WIN PREDICTION ANALYSIS

Kenny D

Table of Contents

Problem Statement

An IT Consulting company needs analysis and prediction of win possibility for potential client deals using variable combinations in order to optimally allocate staff and resources, to mount the company's profitability & meet the growth targets.

- Our company earns revenue by providing IT solutions to clients from various sectors through successful deals.
- In this extremely competitive market, it is vital for company to maximize the successful deals count by exercising caution and prudence in allocation of resources for bidding proposals.
- As a result, our company requires a machine learning (ML) model for making predictive analysis around the defined problems assuring maximum accuracy.

Objective

Objective 1: Predictive Analytics - Build a ML model to predict the probability of win/loss for bidding activities for a potential client.

- In this project through exploratory data analysis & predictive algorithms, we will endow various combinations using the given criteria that enhances the possibility of winning a deal.

Objective 2: Prescriptive Analytics – Identify variable/s that are most likely to help in converting an opportunity into a win.

- The project also aims to predict finest VP & Manager combinations for allocation of deals in order to improve the success rate in bids .
- Further we will evaluate the total amount of Deal Cost Lost due to incorrect prediction which can be used as reference for upcoming deals success predictions.

Data Description

Data set contains 1 target variable (Deal status code), 8 Deal criteria & 10061 closed deals data for period of 8 years.

Column Name	Data Type	Description
Client Category	Categorical	Industry in which the client works
Solution Type	Categorical	The solution group the client requires
Deal Date	Date	The date the opportunity was created (Pitch date)
Sector	Categorical	The sector for which the solution is to be provided
Location	Categorical	Client location
VP Name	Categorical	Vice President – main contact dealing with the client
Manager Name	Categorical	Manager of the team working on the project
Deal Cost	Numerical	The initial cost of the deal (assuming currency as USD)
Deal Status Code (Target)	**Categorical**	**Final status of the deal (Won/Lost)**

Data Cleaning

Before working on the data, few preliminary checks were done to weed out any data discrepancy.

Criteria	Number of Rows	Action Taken	Rationale
Deal Cost = 0	246	Deals removed.	As data is taken from CRM framework, an entry of 0 is not possible. Hence its an error.
Duplicate deals	13	Deals removed.	Duplicate deals might be due to a clerical error. Such entries would incorrectly skew the results.
Client Category = NA	73	Replaced with “Unknown”	Imputing category as mode might skew the data towards that category. Hence, better to rename it to avoid loss of data.
Remaining Number of Deals	9802	EDA & Model Building	Data set with 9802 rows of closed deals was used in ML model building.

EXPLORATORY DATA ANALYSIS

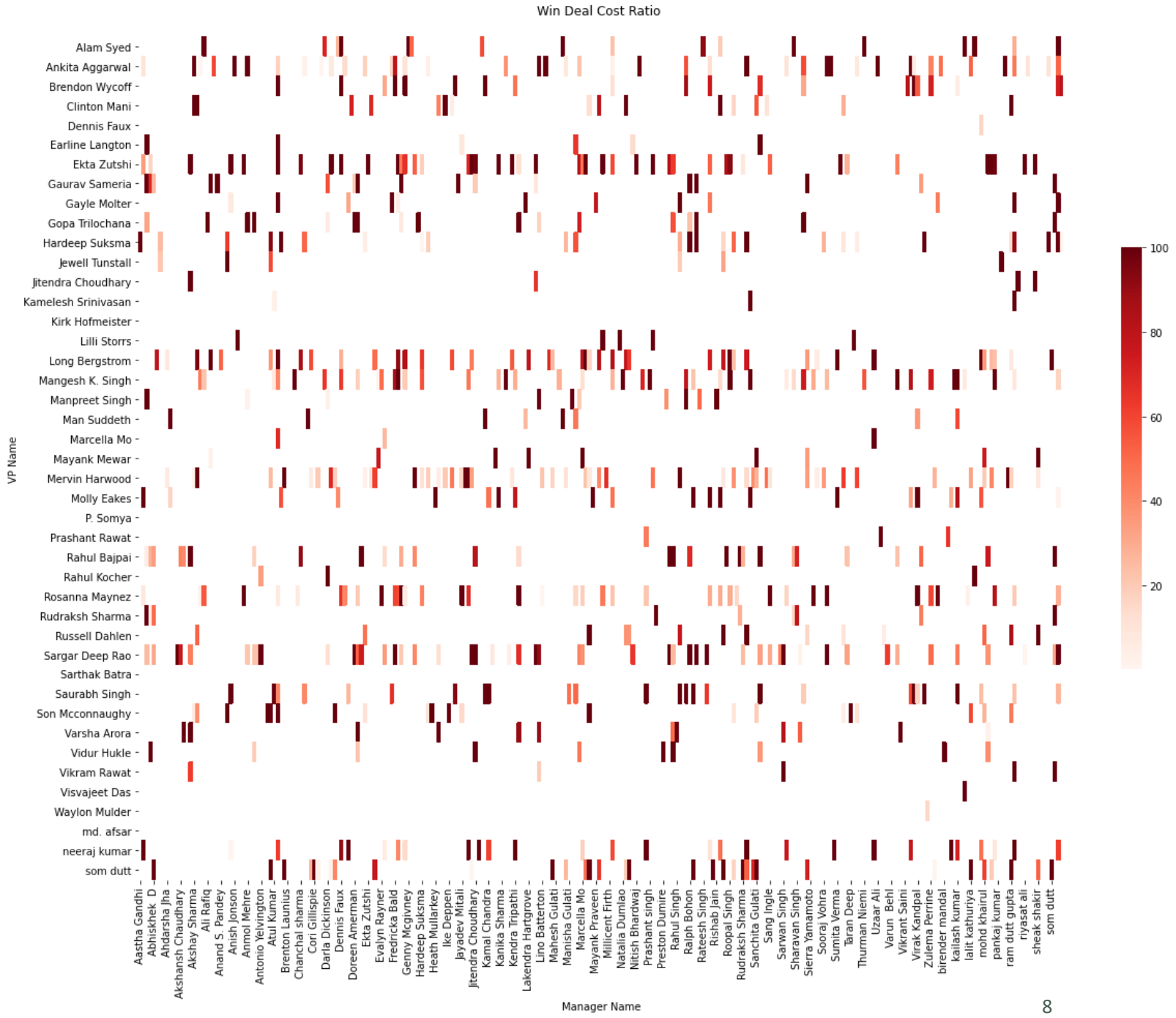
Column Name	Unique Values
Client Category	42
Solution Type	67
Deal Date	2530
Sector	25
Location	13
VP Name	43
Manager Name	276
Deal Cost	1468
Deal Status Code	2

Won Deals Analysis

- Adjacent Heatmap derives there are numerous combinations who have positively converted more than 50% of their deals.
- Total number of deals won are 3751 & Sum of Won deals cost is $ 2,947,958,825 which is 38.25 % of total deals .
- Total Sum of deals won is further bifurcated in to 4 ranges for detail analysis.

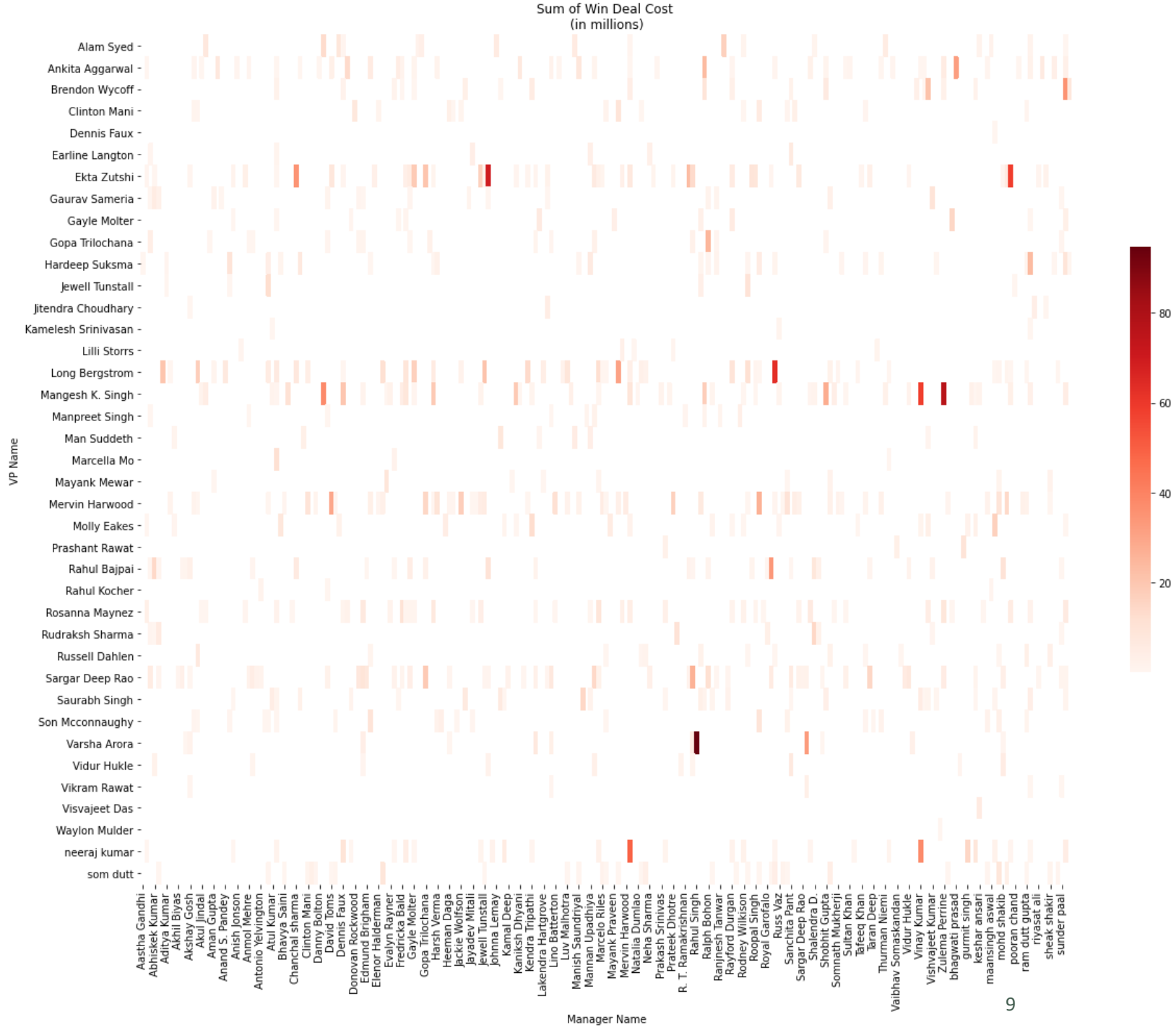
Sum of won deals cost range	VP & Manager Combination	Deals Won
0-25	423	836
26-50	126	1012
51-75	96	937
76-100	307	966

- The combinations with ratio above 0.5 can be given more deals as they have potential.



VP & Manager Combinations

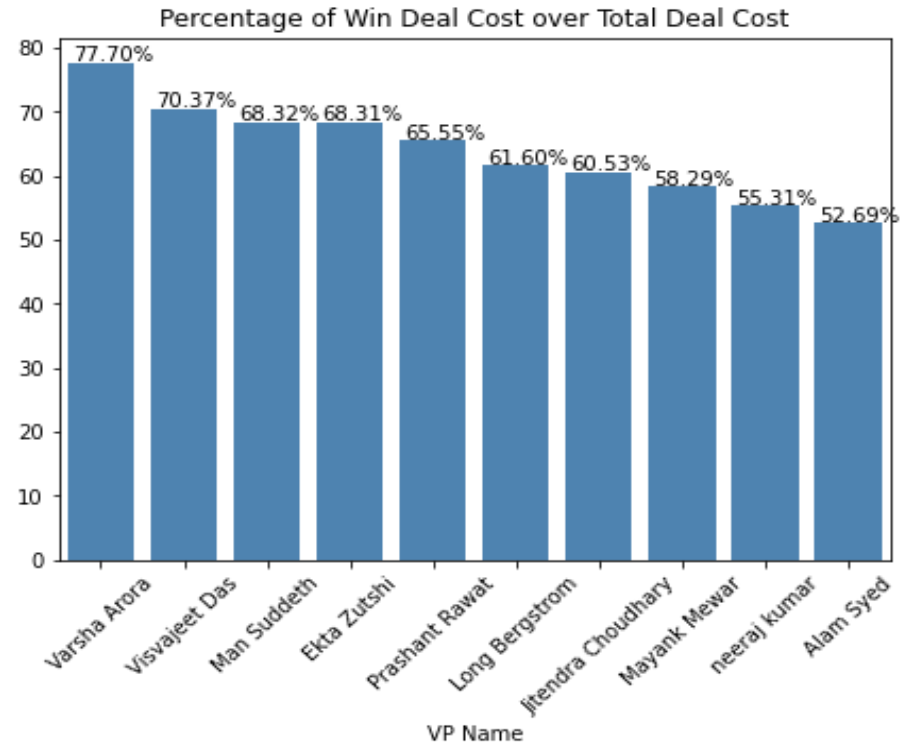
- Data Contains a total of 960 unique combinations of VP & Managers.
- Our Company hired 43 VP's & 276 Managers for contribution of company revenue.
- Data reveals very few VP and Manager combinations are highly successful.
- The Dark red bars are VP Manager combinations who have brought high Deal Revenue to the company.
- Top 5 combination of VP & Managers contribute to 3.15% of won deals.
- From the horizontal bars, we see that VP have more weight over Winning a deal than does Manager.



Win Deal Cost Percentage of VP Name

- We have further singled out the top 5 VP's on basis of highest revenue generation.

VP Name	Manager Name	Win Deal Total Costs $ ' Mn	Number of Win Deals
Varsha Arora	Rahul Singh	94.70	6
Mangesh K. Singh	Zulema Perrine	76.65	9
Ekta Zutshi	Jitendra Choudhary	69.73	11
Long Bergstrom	Russell Dahlen	63.15	75
Ekta Zutshi	Neeraj Kumar	58.52	40



VP Name and Year Combinations

Some VPs have made no deals in year 2018 and 2019.

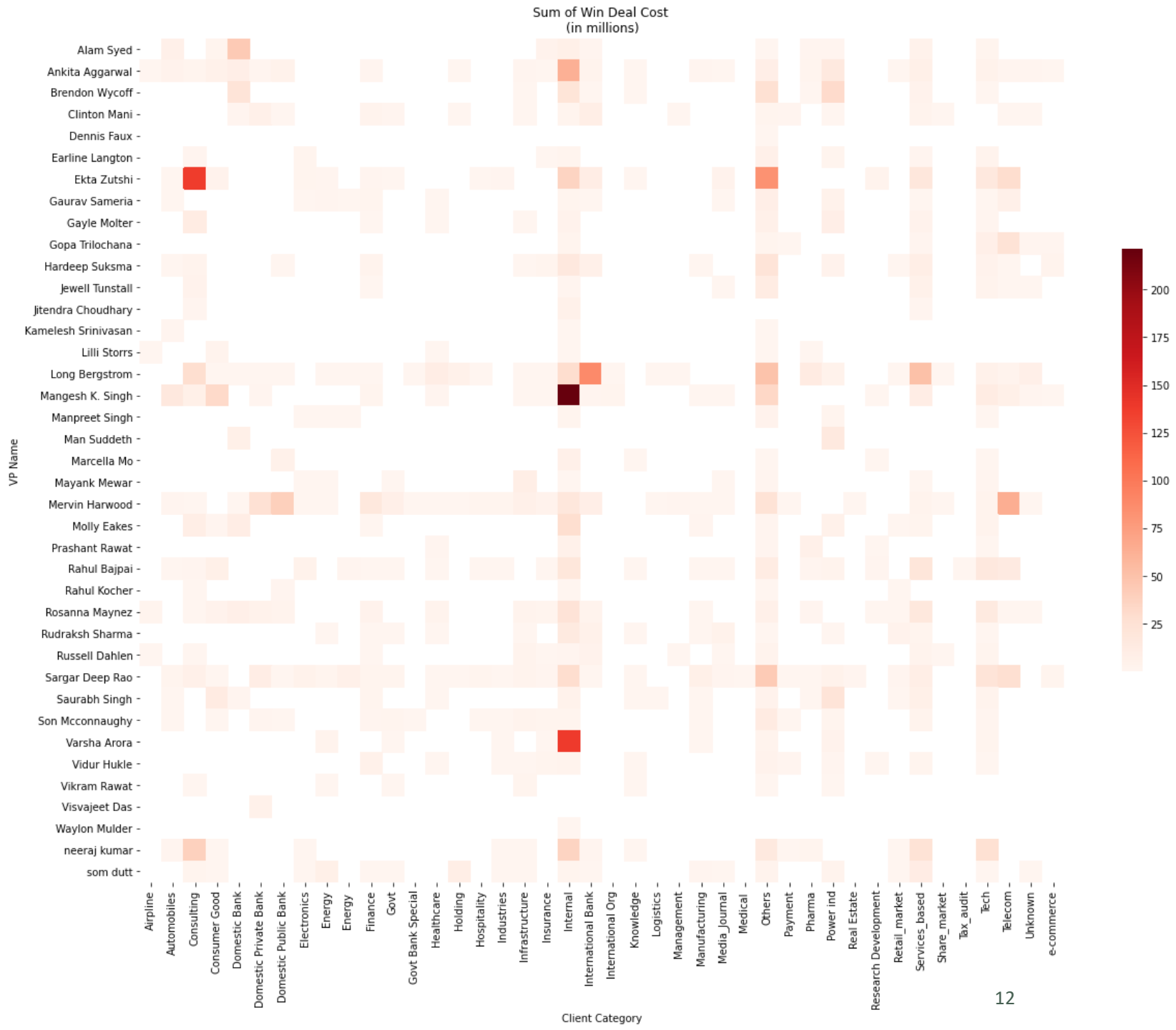
Two inferences

- VP has left the company
- VP is on sabbatical

VP Name	Year	Win Deal Total Cost $ ' Mn	Number of Win Deals
Varsha Arora	2018	111.72	24
Ekta Zutshi	2013	80.00	88
Mangesh K. Singh	2015	75.59	53
	2016	73.96	34
Ekta Zutshi	2014	70.04	73

VP Name & Client Category Combinations

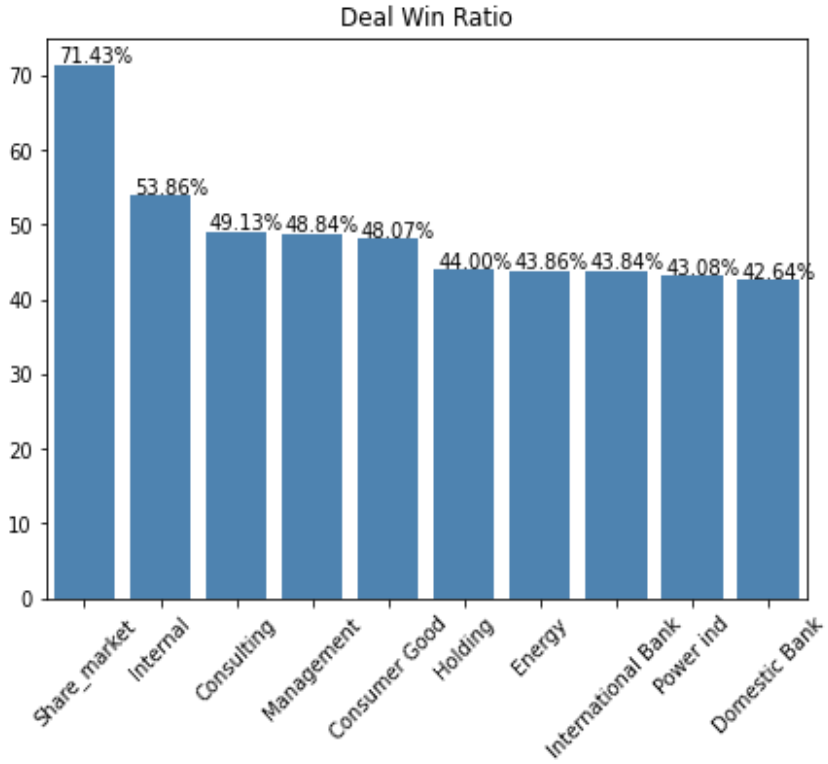
- Our company successfully provides solutions to 42 unique clients.
- Client Categories such as Internal, Others, Tech, Consulting and Service Based have brought in high revenue to our organization.
- Mangesh K Singh has brought in most Revenue from Internal category deals which has a deal win ration of 53%.
- Further to the analysis, we recommend our top performing VP's to be delivered to these client bids to increase the chance of winning.



Win Deal Cost Percentage of Top 5 VP in Client Category

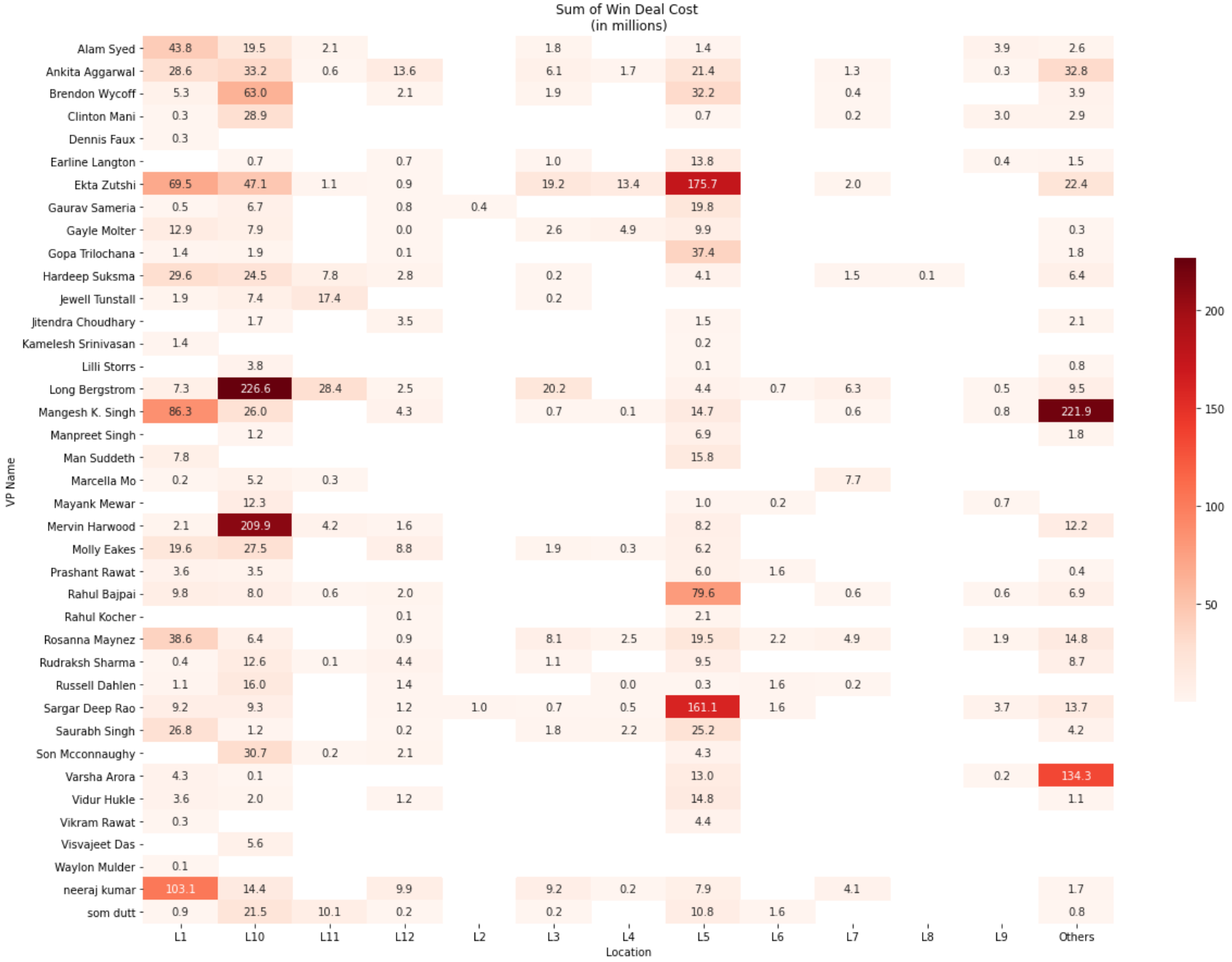
➢ Below mentioned are the clients who have generated high revenue to company along with respective VP's and percentage of their deals on basis of total won deals .

VP Name	Client Category	Win Deal Total Costs $ ' Mn	Number of Win Deals
Mangesh K. Singh	Internal	221.12	107
Varsha Arora	Internal	138.55	19
Ekta Zutshi	Consulting	137.87	52
Long Bergstrom	International Bank	88.56	84
Ekta Zutshi	Others	83.04	108



VP Name and Location Combinations

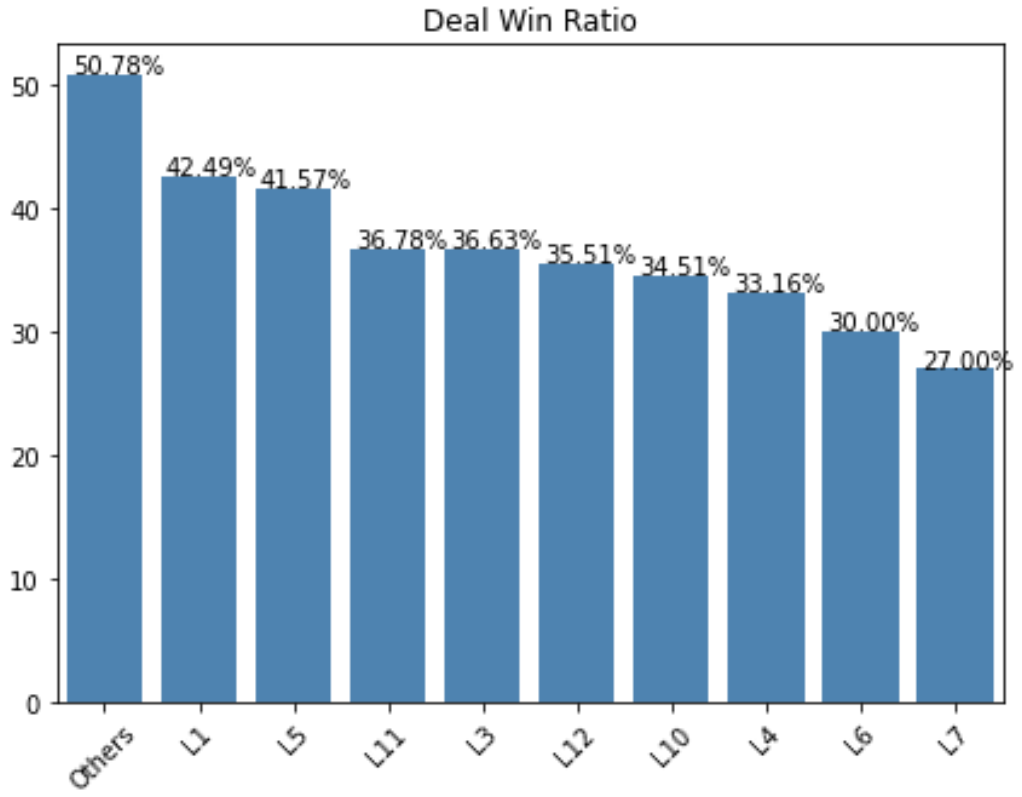
- L1, L10 and L5 seem to be the Win Deal Hot spots.
- We have count of 2726 deals won in these sectors out of 3715 total won deals which contributes to 73% success rate.
- Year 2018 stands out with successful deal count of 627 from these sectors.



Win Deal Cost Percentage of Top 5 VP in Location

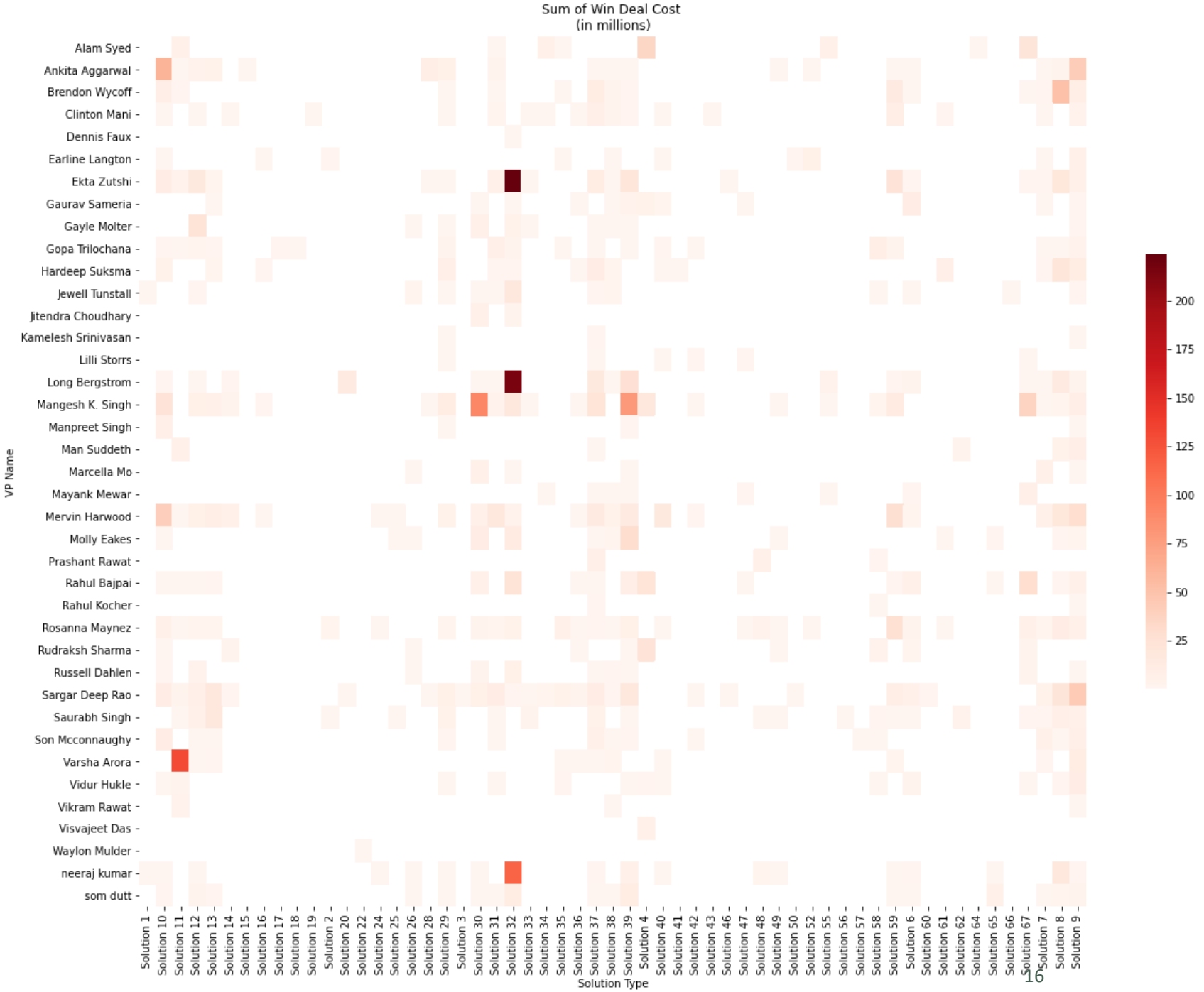
➢ We have further singled out the top 5 VP's with respect to Location on basis of highest revenue generation.

VP Name	Location	Win Deal Total Cost $ ' Mn	Number of Win Deals
Long Bergstrom	L10	226.56	270
Mangesh K. Singh	Others	22.89	94
Mervin Harwood	L10	209.94	312
Ekta Zutshi	L5	175.68	119
Sargar Deep Rao	L5	161.06	249



VP Name and Solution Type Combinations

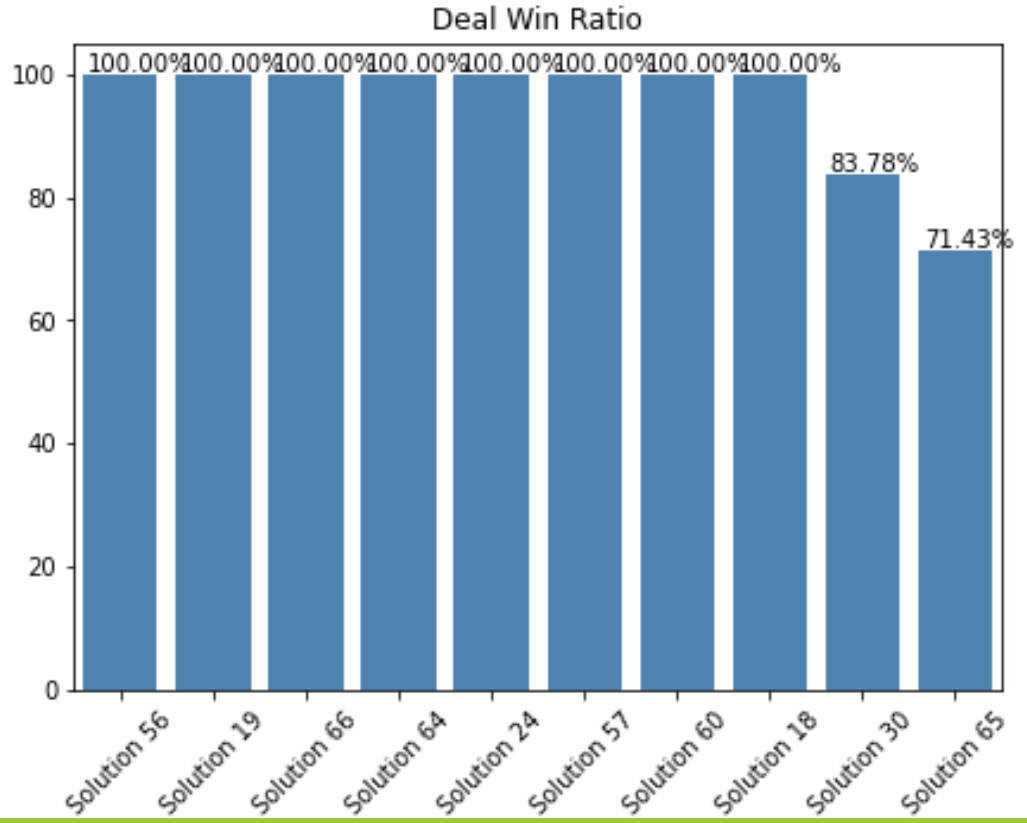
- 57 unique solutions have win deals on a total of 67 solution types.
- Solution 32 has brought in the highest Revenue.
- It has the highest count of 813 win deals over a period of 8 years.
- Solutions from 30 to 32, 39 to 39 and Solution 7 to 9 are better performers than others.



Win Deal Cost Percentage of Top 5 VP in Solution Type

➢ We have further singled out the top 5 VP's with respect to Solution on basis of highest revenue generation.

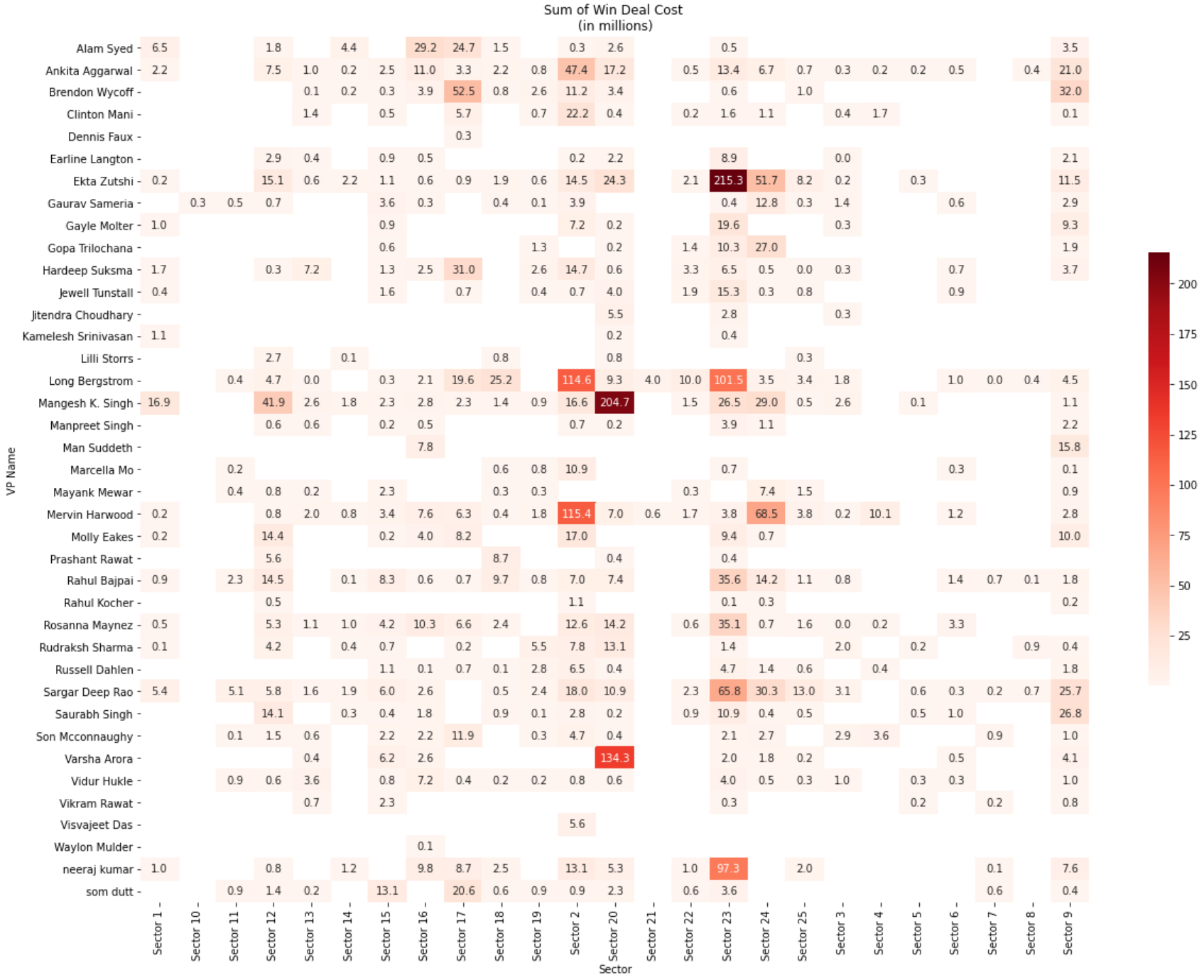
VP Name	Solution Type	Deal Cost $ ' Mn	Number of Deals
Ekta Zutshi	Solution 32	224.17	197
Long Bergstrom	Solution 32	215.58	261
Varsha Arora	Solution 11	130.77	19
neeraj kumar	Solution 32	116.30	172
Mangesh K. Singh	Solution 30	91.31	14



VP Name & Sector Combinations

- All the 25 sectors have given successful bids in providing solutions to the clients
- Top 5 sectors and their win deals analysed from the adjacent heat map are as below:

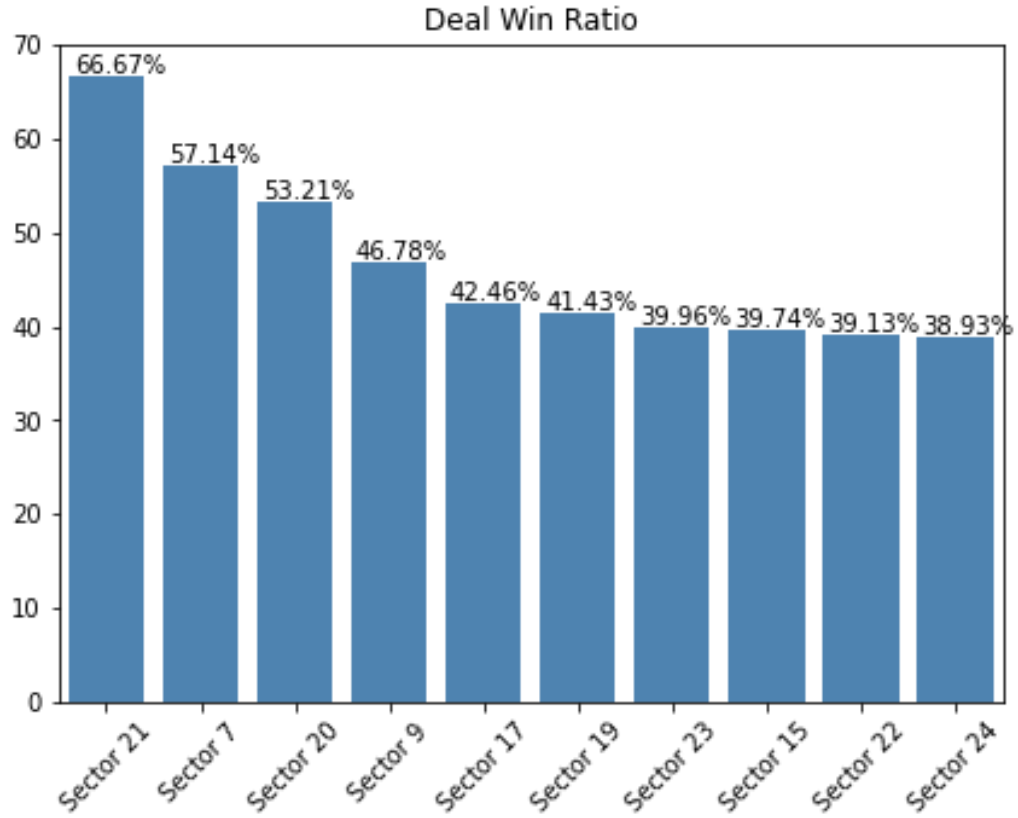
Sector	Won Deals Count
Sector 23	1053
Sector 2	596
Sector 20	389
Sector 24	217
Sector 12	201



Win Deal Cost Percentage of Top 5 VP in Sector

➢ We have further singled out the top 5 VP's with respect to Sector on basis of highest revenue generation.

VP Name	Sector	Deal Cost $ ' Mn	Number of Deals
Ekta Zutshi	Sector 23	215.30	191
Mangesh K. Singh	Sector 20	204.69	80
Varsha Arora	Sector 20	134.28	17
Mervin Harwood	Sector 2	115.43	184
Long Bergstrom	Sector 2	114.63	125



Top 5 VP- Manager Combinations

- Combining inferences from the EDA, we analyze major factors contribute to the efficiency of a VP-Manager Head combinations.
- These are the number of deals transacted , Sum of Cost of Deals Won & Time period .
- Below table contains those top 5 VP-Managers that balance the criteria

VP - SBU Head Combination	Number of Deals Won	Won Deal Cost ($' Mn)
Long Bergstrom & Russell Dahlen	75	63.2
Neeraj Kumar & Molly Eakes	62	49.0
Mangesh K.Singh & Darin Vath	35	37.9
Neeraj Kumar & Vinay Kumar	51	37.5
Rahul Bajpai & Rudraksh Sharma	72	33.8

Model Building

- Feature Engineering
  - Frequency Encoding - All variables are frequency encoded considering the high number of sub categories within each categorical variable.
  - Standard Scaling - Variables are scaled to similar range to avoid skew.

- Train Test Split - Data frame is split into Train and Test in 75% : 25% ratio.

- Models Built - Multiple Algorithms are used to check its respective performance.
  Few of these are – Random Forest, Decision Tree , K Nearest Neighbour & SVM

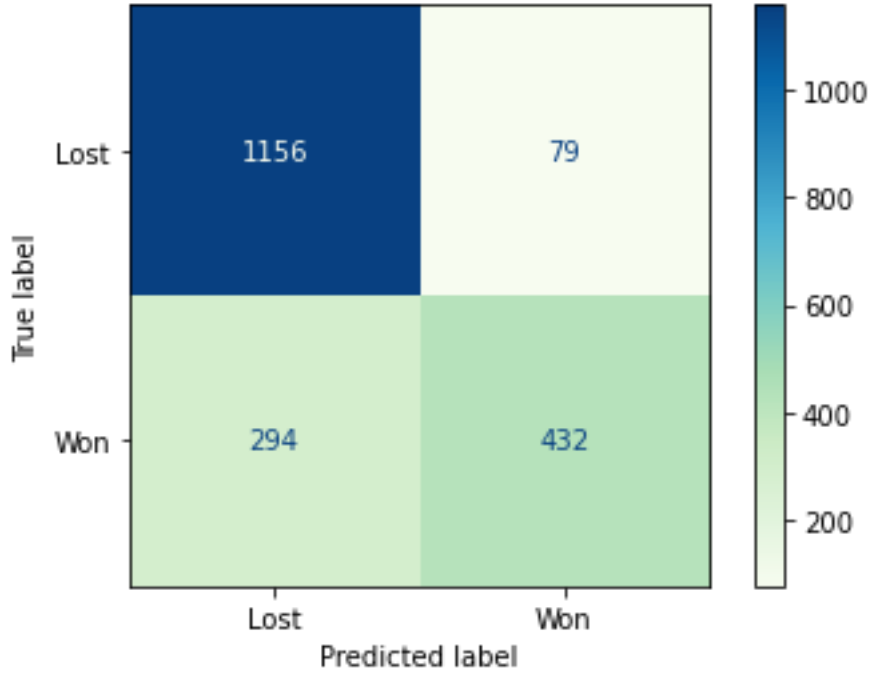
- Performance Validation - To gauge performance, accuracy score, confusion matrix, and classification report were calculated.

Random Forest Classifier

The accuracy of RandomForestClassifier() model is 0.81

CLASSIFICATION REPORT

	precision	recall	f1-score	support
Lost	0.80	0.94	0.86	1235
Won	0.85	0.60	0.70	726
accuracy			0.81	1961
macro avg	0.82	0.77	0.78	1961
weighted avg	0.82	0.81	0.80	1961



Model Cross Validation Summary

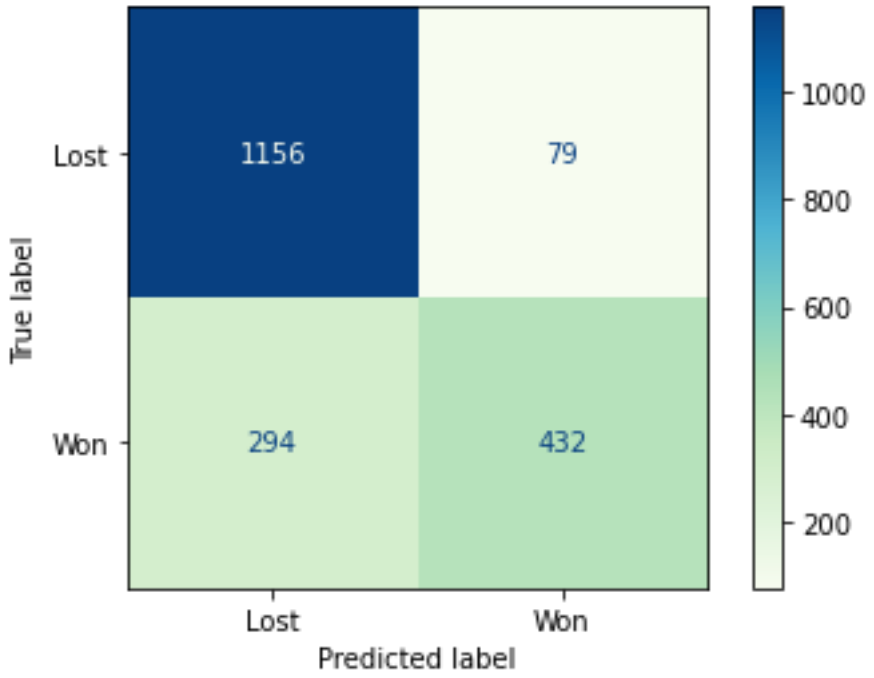
Model Name	Precision	Recall	Accuracy	AUC	Type II Error
Decision Tree Model	72%	62%	74%	0.74	272
Random Forest Classifier	84%	60%	**81%**	0.77	294
K Nearest Neighbour	57%	47%	67%	0.63	387
GaussianNB	60%	11%	64%	0.53	646
Logistic Regression	39%	2%	63%	0.50	713
Support Vector Machine	0%	0%	63%	0.50	726
SGDClassifier	0%	0%	63%	0.50	726
Perceptron	0%	0%	63%	0.50	726

Summary:

False Prediction Loss Calculation

- If a Deal Status is actually “Won”, but is predicted as “Lost”, then company will lose that potential deal amount.
- This is a Type II Error.
- In Random Forest Classifier, the number of False Negative deals are 294.
- Summing the Deal Costs of such deals, we get an opportunity cost loss of $ 222.05 million

- <u>FALSE PREDICTION LOSS = $ 222.05 million</u>



Results

- Our aim is to build a Machine Learning model to predict the win/loss probability for IT Consulting Company .
- We performed exploratory data analysis to identify factors that correlate with winning probability of deals.
- Later, we used these factors to create different models & found Random Forest Classifier with highest accuracy of 81% win prediction probability for the Organization.
- Firstly we conclude stating that quantitative methods should not take precedence over known information or justifiable sentiments about the chances of winning a particular deal.
- However there is a place for win prediction modeling play a competitive edge in securing a deal.

Recommendations

To conclude

- In the absence of a reliable subjective gauge of a deal's likelihood of being won, a quantitative model provides an anchor point for determining the viability of a deal within the opportunities pipeline.

- In this deal win prediction we further identified the opportunity deal loss for **294** deals leading to **$222.05** million such upcoming deals can be turned into profits with added due diligence.

- From analysis, we conclude that Best VP and Manager combinations for any variable is determined by the combination of win ratio and volume of win deals.

Thank You

KENNY D B L DEEPTHI ADITI GOYAL ARJUN S