	0.51	0.19	0.19	0.21
7-Dec-16	0	0.02	-0.30	-0.20	-0.19
8-Dec-16	0	0.59	-0.31	0.03	1.53
12-Dec-16	0	0.70	0.06	0.21	0.23
15-Dec-16	0	0.91	0.07	0.22	0.24
16-Dec-16	0	0.43	0.00	0.21	0.12

Table 8. Sample table with the pricing position calculated for each occupancy date for each hotel

WEIGHTS CALCULATION FOR ARI & MPI

The weights are calculated using the Rankings and Price Positioning values. The idea is to club the Rate Ranking and Price Positioning such that the pricing decisions do not get overly skewed with the rankings as well as rates. This is essential to ensure that the pricing decisions are not very low (or lower than the lowest) in MPI Pricing Algorithm and not very high (or higher than the highest) in ARI Pricing Algorithm. These two algorithms always try to identify the sweet point that has the highest salability. The thumb rule of pricing says, "Lower Rates have higher probability". It is not just that, in today's world of OTAs, online search engines and aggregators, the lower priced hotels in the identified competition set end up appearing on top of the Client Hotel due to their pricing positions (though there are many other factors that decide the hotel positioning on the OTAs but the pricing is one of the major inputs to the hotel positioning on these portal). Thereby, these hotels stand a chance to eat up the demand from the client hotel. Hence, the weights are calculated in such a way that the hotels floating lower rates than the client hotel get slightly higher weightage over the hotels floating higher rates.

The weights are calculated by the formula –

$$f(\omega) = \begin{cases} (\gamma - \rho), & \rho < \gamma \\ (\rho - \gamma), & \rho > \gamma \end{cases} \tag{20}$$

Where

ρ – Price Position

γ – Rate Ranking

Here is an illustration of the output after applying this formula for all the future dates –

ARI Weights					
Occupancy Dates	Hotel Aurora Towers	Crowne Plaza Pune City Centre	Shantai Hotel	Hotel Sagar Plaza	The Central Park Hotel
5-Dec-16	1	4.52	2.33	2.33	3.81
6-Dec-16	1	4.49	2.31	2.31	3.79
7-Dec-16	4	4.98	1.30	2.20	3.19
8-Dec-16	2	3.41	1.31	2.97	3.47
12-Dec-16	1	4.30	1.94	2.79	3.77
15-Dec-16	1	4.09	1.93	2.78	3.76
16-Dec-16	1	4.57	2.00	3.79	2.88

Table 9a. Sample table with the weights for ARI Pricing Algorithm

MPI Weights					
Occupancy Dates	Hotel Aurora Towers	Crowne Plaza Pune City Centre	Shantai Hotel	Hotel Sagar Plaza	The Central Park Hotel
5-Dec-16	5	0.52	3.33	3.33	1.81
6-Dec-16	5	0.49	3.31	3.31	1.79
7-Dec-16	2	0.98	5.30	4.20	3.19
8-Dec-16	4	1.41	5.31	2.97	0.53
12-Dec-16	5	0.30	3.94	2.79	1.77
15-Dec-16	5	0.09	3.93	2.78	1.76
16-Dec-16	5	0.57	4.00	1.79	2.88

Table 9b. Sample table with the weights for MPI Pricing Algorithm

WEIGHTED AVERAGE[5] OF RATES

In order to arrive at the pricing decisions, all the three algorithms use the Weighted Average. The weights calculated at the above step are applied to the hotel rates to calculate the weighted averages.

$$\bar{\mu} = \frac{\sum_{i=1}^{n} w_i r_i}{\sum_{i=1}^{n} w_i} \tag{21}$$

Where

w – Weight for the hotel

r – Rate offered by the hotel

$\bar{\mu}$ - Weighted Average

[5] https://en.wikipedia.org/wiki/Weighted_arithmetic_mean

POISSON DISTRIBUTION[6]

In order to derive the rate for a date, the weighted average is passed though the Poisson distribution. The Poisson distribution is used considering the below given important aspects.

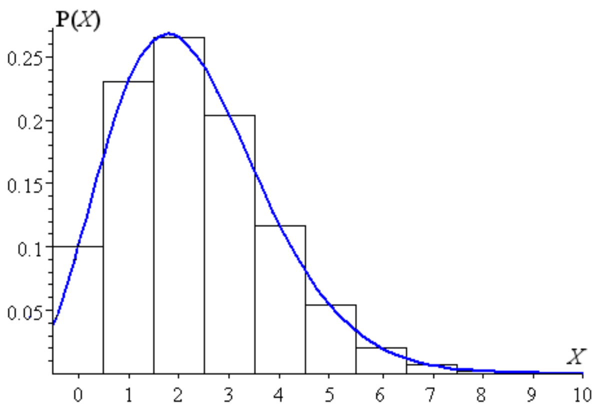


Figure 4. Sample Poisson Distribution

1. A guest booking a hotel for a particular rate does not depend on the fact that other guests have booked the same hotel room for either higher rate or lower rate. Therefore, the guest booking behavior follows the Poisson distribution.
2. The Cumulative Density Function[7] in Poisson distribution does not use the Standard Deviation for calculating the probability. Since there will be very few observations in the distribution of a typical Competition set, it will be very inappropriate to calculate a standard deviation. Hence, the Poisson Distribution is used to arrive at the rate that has higher probability in the given competition set based on their pricing.
3. The rate actualization does not necessarily follow a symmetrical distribution like the Normal Distribution. Most of the time it has been observed that the demand materialization bound on one side and not the other side. Most of the hotel guests tend to book when rates are on the lower side. this phenomenon skews the distribution towards the left.

The cumulative density function for the Poisson Distribution is given by

$$f(x;\mu) = \sum_{x=0}^{n} \frac{e^{-\mu}\mu^{x}}{x!} \qquad (22)$$

Where

e - The base of natural logarithms or Euler's[8] constant (2.7183)

The algorithm uses the inverse CDF in pricing by simultaneously solving the formula 20 to obtain a rate that maximizes the revenue while having maximum probability.

---

[6] https://en.wikipedia.org/wiki/Poisson_distribution
[7] https://en.wikipedia.org/wiki/Cumulative_distribution_function
[8] https://en.wikipedia.org/wiki/E_(mathematical_constant)

Here is an illustration of the results of the formula after solving simultaneously for each of the day in future –

ARI Rate				
Occupancy Dates	Avg	Opt Rate	Rev	Suggested Rate
5-Dec-16	3,454.13	3,307.00	3,288.00	3,299.00
6-Dec-16	3,450.51	3,305.00	3,284.51	3,299.00
7-Dec-16	3,648.59	3,497.00	3,477.22	3,499.00
8-Dec-16	6,244.94	6,039.99	6,012.81	6,049.00
12-Dec-16	3,521.80	3,372.00	3,353.74	3,349.00
15-Dec-16	3,655.40	3,506.00	3,483.84	3,499.00
16-Dec-16	3,216.87	3,077.00	3,057.36	3,099.00

Table 10a. Sample table with the pricing decisions from ARI Pricing Algorithm

MPI Rate				
Occupancy Dates	Avg	Opt Rate	Rev	Suggested Rate
5-Dec-16	3,068.75	2,934.00	2,913.43	2,949.00
6-Dec-16	3,048.29	2,913.00	2,893.60	2,899.00
7-Dec-16	3,228.37	3,086.00	3,068.42	3,099.00
8-Dec-16	4,096.58	3,934.00	3,913.56	3,949.00
12-Dec-16	2,919.26	2,799.00	2,764.57	2,799.00
15-Dec-16	2,887.97	2,757.00	2,737.96	2,749.00
16-Dec-16	2,817.05	2,711.00	2,651.80	2,699.00

Table 10b. Sample table with the pricing decisions from MPI Pricing Algorithm

The figure 5a and 5b compare the Pricing decisions along with the competition rate band. The blue patch is the rate band that the competition is floating. And the yellow line depicts the rates that the pricing algorithms have derived.

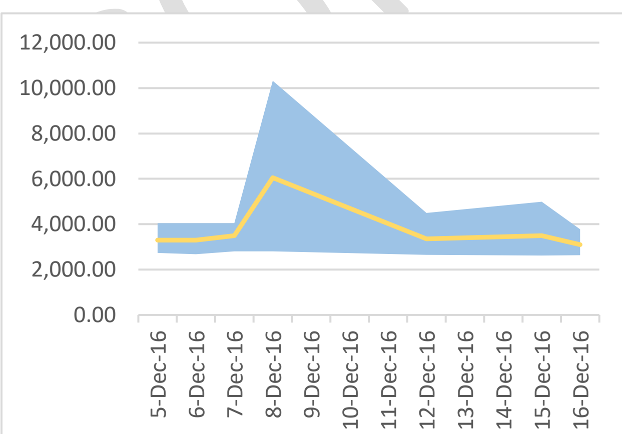


Figure 5a. ARI Pricing decision with competition rates band

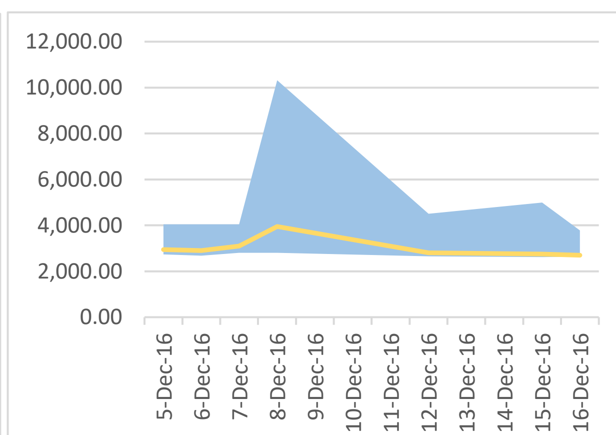


Figure 5b. MPI Pricing decision with competition rates band

From both the graphs, it is very evident that the ARI pricing does not out price the hotel in the market and MPI Pricing does not low price the hotel in the market. In both the cases the hotel prices will be very much within the competition pricing band.

The figure 6 compares the pricing decisions from the three pricing algorithms (i.e. RCP, ARI and MPI)

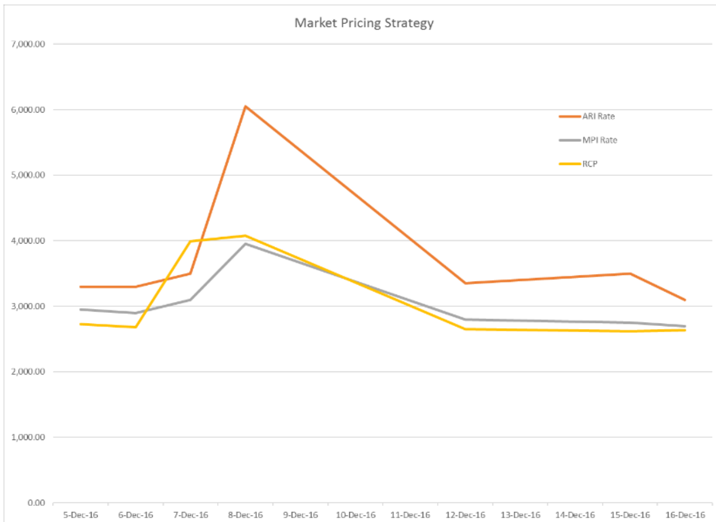


Figure 6. Comparison of Pricing decision

The RCP is more based on the hotel specific demand dynamics. The pricing decisions derived from RCP algorithm is used as an input to the MAP Algorithms. The major difference between the MAP and RCP is the inputs that these two algorithms use. The MAP gives more stress on the market data even though it folds in the hotel specific details.

PQM PRICING

In today's digital world, the social media sentiments can manipulate most of the decision making. Most of the times it has been observed that the hotel customers review the star rankings along with the hotel reviews and experiences of the guest stayed at the hotel before, before making any buying decisions. Therefore, it is very much imperative even in hotel room bookings as well to maintain better social sentiments for the hotel. Hence, the PQM Algorithm considers these social sentiments while deciding the pricing points for the client hotel for all the future dates.

The PQM Pricing is the third algorithm that is part of MAP. The PQM stands for "Price Quality Matrix". This algorithm uses other sentimental inputs like Ranking on each of the OTAs, etc. The sentimental inputs are considered as they influence the buying decisions of the hotel guests. The sentimental inputs are used to derive the "Quality Index" (QI) in this algorithm. For calculating QI the algorithm uses "Star Rating" calculated by each OTAs for each of the hotels (both client hotel and competition hotels) as an input. The Star Rating on various OTAs is based on various factors like –

1. Content Score
    a. The write up about the hotel
    b. The pictures of the hotel in high definition (HD)
2. Review Score
    a. Total number of reviews
    b. Number of favorable and non-favorable reviews
3. Room Availability on the channel
4. Rate Parity – Hotel not having different rates on different OTAs
5. Inventory Parity – Hotel offering same room types across all the OTAs

A simple average of the Star Ratings from all the major participating OTAs[9] is used as the Quality Index.

Even in this algorithm, the pricing position or price distance is calculated in the similar way as in other two MAP algorithms (ARI & MPI). Hence, the details of price position calculations are not discussed here again.

[9] Booking.com, Expedia, MakeMyTrip, Goibibo, ClearTrip, etc. for the Indian hotels

MARKET POSITION

Similar to Price Positioning, the Market Position or QI is also calculated as position of all the identified competition hotels vis-à-vis the Client Hotel Star Rating. The average of the Star Ratings is used for each of the competition hotels.

$$\varphi = \frac{(\mathrm{h} - c_i)}{\mathrm{h}} \tag{23}$$

Where

h – Clint Hotel Star Rating

c_i- Competitor Star Rating

The Client hotel as well as the competition hotels having similar Average Start Rankings will get ZERO value as the market positioning is in relation to the client hotel. After applying this to all the competitors for each day, the output looks something like –

Market Position					
Occupancy Dates	Hotel Aurora Towers	Crowne Plaza Pune City Centre	Shantai Hotel	Hotel Sagar Plaza	The Central Park Hotel
5-Dec-16	0	0.00	-0.25	-0.13	0.00
6-Dec-16	0	0.00	-0.25	-0.13	0.00
7-Dec-16	0	0.00	-0.25	-0.13	0.00
8-Dec-16	0	0.00	-0.25	-0.13	0.00
12-Dec-16	0	0.00	-0.25	-0.13	0.00
15-Dec-16	0	0.00	-0.25	-0.13	0.00
16-Dec-16	0	0.00	-0.25	-0.13	0.00

Table 11. Sample table with the market position calculated for each occupancy date for each hotel

PQM WEIGHTS

The weights are calculated using the QI and Price Positioning values. The idea is to club the sentimental factors that affect the customer buying behavior and Price Positioning such that the pricing decisions are influenced by the market dynamics along with the factors influencing the buying decisions of the customer. As in other two algorithms of MAP, even this algorithm ensures that the pricing decisions are not very low (or lower than the lowest) and not very high (or higher than the highest). It always tries to identify the sweet point that has the highest salability.

As discussed earlier all the OTAs have their own technique of giving a star ranking to each of the hotels, and these rankings play a very significant role in influencing the hotel clients buying behavior. Hence, these rankings are used along with the price positions to calculate the weights in such a way that the market sentiments influence the hotel pricing.

The weights are calculated by the formula –

$$f(\omega) = \varphi - \rho + H \quad (24)$$

Where

ρ – Price Position of each hotel

φ – Market Position of each hotel

H – Client Hotel Start Rating

Here is an illustration of the output after applying this formula for all the future dates –

PQM Weights					
Occupancy Dates	**Hotel Aurora Towers**	**Crowne Plaza Pune City Centre**	**Shantai Hotel**	**Hotel Sagar Plaza**	**The Central Park Hotel**
5-Dec-16	4.00	3.52	3.58	3.70	3.81
6-Dec-16	4.00	3.49	3.56	3.68	3.79
7-Dec-16	4.00	3.98	4.05	4.07	4.19
8-Dec-16	4.00	3.41	4.06	3.85	2.47
12-Dec-16	4.00	3.30	3.69	3.67	3.77
15-Dec-16	4.00	3.09	3.68	3.65	3.76
16-Dec-16	4.00	3.57	3.75	3.66	3.88

Table 12. Sample table with the weights for PQM Pricing Algorithm

In order to arrive at the pricing decision, the weighted average of rates floated by all the identified completion hotels are considered while applying the PQM weights. The algorithm uses the same method of calculating Weighted Averages discussed earlier. The weighted averages are then used as input to the Poisson distribution the same way discussed earlier. The results of these two steps are shown on table 13 and figure 7.

PQM Rate				
Occupancy Dates	Average	Opt Rate	Rev	Suggested Rate
5-Dec-16	3,269.12	3,129.00	3,108.12	3,149.00
6-Dec-16	3,257.45	3,117.00	3,096.80	3,099.00
7-Dec-16	3,452.47	3,310.00	3,286.19	3,299.00
8-Dec-16	5,138.81	4,955.00	4,930.82	4,949.00
12-Dec-16	3,242.21	3,113.00	3,078.77	3,099.00
15-Dec-16	3,306.35	3,278.00	2,266.20	3,299.00
16-Dec-16	3,027.15	2,891.00	2,873.03	2,899.00

Table 13. The sample pricing decisions from PQM Algorithm

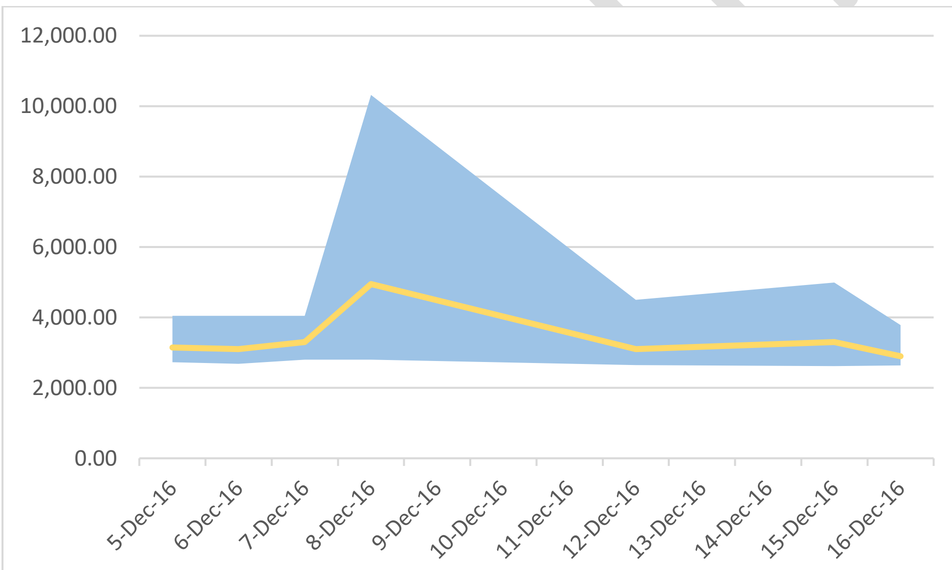


Figure 7. PQM Pricing decision with competition rates band

COMPARISON

All the four pricing algorithms serve different purpose of the pricing tactics.

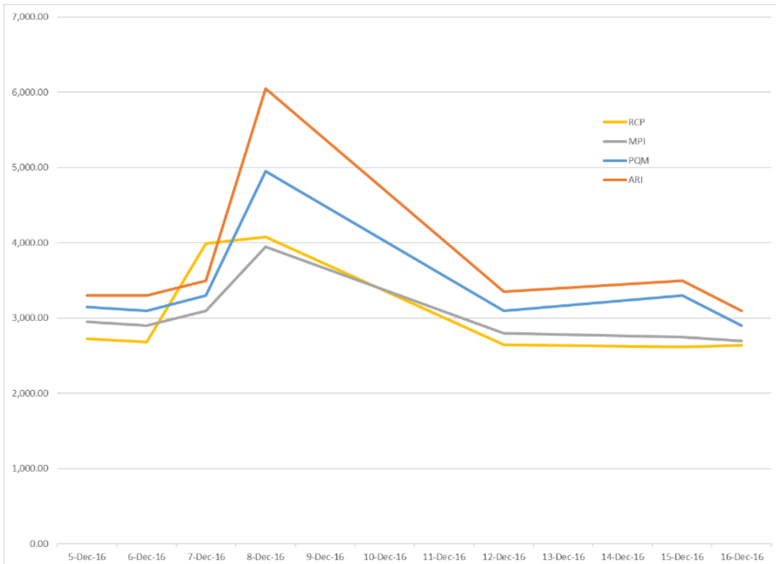


Figure 8. Comparison of all Pricing decisions

1. RCP – The remaining capacity based pricing considers only the hotel condition and decides the pricing. This is typically useful when the prevalent market conditions are not applicable with the hotel demand dynamics. The hotel should adopt to RCP when the hotel is either very busy or very lull as contrary to the competition.
2. MPI – The MPI based pricing is useful when the hotel is not very busy and in line with the market. And, the hotel wants more foot falls or “heads on the beds”.
3. ARI – The ARI based pricing is useful when the hotel is busy and in line with the market. And, the hotel wants to yield on its position.
4. PQM – The PQM based pricing is useful when the hotel wants to retain its pricing in line with its social sentiments.

On figure 9, all the pricing decisions are compared with the competition pricing trends. It can be observed that all the rates from MAP pricing algorithms are influenced by the market trends and the RCP rates are not influenced by the market trends and these pricing decisions are purely governed by the demand conditions existing at the hotel.

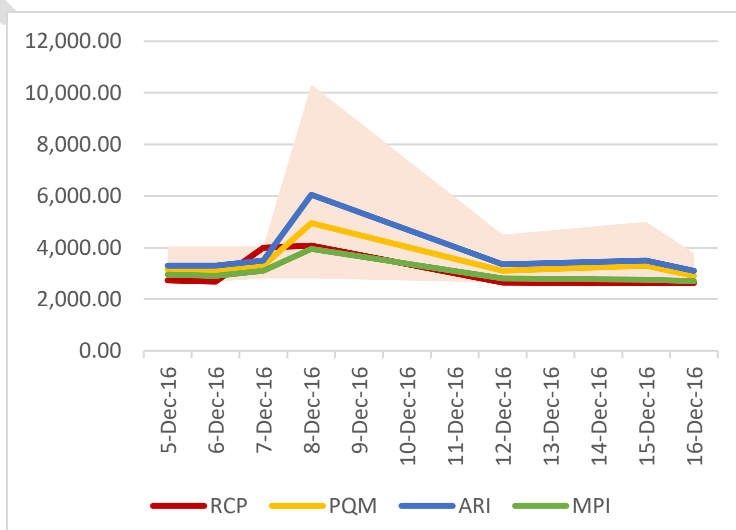


Figure 9. Comparison of all Pricing decisions with Market Trends

CONCLUSIONS

All the four pricing algorithms will be available to the Revenue management team. The Revenue Management team will review and decide which of the pricing recommendations to use for deploying as a rate of the day for any future date. The rate decided based on the PQM algorithm will be used as the default recommendation by the system. This method acts as a median of all other MAP pricing recommendation and in the dynamic world where the hotels have to match with each move of their competitor the approach of considering the price that is neither tool low and nor too high would always be beneficial in the long run. However, the RM team will have an option to override this rate with either one of the recommendations provided by other algorithms or it can override with completely different rate based on the inputs from the hotel and other market intelligence that the team has gathered.