



Vasilis Tillis

Case Study

Task 1

Introduction

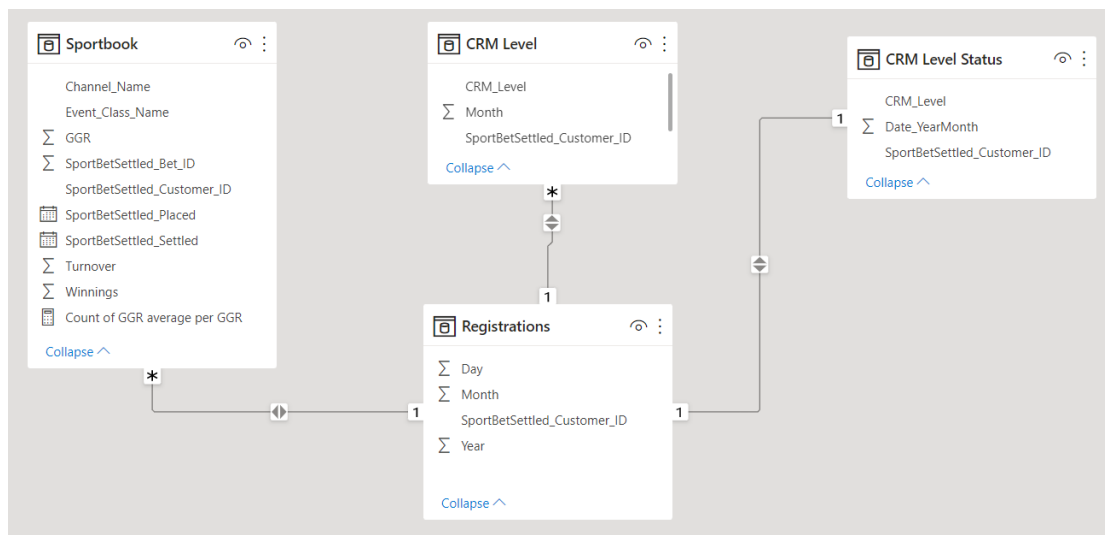
This case study is for the data analyst position at Kaizen Gaming which involves analyzing data to determine whether German sportsbook players bet more on club soccer leagues or international football championships. The results will be used by the sportsbook team to decide whether it is worth investing more in the upcoming World Cup.

Power BI has been selected as the data visualization tool.

Correlation of tables

Initially, we will demonstrate the relationship between the tables.

The correlation between the tables, based on the column "**SportBetSettled_Customer_ID**" is displayed below.



Please take note that the "**CRM Level Status**" table was generated to enable us to determine **the latest status of each customer by date**, which could be bronze, silver, gold, platinum or diamond. For example, if a customer's status was bronze in February 2021 and was upgraded to silver in June 2021, the customer's final status would be silver.

The task is to fill out the following tables with data.

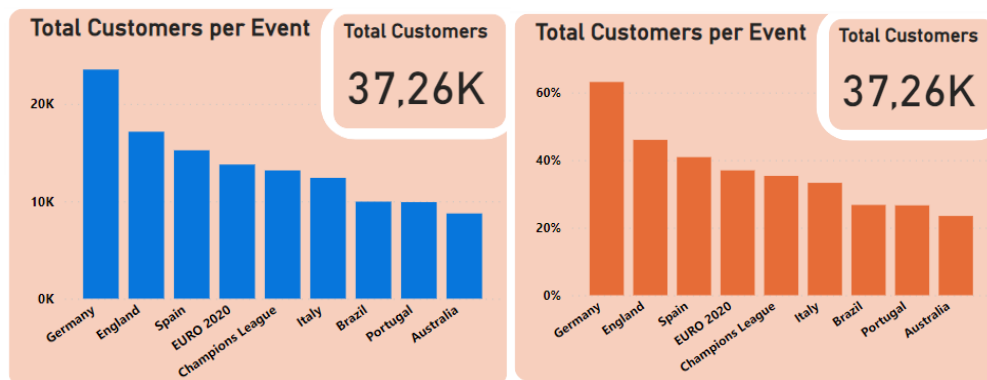
Table 1

KPIs per Championship	Customers	Customers %	TO	TO %	GGR	GGR%	Bets	Bets %	First Bets	First Bets %	Avg Customer Value	Avg Bet Value
Championships												

The objective is to complete the missing elements for each event.

a. "Customers" and "Customers%"

We need to determine the number of customers who participate in each event and their percentage in these two columns. By utilizing Power BI and leveraging the **"registrations"** and **"sportsbook"** tables, we can perform the necessary calculations.

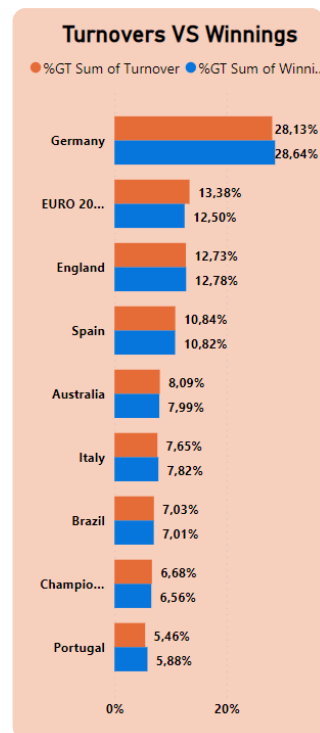


From the graph above, we can observe that the total number of customers is **37,262**. With this information, we can proceed to populate the table below with the relevant data, organized by event.

	A	B	C
1	KPIs per Championship	Customers	Customers %
2	Germany	23526	63,14
3	EURO 2020	13772	39,96
4	England	17141	46
5	Spain	15233	40,88
6	Australia	8746	23,47
7	Italy	12401	33,28
8	Brazil	9961	26,73
9	Champions League	13166	35,33
10	Portugal	9913	26,6

b. "TO" and "TO%"

We need to complete the two columns with the values of the bets per event and their respective percentage, which can be obtained from the "turnover" column in the "Sportsbook" table. By referring to the chart below, we can compare the turnovers against the winnings for each competition.



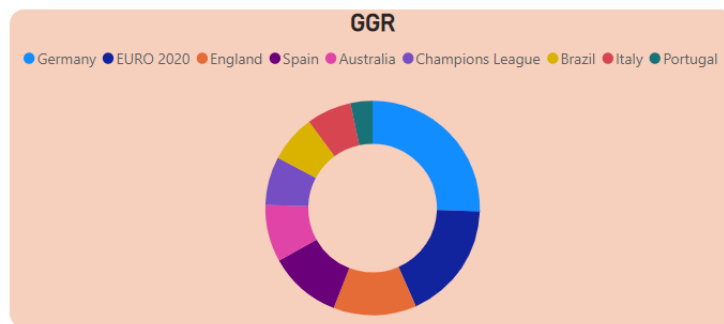
With reference to the data presented above, we can proceed to fill these two columns:

	A	B	C	D	E
1	KPIs per Championship	Customers	Customers %	TO	TO%
2	Germany	23526	63,14	8699368440000000	28,13
3	EURO 2020	13772	39,96	4139191260000000	13,38
4	England	17141	46	3938095128000000	12,73
5	Spain	15233	40,88	3352028112000000	10,84
6	Australia	8746	23,47	2502530616000000	8,09
7	Italy	12401	33,28	2365970424000000	7,65
8	Brazil	9961	26,73	2174384004000000	7,03
9	Champions League	13166	35,33	2065707072000000	6,68
10	Portugal	9913	26,6	1689294024000000	5,46

c. "GGR" and "GGR%"

GGR refers to the term "**Gross Gaming Revenue**" and is calculated by subtracting winnings from turnovers. For this reason, we created a new column named "GGR", which has the following subtraction.

$$\text{GGR} = \text{Turnover} - \text{Winnings}$$

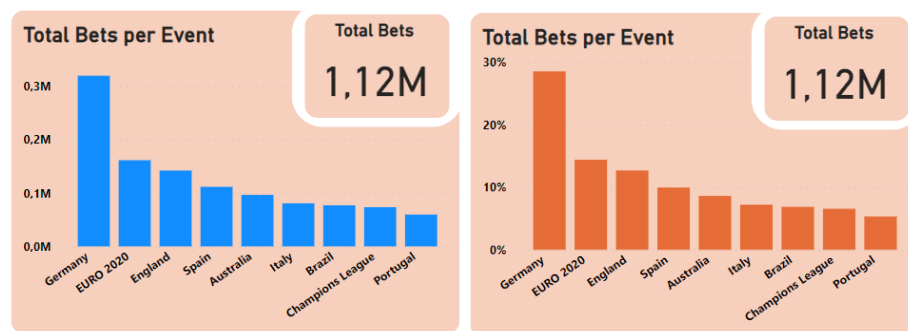


With reference to the graph presented above, which indicates the result between turnovers and winnings, we can proceed to fill in the remaining portion of the table:

	A	B	C	D	E	F	G
1	KPIs per Championship	Customers	Customers %	TO	TO%	GGR	GGR%
2	Germany	23526	63,14	8699368440000000	28,13	1283869249450000	25,48
3	EURO 2020	13772	39,96	4139191260000000	13,38	902776331000000	17,92
4	England	17141	46	3938095128000000	12,73	630695944340000	12,52
5	Spain	15233	40,88	3352028112000000	10,84	551776084130000	10,95
6	Australia	8746	23,47	2502530616000000	8,09	434014579930000	8,61
7	Italy	12401	33,28	2365970424000000	7,65	367372271860000	7,29
8	Brazil	9961	26,73	2174384004000000	7,03	360395837680000	7,15
9	Champions League	13166	35,33	2065707072000000	6,68	340568861530000	6,76
10	Portugal	9913	26,6	1689294024000000	5,46	167336036120000	3,32

d. "Bets" and "Bets%"

Next, we want to calculate the total bets by all customers per event and this can be seen in the graph below:



We can now proceed to fill in the remaining portion of the table1:

	A	B	C	D	E	F	G	H	I
1	KPIs per Championship	Customers	Customers %	TO	TO%	GGR	GGR%	Bets	Bets%
2	Germany	23526	63,14	8699368440000000	28,13	1283869249450000	25,48	319518	28,50
3	EURO 2020	13772	39,96	4139191260000000	13,38	902776331000000	17,92	161264	14,38
4	England	17141	46	3938095128000000	12,73	630695944340000	12,52	142045	12,67
5	Spain	15233	40,88	3352028112000000	10,84	551776084130000	10,95	111481	9,94
6	Australia	8746	23,47	2502530616000000	8,09	434014579930000	8,61	96446	8,60
7	Italy	12401	33,28	2365970424000000	7,65	367372271860000	7,29	80697	7,20
8	Brazil	9961	26,73	2174384004000000	7,03	360395837680000	7,15	76880	6,86
9	Champions League	13166	35,33	2065707072000000	6,68	340568861530000	6,76	73349	6,54
10	Portugal	9913	26,6	1689294024000000	5,46	167336036120000	3,32	59532	5,31

e. "First Bets" and "First Bets%"

First bets should refer to customers that have registered within the period in scope. The period of scope is October 2020- August 2021.

In order to find the number of first bets and their percentage, we will find the customers who registered between October 2020 and August 2021 and find out how many bets were placed per event. The graph presented below illustrates this data:



We can now proceed to fill in the remaining portion of the table1:

	A	B	C	D	E	F	G	H	I	J	K
1	KPIs per Championship	Customers	Customers %	TO	TO%	GGR	GGR%	Bets	Bets%	First Bets	First Bets%
2	Germany	23526	63,14	8699368440000000	28,13	1283869249450000	25,48	319518	28,50	17819	60,3
3	EURO 2020	13772	39,96	4139191260000000	13,38	902776331000000	17,92	161264	14,38	10459	35,39
4	England	17141	46	3938095128000000	12,73	630695944340000	12,52	142045	12,67	12851	43,49
5	Spain	15233	40,88	3352028112000000	10,84	551776084130000	10,95	111481	9,94	11330	38,34
6	Australia	8746	23,47	2502530616000000	8,09	434014579930000	8,61	96446	8,60	6303	21,33
7	Italy	12401	33,28	2365970424000000	7,65	367372271860000	7,29	80697	7,20	8959	30,32
8	Brazil	9961	26,73	2174384004000000	7,03	360395837680000	7,15	76880	6,86	7370	24,94
9	Champions League	13166	35,33	2065707072000000	6,68	340568861530000	6,76	73349	6,54	9511	32,19
10	Portugal	9913	26,6	1689294024000000	5,46	167336036120000	3,32	59532	5,31	7080	23,96

f. "Avg Customer Value"

To find the "Avg Customer Value", we need to divide the turnovers by the total customers per event. Therefore, we can complete the table1 as follows:

	A	B	C	D	E	F	G	H	I	J	K	L
1	KPIs per Championship	Customers	Customers %	TO	TO%	GGR	GGR%	Bets	Bets%	First Bets	First Bets%	Avg Customer Value
2	Germany	23526	63,14	8699368440000000	28,13	1283869249450000	25,48	319518	28,50	17819	60,3	369776776332,57
3	EURO 2020	13772	39,96	4139191260000000	13,38	902776331000000	17,92	161264	14,38	10459	35,39	300551209700,84
4	England	17141	46	3938095128000000	12,73	630695944340000	12,52	142045	12,67	12851	43,49	229747105069,72
5	Spain	15233	40,88	3352028112000000	10,84	551776084130000	10,95	111481	9,94	11330	38,34	220050424210,60
6	Australia	8746	23,47	2502530616000000	8,09	434014579930000	8,61	96446	8,60	6303	21,33	286134303224,33
7	Italy	12401	33,28	2365970424000000	7,65	367372271860000	7,29	80697	7,20	8959	30,32	190788680267,72
8	Brazil	9961	26,73	2174384004000000	7,03	360395837680000	7,15	76880	6,86	7370	24,94	218289730348,36
9	Champions League	13166	35,33	2065707072000000	6,68	340568861530000	6,76	73349	6,54	9511	32,19	156897088865,26
10	Portugal	9913	26,6	1689294024000000	5,46	167336036120000	3,32	59532	5,31	7080	23,96	170411986684,15

g. "Avg Bet Value"

To find the "Avg Bet Value", we must divide the turnovers by the total bets per competition. Below you can see the Table 1 completed

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	KPIs per Championship	Customers	Customers %	TO	TO%	GGR	GGR%	Bets	Bets%	First Bets	First Bets%	Avg Customer Value	Avg Bet Value
2	Germany	23526	63,14	8699368440000000	28,13	1283869249450000	25,48	319518	28,50	17819	60,3	369776776332,57	27226536345
3	EURO 2020	13772	39,96	4139191260000000	13,38	902776331000000	17,92	161264	14,38	10459	35,39	300551209700,84	25667174695
4	England	17141	46	3938095128000000	12,73	630695944340000	12,52	142045	12,67	12851	43,49	229747105069,72	27724278419
5	Spain	15233	40,88	3352028112000000	10,84	551776084130000	10,95	111481	9,94	11330	38,34	220050424210,60	30068156116
6	Australia	8746	23,47	2502530616000000	8,09	434014579930000	8,61	96446	8,60	6303	21,33	286134303224,33	25947479584
7	Italy	12401	33,28	2365970424000000	7,65	367372271860000	7,29	80697	7,20	8959	30,32	190788680267,72	29319186884
8	Brazil	9961	26,73	2174384004000000	7,03	360395837680000	7,15	76880	6,86	7370	24,94	218289730348,36	28282830437
9	Champions League	13166	35,33	2065707072000000	6,68	340568861530000	6,76	73349	6,54	9511	32,19	156897088865,26	28162716220
10	Portugal	9913	26,6	1689294024000000	5,46	167336036120000	3,32	59532	5,31	7080	23,96	170411986684,15	28376235033

Within the Power BI file, there is a sheet named "**Summary**" that displays the analysis above. By utilizing filters, we can explore the graphs to gain different insights.

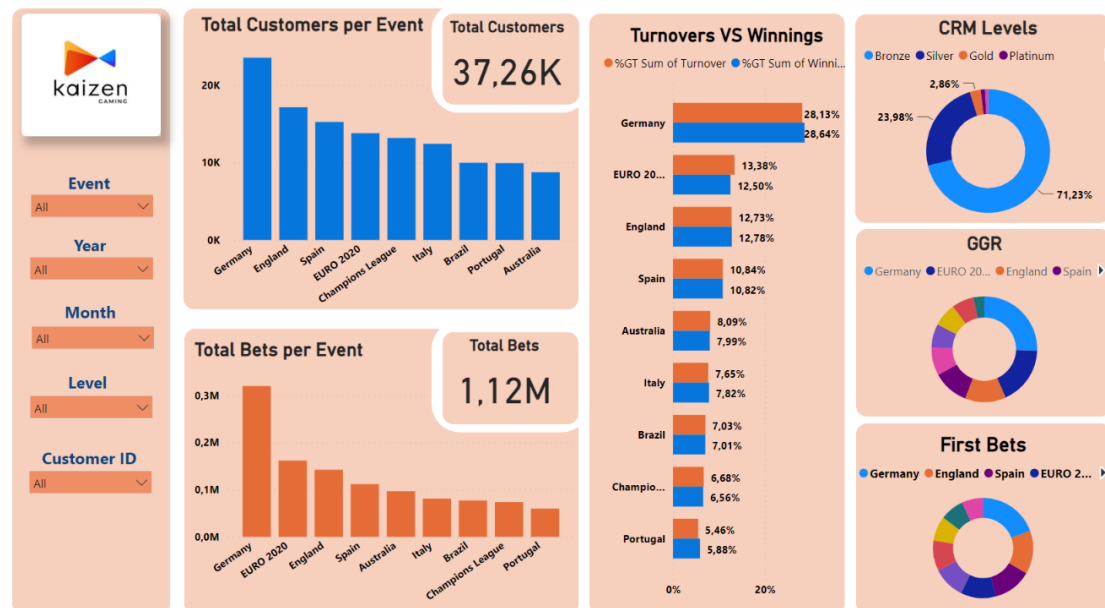


Table 2

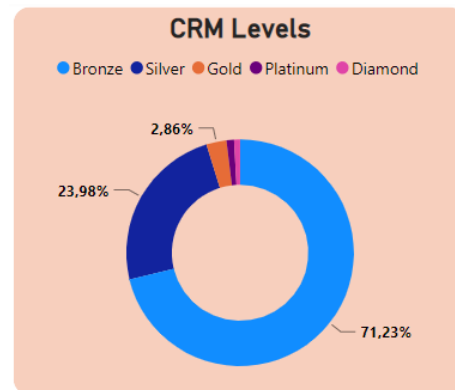
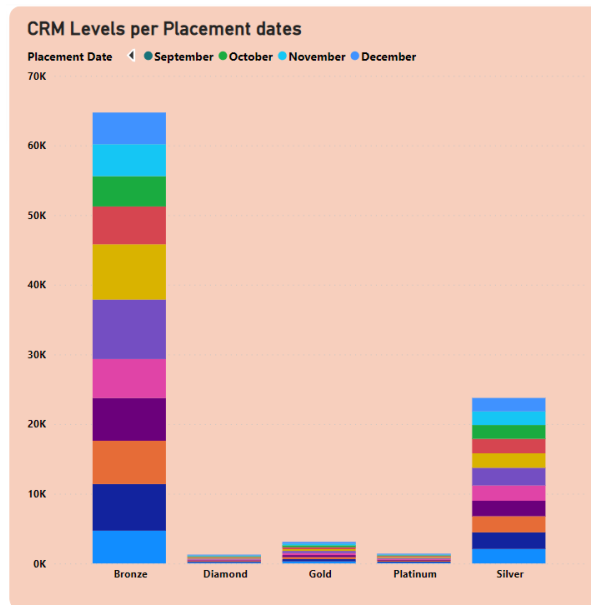
Initially, we generated a new table named "**CRM Level Status**" which maintains a record of each customer's **most recent level**. For instance, if a player was at the bronze level in October 2020 and progressed to silver on January 21, we will retain the silver status.

Our objective is to find the final level and the month of acquisition for all 37,262 customers.

To determine this, we transformed the data in the CRM_Level table.

Firstly, we identified **duplicate records** in the "**SportBetSettled_Customer_ID**" column. Then, we **sorted** the data in **ascending order**, **first by "SportBetSettled_Customer_ID"** column, **followed by the "Date_YearMonth"** column. After the sorting process, we deleted the duplicate records from the "SportBetSettled_Customer_ID" column and kept the newest ones. The new table named "**CRM_Level Status**".

The below graphs exhibit the distribution of players based on level and the month of attainment, along with their respective percentages.



Therefore, we can complete the table2 as follows:

CRM Levels over Month	Bronze	Silver	Gold	Platinum	Diamond
2020 October	4348	1975	263	103	100
2020 Novemner	4586	1952	247	108	117
2020 December	4557	1936	250	126	119
2021 January	4694	2077	286	121	120
2021 February	6699	2377	325	133	119
2021 March	6210	2317	293	137	110
2021 April	6146	2215	293	127	110
2021 May	5591	2196	284	119	109
2021 June	8536	2548	310	144	125
2021 July	7930	2077	254	130	107
2021 August	5423	2073	282	152	121
2021 September	8	9	1	0	5
	64720	23743	3087	1400	1257

The "CRM Levels" sheet in the Power BI file includes an analysis that allows us to view customers by their latest level and the month they obtained it through filters. Additionally, a table has been provided that displays the top 100 customers based on their turnovers. Moreover, it is possible to search for any customer using their "customerID".

Below we can see the complete analysis:

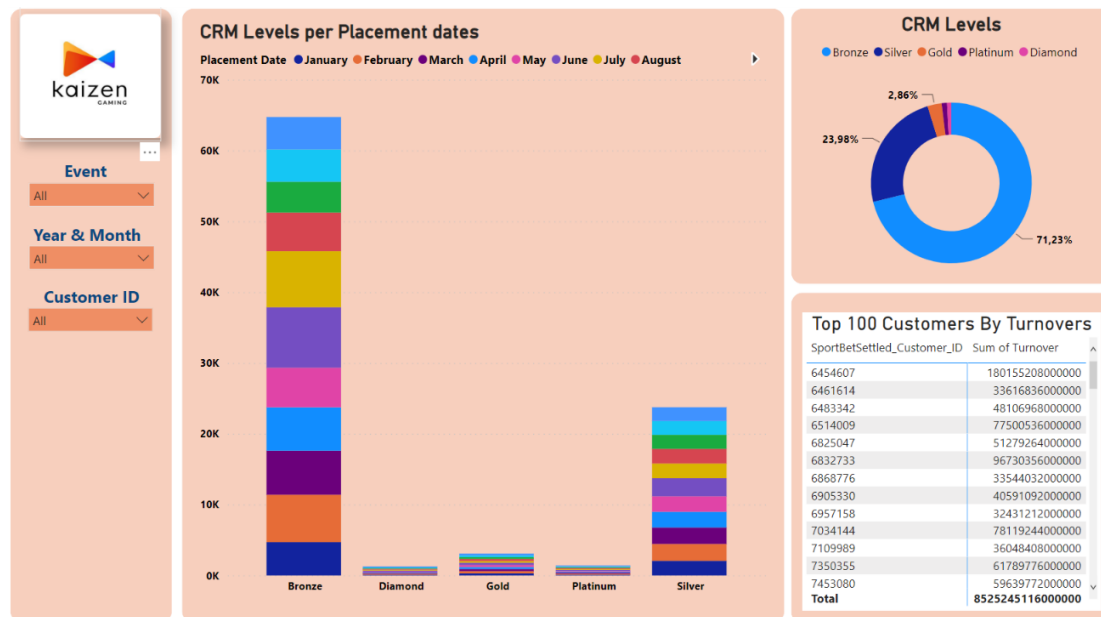
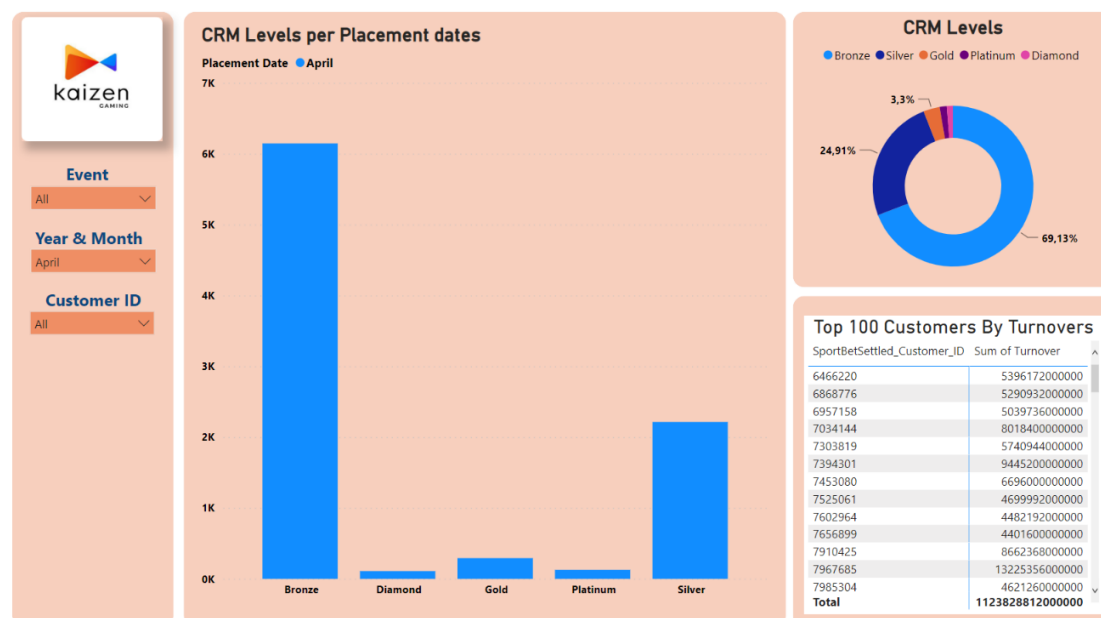


Table 3

Similar to what was done in table 2, we will now apply the same process to table 3, but this time only for customers who registered in **April**. To achieve this, we will utilize the "Year & Month" filters, which are displayed in the above graph. The application of these filters will yield the following outcomes for April:



Therefore, for April we can complete Table 3:

CRM Levels over Month	Bronze	Silver	Gold	Platinum	Diamond
2021 April	6146	2215		293	127