Project :XYZ Mobile Market Entry Analysis

Name: Anuj Sushil Topno

Problem Statement

- This assignment is based on a mobile company 'XYZ Mobiles', a fictional China-based mobile company (manufacturer and supplier).
- The company sees India as a key opportunity to expand its sales. It has been tracking the Indian market for more than a couple of years.
- XYZ Mobiles believes that the Indian market is very similar to China, in which the company currently operates.

Before entering the new market, the company wants to be sure that the whole process will be profitable for them. Hence, you are given the task to check for the following conditions that must be fulfilled in the Indian market for the company to enter:

Sale of minimum 12,000 phones over the sample data in one year

Collection of at least Rs. 20 crores over the sample data in one year

Task1: Justification during Model Development

China_Model

1. Gender was classified into binary data as male (1) and female(0).
2. From the Chinese customer data it is clear that purchase decision depends on 4 factors. Customer age, Gender, last phone life, and annual income. The following changes were made on the raw data
3. The phone life was classified into 4 category as below

Days	Segment
<200	1
200-360	2
360-500	3
>500	4

4. Then after dataset is converted into training and test set with 70:30 Rule and then Logistic RegressionIs Performed On both of them.

Code	Gender
0	Females
1	Males

5. Then ROC Curve, Beta Values and Conversion Matrix (Including Accuracy, Sensitivity, Precision etc.)Is Computed from train and test data through Logistic Regression.

	Using 70/30 Rule
40000x70%=	28000
28000	for training
12000	for test

Task1: Pivot Analysis to get insight

After Data Cleaning

Row Labels	Sum of PURCHASE	Count of ID
0	9836	17715
1	13195	22285
Grand Total	**23031**	**40000**

	Gender Category		
Gender	**Sum Of Purchase**	**Count Of Lead**	**Conversion Rate**
F	9836	17715	55.52%
M	13195	22285	59.21%

Conversion Rate of Male is higher than Female.

Table 1:-

Table 2:-

Row Labels	Sum of PURCHASE Count of ID	
1	2351 6459 2	6987 16452
3	9171 11694 4	4522 5395
Grand Total	**23031**	**40000**

	Phone Age Category		
Phone_Age _Category	**Sum of PURCHASE Count of ID**		**Conversion Rate**
1	2351	6459	36.40%
2	6987	16452	42.47%
3	9171	11694	78.42%
4	4522	5395	83.82%

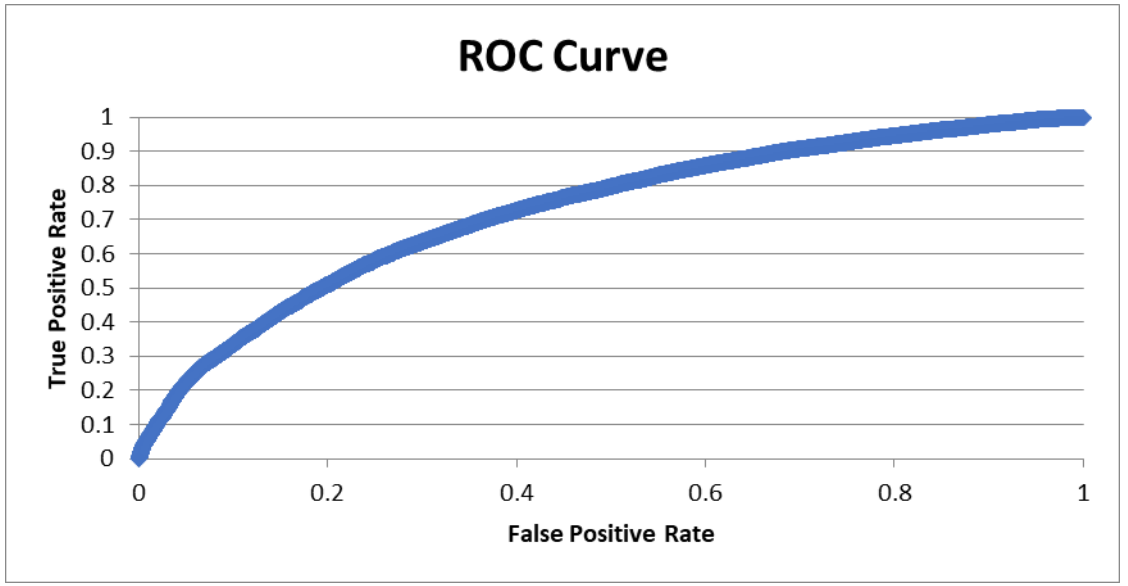
Phone age of 360+ days got high conversion rate .

Task1: Classification Model based on Training Dataset

1. Logistic Regression was made using Real Statistics pack on Cleaned Data after considering factors like age ,annual income , Phone age (age of phone after it's Purchase date) and Gender.
2. The Coefficients, Conversion Matrix and Roc Curve of Training Dataset is given below:

Coefficients	
B0	-1.518413624
B1	-0.011855742
B2	0.217181357
B3	2.2505E-06
B4	0.004185136

Conversion Matrix		
Actual	Predicted	
	Class 0	Class 1
Class 0	6381	5585
Class 1	3612	12422
Accuracy	0.67	
Precision	0.69	
Sensitivity	0.77	
Specificity	0.53	
Recall	0.77	
F1-Score	0.73	
TPR	0.77	
FPR	0.23	

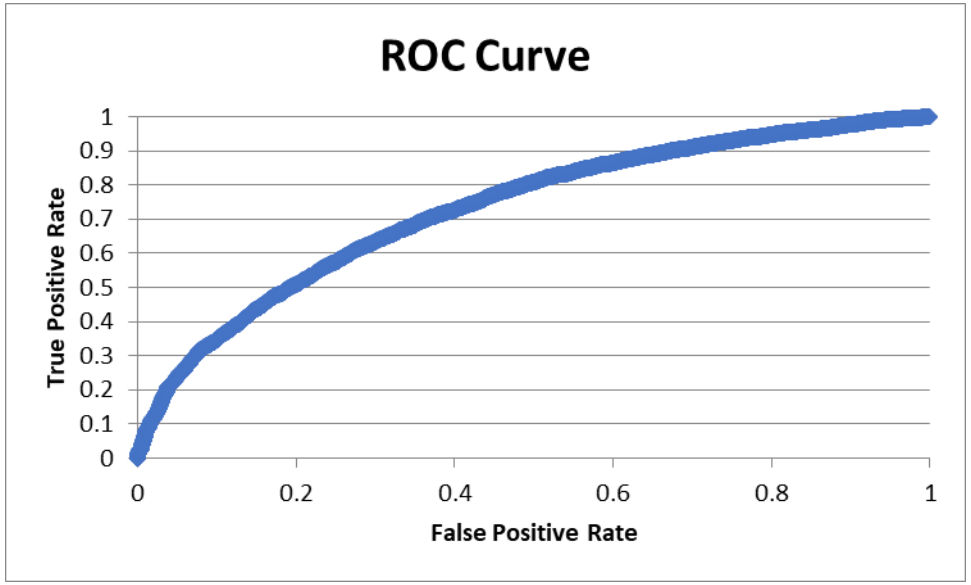


Task1: Classification Model based on Test Dataset And Business Interpretations

1. The Coefficients, Conversion Matrix and Roc Curve of Training Dataset is given below:

Coefficients	
B0	-1.657770239
B1	-0.012385313
B2	0.251535386
B3	2.73119E-06
B4	0.004235263

Conversion Matrix		
Actual	Predicted	
	Class 0	Class 1
Class 0	2622	2381
Class 1	1459	5538
Accuracy	0.68	
Precision	0.70	
Sensitivity	0.79	
Specificity	0.52	
Recall	0.79	
F1-Score	0.74	
TPR	0.79	
FPR	0.21	



2. Through Coefficients we learn.

- As age Increases people are less likely to buy phone .
- Slightly affected by Gender .
- Income plays a important role to buy expensive models but doesnot affect purchase a lot .
- Phone's age plays a major role and its segmentation and category affects the coefficients a lot.

Task1: COUNT OF POTENTIAL CUSTOMERS IN INDIA BASED ON MODEL

1. The purchase date for everyone in the Indian dataset as **1st July 2019.**
2. Gender is converted into Binomial Model.
3. Phone's Age is converted into 4 Segment.

Male	1
Female	0

DAYS	SEGMENT
<200	1
200-360	2
360-500	3
>500	4

4. Purchase status is converted into Binomial Model.

Purchase	
1	Yes
0	No

5. Number Of Potential Customer is 31573 and Conversion Rate is 45.1% .

Task2: Justification Of Customer Segmentation (Clustering)

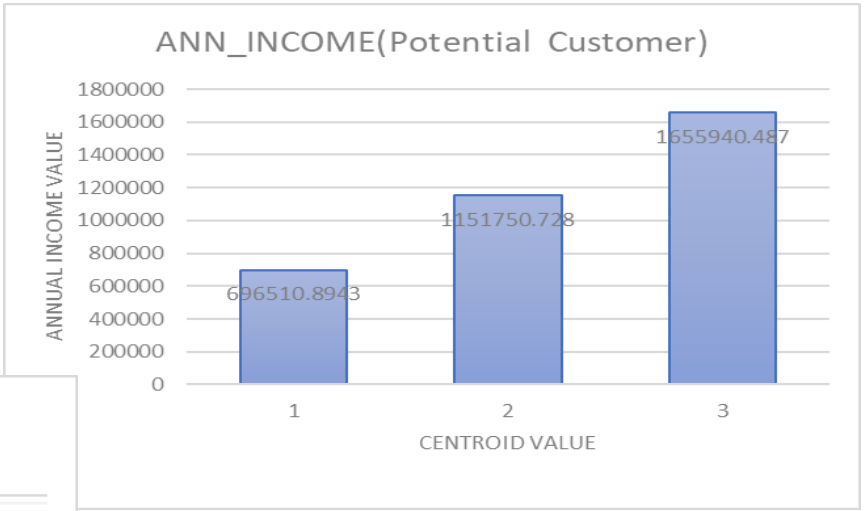
- Clustering is Performed on 3 & 4 Clusters and their error terms (For more scaled and standardized data) is also found and with that centroid values were generated.
- In this analysis 3 clusters results were taken for further analysis and predicting results and then EDA is performed on each cluster & Centroid values and following results were obtained:

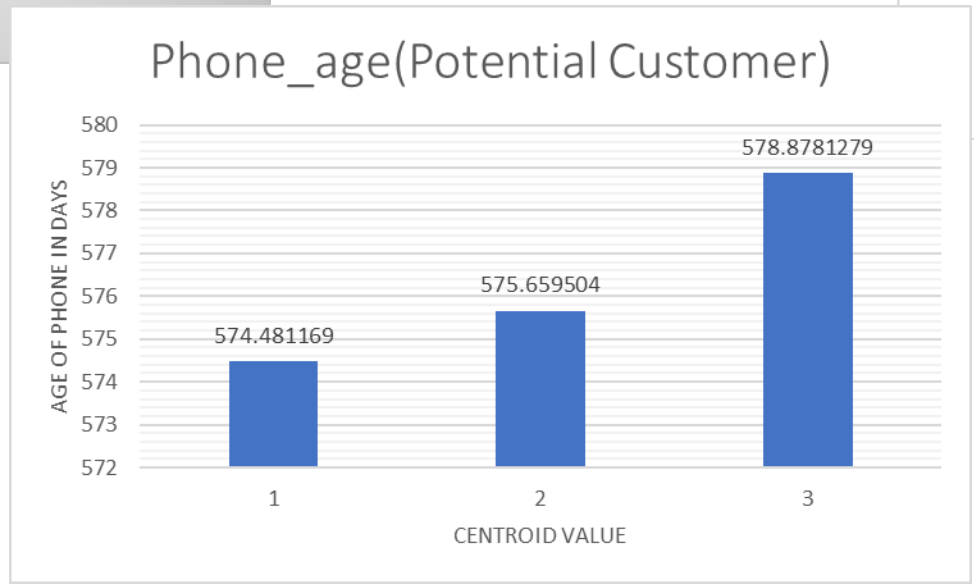
Centroid for Potential Customer			
	1	2	3
CURR_AGE	44.29115	42.92283	49.24305
GENDER	0.437079	0.477511	0.609824
ANN_INCOME	696510.9	1151751	1655940
Phone_age	574.4812	575.6595	578.8781

Centroid Values For Error			
	1	2	3
CURR_AGE	0.962399	-0.95125	0.48633
GENDER	-0.14483	-0.17973	0.337332
ANN_INCOME	-0.8547	-0.34038	1.074958
AGE_PHN	-0.02124	-0.02563	0.04854

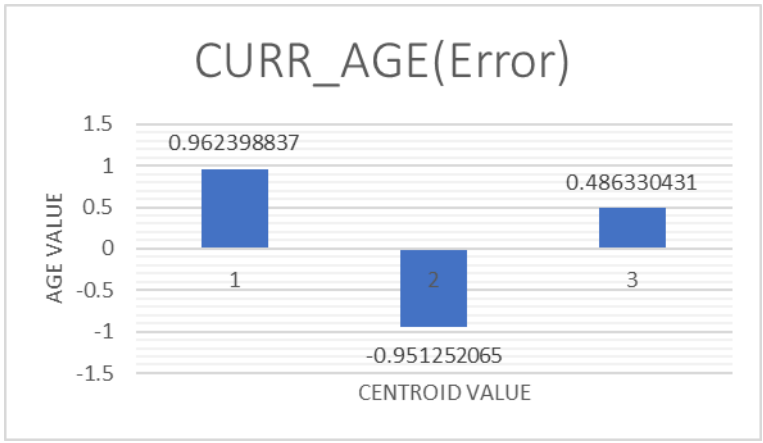
- From the table above, centroid value 2 is to be taken for business decision and and It has error terms in negative as well which is good so most peoples were clustered (i.e., their centroid) around with current age of 42, Annual income around 1151751 and age of phone around 576 days is to target For through our Clustering analysis and EDA analysis of that is also done in further slides.

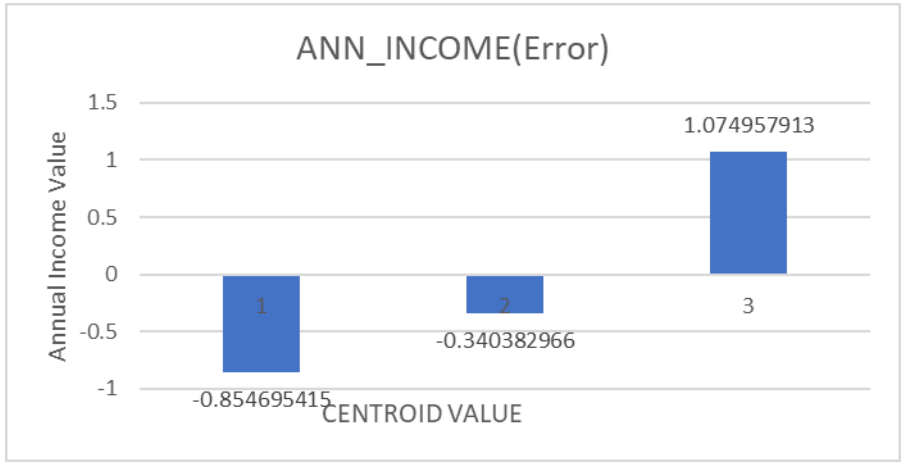
Task2:EDA Analysis Potential Customer

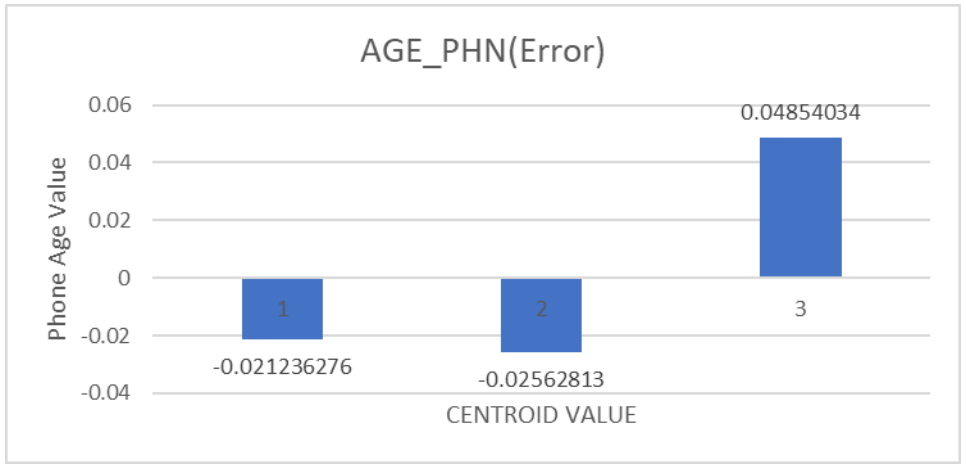




Task2: EDA Analysis Error







Task3: Justification For Business Decision

- Objectives:- Sale of a minimum of 12,000 phones over the sample data in one year .
- Collection of at least Rs. 20 crores over the sample data in one year .

- To segment our customers through their ages to check for final results

	Segments	Age Criteria
Young Age	1	25-35
Mid Age	2	35-55
Old Age	3	55-65

- After Finding Quartile We can Find 3 Financial Range [(Q4-Q0)/(Number Of Ranges You want)] here we took (Q4-Q0)/3 and then that answer is added to Q0 to get Range and added further to get more range.

Q0	300054
Q1	853547
Q2	1123316
Q3	1435209
Q4	1999845

Range1	300054-866651
Range2	866651-1433248
Range3	1433248-199845

- Range1 is Low Income Individual.
- Range2 is Mid Income Individual.
- Range3 is High Income Individual.

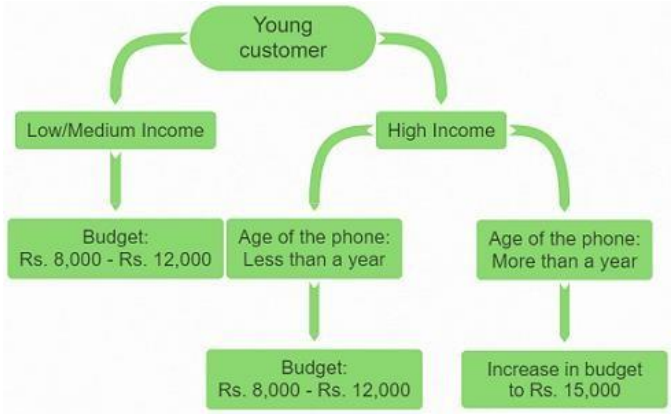
Task3: Justification For Business Decision

Number Of Individual in Age Segment who purchased Mobile

Mid	**M**	**15283**
Old	**O**	**8630**
Young	**Y**	**7660**
	Grand Total	**31573**

For Individuals in Young Age Range (25-35)

It was given as shown below so we separated them into 2 segment Low/Mid and High according to Income so we took average of the income it and found unit sold.



Number of All Young Customer	7660
Number Of High Income Individual	1850
Number Of Low/Mi Icome Individual	5810
High Income ighPhoneage	1850

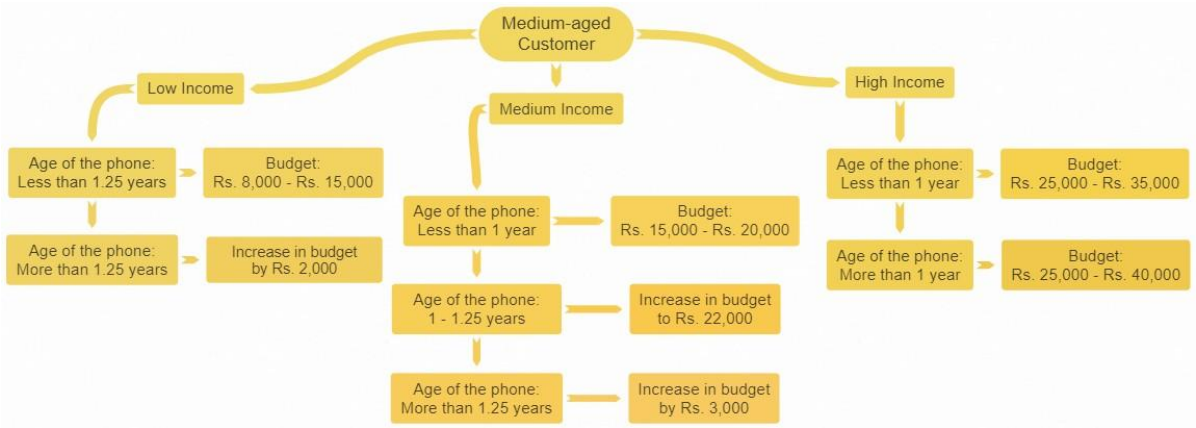
		Expected Revenue
Low/Medium Income		58100000
High Income		27750000
Total		85850000

Task3: Justification For Business Decision

For Individuals in Mid Age Range (35-55)

- Separated Individuals who bought mobile and their income range.

Low	**2185**
Mid	**8242**
High	**4856**



- Then it is further segregated according to phone's age and their budget.

Range	Explanation	Average Of Budget
L1	Low Income Individual with Phone's age less than 1.25 years (456 days)	11500
L2	Low Income Individual with Phone's age more than 1.25 years (456 days)	13500
M1	Mid Income Individual with Phone's age less than a year (365 days)	17000
M2	Mid Income Individual with Phone's age more than a year but less than 1.25 year (365 - 456 days)	22000
M3	Mid Income Individual with Phone's age more than 1.25 years(456days)	25000
H1	High Income Individual with Phone's age less than a year(365days)	30000
H2	High Income Individual with Phone's age more than a year(365days)	32500

Task3: Justification For Business Decision

Then According to the table given we find unit sold under that category and find estimated revenue earned.

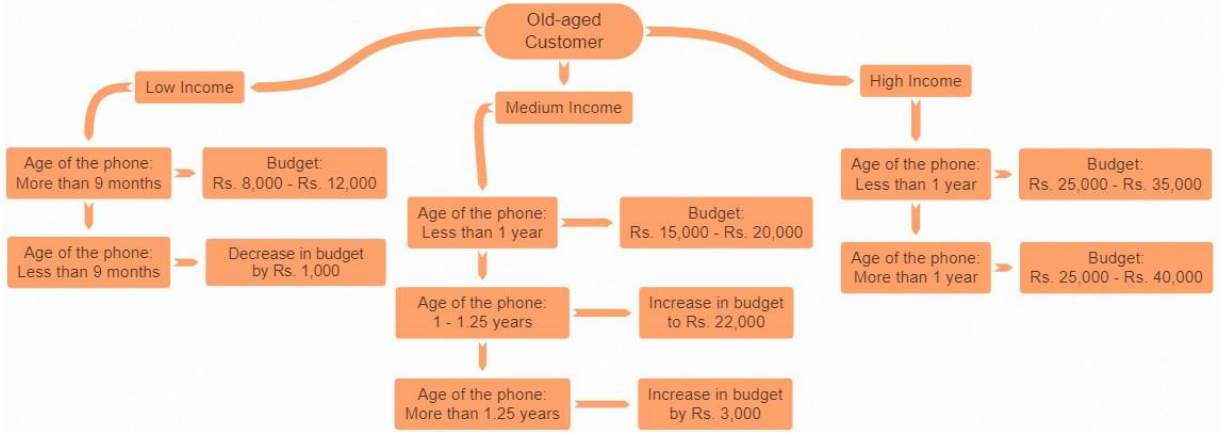
Range	Number	Expected Revenue
L1	813	9349500
L2	1372	18522000
M1	0	0
M2	3043	66946000
M3	5199	129975000
H1	0	0
H2	4856	157820000
Total	15283	382612500

Task3: Justification For Business Decision

For Individuals in Old Age Range (55-65)

- Separated Individuals who bought mobile and their income range.

Low	**2939**
Mid	**2893**
High	**2798**



- Then it is further segregated according to phone's age and their budget.

Range	Explanation	Average Of Budget
L1	Low Income Individual with Phone's age more than 9months(274days)	10000
L2	Low Income Individual with Phone's age less than 9months(274 days)	9000
M1	Mid Income Individual with Phone's age less than a year (365 days)	17500
M2	Mid Income Individual with Phone's age more than a year but less than 1.25 year (365 - 456 days)	22000
M3	Mid Income Individual with Phone's age more than 1.25 years(456days)	25000
H1	High Income Individual with Phone's age less than a year(365days)	30000
H2	High Income Individual with Phone's age more than a year(365days)	32500

Task3: Justification For Business Decision

Then According to the table given we find unit sold under that category and find estimated revenue earned.

Range	Number	Expected Revenue
L1	2939	29390000
L2	0	0
M1	0	0
M2	1078	23716000
M3	1815	45375000
H1	0	0
H2	2798	90935000
Total	8630	189416000

Task3: Justification For Final Business Decision

It was told to us XYZ_Mobiles will capture Minimum of 40% of Market.

- So, after we found out actual sample data of mobile unit sold and revenue generated we further took out 40% of it to see whether it matches the 2 conditions which was given to us which is

- **(1)Unit sold must be more than 12000 and**
- **(2)Revenue generated must be more than 20 crore**

	Unit sold	XYZ_unit	Total_Rev	XYZ_Rev
Young	7660	3064	85850000	34340000
Mid	15283	6113.2	382612500	153045000
Old	8630	3452	189416000	75766400
Total	31573	12629.2	657878500	263151400

Conditions Given

XYZ_Mobiles will atleast take 40% of market
In sample it completes criteria of selling more than 12000 unit and if XYZ takes atleast 40% market it will still sell more than 12000 unit.

In sample it completes criteria of generating revenue of more than 20 crore and if XYZ takes atleast 40% market it will still generate revenue of more than 20 Crore.

Task3: Justification For Final Business Decision

Through the information given in previous slide we can safely say.

XYZ Mobiles can safely enter Indian Market and generate good revenue.

Thank You