greenpanda GAMES

Test for data analyst candidates

Based on the tableau graphs and the sql assignments, which main takeaway would you give our business based on the data?

First look at the data

I worked with three datasets containing information of a mobile game for one week: *"users"* which concerns the installations, *"transactions"* the in-app purchases and *"sessions"* which corresponds to the connections of the players.

The **users table** includes the fields **cohort_date** which is the date of installation of the game, **country_code** which is the identifier of the country where the game has been installed, **network_name** which corresponds to the acquisition channel of the user and a field **user_id** allowing to identify each player.

cohort_date string Date (unparsed)	country_code string Country	network_name string Text	user_id string Integer
2017-09-01	CA	googleadwords_int	1
2017-09-01	CA	applovin_int	2
2017-09-01	MX	googleadwords_int	3
2017-09-01	ES	Organic	4
2017-09-01	AU	adcolony_int	5
2017-09-01	ES	googleadwords_int	6
2017-09-01	AU	Facebook Ads	7
2017-09-01	ES	googleadwords_int	8
2017-09-01	MX	googleadwords_int	9
2017-09-01	AU	Facebook Ads	10

37,719 players from 28 different countries downloaded the game from September 1 to 7 of 2017. The **session table** tells us that among these new players and during this same week, 20,169 users played for a total of 101,652 game sessions.

The **transactions table** informs us that a **revenue of 865.38** was generated from the in-app purchases of **70 of these new players** during this week, with an average spend of 5.6 per order.

Vincent CHARLES

Computing some KPIs

All the SQL queries are in the attached file and the final table is stored in the csv file "GPG_KPIs".

From these three datasets, I performed several SQL queries by creating temporary tables in order to compute different KPIs and to gather them afterwards according to the different days of the week.

```
54  --Retention Rate at D+1
55  CREATE TEMP TABLE rr_d1temp AS
56  SELECT
57  users.cohort_date,
58  COUNT(DISTINCT sessions.user_id) * 1.0 / COUNT(DISTINCT users.user_id) AS RR_D1
59  FROM users LEFT JOIN
60       sessions
61       ON sessions.user_id = users.user_id AND
62          sessions."date"::date = users.cohort_date::date + interval '1 day'
63  GROUP BY users.cohort_date;
64
```

A) KPIs by date

KPI name	Definition	Example
DAU (daily active users)	Number of distinct users who have played the game during a given day	100 for 2017-09-01 means that 100 unique users have played during that day
DNU (daily new users)	Number of people who have installed the game at a given date	100 for 2017-09-01 means that 100 people installed the game that day. In other words, 100 unique users started to play that day
Revenue	Revenue generated by players during a given day with in-app purchases (IAP)	100$ for 2017-09-01 means that IAP generated 100$ of revenue that day
DARPU (Daily Average Revenue per User)	Average revenue generated during a given day by players who had been active that day	0.1$ for 2017-09-01 means that each user who have played during that day has generated on average a revenue of 0.1$

B) Cohort KPIs

KPI name	Definition	Example
Retention rate at D+1 rr_d1	Proportion of users who have played the day after installing the game	40% for 2017-09-01 means that among users who have installed the game on 2017-09-01, 40% have played on 2017-09-02
Cohort Conversion Rate cr	Proportion of users who have paid at least once	1% for 2017-09-01 means that 1% of users who have installed the game on 2017-09-01 have made at least one IAP since then.
Cohort Conversion Rate at D+2 cr_d2	Proportion of users who have paid at least once in the three first days since install	1% for 2017-09-01 means that 1% of users who have installed the game on 2017-09-01 have made at least one IAP between 2017-09-01 and 2017-09-03

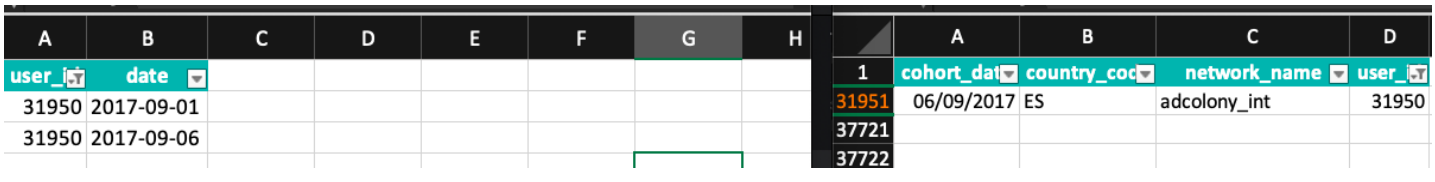
C) Results

Warning

It is important to note that the number of installs is DIFFERENT from the number of daily new users (DNU). Indeed, if a player installs the game but doesn't open it and doesn't run a game session it will be counted as an install but not as a DNU. This is the reason why the number of installs must always be greater than the number of new players.

However, the definition of the DNU is a bit ambiguous as it is defined as both “Number of people who **have installed the game** at a given date” and “unique users **started to play** that day”.

I then decided to choose the number of installs as the number equivalent to the DNU but I cannot guarantee the quality of the DNU indicator as I have observed what I believe to be **potential errors in the dataset sessions**. Actually, for the same user_id we can find game sessions BEFORE the game was even installed (see screenshot below).



A	B	C	D	E	F	G	H
user_i	date						
31950	2017-09-01						
31950	2017-09-06						

	A	B	C	D
1	cohort_dat	country_coc	network_name	user_
31951	06/09/2017	ES	adcolony_int	31950
37721				
37722				

Sessions table on the left and users table on the right

Visualizations and insights

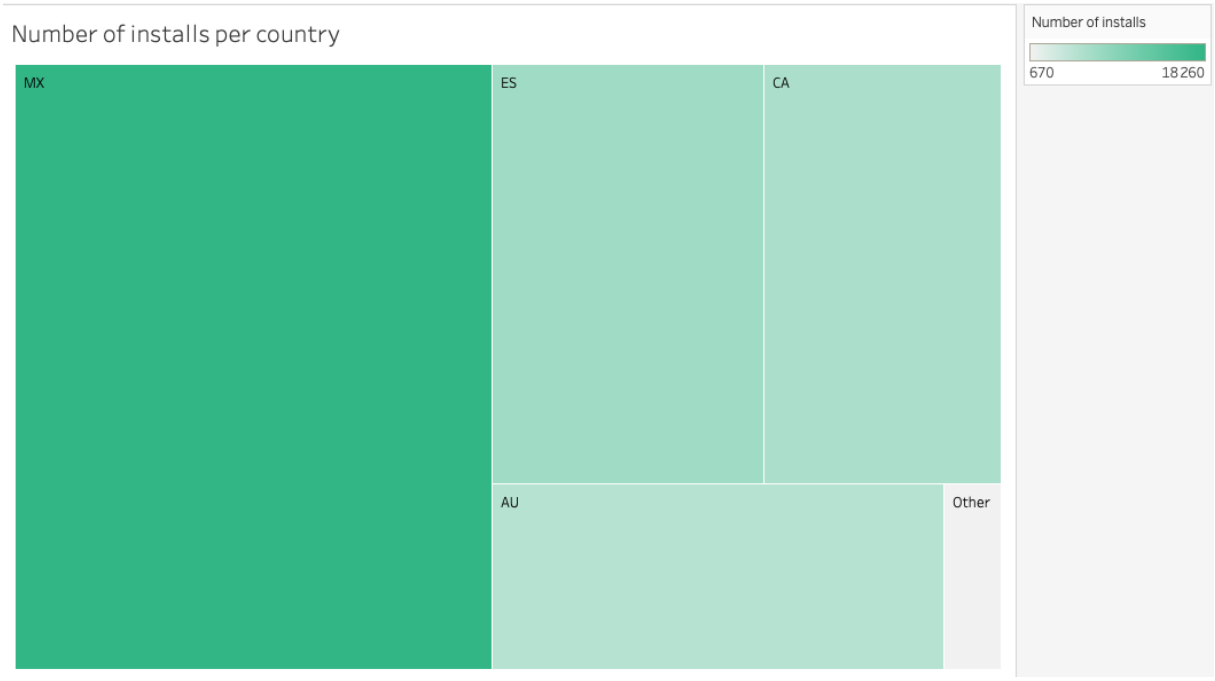
date	revenue	darpu	dau	dnu*	rr_d1	cr	cr_d2
09/01/2017	56,20	0,010	5736	5502	0,23	0,0020	0,0005
09/02/2017	150,06	0,020	7524	6030	0,23	0,0018	0,0010
09/03/2017	184,20	0,022	8210	5855	0,22	0,0015	0,0005
09/04/2017	116,11	0,012	9494	6751	0,22	0,0015	0,0003
09/05/2017	101,24	0,012	8347	5094	0,23	0,0014	0,0002
09/06/2017	111,13	0,013	8349	4854	0,24	0,0006	0,0000
09/07/2017	146,43	0,020	7226	3633	0,00	0,0008	0,0000

Most KPIs tell us about the health of the game through DAU, average income per user, retention rate, etc. If these numbers were to collapse after an update or during an A/B test, we could then detect it right away and try to correct what the players don't like.

Thanks to the **Cohort Conversion Rate** (cr & cr_d2), we can see that the majority of new players who make purchases do so on the day they install the game.

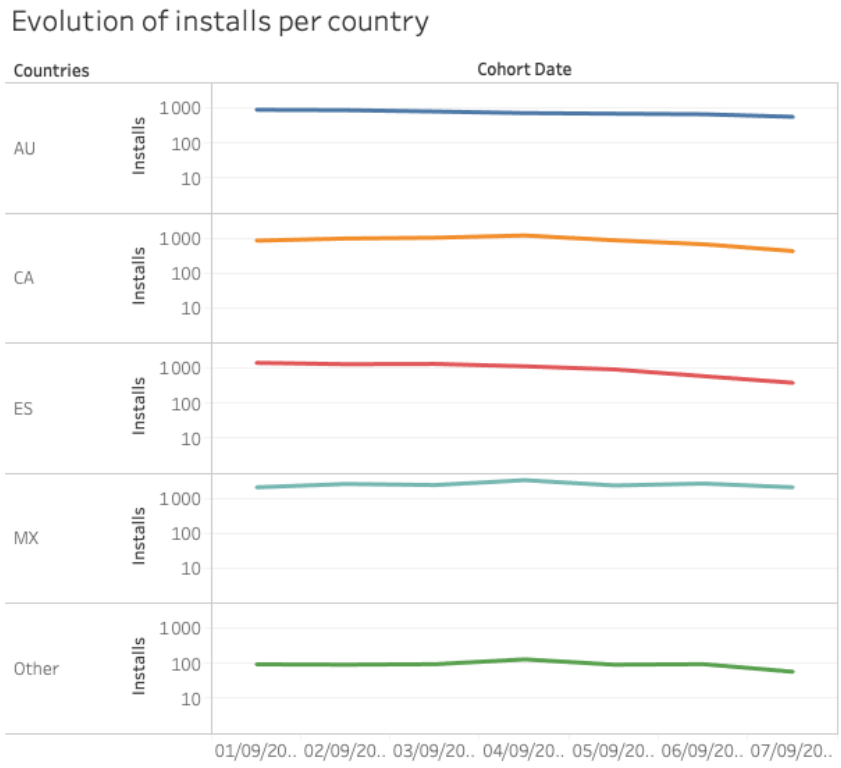
To further stimulate this behavior, it may be interesting to work on a system of promotion and credits offered for any in-app purchase on the first day. Thus, with **A/B testing systems it may be possible to increase the Cohort Conversion Rate over a week** and by extension the volume of buyers and revenue.

Offering welcome credits to all players could **increase the Retention rate at D+1** (rr_d1) as long as it remains balanced and does not have a negative impact on the Cohort Conversion Rate.

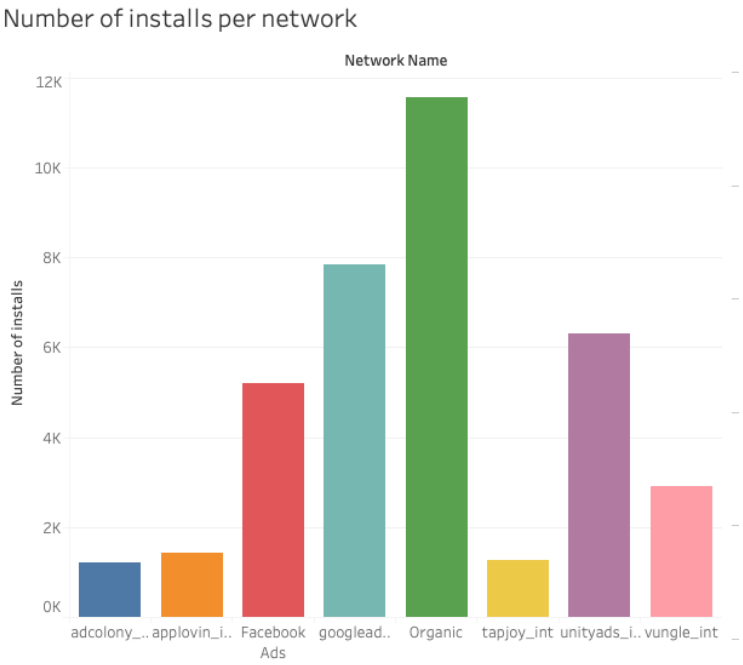


As can be seen on this chart representing the number of installations per country, the largest markets are Mexico, Spain, Canada and Australia. It may be interesting to continue to target

these rather favorable markets without neglecting the "Others". Indeed, the United States with only 178 installations in the week offers a market with a much higher potential of new players.

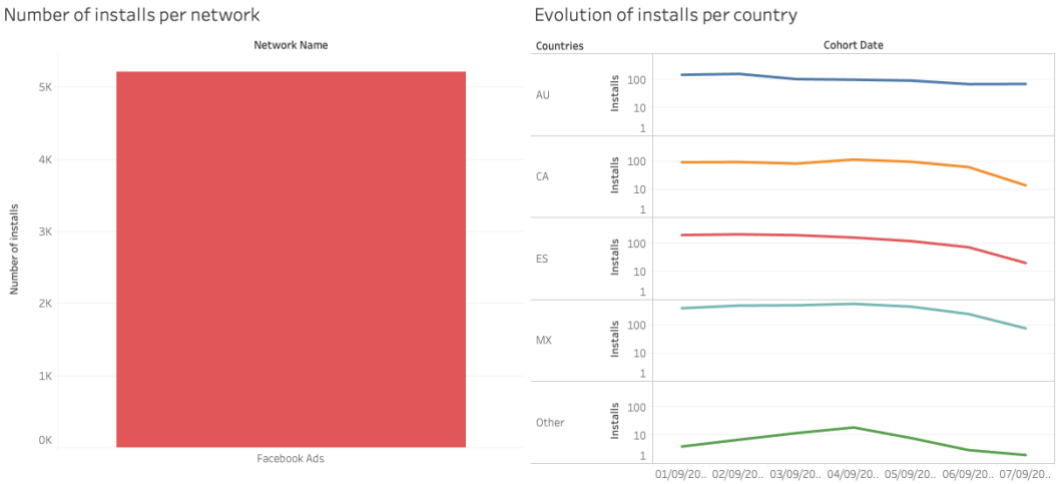


The chronological evolution of the installations informs us of a certain stability among the main markets, a result linked to the good performance of organic growth and of the main acquisition channels: Google Ads, Unity Ads and Facebook Ads.

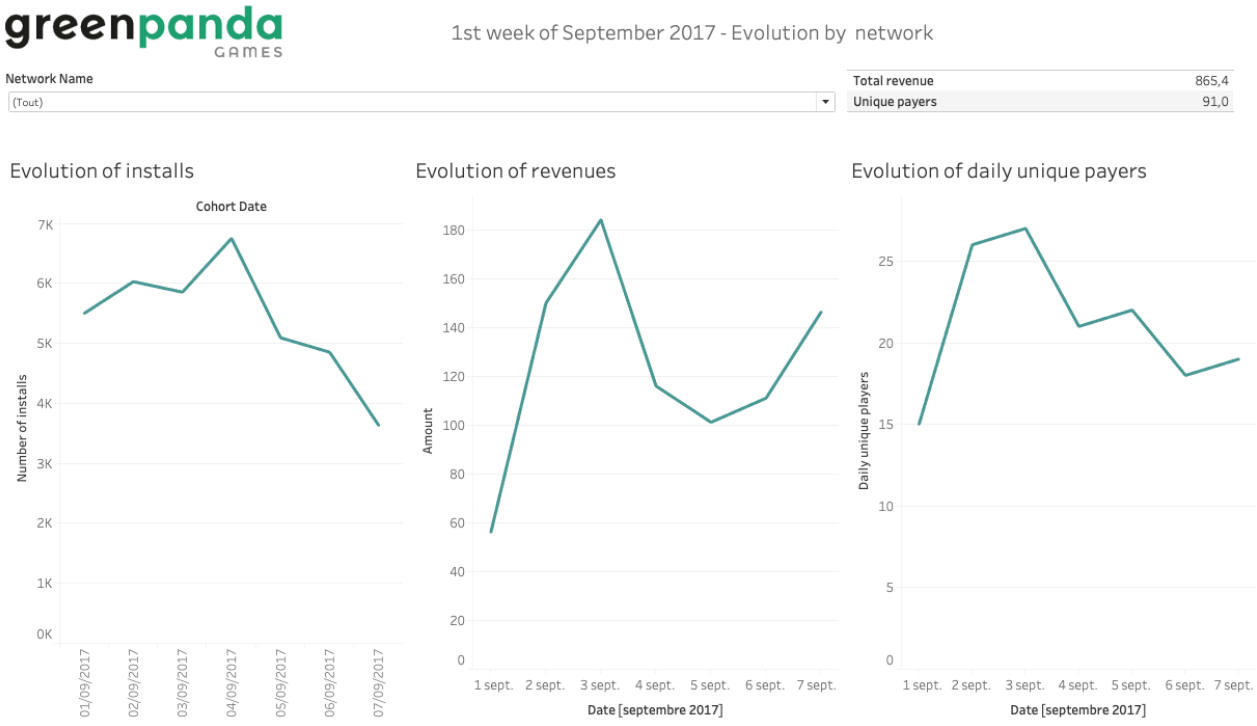


Vincent CHARLES

For example if we zoom by attribution network, we notice a global performance drop for Facebook Ads at the end of the week.



With our interactive dashboards it is possible to explore and compare the performance of the different networks and thus to know how much revenue has been generated as well as the number of unique installs and payers for each acquisition channel.



To go further in analyzing the performance of these different networks, it would have been interesting to have more data related to online advertising costs. Indeed, some KPIs such as the CPC or cost per click, can be used as a basis for the analysis. Perhaps even more precisely, the CPO or cost per order is equivalent to calculating the cost of the advertising campaign for the client / post-click order value (in-app purchases).

Vincent CHARLES

These metrics ensure that the revenues generated with users remain higher than the cost of acquiring them. Thus, it is possible to compare the different networks in order to know which perform best in terms of profitability but also according to the different target markets (countries, user personnas, etc.).