

From analysis to decision making

-data analysis and visualization-

Helena Giménez

Thanks for coming!

Some merchandising

Enjoy our notebooks & pens! :)

Coffee Machine and Water

Do you have 5 cents?

But first... let's relax!

Write in your post-it what was your favourite film when you were a child!

What we are going to need for today's masterclass



REDAH

I need your email to give you access to this tool. It will be used during the session and for the home tasks.

Slack: #decision_making
<https://redash.bdatainstitute.com>



KAHOOT!

Install this APP on your phone, or connect to their website. More instructions will be given.

Working Agreements

- Celebrate our failures and our wins
- No judgment (there are no silly questions)
- Respect for each other
- Interrupt when needed (and use Slack whenever needed)
- Keep a positive & constructive attitude
- Ask before our info on social networks (i.e., check for privacy needs)
- Have fun
- Self-organizing silence, by raising our hands
- Fred joins to help us improve our students' experience

ABA English

who we are



Vision

Empower people through the most effective digital English learning.

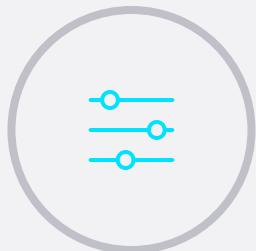
Mission

To teach English combining a unique learning methodology and innovative technology.

company values



Student
Centric



Play to win
as a team



Make it lean,
be fast



Be bold

+25M

users of our applications
we are in trend

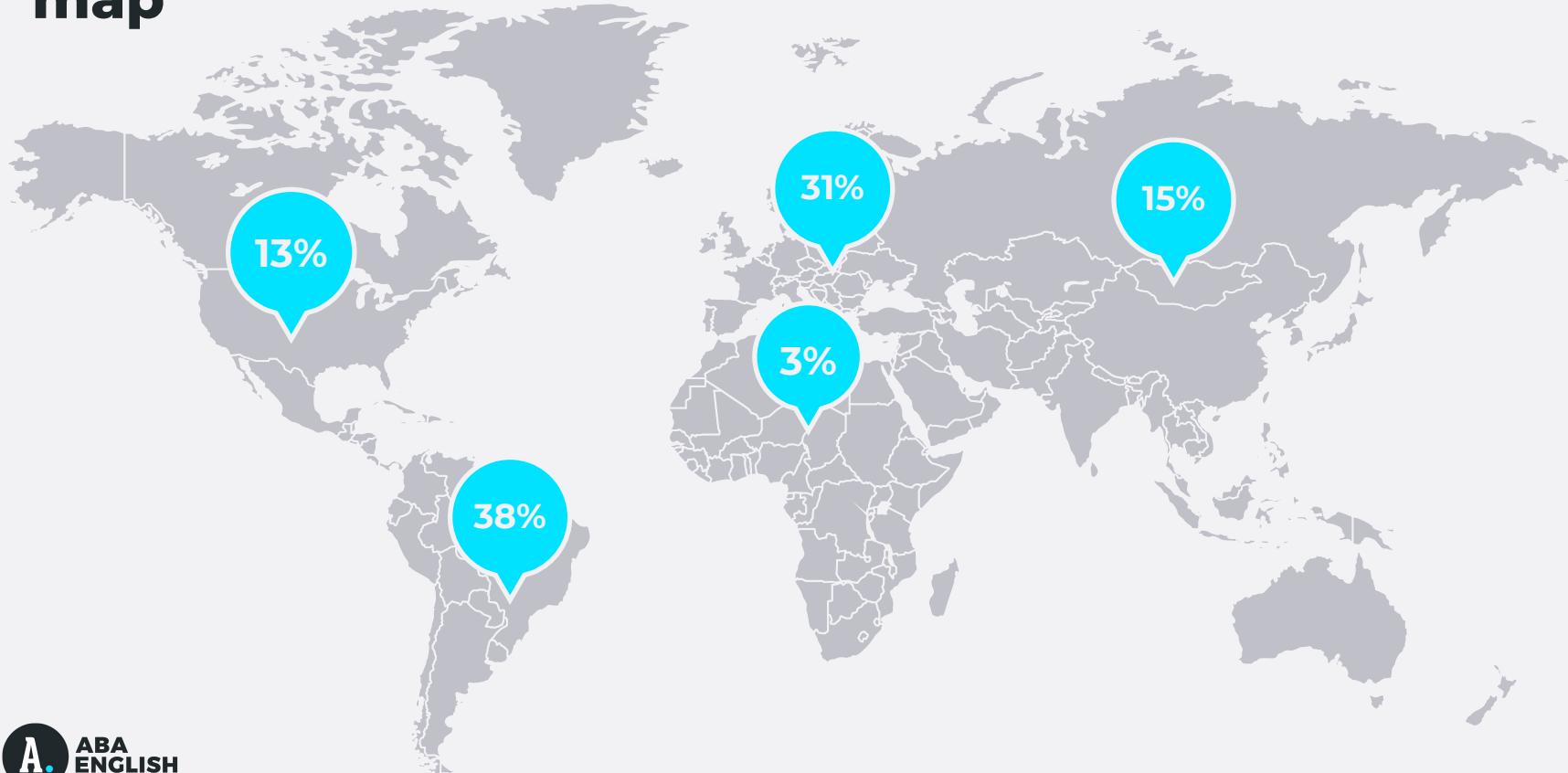
700k

monthly active users
we are useful

12

languages are supported
we are global

world map





Family

You will enjoy working with those that you love.

Talent

We must search those talents that will take the company to the next level.

Attitude

Bring a smile everyday at the office. We spend 10 hours/day there!



Family

You will enjoy working with those that you love.

Talent

We must search those talents that will take the company to the next level.

Attitude

Bring a smile everyday at the office. We spend 10 hours/day there!



Family

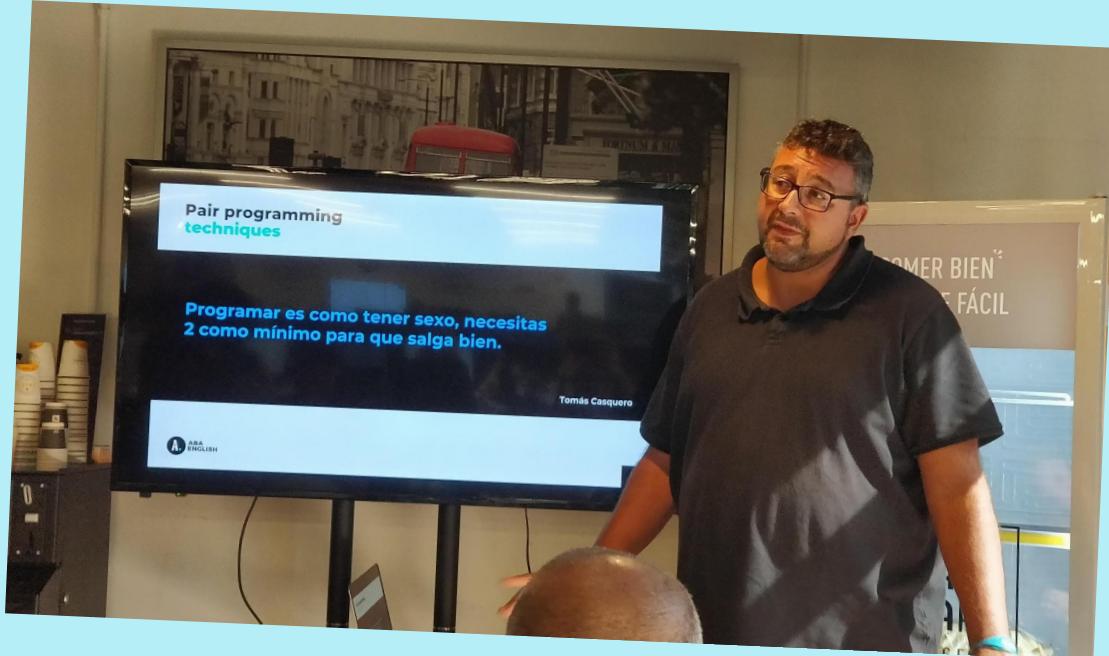
You will enjoy working with those that you love.

Talent

We must search those talents that will take the company to the next level.

Attitude

Bring a smile everyday at the office. We spend 10 hours/day there!



Family

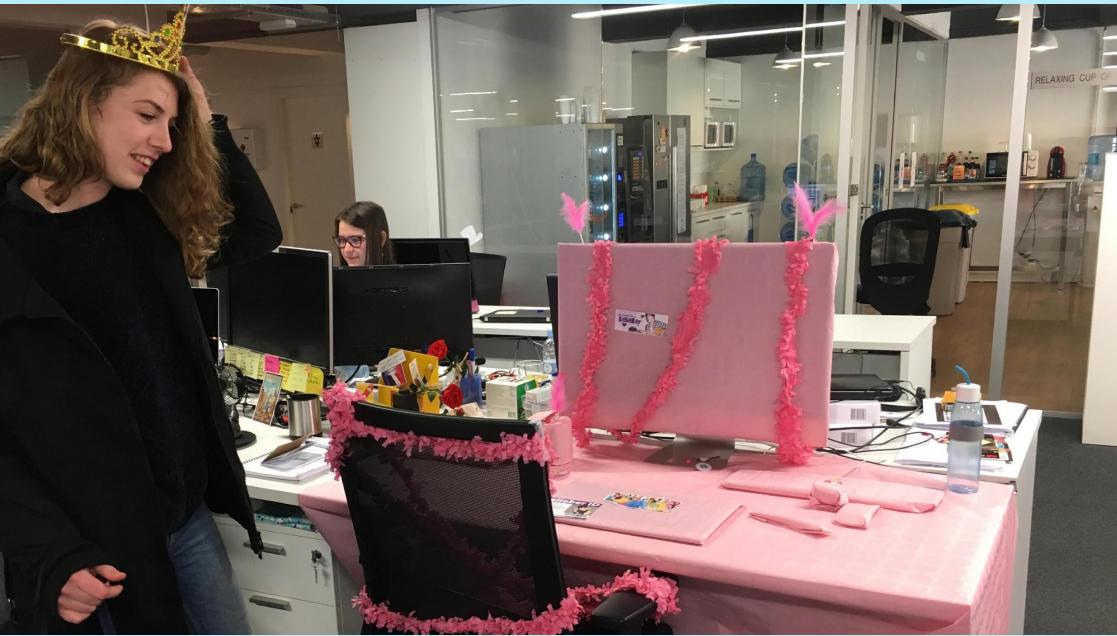
You will enjoy working with those that you love.

Talent

We must search those talents that will take the company to the next level.

Attitude

Bring a smile everyday at the office. We spend 10 hours/day there!



Family

You will enjoy working with those that you love.

Talent

We must search those talents that will take the company to the next level.

Attitude

Bring a smile everyday at the office. We spend 10 hours/day there!

we love our product

A2. Lower intermediate >

25. A GOOD COMBINATION
The verb to have in the affirmative.

26. THE WATCH
To have in negative and interrogative.

CH negative
interrogative.

DESS

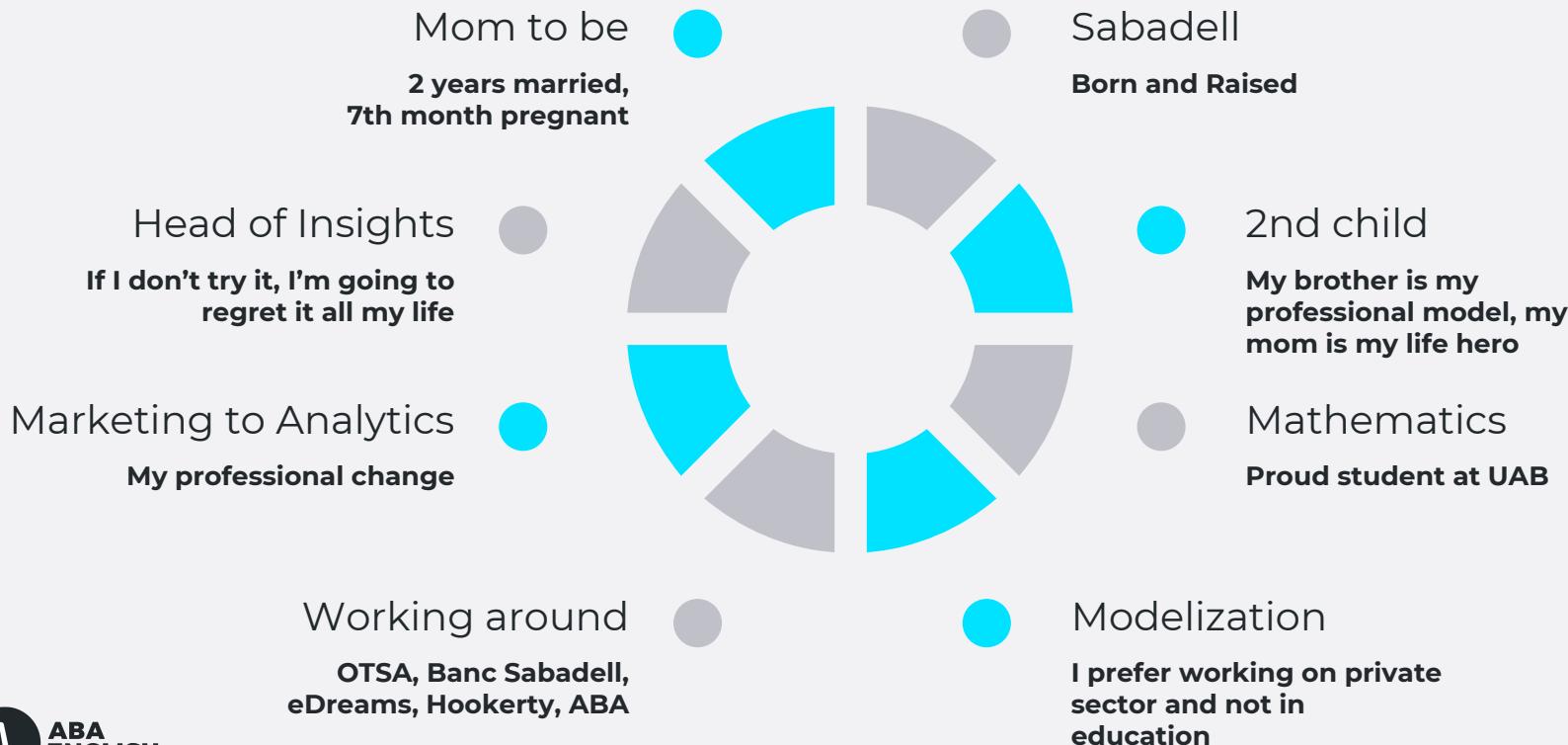
A hat

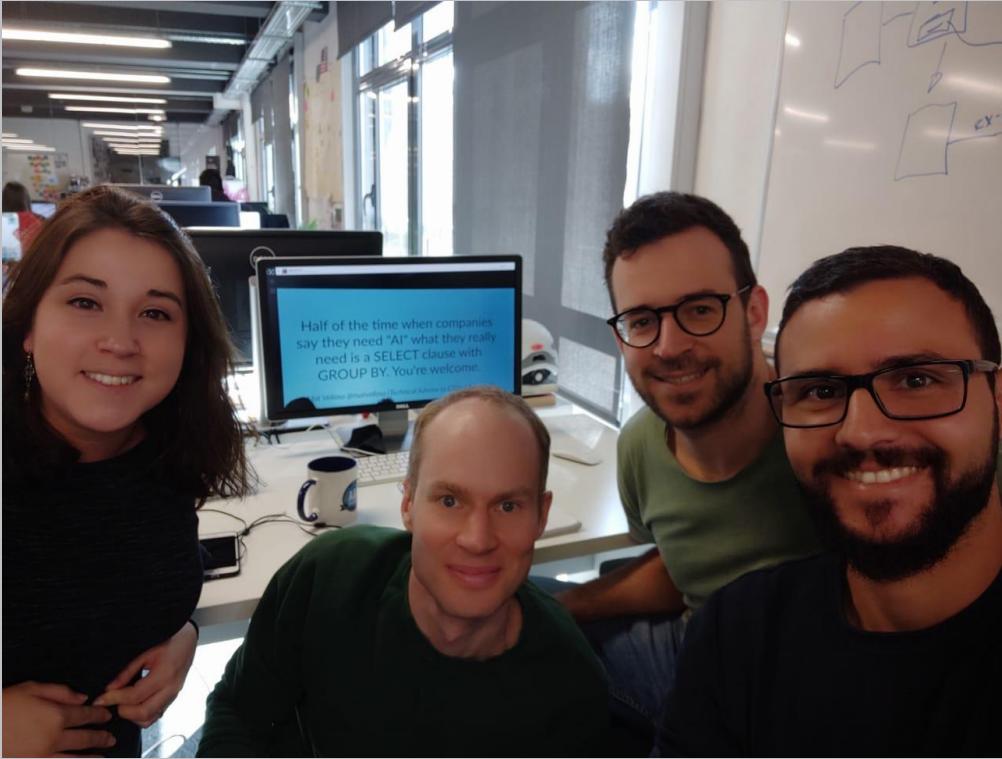
**we love
our product**



Helena Giménez

Who I am





Insights

Data Science
Senior Data Analyst
Data Engineer
Head of Insights

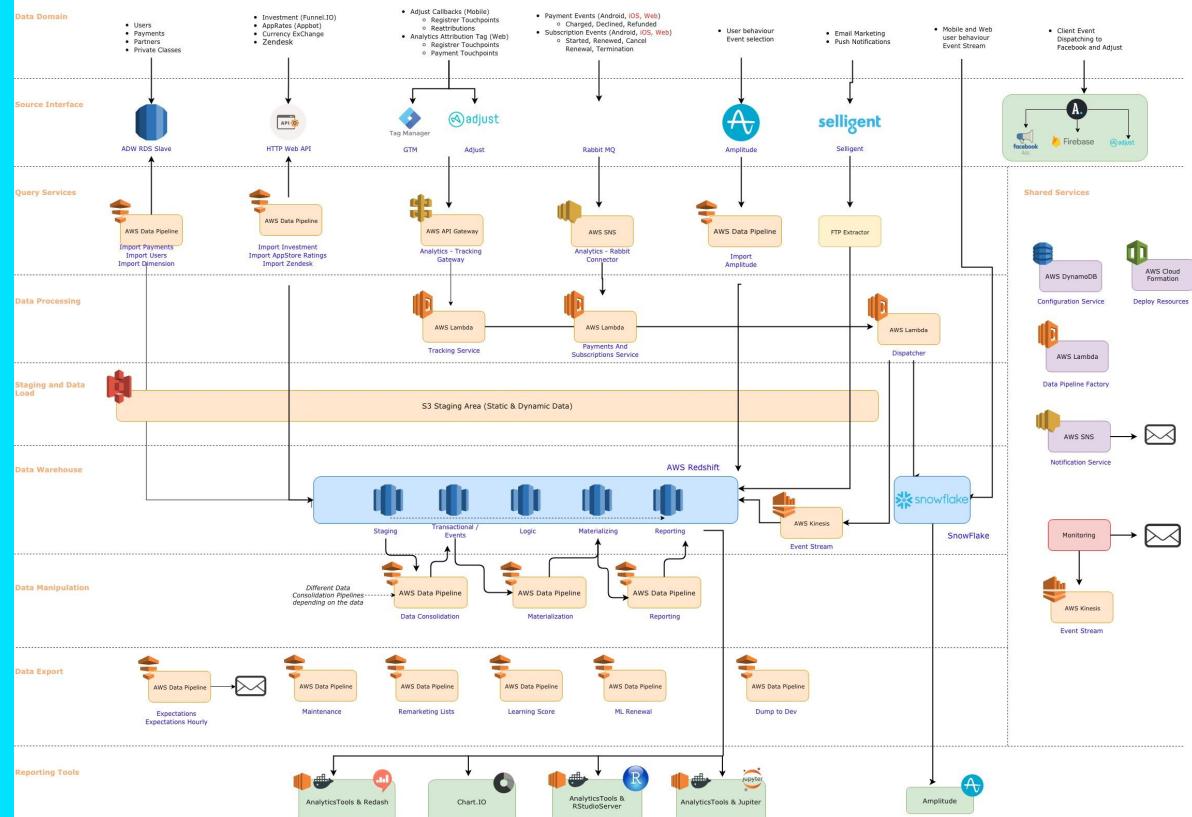
But we want to grow!

Data Engineer
Data Analyst

Our architecture

Constant evolution is our mantra. We must adapt ourselves if we want to bring the best solution.

Create a **maintainable architecture** if you do not want to go crazy.



What we do as an Insights Team

01 Owners of the **data**

We must ensure, that all the data that we show is correct and make sense (not always easy)

02 Data **Driven** Company

We must evangelise the **DATA** as our only way to make decisions

03 **A/B** testing

Do you have an hypothesis? Nice, let's test it! Because I don't trust your spider-sense...

04 Stable **data** ecosystem

We must create a stable DB, that alerts us when something is wrong, and don't create crashes

05 Analyse **EVERYTHING**

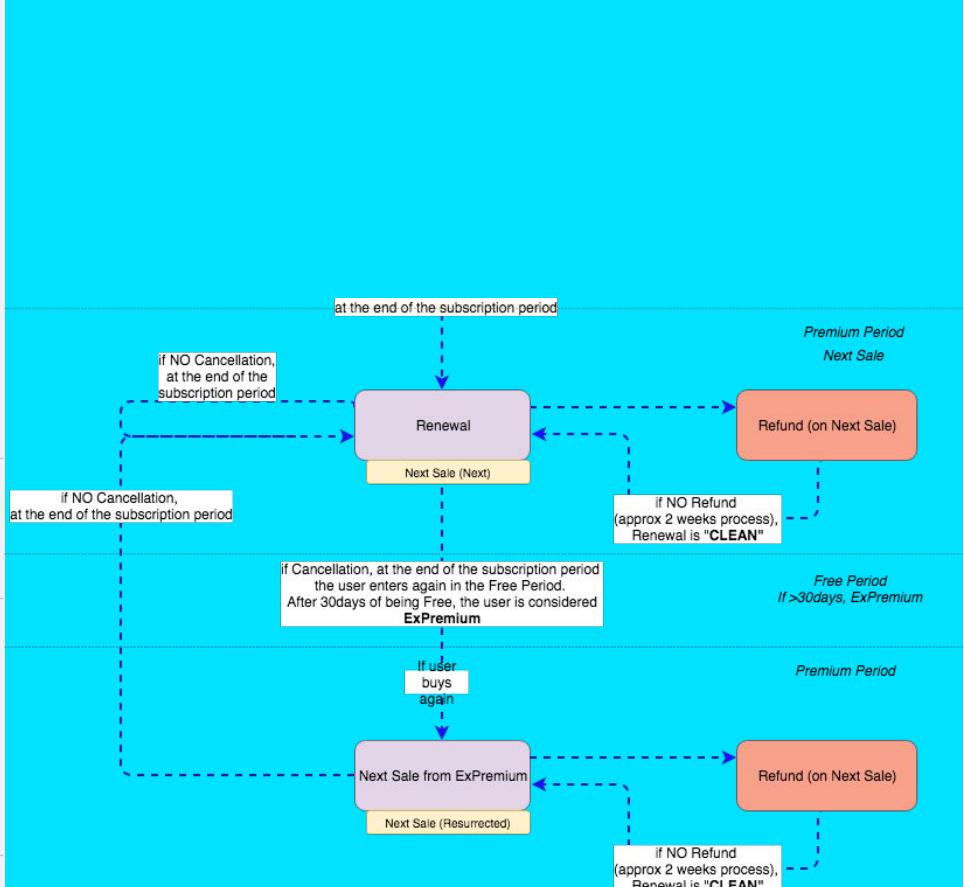
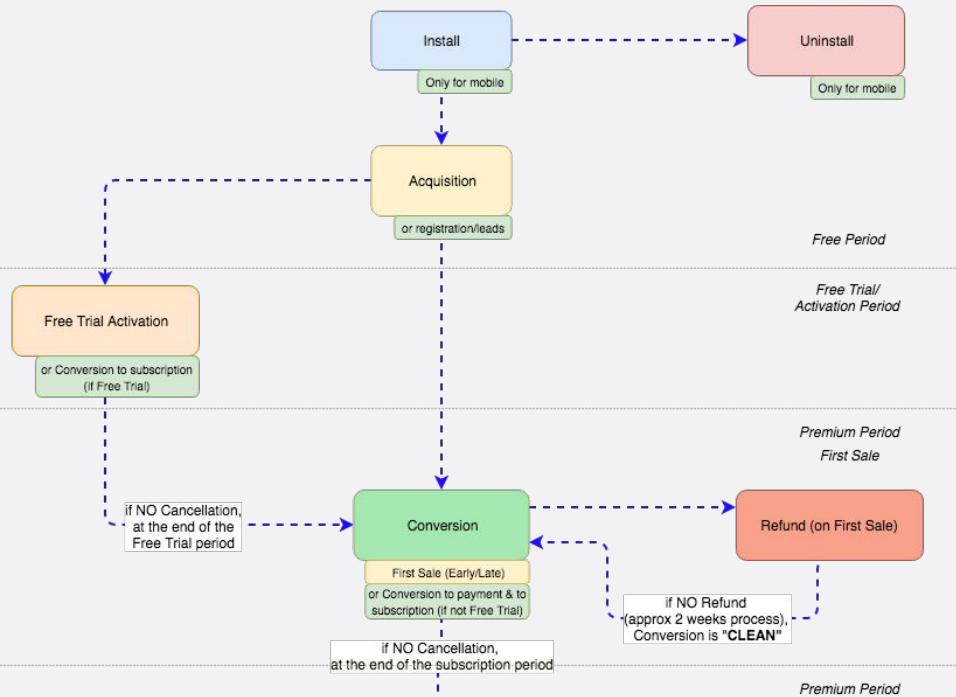
Increase of CPL, company NPS, contact tickets, booking sales, new releases...

06 One Single **point** of truth

All the data that the company use, came from our DB. If not, they will correlate isolated things

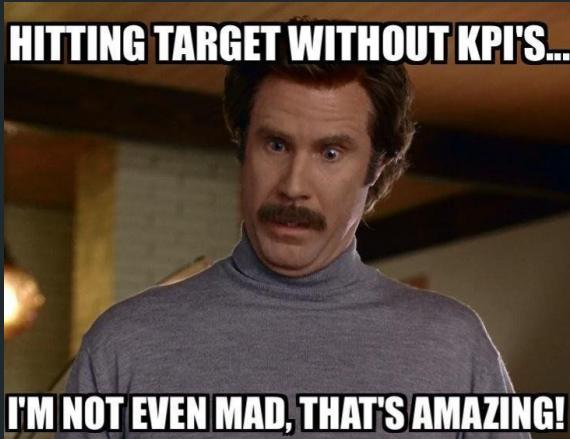
USER FLOW

Or how the user drives around our systems, and if we are lucky he will give us money!



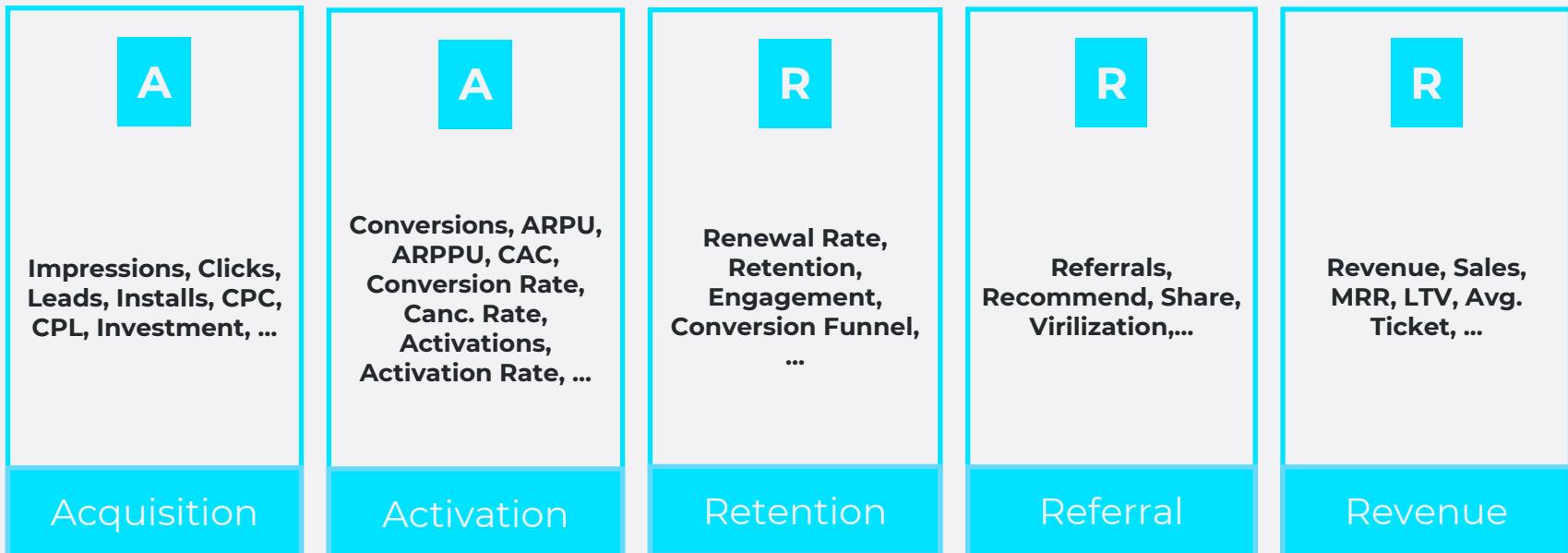
KPI

Definition



A **Key Performance Indicator** is a measurable value that demonstrates how **effectively** a company is achieving **key** business objectives.

PIRATE METRICS



Acquisition KPI

CPC (Cost per Click)
Investment/Clicks = €

CTR (Click Through Rate)
Clicks/Impressions = %

Leads
Number of acquired users

Installs
Number of APP Downloads

CPL (Cost per Lead)
Investment/Leads = €

Acquisition Rate
Leads/Install = %

Paid Share
Proportion of paid leads vs
organic leads
Paid Leads/Organic Leads = %

Uninstall
Number of APP uninstalls

Uninstall Rate
Uninstall/Install = %

Activation KPI

Activations xD (Cost per Click)
Number of subscription act.
during first X days

Conversions xD
Number of conversions (paym)
during first X days

Activation Rate xD
Activations xD/Leads = %

Conversion Rate xD
Conversions xD/Leads = %

Cancellations xD
Those activations that at end,
never converts to payment

Cancellation Rate
Cancellations/Activations = %

ARPU xD
Average Revenue Per User
Revenue xD/Leads

ARPPU xD
Average Revenue per Paying
User
Revenue xD/Conversions xD

CAC xD
Investment/Conversions xD

CPAct xD
Investment/Activation xD

Retention KPI

Renewal Rate

Sales/Potential Sales = %

Retention over time

Activity retention during days in a row

LT

Amount of time that a client stays as a “premium” user

Churn Rate

Relation between those users that were “premium” one certain period, and the next one not.

Revenue KPI

Revenue (Cohorted or Actual)
Amount of € on a certain period

Sales
Numbers of payments

Avg. Ticket
Revenue/Sales = €

ROI xD
(Revenue xD/Investment -1) = %

MRR
Monthly Recurring Revenue

LTV
Amount of revenue, during the lifetime of a client

Refunds
Number of Sales that are returned to the client

Uninstall Rate
Uninstall/Install = %

Analysis workflow

01

First, we need to identify all those affected KPIs

Decide the critic alarm you want to rise

02

03

Try to find the why's not the consequences. Visual correlations sometimes can disorient

On alarm mode

Sometimes we need to analyse data, because is our Business As Usual. A/B testing, new features, new marketing campaigns...

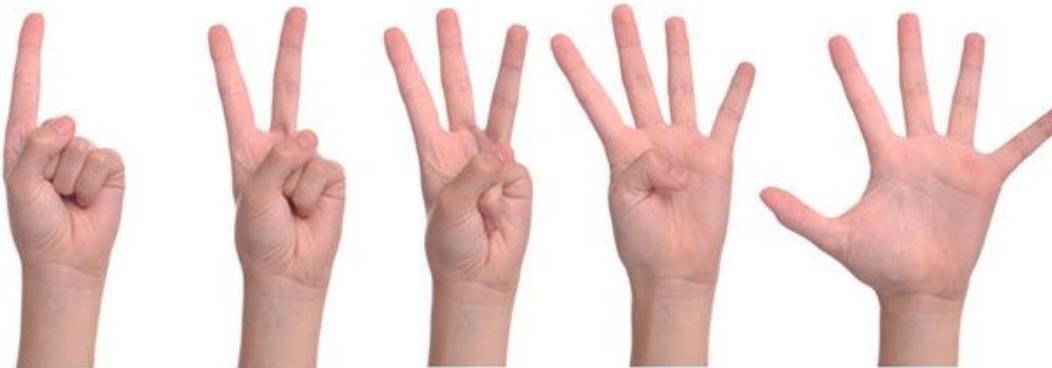
But in other occasions, we will receive an alarm! (Welcome to the company reality)

That's why is so important to have a systematic workflow. Be sure everyone is doing their job. Identify if it's a P1 or just a bug that must be fixed.

Try to stay calm (not always easy).

Do always a retrospective of the issue. Most of the time, a better communication could solve the problem.

Everything Under Control?



I don't
understand
at all

I need to
go over
this again

I think I got
it, but am
not
completely
comfortable

I got it

I can
explain it
to someone
else



coffee break

Let's do some [networking](#)

**Let's
play!**

Kahoot!

**Can you remember
everything about KPIs?**

<https://create.kahoot.it/kahoots/my-kahoots>

Find the WHY

01 Detect the KPI

02 What other metrics
affect this one?

03 Find the origin

04 Do we know if
something has
changed?

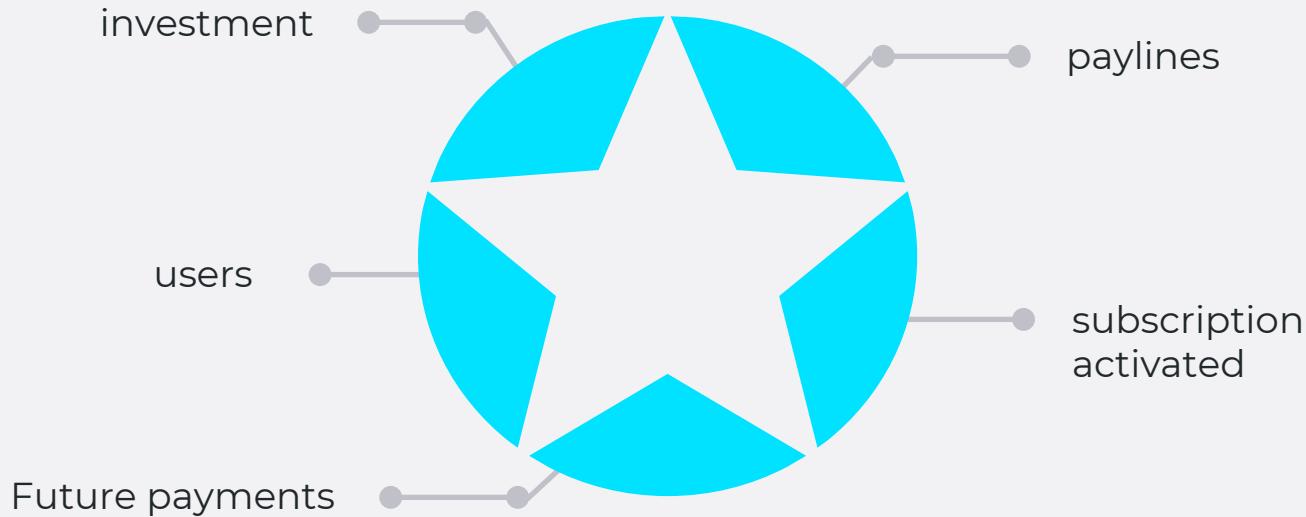
Our main tool



How we are going to
work from now on?

<http://redash.bdatainstitute.com>

our data base

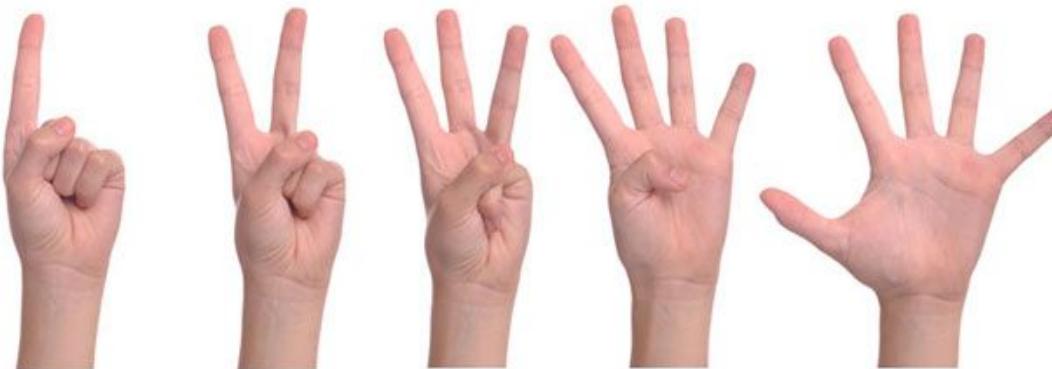


**Play with our DB and give
us some information**

Create a **graph** that represents the **evolution**
on time of the **CPL**

01

Everything Under Control?



I don't
understand
at all

I need to
go over
this again

I think I got
it, but am
not
completely
comfortable

I got it

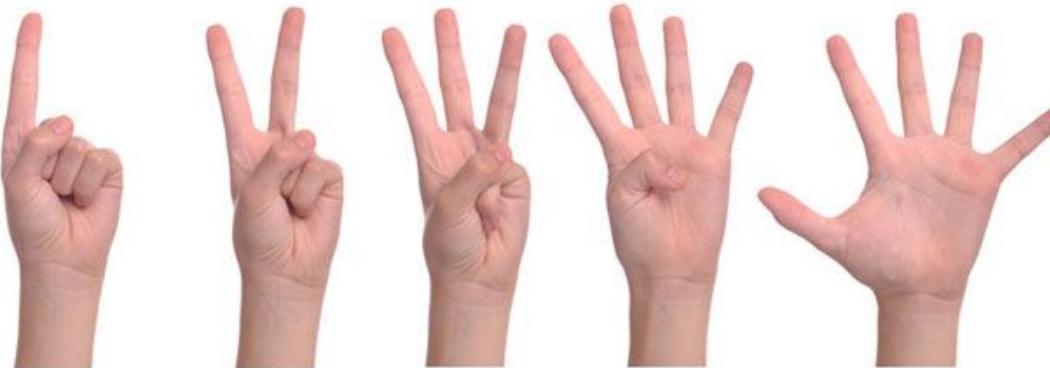
I can
explain it
to someone
else

Play with our DB and give us some information

Compute the **Revenue 8D** for Android
during **August**.

02

Everything Under Control?



I don't
understand
at all

I need to
go over
this again

I think I got
it, but am
not
completely
comfortable

I got it

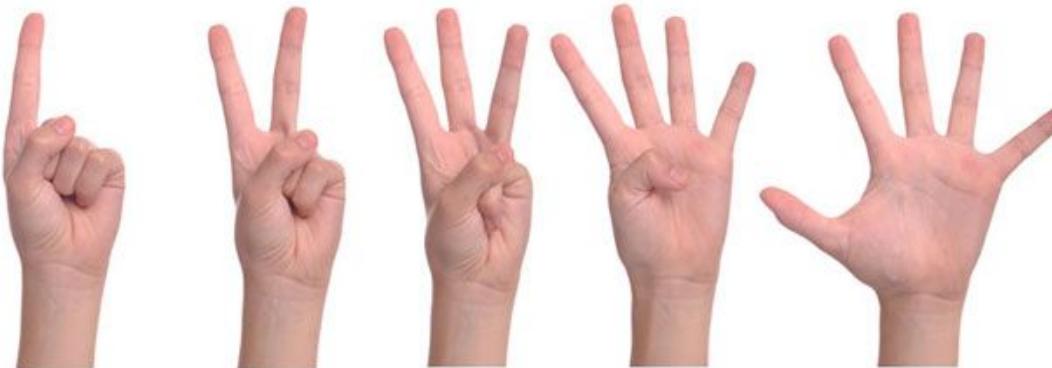
I can
explain it
to someone
else

03

Which **campaign** has the better **ROI**
8D during **July**?

**Play with our DB and give
us some information**

Everything Under Control?



I don't
understand
at all

I need to
go over
this again

I think I got
it, but am
not
completely
comfortable

I got it

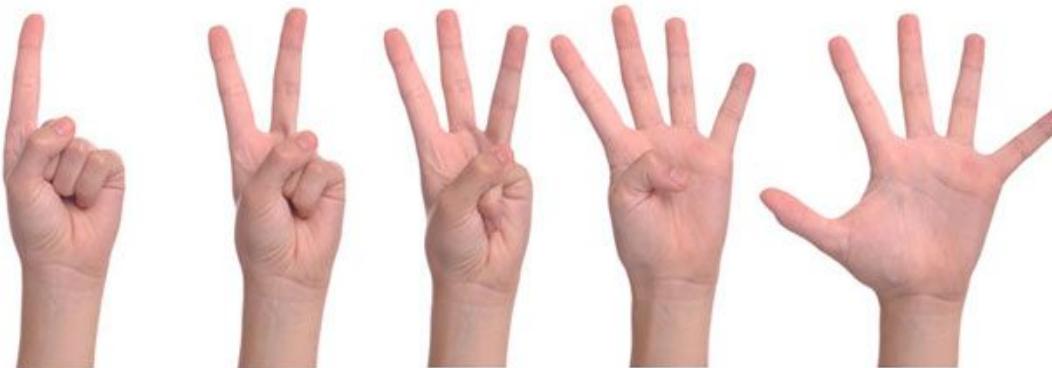
I can
explain it
to someone
else

04

Can you give us the **evolution** for
Conversion Rate for iOS during last year?

**statement
title**

Everything Under Control?



I don't
understand
at all

I need to
go over
this again

I think I got
it, but am
not
completely
comfortable

I got it

I can
explain it
to someone
else

A **business dashboard** is an information management tool that is used to track **KPIs**, metrics, and other key data points relevant to a business, department, or specific process.

Through the use of data visualizations, dashboards simplify complex data sets to provide users with at a glance awareness of current performance.

Our clients don't talk data

01 Listen what they want

02 Don't talk to them like they are data experts

03 Make sure you can do what you promise

04 Remember that your clients MUST understand the data

Our clients are different

Hola Helena,

En la reunión que hemos tenido hoy de weekly KPIs me he liado un poco, porque para mi el CAC es lo equivalente al CPA [aquí](#). En este ejemplo la semana 41 el CPA en iOS es de 85,79€ vs 94,54€ lo que dijiste tu. No entiendo muy bien porque hay esas diferencias.

Creo que tendríamos que poner todo en un sólo dashboard de marketing aunque sea largo y que todo lo relacionado con MKT esté ahí , sino, al menos para mi sigue siendo un poco lio.

Gracias!

Our clients are different

Hi Helena,

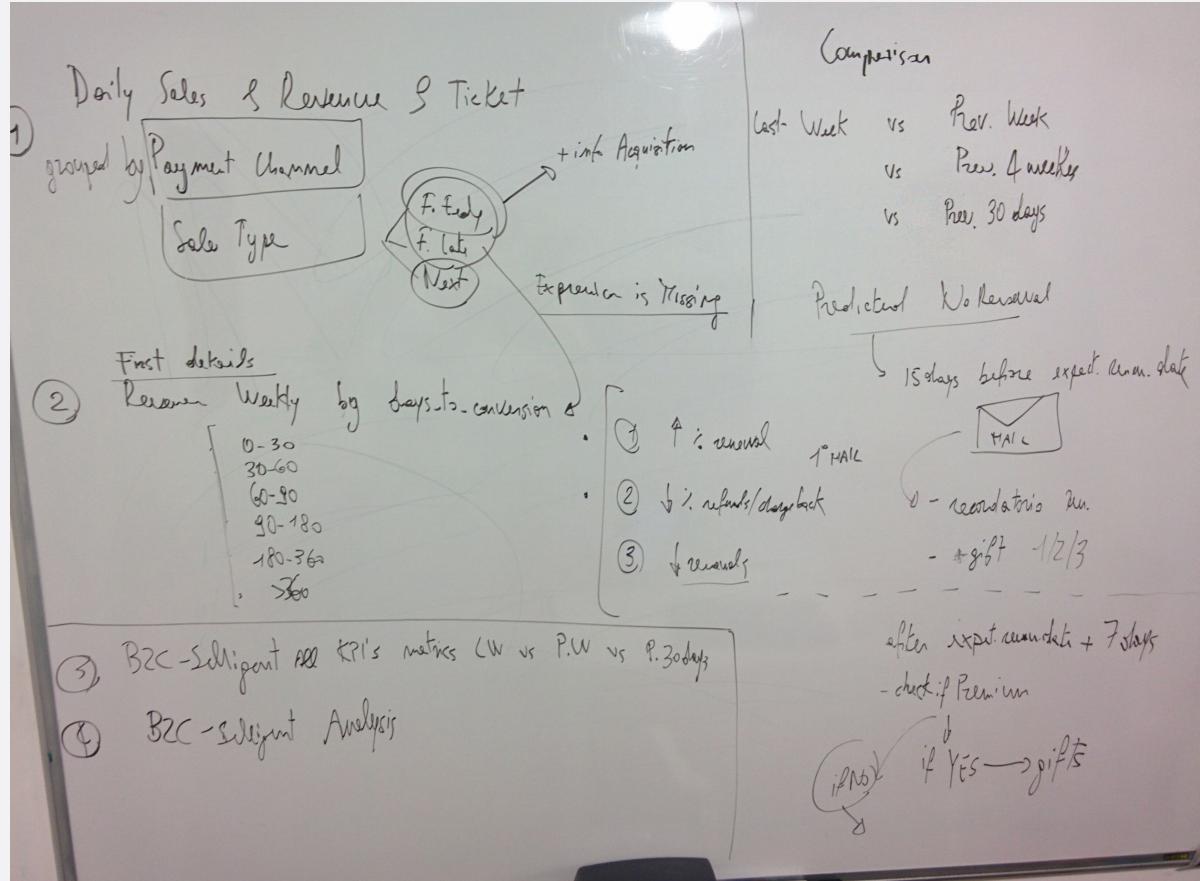
Some ideas to evolve DARIA into V2 turning it better:

- 1- All numbers in "extrapolated" => Kill it. No need. I do not use them at all. I don't think anyone looks at them at all.
- 2- Instead: As today we have the mini-charts at daily basis and a drop-down to choose the period. We remove the drop-down and create two charts. One weekly and another daily. and we add two lines. This vs last year :)
- 3- Yesterday numbers. The % vs previous period. I think it compares vs yesterday. Can we change it YoY? so vs same day last year.
- 4- last 7 days: Same than previous. We leave it but show the % comparison vs last 7 days last year, instead of 7 previous days.
- 5- Acquisition: We add ROI, CAC and Activation rate (8d) in the table. I know I have the charts below with the cohorts, but makes it harder to compare apple with apple vs YoY so this way we can easily see correlations.
- 6- We add a chart comparing late vs early first revenue this year vs last year.
- 7- At the end: We add LTV vs CAC. We also add active total subscriptions (even if you only have it at monthly basis).
- 8- At the end, we add a graphic showing daily revenue evolution by country this year vs same days last year.

Can you please look at it and tell me by when this can be done? Of course if you have other ideas how to improve it, feel free.

Thanks

Our clients are different



Our clients are different

Hello Helena,

We would like to create an ASO dashboard in order to every day and every week be able to understand what is happening in each market. We already have some dashboards we can follow some metrics for ASO, but there dashboards are missing several important elements.

I would therefore kindly like you for the support.

The dashboard should include:

- Organic installs
- % of organic vs- paid
- Install ratio (organic)
- Rating
- Category ranking
- Ranking of 10 top keywords in 3 core markets (needs to be defined)
- Crash rate
- Uninstall ratio

All KPIs should be DoD and WoW (most importantly Wow).

It should be possible to filter by market and operative system.

Do you think this is possible? If you want we can sit together and draw a draft in order to make sure that we are aligned.

I will define the top 10 keywords in the main markets shortly.

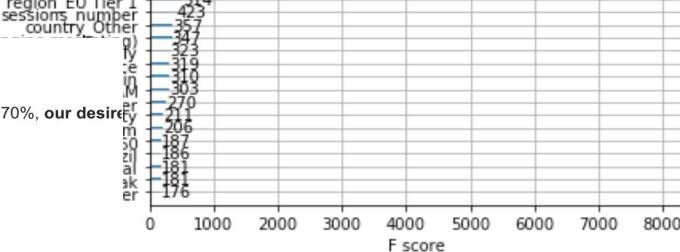
I also wondered if you know if it could be possible to add any information about the competitors in a chartIO dashboard?

6) RESULTS

The overall accuracy of the model is **67%** (that is, the probability of correctly predict if a user renews or not). Our expectations were about to 70%, our desirey

More specifically:

- the probability of correctly predict a No Renewal User is **74%**
- the probability of correctly predict a Renewal User is **60%**



	precision	recall	f1-score	support
0	0.67	0.74	0.70	2509
1	0.67	0.60	0.64	2276
avg / total	0.67	0.67	0.67	4785

D IMPROVE THE ACCURACY

7) FEATURE IMPORTANCE

The features/variables which resulted to be more important for the prediction are those mostly related to **Payment, Demographic and Marketing Information**, like the amount spent, the date difference between acquisition and payment, the number of sessions, the number of study hours, and the number of exercises.



aggregating categorical features, dropping high correlated features

· the first absolute payment of the user

d by using **Web, Android** and **iOS** data from 2016-06-01, trying to structure the behaviour inform

- proceed with the "Productization Phase". It will involve:

- a) more support from our team in terms of Data Engineer
- b) collaboration with Product team and CRM team in order to create a new mail flow/ push notification/ new alert in the app

s from the expected Renewal Date; the output is a value for each user [0 if No Renew; 1 if Renew] marketing action (e.g. a promo offer for users labelled with 0 value).

Android and iOS users.

- take advantage of the work done and face an actual problem for the company, which is the Cancellation Rate. Our idea is to work on a machine learning

predict the probability of a user to CANCEL a FREE TRIAL during his first hours/days of FT use. Our experience in the Renewal Model allows us to data: new ML project.

, channel, partner type, etc...)

:count, the days difference between acquisition and payment, the payment_gateway, the cancellation

e.g. the number of sessions, study sessions, study time, number of exercises, number of finished

We can discuss it in more details,
let us know for any questions
thank you

Not always the best analysis is the best solution.

**Pragmatism, speed, work optimization, empathy and
be open minded are really important aspects for a Data
Analyst.**

Dashboards: best practices

London 51.51 N, 0.13 W Mon 9 Dec @ 16:02:42

[Go to Map](#) - [Go to Grid](#) - [Change City](#)

WEATHER STATIONS (MULTIPLE SOURCES)									5	WEATHER (METAR)			848	FORECAST (YAHOO! WTH)			
STATION	WIND SPEED	WIND GUSTS	DIRECTION	TEMPERATURE	HUMIDITY	RAIN TODAY	PRESSURE	FORECAST		London City Airport	SW at 3 mph	11 C	Mon	Tue	1748		
CASA Office: Bloomsbury W1	8 mph	9 mph	SE ↗	11.5 °C	76%	0.0 mm	1027.9 mbar	Clear Night		Mostly clear			10 C	9 C			
Lambeth Meters: Brixton SW9	4.3 mph	4.3 mph	SW ↘	11.0 °C	83%	0.0 mm	1026.4 mbar	Clear Night									
Hampstead NW3	3.6 mph	3.6 mph	S ↑	9.8 °C	84%	0.0 mm	1029.0 mbar	Clear Night									

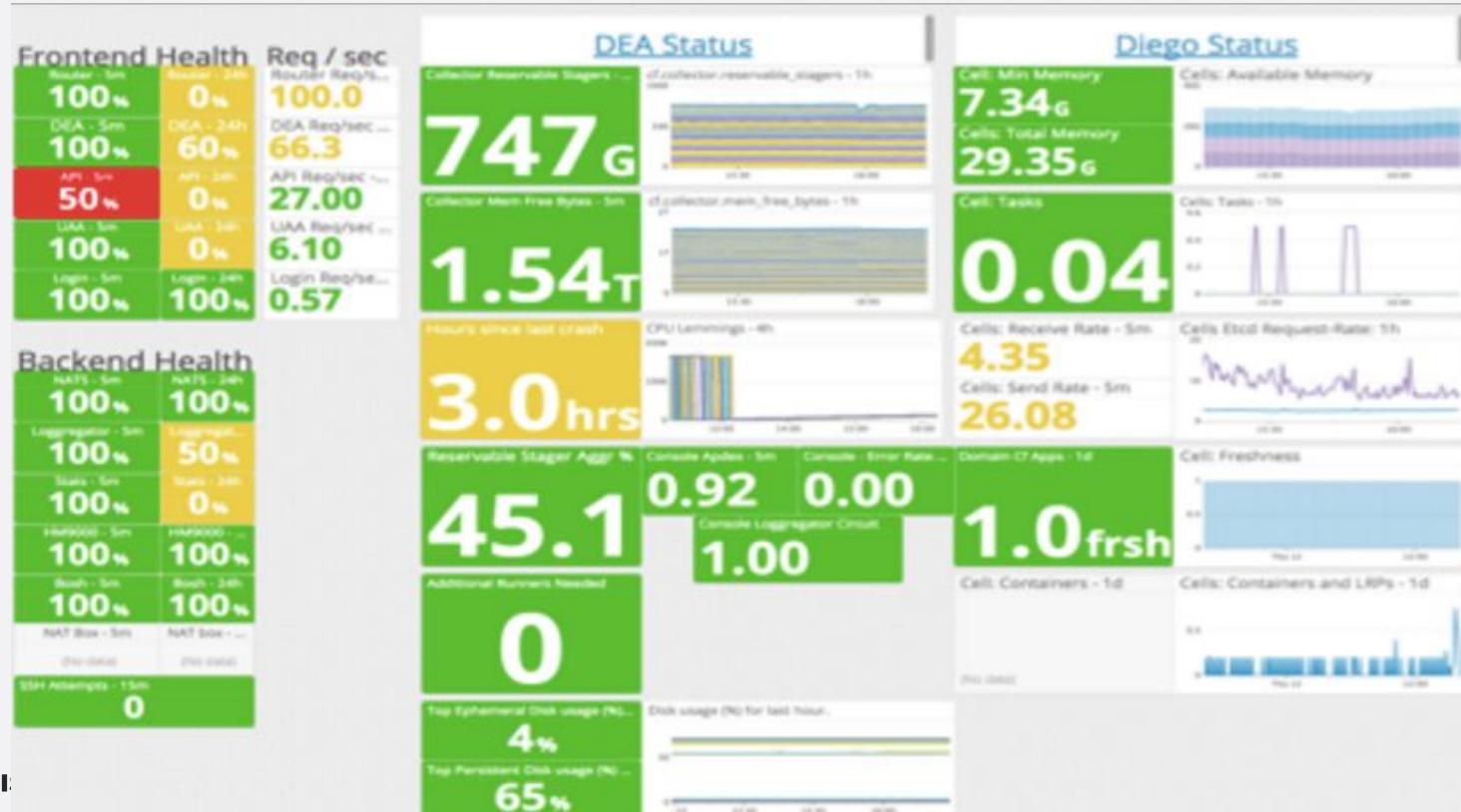
TUBE LINE STATUS (TFL)										39	BIKE SHARING (TFL)					38	IN SERVICE (TFL)					9	AIR POLLUTION (DEFRA)					1748
Bakerloo	Good Service	7354	4.3 %	Stations Full	4.9 %	Stations Empty	378	7197	London buses	7197	μg/m³ TIME AVERAGED	OZONE	NO ₂	SO ₂	PM _{2.5}	PM ₁₀												
Central	Good Service	430					Underground trains	378				Bloomsbury	13	38	4	9	10											
Circle	Good Service											Marylebone Rd	9	16	26	22	34											
District	Good Service											N Kensington	14	40	?	12	18											
H & C	Good Service																											
Jubilee	Good Service																											
Metropolitan	Good Service																											
Northern	Good Service																											
Piccadilly	Good Service																											
Victoria	Good Service																											
W & C	Good Service																											
Overground	Good Service																											
DLR	Good Service																											

RANDOM TRAFFIC CAMERAS (TFL)										10	BBC LONDON NEWS (BBC)					48	OPENSTREETMAP UPDATES (OSM)					248
Old Kent Rd/Asylum Rd	High St/Grosvenor Rd W Wickham											Rigby killer 'a soldier of Allah' Mayor bike 'scaring' claim withdrawn Murder police found grave in garden Cameron praises 'towering' Mandela Police crackdown on pirate site ads. Why do we value gold?										

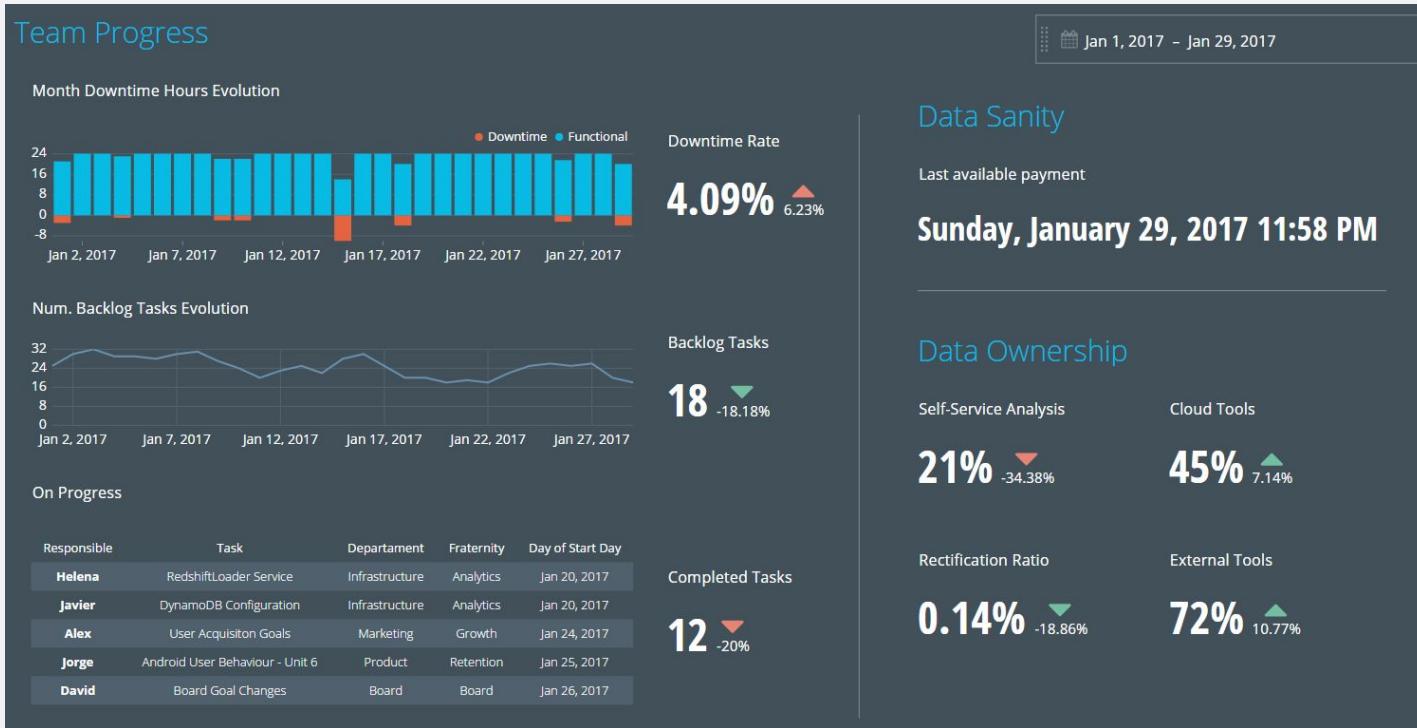
ELECTRICITY (N.GRID)					29	MOOD (LSE HAPPINESS)					38	TWITTER TRENDS FOR LONDON										198
Demand (Great Britain)	48211	MW	8%	unhappier	than the long term average for here	13%	happier	than the whole country right now														

MPs #NFL Christmas #Confident Xmas #ashes London #RIPAlexTurner #12DaysOfJonesDAY9 Waca										198							
0.1 km	The Green	4 Sydney Terrace	20m	0.1 km													

Dashboards: best practices



Dashboards: best practices



Dashboards: best practices

01 Makes simplest those complex data / KPIs

02 It tells a story

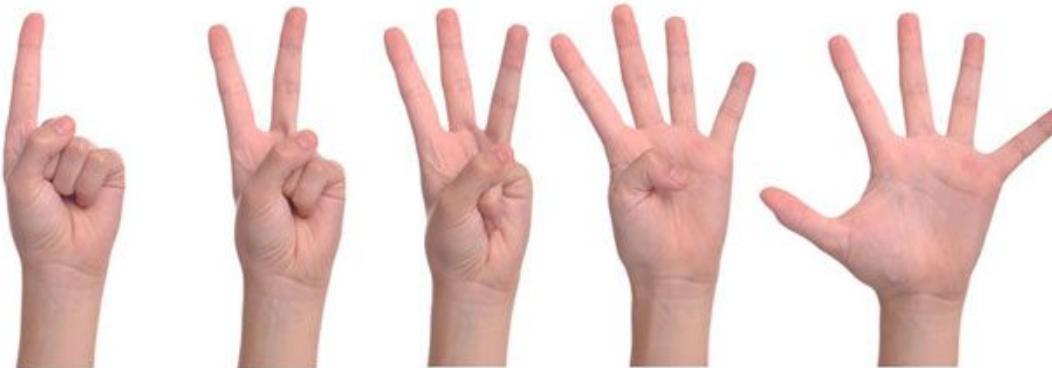
03 Choose visualizations that express the meaning of the data

04 Don't show everything if it's not needed

Dashboards: best practices

<https://www.sisense.com/blog/4-design-principles-creating-better-dashboards/>

Everything Under Control?



I don't
understand
at all

I need to
go over
this again

I think I got
it, but am
not
completely
comfortable

I got it

I can
explain it
to someone
else



coffee break

Let's do some [networking](#)

Our main tool



Do you want to check
how to create a
dashboard in redash?

<http://redash.bdatainstitute.com>

we have a new product in the city

Free Trial

We want to know, if our MVP has made some business improvements.

The revenues are higher? Do we have more retention?

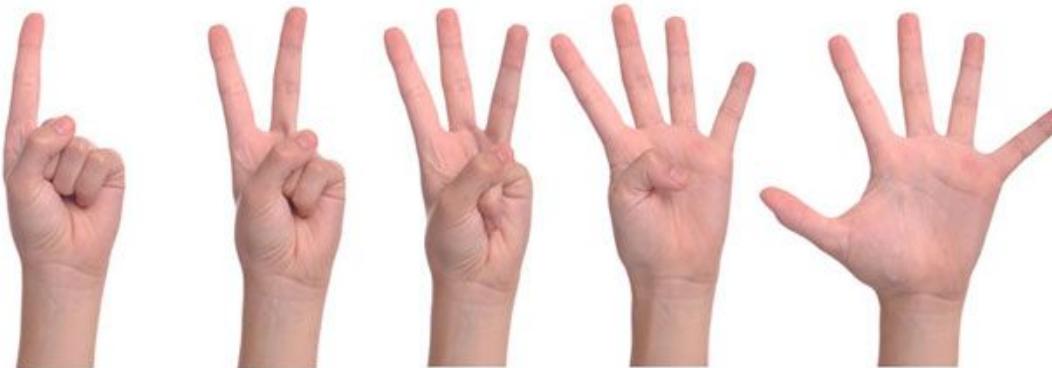
OUR FIRST INVESTIGATION HAS BEGAN!
Ask me whatever you need!



01



Everything Under Control?



I don't
understand
at all

I need to
go over
this again

I think I got
it, but am
not
completely
comfortable

I got it

I can
explain it
to someone
else

**we have a
problem**



02



Email alarm

Our Data Checker has sent an email during the weekend to our Directors. Of course, you are included on the email. We need to answer as soon as possible, because nobody wants that your CEO is worried on a Saturday!

Data Checker

Data Checker <insights@abaenglish.com>

Sat, 29 Sep, 07:45

to codir ▾

Hello,

The following anomaly was detected on **2018-09-29 05:45** as part of the test **acquisition_cpl_daily**:

Check that acquisition and CPL values are normal

Observed value of Acquisition (acquisition) was too low: 7476.000

It is lower than the expected minimum of **7545.356**

Please perform the proper actions if needed.

Have a nice day,

Insights automatic data checker.

Everything Under Control?



I don't
understand
at all

I need to
go over
this again

I think I got
it, but am
not
completely
comfortable

I got it

I can
explain it
to someone
else

we have a **PROBLEM**

 **03**

SOS

Our Data Checker has detected that one KPI has decreased. We need to check, if everything is OK. Related KPIs? Who we need to inform? Do we have a crisis?



Data Checker

acquisition.yaml :: paid_share_no_drop [iOS]

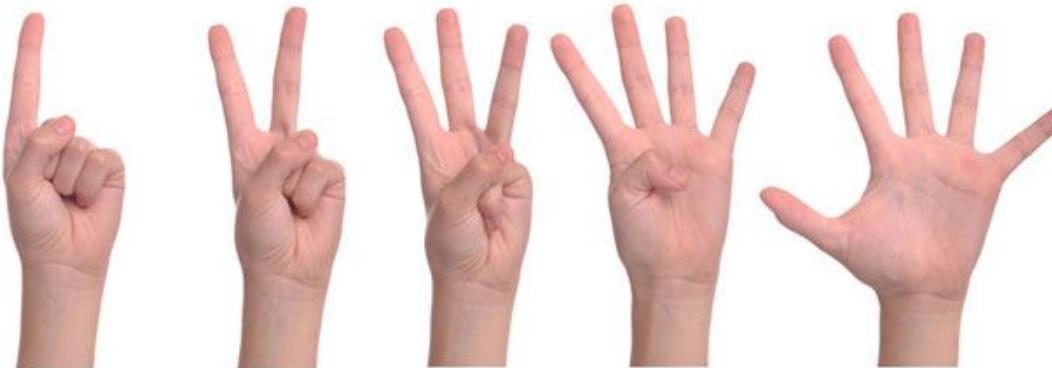
Test case: paid_share_no_drop
Time: 2018-09-22 05:50
Description: Test there's no drop in paid share of acquisition
Run id: 68e5d7f6-be2b-11e8-bc15-02d26a075182

```
SELECT
    date_of_activity,
    SUM(CASE WHEN partner_category = 'Paid'
        THEN acquisition
        ELSE 0 END) / NULLIF(SUM(acquisition), 0) :: NUMERIC AS paid_share
FROM reporting.cohort_analysis_daily
WHERE date_of_activity < CURRENT_DATE
    AND channel = 'ios'
GROUP BY 1
ORDER BY 1 DESC
LIMIT 31;
```

details: {deviation: low, max: 0.6713224637431152, mean: 0.5685909671037546, min: 0.46585947046439385,
std: 0.03424383221312024}
element_count: 31
missing_count: 0
missing_percent: 0.0
observed_value: 0.4203084832904884

	date_of_activity	paid_share
0	2018-09-21	0.420308
1	2018-09-20	0.511257
2	2018-09-19	0.534702
3	2018-09-18	0.477847
4	2018-09-17	0.539198
5	2018-09-16	0.545040
6	2018-09-15	0.515496
7	2018-09-14	0.544058
8	2018-09-13	0.567691
9	2018-09-12	0.597384
10	2018-09-11	0.617708
..	2018-09-10	0.620426

Everything Under Control?



I don't
understand
at all

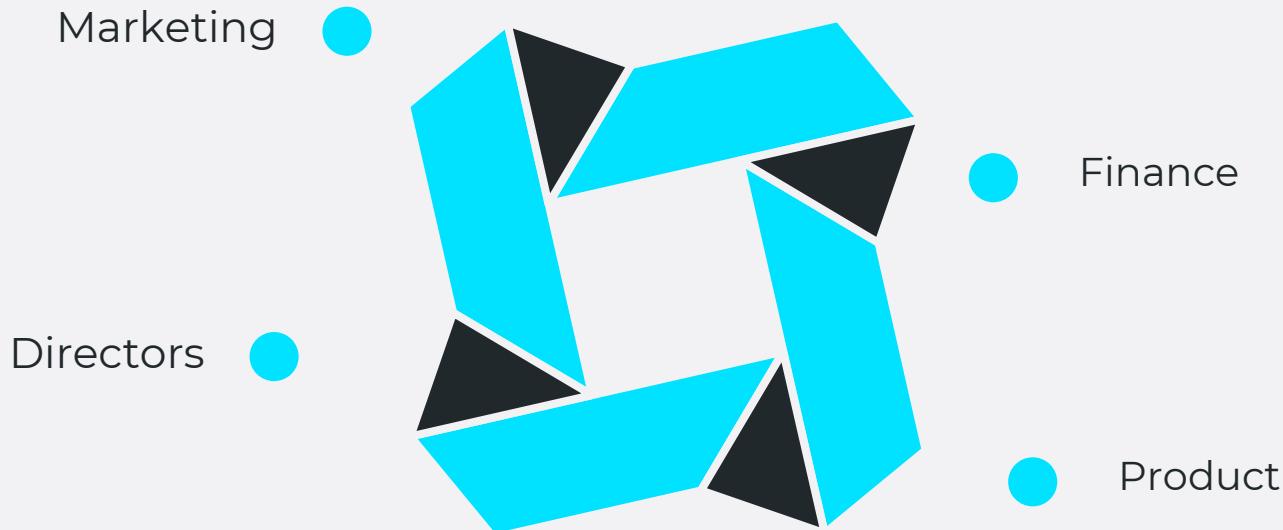
I need to
go over
this again

I think I got
it, but am
not
completely
comfortable

I got it

I can
explain it
to someone
else

Homework for next session



What you need to prepare

01

Present **the** problem

The rest of the class don't know what are the needs of the petition. You need to explain, the better way you can.

We need to understand what are those principal KPIs that you detect are the more representative for this client.

Define **main** KPI

02

03

Present **the** dashboard

Show us the most impressive dashboard that you are able to create. We will evaluate this dashboard with different criterias.

our clients testimony

“Queremos ver todas las métricas importantes de negocio para hacer un análisis semana a semana. Queremos poder verlo el trend de esta métrica durante este año. Además, nos gustaría poder tener el comparativo respecto al mes anterior. Lo puedes tener para la reunión del comité de la semana que viene?”

July
CEO

“No entiendo porque no puedes actualizarme el Excel de ventas que tengo, solo te costará 15 min al día. Es muy importante hacer el seguimiento de las ventas día a día. Me dicen que tengo que usar la herramienta de datos nueva, así que espero que puedas darme el recuento de ventas, el ticket gross y net, y tenerlo clasificado por país y por divisa. Ah y por método de pago.”

Matt
Director Financiero

our clients testimony

“Buenos días, necesitaría un dashboard con todas las métricas de marketing concentradas. Tenemos que poder ver su evolución diaria, que podamos comprar subidas y bajadas de investment, y que sea muy sencillo de entender. Además, queremos poder analizar las campañas de UA. ¿Lo tendríamos para mañana?”

Victoria

Director de Marketing

“Lamento molestaros, pero nos acabamos de dar cuenta que no tenemos un cuadro de control para el MVP que queremos lanzar la semana que viene! Es imprescindible tener controlados los KPIs que vamos a monitorizar, teniendo en cuenta que afectará tanto a la conversión como a la activación del producto. ¿Qué nos podéis ofrecer?”

Anthony

Director de Producto

thank you.