Cooladata

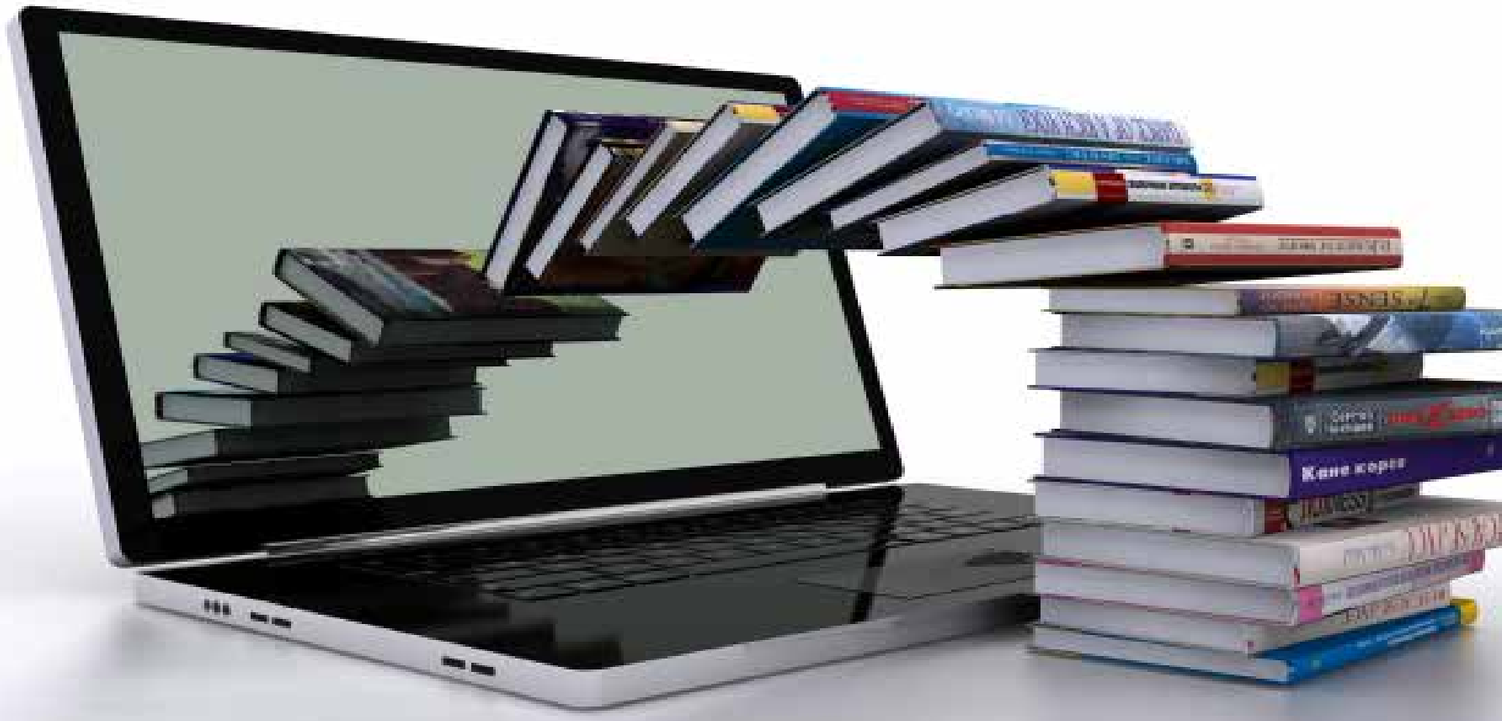
ABA English:

An eLearning Leader Learns from Its Users

Cooladata

ABA English is an online, subscription-based, distance learning platform that helps adults in over 170 countries learn English. ABA faced an interesting challenge - an abundance of data about their users, which they were unable to effectively harness. Because the learning platform is registration-only, ABA can name every user, from their first visit to the platform to the point they decide to upgrade to a paid premium account, and beyond.

Combined with ABA's large user base, this created an enormous "wealth" of data. It was clear to ABA's data science team that they had the data to make very sophisticated analyses about their user base. They identified several critical reports they could construct based on their data. Some focused on monetization and growth opportunities, such as helping identify and better understand their most valuable users. Some focused on areas in need of improvement, such as understanding what causes users to disengage or churn.



Cooladata

ABA's Wish List of Analytics Reports

Growth-Oriented

- **Premium user behavior** - compare the behavior of users just before they upgraded to premium, compared to other users who did not upgrade.
- **Probability of upgrade to premium** - use predictors such as number of clicks on pricing page, engagement in learning units, number of videos viewed, and CRM data, to predict the likelihood a user would upgrade to premium and renew their subscription. This would allow ABA to run targeted campaigns on the users most likely to result in revenue.
- **Engagement impact of new features (A/B testing)** - do new features cause users to engage more and eventually upgrade to premium, or the opposite?

Cooladata

Improvement-Oriented

- **Engagement drop** -detect drop in engagement in a specific part of the software, to identify bugs or UX issues.
- **Stuck in the process** - are users stuck in the process of the learning platform, and in which scenario or device? ABA English offers a series of learning units users must complete from beginning to end. It is critical to see in which units users get stuck and do not proceed to the next learning units, and why.

Cooladata

Five Missing Pieces

The data was there. The data science expertise and analysis skills were there. But the technology wasn't there to make these reports happen. What was missing?

Desired Report	What Was Missing?
GROWTH-ORIENTED (POSITIVE MEASURES)	
Premium user behavior	▪ **Engagement data stored separately from user data**, so there was no easy way to segment users based on their behavior. ▪ **Unable to create segments for** (a) users who never upgraded, (b) upgraded first time, (c) upgraded and renewed subscription. ▪ **Time-series analysis** - for example, being able to look at a user who eventually upgraded to premium, and see how she behaved as a new user in her first learning units.
Probability of upgrade to premium	▪ **Engagement data stored separately from user data** - in order to create a predictive model, it was necessary to use these two data sets together. But there was no indication, when exporting user engagement data, which specific user performed each engagement action.
Engagement impact of new features (A/B testing)	▪ **Unable to create segments for** users who had the new feature turned on vs. off ▪ **Time-series analysis** - being able to analyze whether a user who interacted with the new feature, eventually ended up becoming more engaged and upgraded to premium.

Desired Report	What Was Missing?
IMPROVEMENT-ORIENTED (NEGATIVE MEASURES)	
Engagement drop	▪ **No quick access to behavioral metrics** - it was possible to identify simple things like a drop in time spent on a certain section of a learning unit. But ABA was missing more sophisticated behavioral metrics, like a drop in a likelihood to upgrade or in the engagement over the following days or weeks. They also couldn't easily slice the data by fields such as device, browser information, version or app type, to identify issues in certain environments.
Stuck in the process	▪ **No way to track and analyze events** on every transition from section to section within a learning unit and between learning units. This creates a huge volume of data which is very difficult to process and analyze using traditional methods.

The challenges boiled down to five missing pieces:

- Engagement data stored separately from user CRM data
- Limited ability to segment users for analysis
- No time series analysis
- Behavioral metrics not readily available (required heavy manual analysis)
- No way to easily analyze large volume of events

Cooladata

Cooladata: Engagement and User Data, Combined and Analyzed

Enter Cooladata, an end-to-end big data behavioral analytics platform, which provides the data warehouse, instant data integration and powerful behavioral analytics. ABA English implemented Cooladata's end-to-end big data analytics platform, and started tracking all user interaction with Cooladata. User data from the back-end CRM was also moved into Cooladata, and the platform started correlating each engagement event with a specific user.

Once individual user engagement data was correlated with that same user's back-end data, such as subscription to premium and subscription renewals, half of the problem was solved.

Cooladata

To solve the other half, Cooladata helped perform advanced analysis on the fly:

- **Flexible user segmentation,** making it possible to look at premium user behavior separate from non-premium users, and segment the engagement impact of new features.
- **Time-series analysis,** making it possible to see how premium users behave in earlier stages, and subsequent behavior of users experimenting with new features.
- **Fast access to behavioral metrics,** making it possible to see a drop in engagement beyond the basic metrics. For example, Cooladata can show that users experience a certain learning unit show much lower engagement in the app for the next week and much lower likelihood to upgrade to premium. Indicating a major bug or interaction problem.
- **Event-based analysis -** Cooladata is an event-based platform which tracks and analyzes all clicks and transitions with an application. This made it possible to understand which users are getting "stuck", in which sections of which learning units.

Cooladata

Another important capability was Cooladata' JDBC connector, enabling ABA to export all the data easily for analysis in R by the data science team. The Cooladata export made it much easier to:

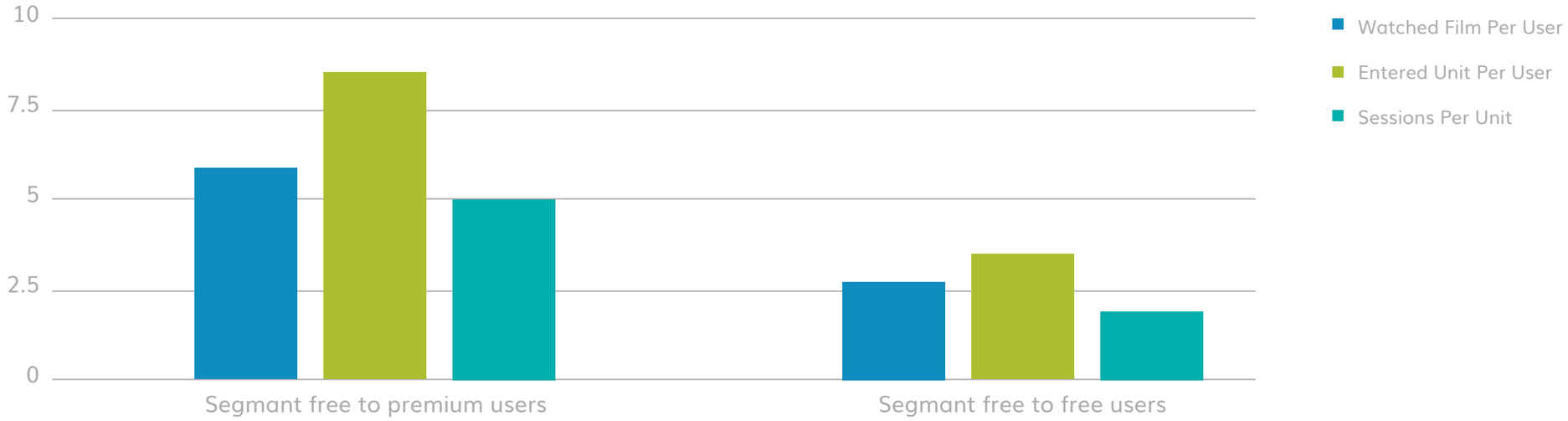
- **Build a predictive model for probability of upgrade to premium**
- **Analyze where users are stuck** in the process and how to improve it

These two analyses were done outside Cooladata, but Cooladata provided everything needed for the analysis, without requiring complex data preparation and correlation.

Wish List of Reports, Delivered

Here are examples of reports ABA English built in Cooladata to address each of their report scenarios.

Premium behavior



This report compares behavior of free users vs. premium users just before their upgrade. The report shows a very marked difference in engagement metrics.

Cooladata

Probability of upgrade to premium

userid	Registration day	Total session #	Total time usage	Total day usage	Total level #	Total unit #	Watched film	Listened Audio	Max prog update	Converted user	Conversion days
11073570	29/08/2016	0	0	0	0	0	0	0	0	0	NA
11068413	29/08/2016	0	0	0	0	0	0	0	0	0	NA
10597625	07/08/2016	0	0	0	0	0	0	0	0	0	NA
10041725	12/07/2016	7	6615415	34	1	1	2	4139	74	0	NA
9889540	03/07/2016	1	1063226	1	1	1	1	55	25	0	NA
9916325	04/07/2016	3	1283128	10	1	1	1	54	37	0	NA
9894530	03/07/2016	11	91909249	1	1	1	3	442	100	1	1
9901931	03/07/2016	8	109648843	2	1	1	6	205	100	1	3
10041570	12/07/2016	19	45712146	29	1	1	5	300	100	1	34

2 classes of users: users that do not convert / users that convert

Generate 2 sets of data: training set (to build the model) / test set (to test the accuracy)

Model: SVM

- Type: *classification*
- Kernel (type of algorithm that generate the feature space): *radial / polynomial / linear / ...*
- Some parameters for the kernel algoritm (cost, gamma, ...)

Output: Accuracy on Training set and Testing set in assigning correctly the % of not converted and the % of converted

This model, extracted from Cooladata and calculated in RI, measures the probability that specific users would upgrade to a premium account.

Cooladata

Engagement impact of new features (A/B testing)

Starting Full Roll Out: 2017-06-01

day	active users	opend moment type	opened moment	started moment	started moment exercise	selected moment answer
2017-0..	13,682	18.8	17.8	17.5	17.5	17.4
2017-0..	13,757	18.3	17	16.9	16.8	16.7
2017-0..	16,454	21.9	21.2	21	21	20.8
2017-0..	18,544	24	23.2	23	23	22.7
2017-0..	19,335	31.4	30.1	29.5	29.5	29
2017-0..	20,081	26.4	25.5	25.1	25.1	24.8
2017-0..	20,861	31.1	30	29.5	29.5	29.1

This is how ABA English analyzed the impact of a major new feature, ABA Moment. The feature allows users to jump between learning modules, and the ABA Team wanted to verify that this had an overall positive effect on engagement. The table above shows the number of users who accessed the feature and performed specific actions.

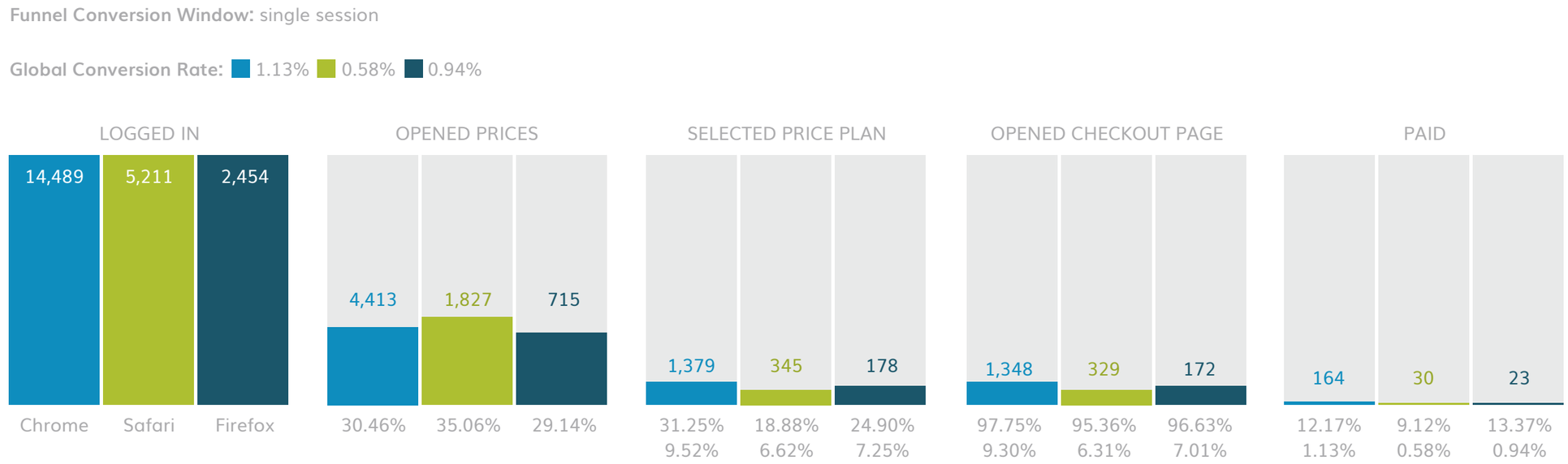
Cooladata

ABA Moment: Number of Premium Active Users sending an ABA Moment event by day

This chart shows the growth in consumption of the new feature, and the frequency of specific interactions, over a two-month period.

Cooladata

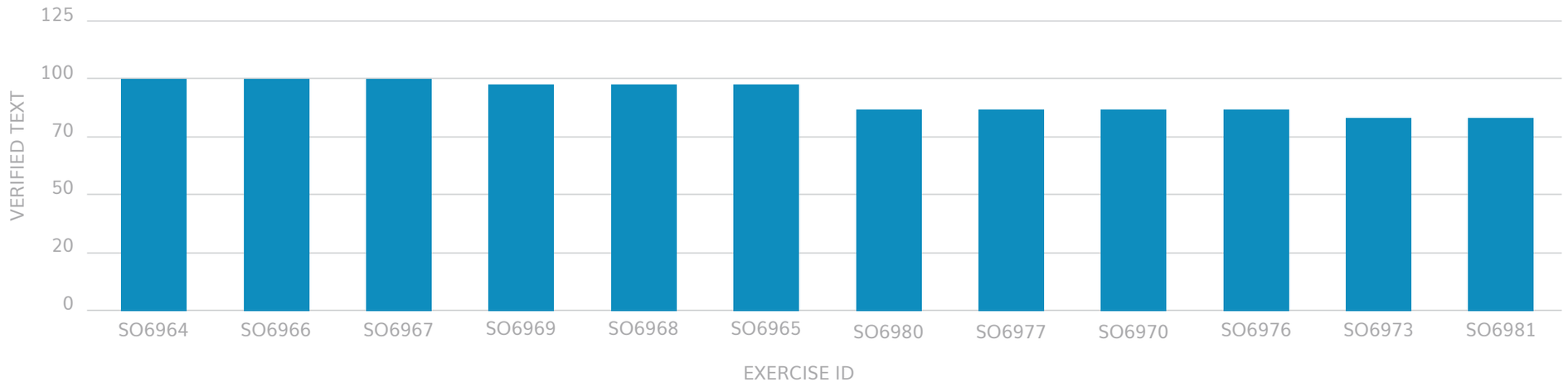
Detecting engagement drop



ABA English evaluated up-to-the-minute conversion data across the funnel, to compare behavior between different versions of the same app on different devices to. This helped detect platform-specific issues, defects and problems.

Cooladata

Stuck in the process



ABA English used the chart above to better understand where users were abandoning the learning process. The chart helped them problems or issues related to specific units of the curriculum.

For example, are specific lessons too hard or too boring, causing users to abandon the process without completing it? This would be indicated by users being stuck in specific learning units for too long.

Cooladata

What Did ABA Discover?

The point of the entire exercise, of course, was to derive insights for the business. What did ABA English learn from the new reports and data Cooladata put at their disposal?

Helping users on their path to premium

Based on Cooladata's user segmentation, ABA discovered that users who ended up upgrading to premium, or renewing their premium subscription went more in-depth in their studies, going through more learning units and viewing more videos. This showed the ABA team that encouraging higher engagement and progress through the learning units would also encourage users to subscribe to premium, and later renew their subscription.

Cooladata

Validating a major feature: ABA Moments

ABA introduced an important feature, ABA Moments, which allows users to turn every available moment into an opportunity to learn English. A controversial aspect of this feature was that it allowed users to go outside their current learning unit and interact with other units (going outside the regular learning sequence).

ABA needed to understand if this new feature helped users enrich their learning experience and made them more likely to engage and upgrade, or interrupted the sequence of learning units and decreases engagement.

Using Cooladata, ABA monitored and compared the behavior of a group of users for which this feature had been opened, vs. regular users who didn't have it. They used the data to improve the ABA Moments feature, to a point that led to higher engagement. Based on the data, they decided to go with the feature and launch it to all users.

Cooladata

Visibility of the sale cycle

Using Cooladata's time series analysis, ABA English was easily able to understand the time needed for users to upgrade to premium, across different segments. This enabled the executive and marketing teams to make more accurate revenue projections. For example, how long would it take from marketing spend and new acquisition of users to actual revenues.

Solid model of high-premium-potential users

Using the data exported from CoolaData, ABA's data science team was able to build a solid machine learning model predicting the likelihood of a user to upgrade to premium, and the likelihood of a premium user to renew their subscription. This helped ABA's product team find and analyze the weak points in the ABA application which were causing users to churn and required improvement.

Cooladata

Understanding why mobile conversion was broken

An endemic problem in ABA's platform was that on mobile devices, conversion to premium was substantially lower than on desktop. Using Cooladata, the data science team analyzed behavior in and between learning units, segmented for mobile and desktop users. They gathered a month of data and exported daily aggregates at the per-section and per-user level. Then they performed an in-depth statistical analysis in R, which tracked the progress of different users through the learning units.

The analysis resulted in an interesting finding: mobile users often get "stuck" in a specific learning unit, then leave the site.

A closer investigation showed that, while on the desktop site there is an easy way to "skip" a difficult section or unit to move to another unit, in mobile the process was linear. The UX on mobile was designed in such a way that users who got stuck in a difficult part of a learning unit had no option to skip, and just left. Those users never reached the more advanced premium learning units, and so would not convert.

Fixing this experience on mobile and enabling a non-linear experience, like on desktop, increased conversion rate on mobile dramatically, bringing it close to desktop levels.

Cooladata

Real time bug resolution

Cooladata helped ABA identify negative changes to behavioral metrics which were focused on specific parts of the application, and immediately notify relevant staff about those changes. This helped identify bugs very near to real time, identify on which device the bug occurs, and fix them before they hurt conversion and revenues. In addition, this capability helped ABA identify the impact of UX changes or improvements, and in case an interface change was not helpful to users, roll it back.

Cooladata

ABA Helps Users Learn - and Now Learns from its Users

ABA always had rich data about its users, but found it difficult to process and make sense of the data. By implementing Cooladata, ABA got access to a series of "dream reports" with new, crucial insights about its users. Having access to these reports helped the data science team learn much more from how users interact with the learning platform.

Beyond reports, Cooladata was able to generate much more useful raw data, that was fed into a deeper analysis and yielded surprising results. Most importantly, it helped ABA fix a major problem in their mobile conversion.

Having the data is never enough. Just like ABA's users cannot take a dictionary with English words and phrases and turn it into a spoken language, ABA itself wasn't able to take its user data and convert to insights. ABA helps users break down the problem of learning a language, by parsing the language sequentially, learning unit by learning unit; Cooladata was helpful in turning their mass of data into a series of reports that provide real business insight.