SALES PIPELINE ANALYSIS OF SUZHI TECH. INC. FOR THEIR INDIAN MARKET

Submitted by: Group 5

Teammates:	**Contribution**
Sakshi Khadayate	1/3
Subash Rajaseelan	1/3
Shubham Bhutada	1/3

TABLE OF CONTENT

ABSTRACT

This report covers the analysis performed on sales data for SUZHI TECH. INC. (SAAS startup) using R language. The dataset was obtained from Kaggle, and it has 78025 business opportunities the startup had in the previous financial year. Different statistical analysis along with goodness of fit and hypothesis testing have been performed in this report.

Few additional columns were defined based on the existing data:

Quarter of closing, Month of closing, Business size, Cumulative win %, Revenue Bin, Employee size Bin

OBJECTIVE

The main purpose of this project is to analyze the sales data collected by a SAAS startup in a given year for various opportunities based on different technology primaries. Objective is to perform analysis based on client revenue size, client employee size, previous business with the client, sales velocity, opportunity size, location etc. using histograms, scatter plots to get clearer picture.

The Identifier of performance is opportunity status. This indicator tells us if the startup was able to get the business from the opportunity generated with the available combination of different parameters/indicators of the opportunity.

SCOPE

The data compares opportunities generated for 4 products/technology primaries from the startup.

Table 1

Technology primary	Product Description
Suzhi Analytics	Application Suite for Analytics
Suzhi ERP	Used for applications in Enterprise resource planning.
Suzhi PLM	Deals with the aspects of Product lifecycle management.
Suzhi SCMS	Has applications in the domain of supply chain management.

We have 5 different sales mediums in consideration

Table 2

B2B medium	Description
Outbound	Opportunities created from campaigns that are run by the marketing team.
Cold calling	Opportunities created from leads generated through unsolicited phone outreach
Channel partners	Opportunities created by partnerships from third party sources
Events	Opportunities created through gatherings and interactions in a social setup
Inbounds	Opportunities created by prospects reaching out to evaluate the product

Our data Is limited to **7 locations**:

Bengaluru, Chennai, Delhi, Hyderabad, Kolkata, Mumbai, Pune.

DATA DESCRIPTION

The dataset has information on different attributes related to sales. The information collected is divided in columns like Opportunity ID, Technology Primary, Client Location, B2B Sales medium, Sales Velocity, Opportunity Status, Sales stage iteration, Sale cycle (days), Opportunity Size (USD), Discounts %, Business from client last year, Competitor insight, Client revenue sizing, Business size, Client employee size, Cumulative Win % *(For a won opportunity, win% is 100% and for a lost opportunity, win% is 0%.)*, Start date, Close date, Closing month, Closing quarter.

Business size: Based on the number of employees a client has, business size is categorized in 4 types.

Table 3

Business Size	Number of Employees
SMB (Small and medium business)	1 – 1000
Commercial business	1000 – 1500
Midmarket	1500 – 5000
Enterprise Business	5000+

EXPLORATORY DATA ANALYSIS

GENERAL STATS OF THE DATA

Total opportunities: 78025 Won Opportunities: 17627
Total lost opportunities: 60398 Average Sales cycle length: 85.98633 days
Highest won opportunity value: 149695 USD Smallest won Opportunity value: 0 USD
Total Revenue: (Only won opportunities) 42873770 USD
Win rate: number of won opportunities /Total number of opportunities = 17627 / 78025 = 22.59147

SALES VELOCITY

Sales velocity is **a measurement of how fast you're making money**. It looks at how quickly leads are moving through your pipeline and how much value new customers provide over a given period.

Fig 1

Sales velocity is affected by won opportunities. So, while calculating we will take in account only the values related to won opportunities.
In the above formula,
Average Deal Size = Total Deal Size of Won Opportunities / Total number of won deals = 24322.78329
Won Opportunities = 17627 Average Sales cycle length of Won deals = 85.75463

$$\text{Sales Velocity} = \frac{17627 \text{ X } 24322.78329 \text{ X } 22.5914\%}{85.75463}$$

Sales Velocity = 1,129,476.61246 USD / day ~ **1.13 Million USD / day**

CONFIDENCE INTERVAL FOR DISCOUNT OFFERED

In many cases, Discount offered to the client for an opportunity act as an incentive for converting a deal. Discount offered to clients ranges between 0 to 30. When a deal is on the verge the sales team may negotiate the deal by offering discount. These negotiations start from the discount values as small as possible. That's why most of the won opportunities have discount values less than 10%.

Let's take simple random sample of 100 discount values. In this sample we can see that 73 opportunities were won with the discount value between 0 to 10%

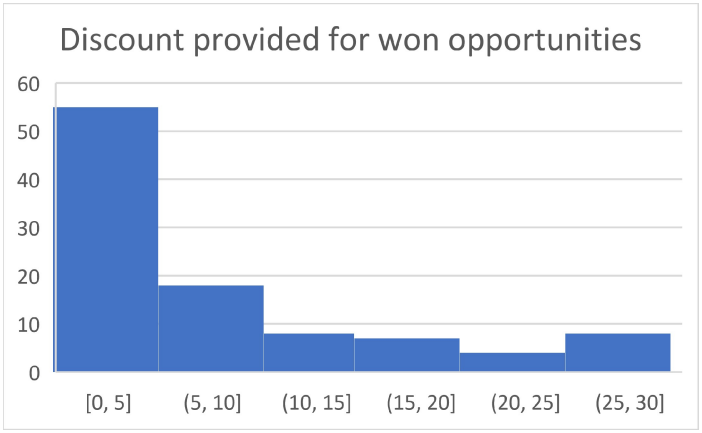


Fig 2

That means,

Success (p) = 73/100 Failure (q)= 100 – success = 27%

X ~ b(n,p) X = R.V of discount given for won opportunities between 0 – 10 %

Considering 95 percent confidence interval,

$\alpha = 1 - 0.95 = 0.05$ $\alpha/2 = 0.025$ $Z_{\alpha/2} = 1.96$ (From table)

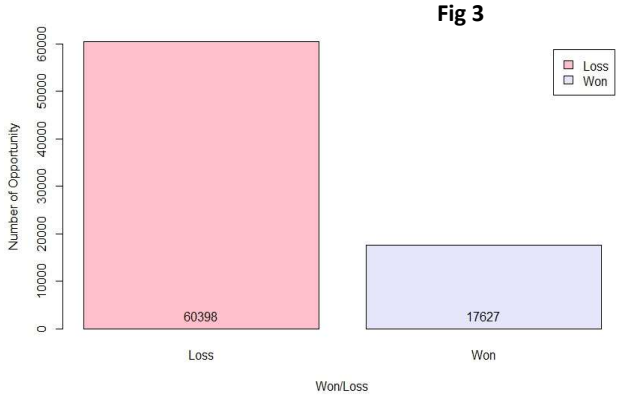
$e = Z_{\alpha/2} (p * q / n)^{1/2} =$ **8.70%**

With 95% confidence interval, if an **opportunity is won** then **73 ± 8.70%** of these opportunities were offered **discount between 0 to 10%.** This error can be reduced by increasing the sample size.

PART 1: OPPORTUNITY STATUS (WIN/LOSS) BASED COMPARISON

OVERALL WON VS LOST OPPORTUNITIES

Fig 3



Out of the total 78025 opportunities generated, 17627 opportunities were successfully converted. These numbers of opportunities are the combined opportunities generated for the clients with whom we had some business last year or newer client leads generated this year. From this data we get the ratio of 22.59% success rate. The won opportunities contribute towards the revenue of the firm with the opportunity size.

QUARTER WISE OPPORTUNITY CONVERSION

All the generated opportunities throughout the year can be segregated in four quarters of the financial year based upon the closing date.

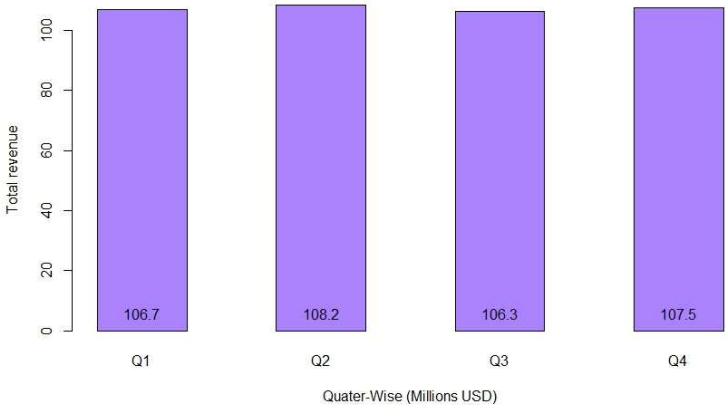
When divided in 4 quarters, we can see that the opportunities generated in different quarters of the year are nearly same**.**

The conversion rate in all the quarters is also nearly equal.

The success /conversion rate in all 4 quarters is ~ 30%

Fig 4

QUARTER WISE REVENUE GENERATED FROM THE WON OPPORTUNITIES



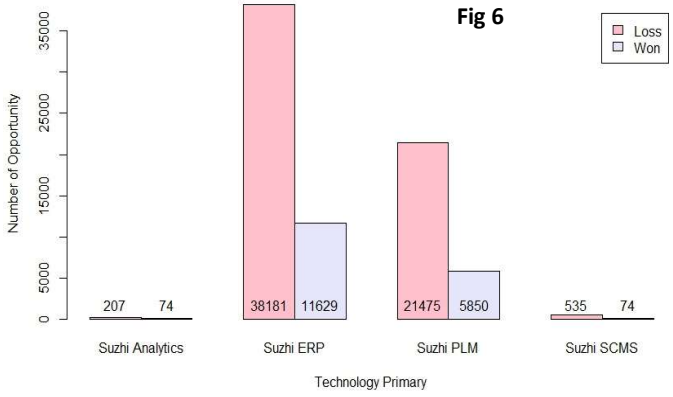
As mentioned above, the won opportunities contributed towards the revenue of the firm. Revenue generated is almost same in all the quarters.

Total revenue generated = 106.7 + 108.2 + 106.3 + 107.5 **= 428.7 million USD**

Fig 5

TECHNOLOGY PRIMARY WISE OPPORTUNITY CONVERSION

Fig 6



The four different products from the SUZHI analytics have established their own market and have different conversions for their respective opportunities.

We can see the ratio of conversion of opportunities for these technologies in the image below.

It is evident that SUZHI analytics has the highest conversion ratio and SUZHI SCMS has the lowest conversion ratio

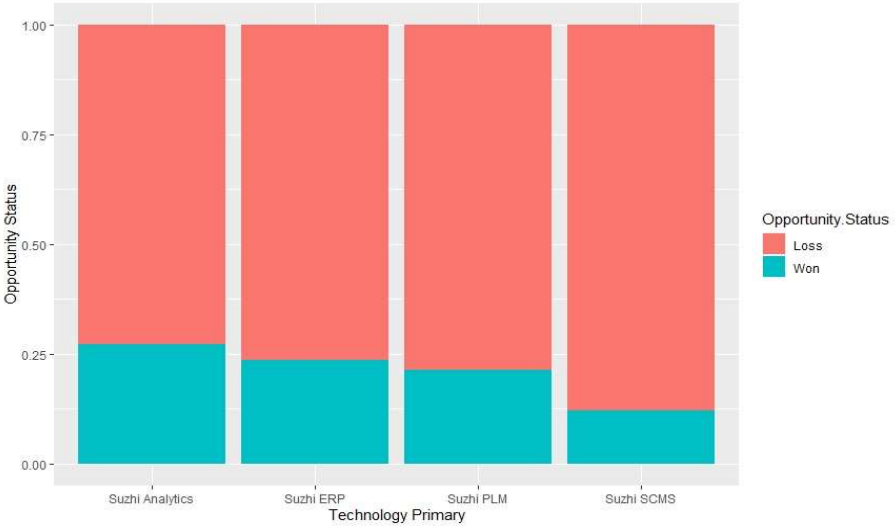
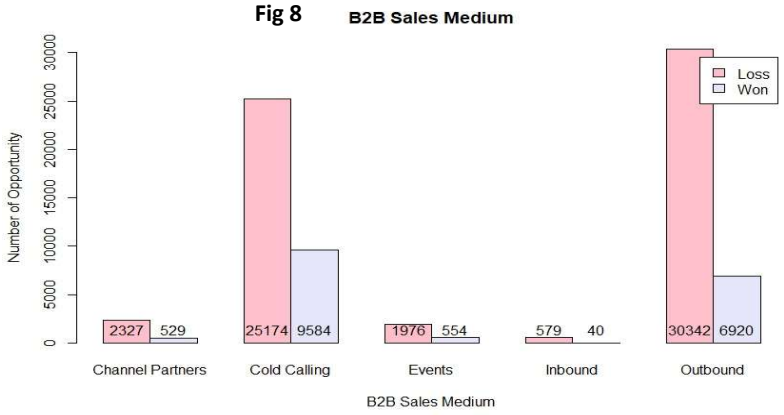


Fig 7

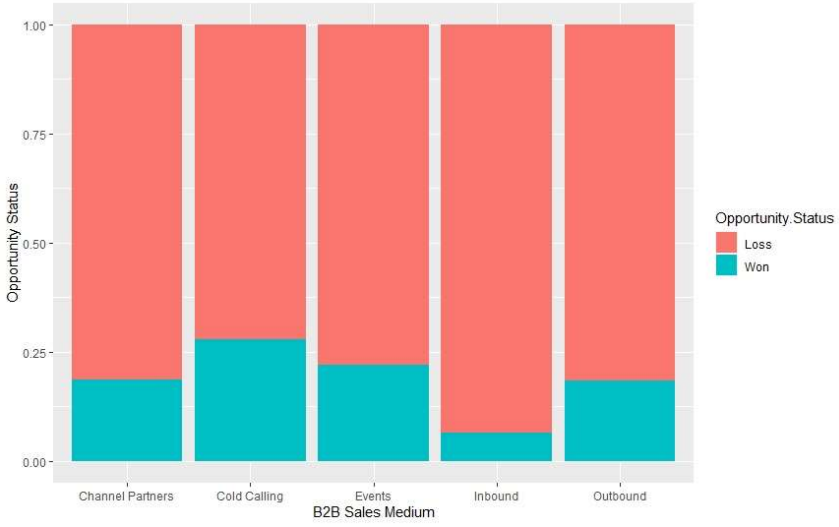
OPPORTUNITY CONVERSION THROUGH DIFFERENT B2B SALES MEDIUM

As we know, we have 5 different sales mediums dealing with the opportunity's generations for our 4 products. These medium combinedly generated 78025 opportunities out of which 17627 opportunities were successfully converted. But if we observe the data, we can see that cold calling medium has given the best conversion rate when compared with other mediums of the B2B sales.

Fig 8



From the above graph we can see that **72020 opportunities (92%)** are generated only from cold calling and outbound B2B sales medium. The other 3 mediums of B2B sales contribute for the **remaining 8% opportunity generation.**



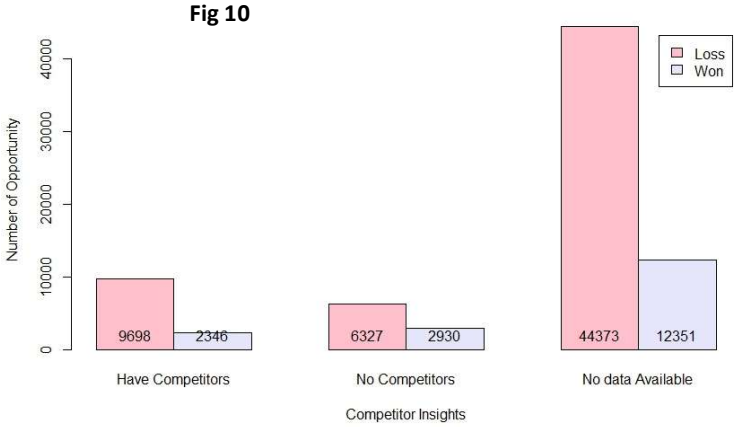
Out of all the mediums, **Cold calling** resulted in the **maximum conversion (27.5%) of opportunities**.

Fig 9

COMPETITOR INSIGHT

For all the opportunities the conversion is affected when we have a competitor offering a similar product or service.

Fig 10



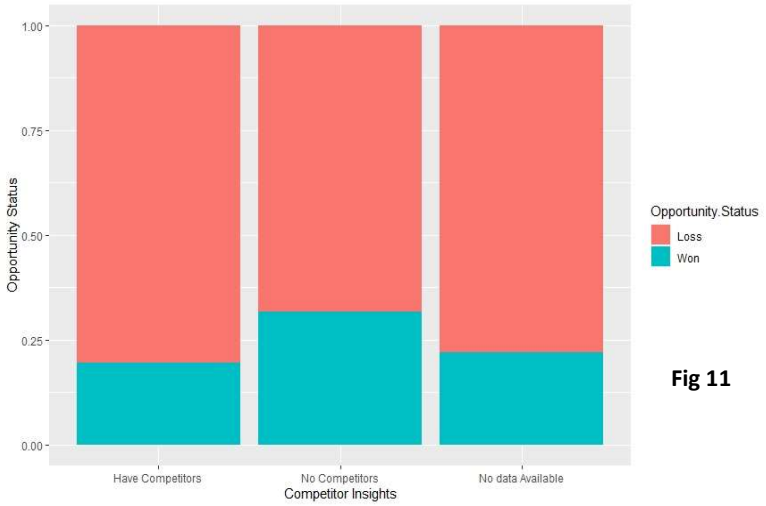


Fig 11

For most of the opportunities, we didn't have insight about if we have any other competition for the opportunity.

We have the maximum conversion of opportunities when we don't have any competition.

When we have competitors, our chances of converting an opportunity go down.

It can be seen from the ratios of conversion compared to competitor insight.

Conversion when no competition is there = **31.65%**

Conversion when competition is there = **19.47%**

When we don't have intel about the available competition, conversion = **21.77%**

CONVERSION BASED ON THE CLIENT LOCATION

Fig 12

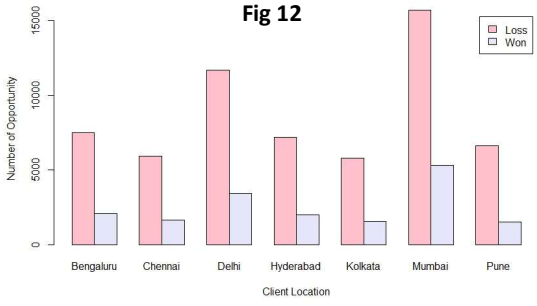
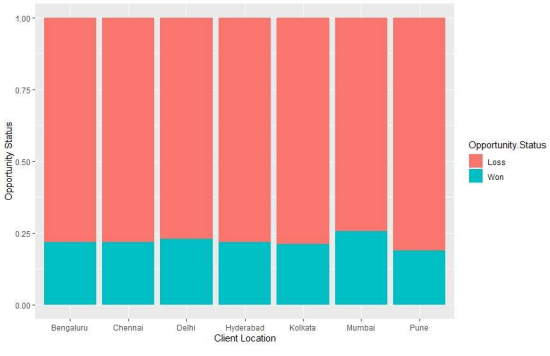


Fig 13



All the generated opportunities for SUZHI tech. re divided in 7 locations across India. These locations with their generated opportunities have different conversion rates. From the available data, we can conclude that Mumbai has the highest conversion rate compared to other cities. As for other cities the conversion percentage is almost similar with the lowest conversion rate for Pune.

OPPORTUNITY CONVERSION BASED ON CLIENT BUSINESS SIZE

As we have already described 4 categories for the client business size based on the number of employees the client has, we checked the conversion of opportunities in these categories.

Fig 14

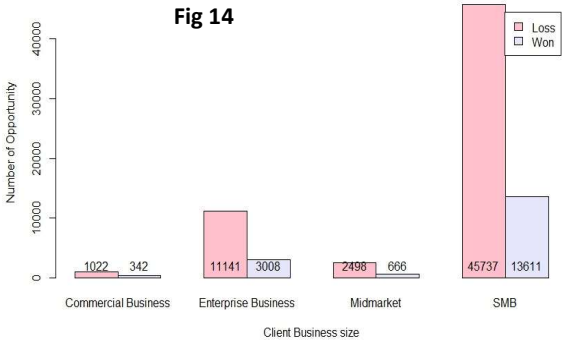
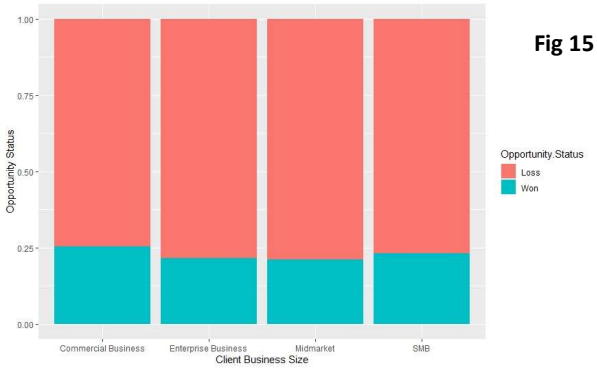


Fig 15



Out of the 4 categories of business size, most of the opportunities generated are for the SMB category. Though the number of opportunities is highest for SMB, the highest conversion rate is still obtained from commercial business category.

CUSTOMER RETENTON

We can a customer is retained If we had some business from them last year and have successfully converted an opportunity with them this year as well.

Fig 16

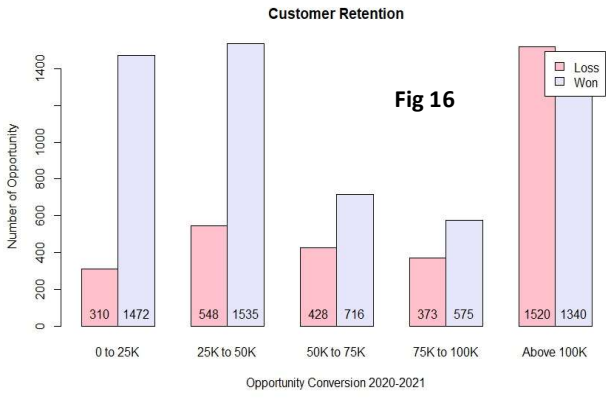
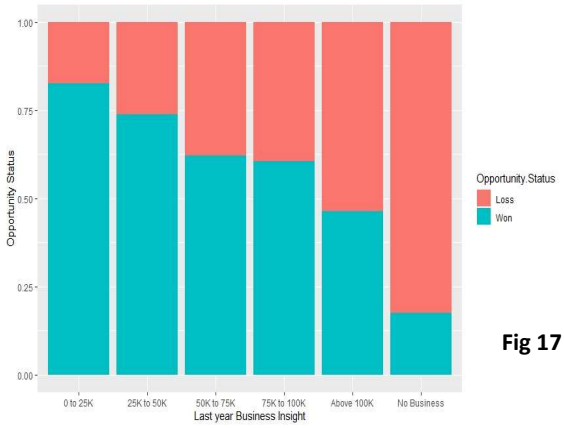


Fig 17



In Fig 16, We have taken only the opportunities which had some business with us last year. After dividing the opportunities based on the last years business value, we can see that if we had smaller opportunities with a client, they are more likely to retain with us than the clients with whom we had secured bigger opportunities in last year.

It can be seen from the Fig 17 that customers with whom we had business of less than 25k USD are most likely to be retained with us and the customers with whom we had business of more than 100k USD are least likely to be retained.

NEW DEAL ACQUISITION

The last column of this graph shows the newly won opportunities.

It's very evident from the graph that the firm is struggling in winning new opportunities.

Only around **17.32%** new targeted opportunities are successfully converted.

CLIENT REVENUE VS OUR REVENUE

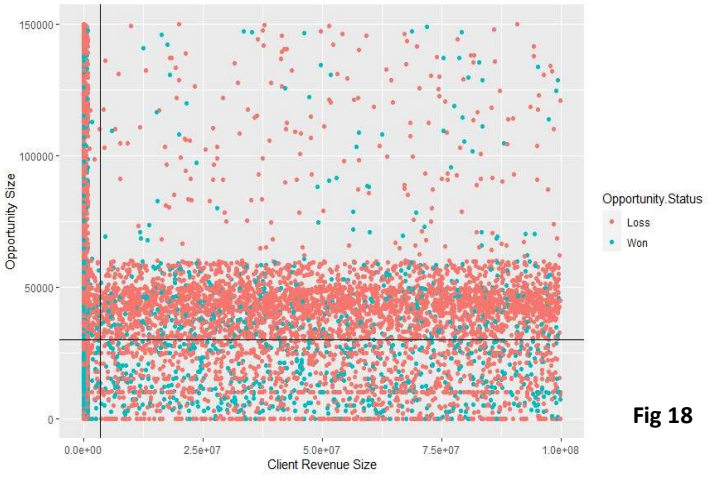


Fig 18

Opportunities available as well as converted are equally distributed among range of Clients Revenue Size.

Win% is more with opportunity Revenue less than mean value even though Number of Opportunities available are more right above the Mean of opportunity Revenue.

PART 2: CORELATION AND JOINT PROBABILITY

CORRELATION

Fig 19

Table 4

	CORRELATION OF CLIENT EMPLOYEE SIZING AND CLIENT REVENUE SIZING	CORRELATION OF OPPORTUNITY SIZE VS CLIENT REVENUE SIZE	CORELATION OF SALES CYCLE LENGTH VS WIN RATE
Correlation coefficient	0.20	0.08	-0.01
Description	This correlation coefficient compares the size of client revenue with the number of employees the client has. When calculated for all the opportunities, we understand that weak positive correlation exists between these two parameters.	When we checked the correlation for opportunity size with the client revenue size, we get a very weak positive corelation between the revenue and the opportunity available with the client.	The number of days elapsed before the sales cycle reaches its final status is sales cycle length. Correlation between the number of days of sales cycle and the win rate associated is very weak and is negative. Negative correlation signifies that they are inversely related and increase in one will lead to decrease in other.

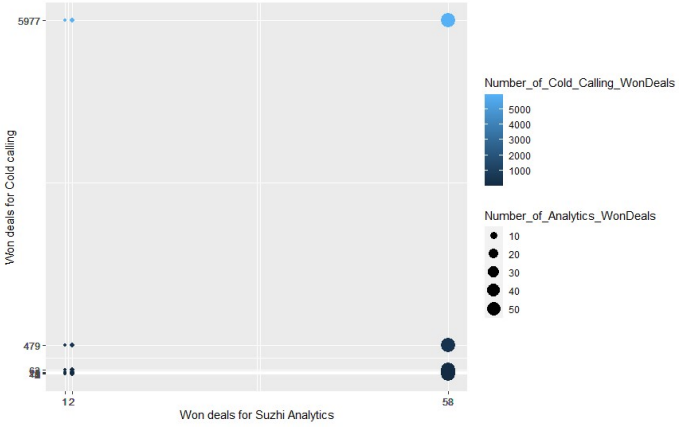
INFERENCE FROM CORELATIONS

The weak corelation between the client employee sizing with the client revenue size and the opportunity size discards a common misconception that a client with huge employee size will always have huge revenue or bigger opportunities.

The corelation value is so small, we can even say that these parameters are very weakly corelated.

The negative corelation value we get for sales cycle length and opportunity win rate shows an inverse relation between these parameters. But as the corelation coefficient is very small, these parameters have a very weak corelation. This weak n corelation debunks the common assumption that increase or decrease in the sale cycle length has an impact on the chances of winning or losing the opportunity. Weak corelation means that the parameters are not affected by each other.

JOINT PROBABILITY FOR SUZHI ANALYTICS AND COLD CALLING



Joint Probability for Won Deals for Technology Suzhi Analytics Vs B2B Sales Medium Cold Calling

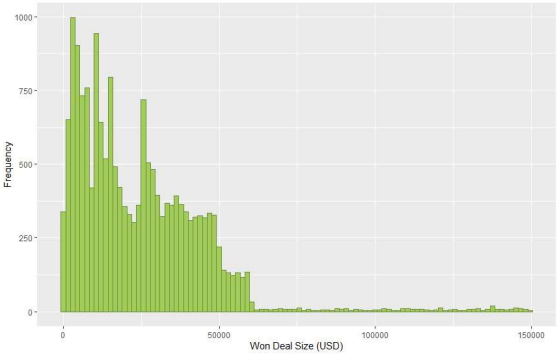
Correlation Coefficient is 0.50

Fig 20

PART 3 PROBABILITY DISTRIBUTION

DISTRIBUTION OF REVENUE FROM WON OPPORTUNITIES

Fig 22




```
> mean(Sales_Won$Opportunity.Size..USD.)
[1] 24322.78
> sd(Sales_Won$Opportunity.Size..USD.)
[1] 21141.46
> #Coefficient of Variance
> mean(Sales_Won$Opportunity.Size..USD.)/sd(Sales_Won$Opportunity.Size..USD.)*100
[1] 115.0478
> skewness(Sales_Won$Opportunity.Size..USD.)
[1] 2.05169
> kurtosis(Sales_Won$Opportunity.Size..USD.)
[1] 7.483102
```

Fig 21

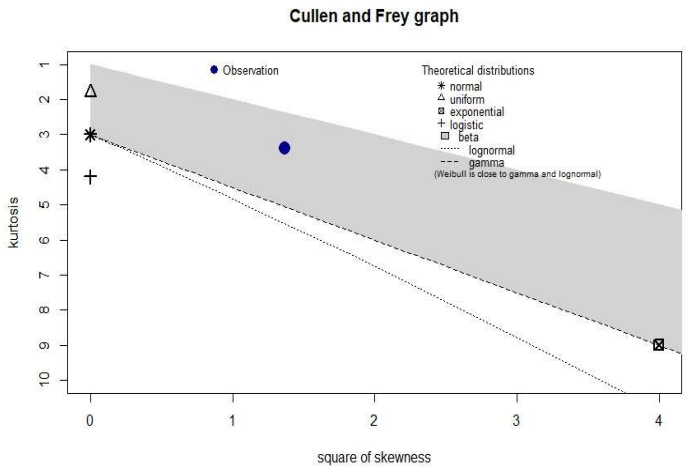
The above graph shows the distribution for the revenue generated from the successfully converted opportunities. The shape of the graph shows that the bigger deals are less frequent than the smaller deals.

The positive skewness shows that it is right skewed and is heavily skewed, as the skewness value is greater than 1.

Because of the skewness to right → Mode<Median<Mean

Also, the positive kurtosis value signifies that there are outliers in the data and data is tailed.

DISTRIBUTION OF DISCOUNTS OFFERED IN THE WON OPPORTUNITIES



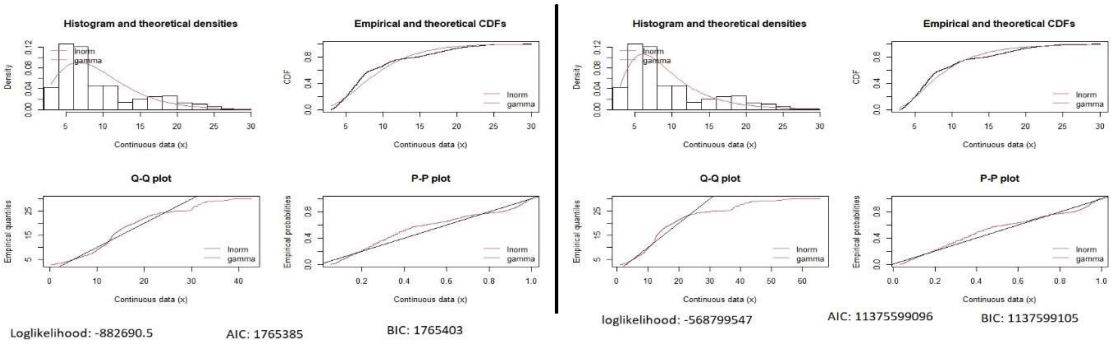
The Culley & Frey graph below shows that this plot is close to gamma, and lognormal distributions.

The best of fit analysis of the two distribution types below shows that gamma is the best fit since it has the highest loglikelihood value, the lowest AIC and BIC values, and its theoretical P-P and Q-Q plots best match the empirical plots

Fig 23

GOODNESS OF FIT FOR DISCOUNT OFFERED

Fig 24 & 25



Loglikelihood: -882690.5 AIC: 1765385 BIC: 1765403

loglikelihood: -568799547 AIC: 11375599096 BIC: 1137599105

PART 4: HYPOTHESIS TESTING

PROBABILITY OF WINNING

When we consider all the 78025 opportunities, we have 17627 won opportunities. It gives us the win rate of 23.1%.

For this hypothesis we will be taking in account randomly selected 20% of opportunities from the entire population and checking if the win rate we get from this sample is similar to the population win rate. Let,

H0 = There is no significant difference between the Sample Won percentage and the population Won percentage

H1 = There is a significant difference between the Sample Won percentage and the population Won percentage

By proportion test,

We get P value = 0.16

From the sample we get the win rate as 22.59%, which is almost equal to the win rate of the population which is 23.1%.

Therefore, we fail to reject the null hypothesis which states that there is no significance difference in the sample won percentage and population won percentage.

MEAN OF OPPORTUNITY SIZE OF WON OPPORTUNITIES

The mean of the overall won opportunities was found to be 24322.78 USD.
For this hypothesis, we are taking 20% data as a sample for testing. We need to check if the mean of won opportunity sizes differ significantly on the confidence interval of 95%. Let,
H0 = There is no significant difference between the Sample mean opportunity size
H1 = There is significant difference between the Sample mean opportunity size

From the Z test
P value = 0.24

From T test,
p value = 0.48
sample mean revenue = 24080.47 USD.

As the P value is > 0.05 for both Z test and T test,
We fail to reject null hypothesis which states that there is no significant difference between the sample mean and population mean of the opportunity size.

MEAN OF CLIENT REVENUE SIZE FOR WON OPPORTUNITIES

The average client revenue for the won opportunities was found to be 2985636 USD. Let's check if this mean revenue stands true when we consider a sample of the population when 95% confidence interval is considered.
Considering sample consisting of 20% opportunities from the won opportunities.
Let, H0 = There is no significant difference between the mean client revenue of the sample
H1 = There is significant difference in the mean of the client revenue of sample and mean of client revenue of population.

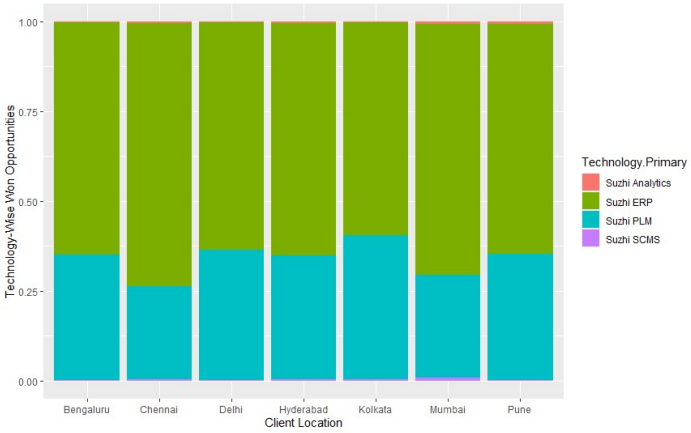
From the Z test
P value = 0.64

From the T test,
P value = 0.7144
Estimated sample mean = 3068959 USD

As the P value > 0.05 for both Z and T test,
We fail to reject null hypothesis which states that there is no significant difference between the sample mean.

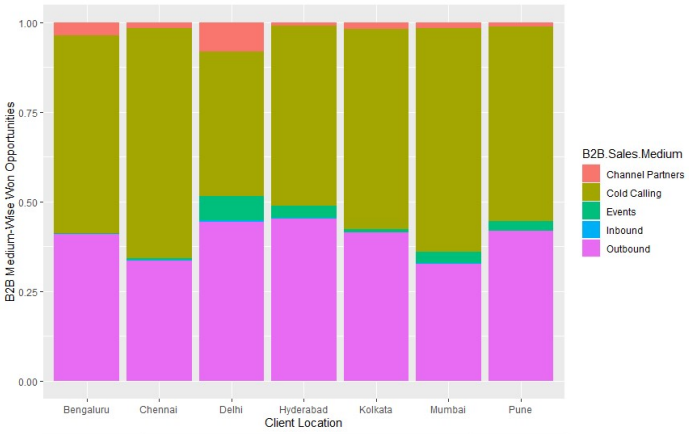
PART 5: INSIGHTS

ANALYSIS FOR BUSINESS GROWTH



As inferred previously that Suzhi Analytics has the highest % Won but the Number of Opportunities is very less

Fig 26



Cold Calling followed by Outbound Sales Medium Has Both high % Won and Number of Opportunities as well.

Fig 27

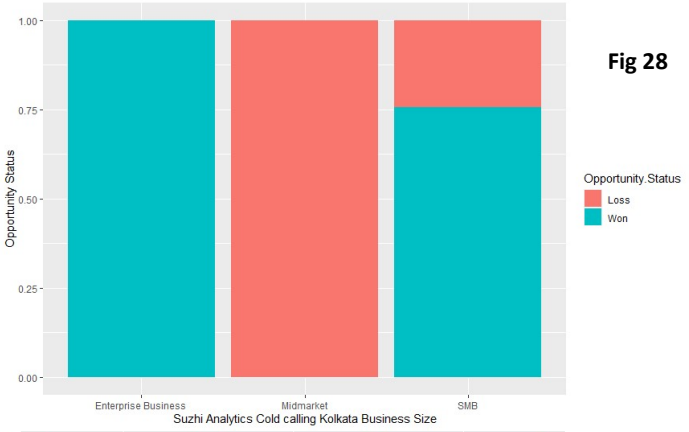


Fig 28

100% conversion rate - Suzhi Analytics with Cold Calling for Client Location Kolkata for Enterprise Business. To grow overall Win % Company should work on generating more leads in given criteria as this best performing area

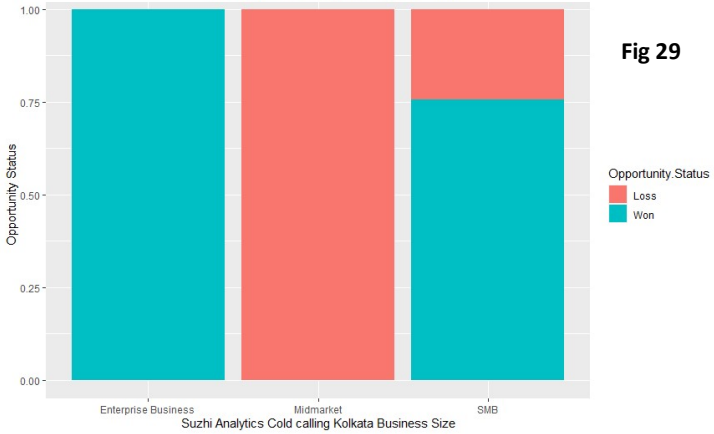


Fig 29

Inbound B2B sales medium has low customer conversion rates at all Client locations except Chennai, other locations teams should connect with the Chennai team to improve % Win.

one way could be improving local SEO to increase conversions via organic lead generation which is major contributor to Inbound Sales Medium.

SUGGESTIONS FOR IMPROVING Q4 REVENUE OF 2021-2022

1. **Generating more organic leads -** Focus more on inbound lead generation as inbound leads cost 61% less to acquire a customer and has 35% higher chances of converting an opportunity
2. **Start-Up programs -** Focus on start-up markets like Bengaluru with an opportunity ecosystem of $42.7 billion for SaaS Products.
   Coming up with Start-up Programs with software credits is an effective way of penetrating this market.
3. **Emerging Products Focus -** Focus on selling Suzhi Analytics as the Analytics Market in India is worth $25.5 Billion USD with high growth potential
4. **Sales Team Training -** Train the sales team to tackle and win deals that has competitor insights by providing Competitor battle cards and objection handing workshops
5. **Maintain Lower Discounts -** 67% of our total revenue discounts are in the range 0-10 and hence offer lower discounts to improve ARPU

CONCLUSION

By the analysis using R of the sales pipeline data for a startup, we were able to draw some meaningful insights which enabled us to reach at a state from where we could come up with suggestions for improving the results in the upcoming quarter. Using R, we could address the concern for a firm so that they can improvise and prioritize their resource utilization. It also highlighted the poor performance of some areas and excellent performance of other areas, so that the firm can work upon them with respect to their long-term goal.

REFERENCES

- https://www.kaggle.com/gauravduttakiit/sales-pipeline-conversion-at-a-saas-startup
- https://www.r-graph-gallery.com/index.html
- https://medium.com/@mygreatlearning/r-programming-tutorial-ac1ccddffda9
- https://www.displayr.com/how-to-create-a-correlation-matrix-in-r/
- https://www.youtube.com/watch?v=49fADBfcDD4&t=767s
- Lab files from IE6200