

# Assignment - Market Entry Analysis

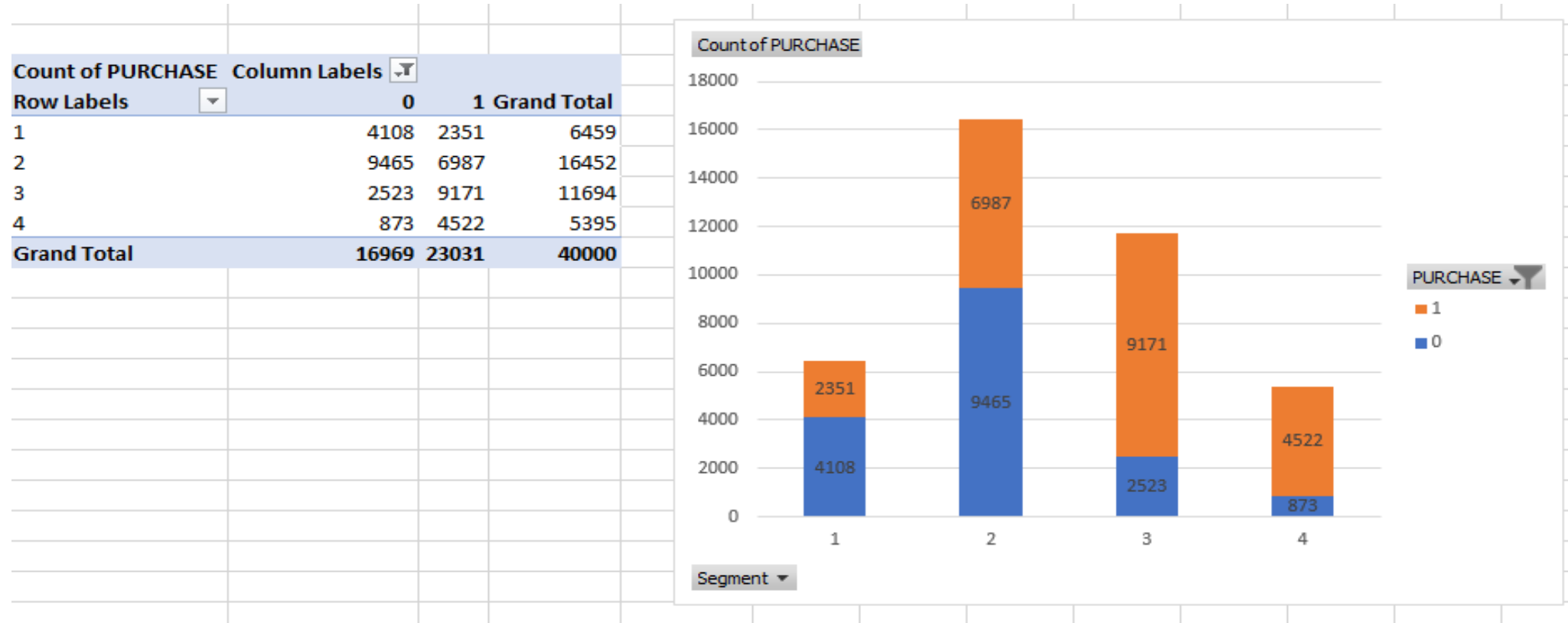
## Chinese Data Set

A	B	C	D	E	F	G	H	I	J	K	L
ID	CURR_AGE	Age Group	GENDER	ANN_INCOME	AGE_PHN	PURCHASE	Segment	INCOME Segment			
00001Q15YJ	50	50-59	M	4,45,344	439	0	3	4			
00003I71CQ	35	30-39	M	1,07,634	283	0	2	1			
00003N47FS	59	50-59	F	5,02,787	390	1	3	5			
00005H41DE	43	40-49	M	5,85,664	475	0	3	5			
00007E17UM	39	30-39	F	7,05,723	497	1	3	7			
00007I26OR	28	20-29	F	3,89,995	443	1	3	3		Age	Group
00015B11UO	54	50-59	M	85,056	425	0	3	0		20	20-29
00020K99TA	28	20-29	F	4,53,584	173	0	1	4		30	30-39
00020W72QC	25	20-29	F	3,24,575	300	0	2	3		40	40-49
00022F48XA	47	40-49	M	3,63,206	474	1	3	3		50	50-59
00026X43XZ	61	60-69	M	2,92,983	210	0	2	2		60	60-69
00031Q27QZ	33	30-39	F	4,38,941	100	1	1	4			
00032B38ZX	50	50-59	M	4,78,445	329	1	2	4			
00033C02IM	25	20-29	M	2,77,729	430	1	3	2			
00034P01OK	45	40-49	M	3,85,604	492	1	3	3			
00038B31VO	51	50-59	F	1,59,413	585	1	4	1			
00039X03RX	25	20-29	M	2,27,439	23	0	1	2			
00040B49KN	43	40-49	M	5,72,440	356	0	2	5			
00040O73KD	35	30-39	M	2,40,251	536	1	4	2			

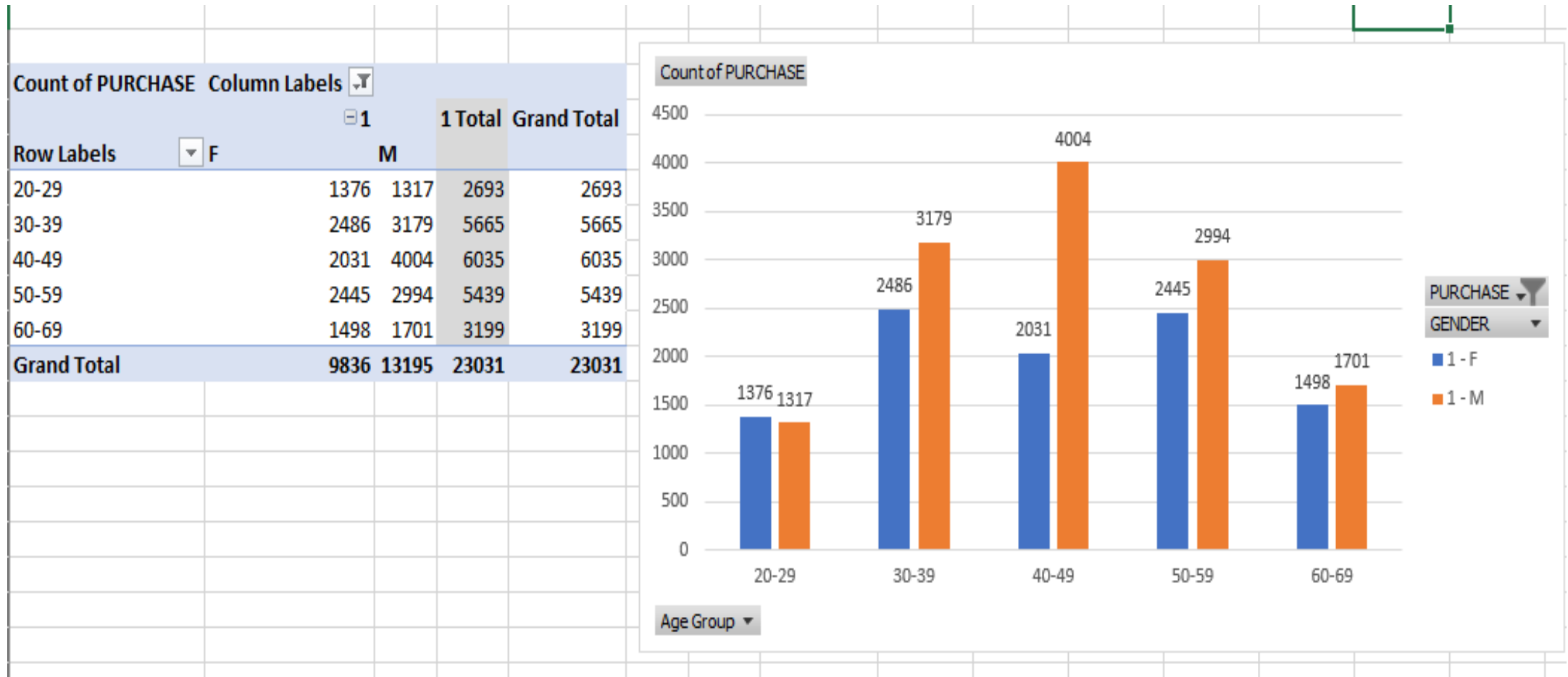
- Data set has been cleaned for analysis
- New Column Age group has been created to make analysis
- New column Income Segment has been created
- New Column Segment has been created by using AGE PHN column based on following conditions

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

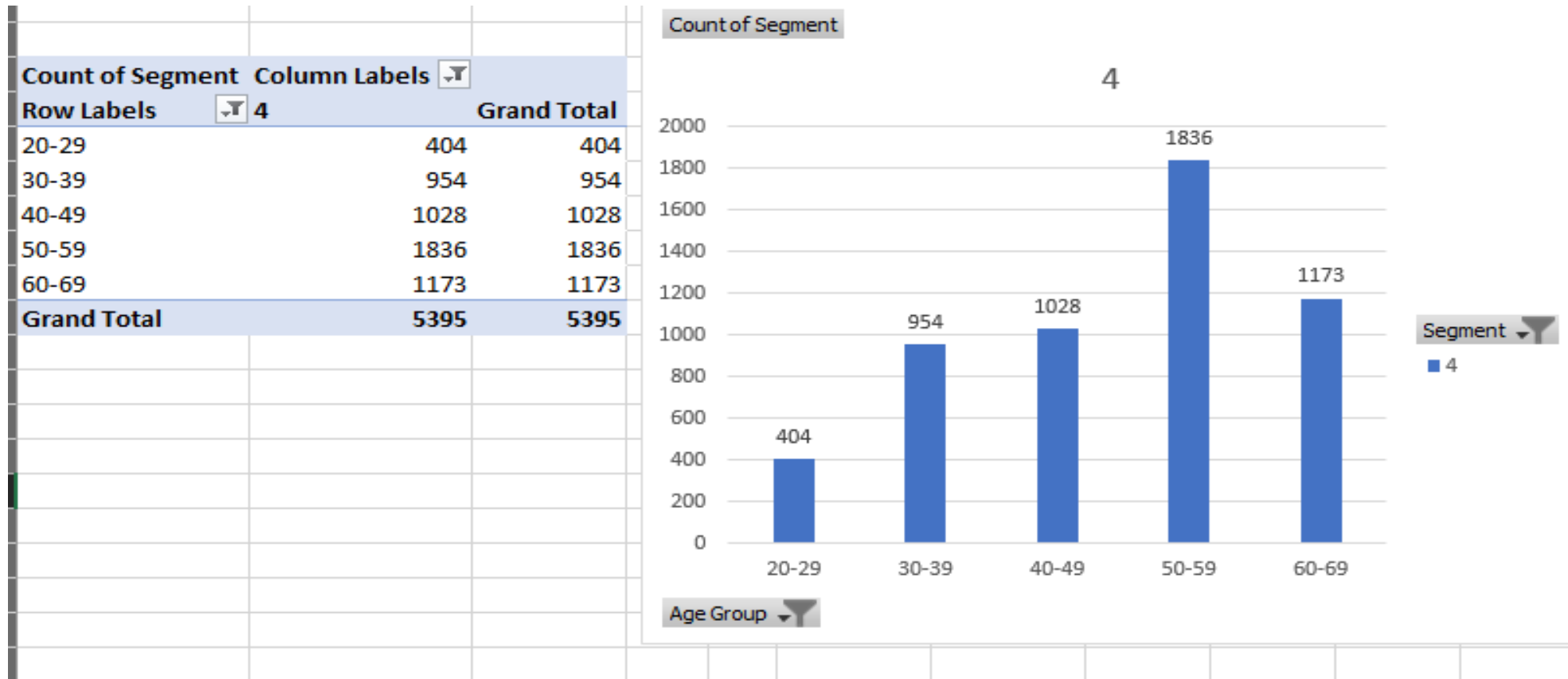
- When comparing and visualizing Age of the phone segments and Purchase rate , its is clear that people with the phone age between 360-500 has done maximum number of purchases.



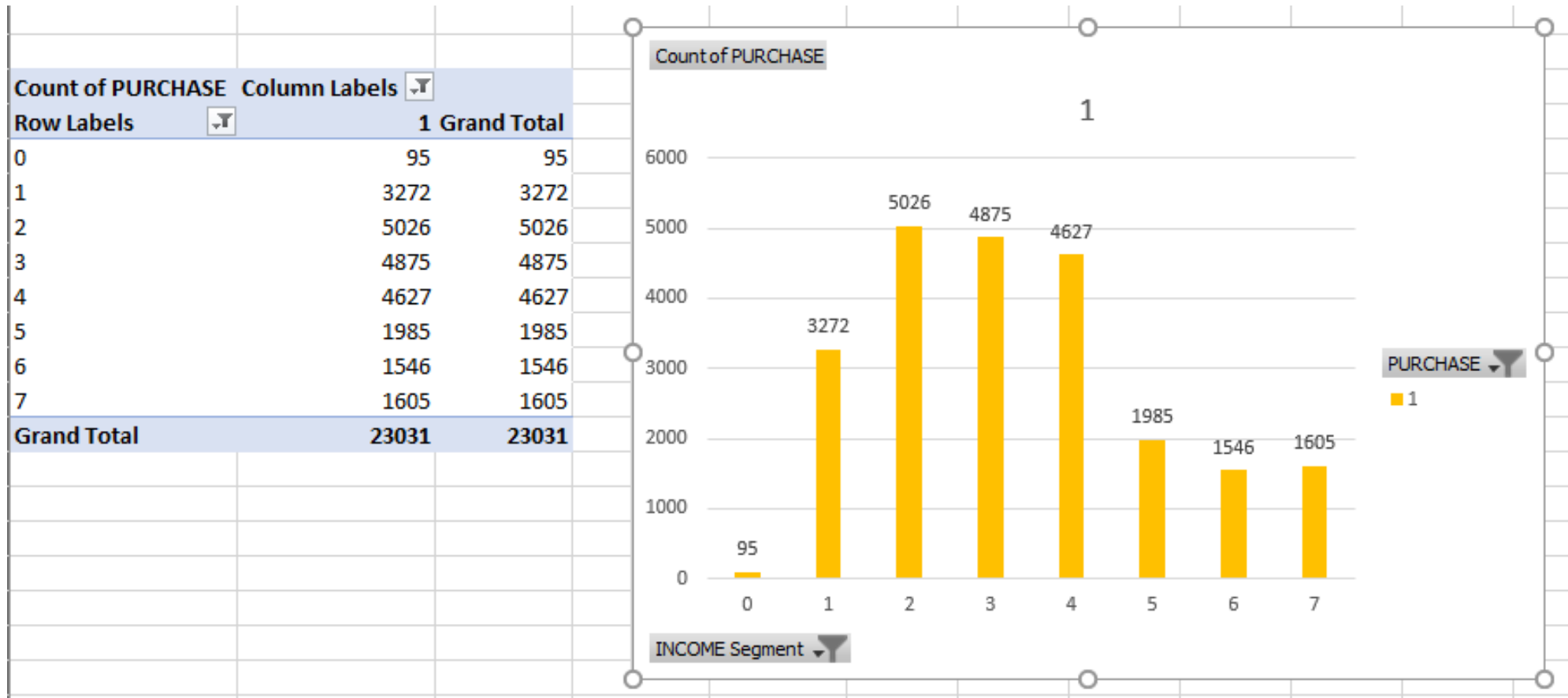
- Age group between 40-49 males have high number of purchase rate and 30-39 aged Females have high number of Purchase rate



- People aged between 50-59 are using their old phones for Maximum number of days when compared to other age groups



- People ranging between 2 to 3 lakhs of income have maximum number of purchases



	A	B	C	D	E
		<i>CURR_AGE</i>	<i>INCOME Segment</i>	<i>Segment</i>	<i>PURCHASE</i>
<i>CURR_AGE</i>		1			
<i>INCOME Segment</i>		0.185625375	1		
<i>Segment</i>		0.043469568	0.032797303	1	
<i>PURCHASE</i>		-0.012129471	0.16459209	0.367895	1

- On doing bivariate analysis it has been found that age and income has high Correlation , so higher the age high is the income . But Purchase does not depend upon age because it has negative correlation
- Phone Age and purchase rate are highly correlated.
- People with high income have more purchase rate , it has positive correlation



## Statistics of Chinese Data Set

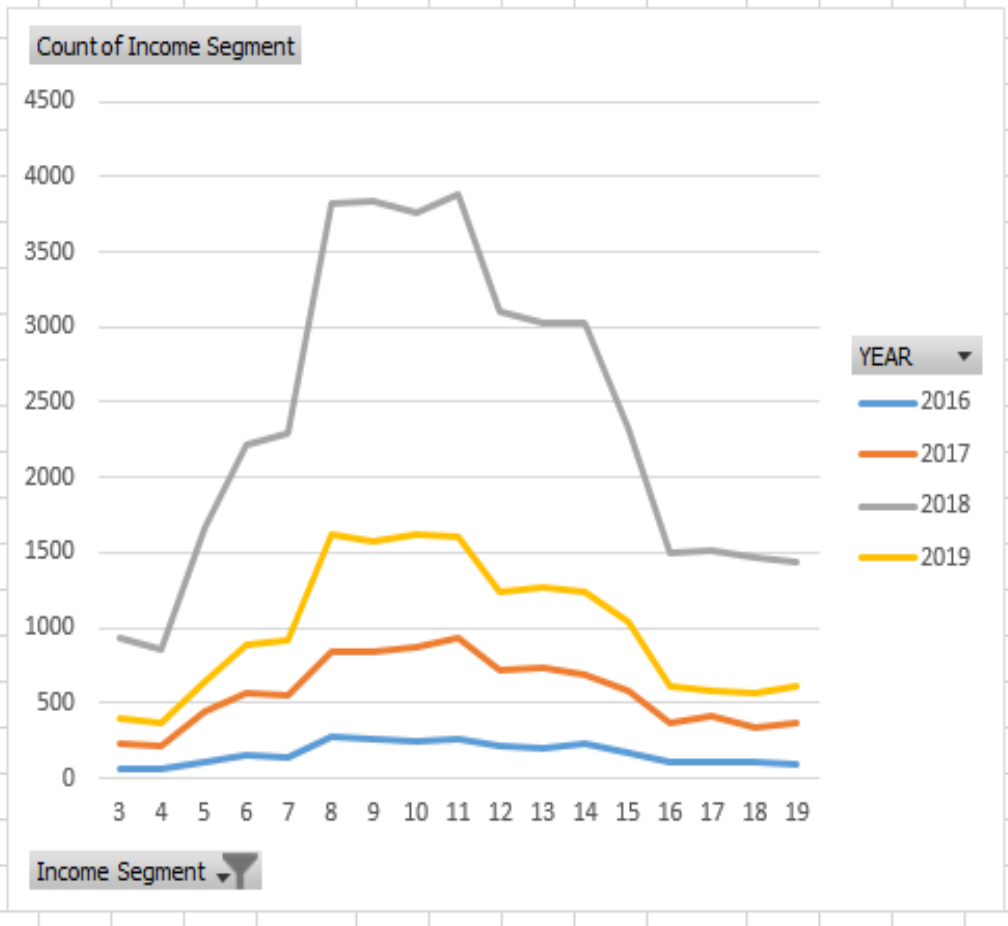
Descriptive Statistics				Multiplier	2.2	Outliers and Missing Data						
	50	4	3		50	4	3		50	4	3	
Mean	44.99732	3.091627	2.40061	Min	25	0	1	mean	44.99732	3.091627	2.40061	
Standard Error	0.059102	0.008743	0.004563	Q1-Min	10	2	1	stdev	11.8202	1.748509	0.91261	
Median	45	3	2	Med-Q1	10	1	0	# outliers	0	0	0	
Mode	46	2	2	Q3-Med	10	1	1	# blank	0	0	0	
Standard Deviation	11.8202	1.748509	0.91261	Max-Q3	10	3	1	# non-num	0	0	0	
Sample Variance	139.7172	3.057283	0.832857	Mean	44.99732	3.091627	2.40061	1	-0.84578	-1.19624	-0.43897	
Kurtosis	-1.2019	-0.40391	-0.76191					2	1.184639	1.091429	0.656786	
Skewness	0.0061	0.572997	0.190337	Min	25	0	1	3	-0.16898	1.091429	0.656786	
Range	40	7	3	Q1	35	2	2	4	-0.50738	2.235261	0.656786	
Maximum	65	7	4	Median	45	3	2	5	-1.43799	-0.0524	0.656786	
Minimum	25	0	1	Q3	55	4	3	6	0.761635	-1.76815	0.656786	
Sum	1799848	123662	96022	Max	65	7	4	7	-1.43799	0.519513	-1.53473	
Count	39999	39999	39999	Mean	44.99732	3.091627	2.40061	8	-1.69179	-0.0524	-0.43897	
AAD	10.23304	1.413827	0.781888					9	0.169428	-0.0524	0.656786	
MAD	10	1	1	Grand Min	0			10	1.353841	-0.62432	-0.43897	
IQR	20	2	1					11	-1.01498	0.519513	-1.53473	
				Outliers	None	None	None	12	0.423231	0.519513	-0.43897	
								13	-1.69179	-0.62432	0.656786	
								14	0.000226	-0.0524	0.656786	

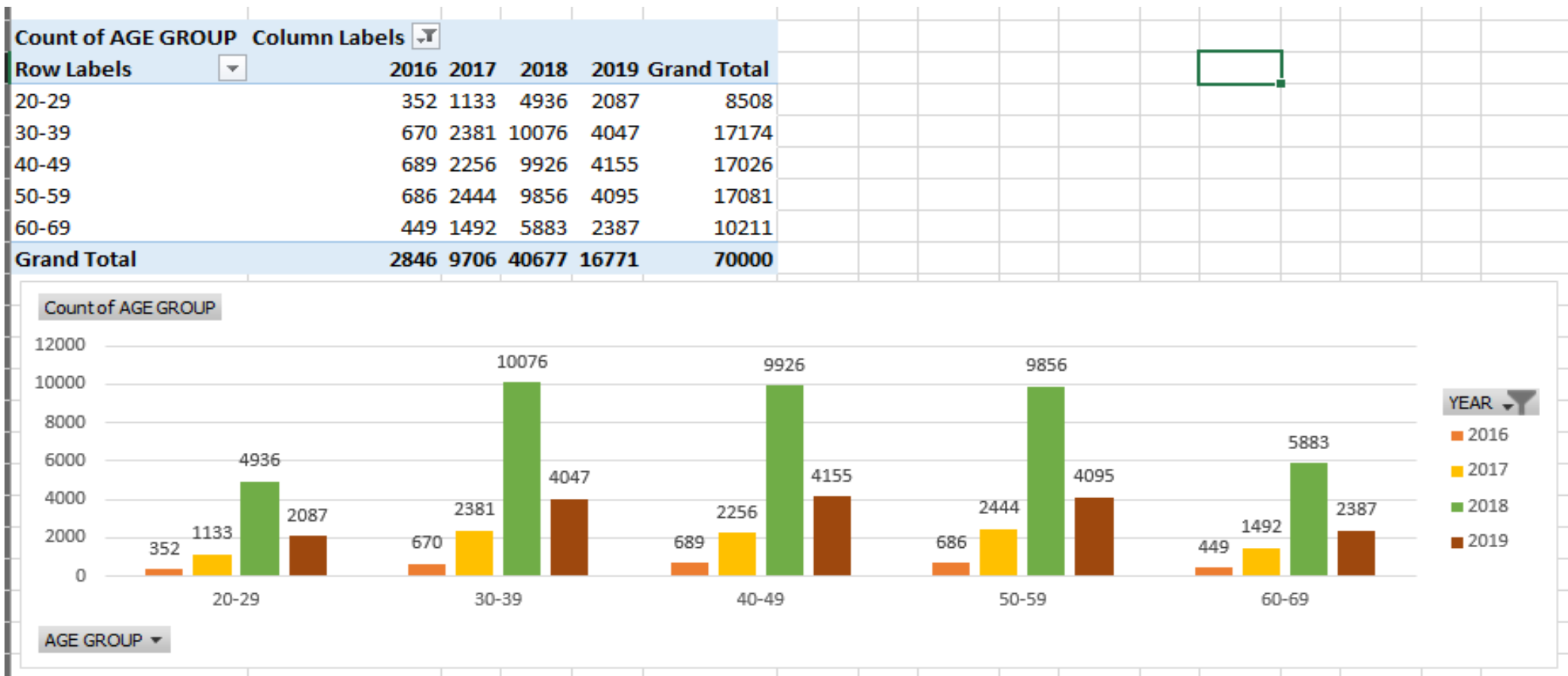
- Indian Data set

Cleaned data set with new columns Income segment and Year of Old Purchase

A	B	C	D	E	F	G	H	I
ID	CURR_AGE	AGE GROUP	GENDER	ANN_INCOME	Income Segment	DT_OLD_PURCHASE	YEAR	
20710B05XL	54	50-59	M	14,25,390	14	20-04-2018	2018	
89602T51HX	47	40-49	M	16,78,954	16	08-06-2018	2018	
70190Z52IP	60	60-69	M	9,31,624	9	31-07-2017	2017	
25623V15MU	55	50-59	F	11,06,320	11	31-07-2017	2017	
36230I68CE	32	30-39	F	7,48,465	7	27-01-2019	2019	
11264G01HZ	48	40-49	F	10,51,927	10	24-11-2018	2018	
74250S23UO	26	20-29	F	10,76,402	10	22-09-2018	2018	
26735J66DB	45	40-49	F	14,81,949	14	05-04-2018	2018	
93404P60ED	55	50-59	M	17,25,607	17	02-01-2018	2018	
56557A36QV	64	60-69	F	3,12,323	3	23-04-2018	2018	
38353F50LZ	53	50-59	M	5,46,574	5	05-06-2019	2019	
54684T21RX	44	40-49	F	12,03,691	12	12-07-2017	2017	
46929E04HS	59	50-59	F	7,24,688	7	22-06-2019	2019	
20647X82EQ	27	20-29	F	9,75,130	9	05-03-2019	2019	
34956P25RT	57	50-59	F	14,22,399	14	14-04-2019	2019	
07090V20JQ	40	40-49	F	15,58,045	15	06-02-2018	2018	
78392T89DQ	33	30-39	M	6,69,737	6	31-05-2019	2019	
07257K04CB	57	50-59	F	7,74,593	7	31-12-2018	2018	
65658K80PS	59	50-59	F	9,93,201	9	05-03-2019	2019	
69803K32CS	42	40-49	F	10,50,793	10	21-10-2016	2016	
49525O29YH	40	40-49	M	15,98,014	15	18-03-2018	2018	
25740R14MI	63	60-69	M	7,76,801	7	25-07-2017	2017	

Count of Income Segment	Column Labels				
Row Labels	2016	2017	2018	2019	Grand Total
3	65	226	929	399	1619
4	65	218	860	372	1515
5	112	440	1669	650	2871
6	162	564	2211	884	3821
7	142	549	2298	913	3902
8	277	846	3815	1618	6556
9	270	839	3839	1569	6517
10	253	881	3765	1615	6514
11	258	940	3889	1599	6686
12	223	715	3105	1240	5283
13	208	733	3023	1268	5232
14	226	697	3027	1240	5190
15	165	581	2322	1034	4102
16	115	365	1496	607	2583
17	114	410	1519	578	2621
18	104	334	1473	567	2478
19	87	368	1437	618	2510
<b>Grand Total</b>	<b>2846</b>	<b>9706</b>	<b>40677</b>	<b>16771</b>	<b>70000</b>





- These Visualizations clearly shows that maximum number of the old phones are purchased in the year 2018

A	B	C	D
	<i>CURR_AGE</i>	<i>Income Segment</i>	<i>YEAR</i>
<i>CURR_AGE</i>	1		
<i>Income Segment</i>	0.137319393	1	
<i>YEAR</i>	-0.008345891	-0.001985581	1

- From the correlation it has been found that income does not affect the purchase rate of people because it is negatively correlated.

	A	B	C	D
1	ID	PURCHASE	Age Category	Income Category
2	00001Q15YJ	N	Old	Medium
3	00003I71CQ	N	Medium-Aged	Low
4	00003N47FS	Y	Old	High
5	00005H41DE	N	Medium-Aged	High
6	00007E17UM	Y	Medium-Aged	High
7	00007I26OR	Y	Young	Medium
8	00015B11UO	N	Old	Low
9	00020K99TA	N	Young	Medium
10	00020W72QC	N	Young	Medium
11	00022F48XA	Y	Medium-Aged	Medium
12	00026X43XZ	N	Old	Medium
13	00031Q27QZ	Y	Medium-Aged	Medium
14	00032B38ZX	Y	Old	Medium
15	00033C02IM	Y	Young	Medium
16	00034P01OK	Y	Medium-Aged	Medium
17	00038B31VO	Y	Old	Low
18	00039X03RX	N	Young	Medium
19	00040B49KN	N	Medium-Aged	High
20	00040O73KD	Y	Medium-Aged	Medium

- Categorized data set

- Transformed Data

A	B	C	D	E
Y		Income		
		Low	Medium	High
Age	Young	692	2001	0
	Medium-Aged	1567	7503	2630
	Old	1108	5024	2506
N		Income		
		Low	Medium	High
Age	Young	811	1318	0
	Medium-Aged	1997	4786	1104
	Old	2191	3387	1375
Total		Income		
		Low	Medium	High
Age	Young	1503	3319	0
	Medium-Aged	3564	12289	3734
	Old	3299	8411	3881

- Accuracy , Gini Score , Entropy and Info Gain

Y/N		Income			Accuracy
		Low	Medium	High	
Age	Young	0.54	0.60	0.00	
	Medium-Aged	0.56	0.61	0.70	
	Old	0.66	0.60	0.65	

Y/N		Income			Gini Score
		Low	Medium	High	
Age	Young	0.50	0.52	0.00	
	Medium-Aged	0.51	0.52	0.58	
	Old	0.55	0.52	0.54	

Y/N		Income			Entropy
		Low	Medium	High	
Age	Young	0.52	0.44	0.00	
	Medium-Aged	0.52	0.43	0.36	
	Old	0.53	0.44	0.41	

Y/N		Income			Info Gain
		Low	Medium	High	
Age	Young	0.48	0.56	1.00	
	Medium-Aged	0.48	0.57	0.64	
	Old	0.47	0.56	0.59	



- Featured Score

Score	Low	Medium	High
Accuracy	4999.0	14528.0	5136.0
Gini	4391.4	12539.4	4284.1
IG	3989.9	13482.3	4703.7
Total Count	8366	24019	7615
Score (Income)	429880306.0	Accuracy	
	370546722.8	Gini Score	
	393030388.4	IG	

[

- With the analysis done by using Indian dataset maximum number of old phone are bought in the year 2018
- On analysing Chinese data set most number of people are likely to change their phones within 360 – 500 days
- The purchase rate of people who bought new phones has been 57%
- Based on analysis done, **Xyz mobiles** can enter Indian Market.

