

The Lean Product Waltz:

Data-Driven User Research & Research-Driven Data Analysis

Digital products and channels are generating exponentially more data each day. Enhancing user research with data analysis has become less expensive and more practical than ever before, fostering the adoption of lean research approaches. However, a deep understanding of the context in which the data were generated is imperative to navigating the data during analysis, and ultimately to providing interpretable and actionable results. This context must be provided in large part by carefully designed conversations with the user through research techniques: interviews, workshops, focus groups, surveys, and the like.

The interplay between data analysis and user research is not just a one-time occurrence, but rather an iterative process of extracting knowledge at each step and applying it in the next. In some very real sense these two practices can be considered as engaged in a carefully choreographed dance, for which finding the rhythm has the potential to produce some notable results – for example, in the case of FIBR partner <u>Nomanini</u> described below.

When applied to lean product development in particular, the result of this "dance" is typically more efficient progress along the product roadmap. This means the product more rapidly approaches the optimal state for the target users, typically translating to faster development, higher revenues, increased engagement and retention, and more data to feed back into the cycle.

In this briefing note, we explore this process in the following sections:

- 1. Motivation: Setting the Tone with Lean Product Development
- 2. FIBR Case Study: "Dancing" with Nomanini's POS and Airtime Loans
- 3. Conclusions & Takeaways: Catalyzing and Continuing the Waltz

Lastly, we've included as an appendix some guidance around some frequently asked questions regarding lean product development. While this is not an exhaustive list, the goal is less to provide a strict set of instructions than to present a framework for ideas to organically arise and be methodically treated.

Motivation: Setting the Tone with Lean Product Development

Lean product development¹ is a methodology for rapidly introducing and continuously optimizing a product in an untested, dynamic environment. Where more traditional approaches focused on careful design in a controlled environment prior to introduction to the target users, lean product development focuses on iterative learning via the following set of axioms that we put to action within FIBR. To do so, lean development typically combines agile software development techniques with user centric research activities like Human Centered Design or Customer Development.

First, we focused on outcomes, not outputs. Rather than dwelling on the exact specifics of how we were planning to build a product or feature, we started with our intended effects of the product on the users – like alleviation of pains, opportunities for gains – and worked backward from there. We prioritized observation of

¹ For a deeper dive into the principles of lean product development, 18F's <u>Lean Product Development Guide</u> and <u>The Lean Product Playbook</u> are two useful resources.





the targets' behavior over features or their responses, closely observing what they do rather than what they say.

This approach provided us with the ability to quickly learn and respond through a series of collaborative experiments with users. Assumptions make their way into designing any product, and the sooner those assumptions can be tested and either verified or invalidated, the better. A major strength of lean product development is that the core includes these experiments to generate data to be put into context using insights from user research, teaching stakeholders how to respond with the next iteration.

In order to prioritize these lessons into actionable plans for next steps, we had to make sure to involve individuals with the right mixture of skills on the product team. These skills may vary depending on the product, but typical examples in fintech products may include: user experience (UX) design, software development,

Lean Principles

- Outcomes, not Outputs
- Always Think Problem First
- Make Assumptions Explicit, Test Most Critical First
- Involve a Cross-Disciplinary Team
- Keep Experiments Light
- Learn and Respond
- Build Shared Understanding
- Pay Attention to Prioritization

data analytics, topical expertise, business modeling, marketing, project management, and more.

In regular meetings, like sprint kickoffs and wrap-ups, each participant can provide input from their own perspective, allowing the team as a whole to triangulate the next problem to address first and design the next set of lightweight, knowledge-generating experiments.

FIBR Case Study: Dancing with Nomanini's POS and Airtime Loans

Putting these lean principles into practice has been an integral part of the FIBR partnerships to this point. One particular example is in our ongoing engagement with Nomanini, where we utilized historical data, digital analytics, and several methods of user research to learn about how to utilize their platform and existing products to best serve the needs of SME merchants and banking/mobile money agents in Ghana.

Background

Prior to becoming a FIBR partner, Nomanini already had an existing Point of Sale (POS) system (among other fintech products) on top of their payments platform. The POS allowed merchants to top up their account, then sell airtime to retail customers. The merchant accepts cash payments and the POS prints a voucher with an airtime code that the customer uses in



lieu of a scratch-card to complete the transaction. Nomanini had previously deployed quite successfully with a field partner in Mozambique, and was looking to replicate that success in Ghana when they joined FIBR.



Thus they partnered with <u>InCharge</u>, an aggregator with over 2000 agents in Ghana, to launch a pilot program of 100 POS devices for merchants. With this program, they sought not just to use the Mozambican product directly, but to customize to the needs of Ghanaian merchants. For example, one of the major pains that was hypothesized initially was that merchants were running out of float in the middle of their operating hours. In other words, they could be forced to close their business and stop selling airtime in order to travel to a location where they could exchange collected cash for more electronic float.

The major intended outcome (not output) was to allow the merchant to remain available for business with their customers, even after exhausting their float. One mechanism for making this happen could be to allow the merchant to request a short-term top up loan that they could pay back during a lull in business or after hours.

By following our lean principles of thinking *problem first* (rather than solution) and *testing critical assumptions*², it was clear that the place to start was to examine what evidence we had on hand. We assembled our *cross-disciplinary team* – data scientists, user research experts, business modelers, and POS experts – and got to work.

Initial Data Exploration by Proxy

Initially we had no data in Ghana to work with and hadn't yet arranged interviews with Ghanaian merchants, but decided the Mozambique data could serve as a potential proxy for a *lightweight experiment* in the form of data analysis. In other words, we might already have evidence about the severity of the pain caused by stocking out of float. The data was pretty bare bones, but included parameters of 5600+ transactions across 1300+ devices.

Upon examining the data, BFA team members Matt Grasser and Javier Linares confirmed that at least in that environment merchants did in fact run out of liquidity early in the day, followed by a significant delay in topping up and starting sales again. We can see this in a set of charts that illustrate each of these conclusions:



Figure 1: Histograms from the nano-project analysis of (left) the time of day at which merchant accounts hit zero balance, and (right) the delay in subsequently topping up the float. Many merchants hit low balance; most often early

² Lean Product Principles utilized in the Nomanini engagement are listed in *italics* in this section.





in the morning, and halfway through the day. We also see that there is a significant number of devices that take hours to top up.

While it is obvious that running out of digital stock will hurt a merchant's sales, we did not know conclusively prior to this analysis how frequently merchants were actually running out, or if it was happening at all. By *testing our assumption* that this was occurring, we were able to first confirm this in an inexpensive manner by sitting at a laptop, thereby increasing our confidence and reducing the risk of investing in more detailed follow-up research.





Merchant Interviews

With this confidence in our hypothesis following the confirmation by the Mozambique data, it seemed reasonable to invest a bit of time in exploring the Ghanaian case, to see whether there could be nuances. BFA team members Asa Nyaga and Brendan Ahern conducted an initial set of six merchant interviews, with their explicit goal stated as follows:

"To assess the value proposition of Nomanini's offering a 'credit ladder' to InCharge merchants on the basis of credit scoring linked to merchants' sales of airtime and other virtual goods."



Figure 2: Gina, a merchant from the Nomanini User Research Report.

In this step, we verified that the problems we inferred from the Mozambique data set in fact do apply to merchants in Ghana. In fact, the research report noted that "all 6 of the merchants interviewed indicated that running out of float – and suspending airtime sales – had become an issue."

Liquidity was the major problem, and we were now confident we had the right resources and channels to put together a solution.





Follow-Up Interviews

Ideally, at this point we could have taken the interview results to learn and respond with a deeper dive into Ghanaian merchant POS data. We had planned to conduct segmentation of existing data first, perhaps informing a prototype credit model, then do the deeper dive. However, due to unforeseen delays in hardware procurement and distribution to merchants, our plan to collect hard data on transactions was delayed. In the process of adjusting our plan of action accordingly, we successfully turned to our lean product principles for an answer.



Figure 3: FIBR team members from BFA and Nomanini conduct an interview with one of the merchants.

By paying attention to our *prioritized goals* and leaning on members of our *cross-disciplinary team* – namely our user research expert, product owner, and technology expert – we conducted follow-up interviews. The results served to supplement the delayed acquisition of transactional data, and ultimately to come to a better understanding of the cash flows of merchants and to identify other credit-related needs, such as credit to purchase physical goods and expand stock in addition to virtual goods.

Airtime Data Analysis

Shortly after the follow-up interviews, the hardware arrived. Now that we'd heard about what Ghanaian merchants said it was time to find out what they did. InCharge provided 10 of their merchants with Nomanini's POS systems that included an airtime sales feature, but that did not yet have a loan feature. This lightweight experiment produced results which afforded us the ability to observe and analyze merchant behaviors without having to be on-site. Some of the key results are displayed in the accompanying figure.







Figure 4: charts of proactive and reactive merchants from the airtime data analysis

We identified two main segments among merchants based on what we termed "proactive" and "reactive" top up behavior. The proactive merchants seemed to know when they were about to hit low balance and would top up without a break. The reactive merchants would drop below a threshold balance, then have a break in sales while they found the opportunity to go and top up. When we followed up briefly with some of the agents, they confirmed this behavior. For instance, one proactive merchant responded, "50 cedis balance means low. I don't like to say NO to customers. It's annoying." For this particular merchant, "annoying" meant transferring funds from a personal mobile money account into the InCharge account while waiting for an opportunity to deposit cash into the bank. If, however, no personal funds were available, this would mean that sales would need to cease, which would be more than annoying.

In addition to the analysis of temporal behaviors, we also looked at basic summary statistics of these transactions. This informed us of exactly what the average volume and total value was across all merchants, as well as the variance across the set.

With this data in hand, we were finally able to design a loan product that would fit the behavior and needs of actual Ghanaian merchants. The behavioral analysis led directly to a tiered credit system, and the statistical summary led to the initial loan amounts for each tier. All that was left was to get feedback on the loan design, which we did in as inexpensive a manner as possible.

Merchant Survey

To save time and money, rather than traveling to the merchants with this design Javier and Werner initially used WhatsApp and Google Forms to conduct surveys across the merchants we had data on. The results were immediate and encouraging. When asked to rate their interest in a POS-based loan on a scale of 0-5, every single merchant responded with a 4 or 5. Even more promising, one merchant asked over WhatsApp if she could get a 200 GHc credit that day (she later repaid on time):







Figure 5: WhatsApp conversation in which a merchant requested and received a loan.

With these encouraging results, they then followed up with an in-field product survey to test the initial loan design and interface implementation. This led to a list of products improvements/changes, that were incorporated as a loan request feature right into the POS, then distributed for further experimentation and data generation.







Figure 6: Screenshots of the loan request/disbursement process that was built into the POS as a result of the lean waltz.

Next Steps

At the time of writing, we're awaiting the next set of data from a broader set of merchants. We're looking forward to the potential of applying machine learning techniques that could give us the ability to improve the product even further, for instance:

- Determining when a merchant is about to run out of float, and pre-emptively remind them of the loan feature
- Determining how much a merchant is selling, and suggesting an amount that may get them through the hour/day/week
- Determining what sort of usage history we need in order to safely authorize the overnight loans
- Predicting if a merchant would repay a larger loan with a week term (the second step in the "credit ladder")

As we await the production of this data, we may also continue to look more deeply at other larger related data sets as a proxy. In the meantime, we continue to observe how the merchants behave as they request and repay loans. We also want to complement that data stream with a set of demographic and business parameters that could also be predictive of financial behavior.

Conclusions and Takeaways: Catalyzing and Continuing the Dance

Through the experiences resulting from following lean principles and combining interview-based research and usage data as described above, the FIBR team was able to quickly come to a solution that solves real problems and gives merchants "super powers" they may not have had otherwise. In the case of Nomanini's credit product, merchants could "time travel" to get airtime credit from the future, in order to best serve their customers now.

Beyond reinforcing the value of traditional lean approaches, we also noted three emerging takeaways: a framework for deciding whether to start with research or data analysis, an understanding of how these two techniques feed into each other iteratively, and a deeper understanding of some of the similarities and differences in the application of lean principles in African fintech versus the Silicon Valley environment in which this framework was forged.



Learning to Dance: Determining Which Step Comes First

While the gains realized in this particular approach were specific to Nomanini merchants, the manner in which we approached the problem was not. As a result of subsequent discussions, we developed the following working framework around this interplay of data analysis and user research when approaching the lean development of a new product or feature:



The Endless Waltz: Lean Product Design is Tirelessly Iterative

In conclusion, we saw that our adherence to lean principles naturally gave rise to this interesting phenomenon of data and research working in harmony with each other, spurring on innovation at each step. In fact, the synthesis of Nomanini's financial product in the form of the perfect credit model will never be complete. The merchants will evolve their practices as a result of the new features, features will evolve to fit the merchants' new behaviors, and the interplay of results from data analysis and those from user research will continue to feed each other iteratively and indefinitely in an endless dance.

This approach can be generalized and applied to any scenario where it is possible to mix information from user interviews (hear what they say) with data from user behavior (see what they do). Once a digital product gets used, it can start a virtuous loop with the usage data broadening our understanding while highlighting areas of enquiry for further user research. This effect is particularly desirable when resources for user research are limited, when the users are poorly understood to start with and when it is reasonable to expect a bias from early adopters. In other words, the sooner you can put a digital product in the hands of your low-income users, the faster you will learn how to make it great for them.



Finding the Rhythm: Lean Products for FIBR versus Silicon Valley

In our FIBR engagements, both with Nomanini and our other two Ghanaian partners, we were pleased to find some commonalities with the effectiveness of the SV approach to our lean product design. We were able to use many of the same product analytics tools, basic premises, and principles to produce comparable results in the product process.

That said, there were also some notable differences working with these principles in emerging markets, and often doing so from a remote location.

Traditional testing tools may not function properly, or even exist: First, many of the digital tools that SV can take for granted as part of their toolbox exist only in nascent state in emerging markets, or sometimes don't exist at all. For instance, it is harder to set up interviews with potential users of a product without a craigslist "gigs" section to find subjects, or a saturated adoption of facebook to make the most of landing pages and to push ads through. Tools like <u>UserTesting</u> – a product commonly used to collect more subtle, qualitative feedback from individuals while they use the product – don't quite work in emerging markets either due to lower adoption and less reliable and lower bandwidth networks³.

Physical distance between teammates & users: Next, there are some issues that arise due to the physical distance between teammates. Since our product designers are very remote relative to the users (for instance in NYC and Accra, respectively), there's less availability for direct user research and thus more of a need for innovative means of collecting cheap, reliable data. Likewise, our ability to "scratch our own itches" was adversely affected. If an issue arose, it meant carefully communicating the solution through a chain of teammates, increasing the risk of miscommunication versus an SV office setting.

Differences in ecosystem maturity: Third, we noted some key points in our process that were affected by differences in the maturity of each environment's respective fintech ecosystem. In developing markets, less data is available about users to start with, which means more work has to be done to either generate that data, or find sufficient proxies. Equally important, fintech products in emerging markets need to be more robust than those in SV before they can be put in the user's hands. While in either case you don't want to lose anyone's money, the potential for mistakes that lead to devastating consequences can be higher in emerging markets, and thus financial service providers and developers can be more risk-averse.

Stronger value propositions can mean eager early adopters: Finally, and perhaps most notably, because in these markets there are often basic financial problems lacking sufficient technical solutions, the users' appetites to try untested new tools are still strong. With the right product, marketing materials, and value proposition, early adopters abound. We've found potential users quite ready to try out new and innovative solutions, which often can lead to an exciting leap-frogging of even some cutting-edge SV solutions.

³ At the time of writing, UserTesting has recently launched an Indian user pool, which could be an encouraging sign of a push into emerging markets.



Appendix: Lean Product Design FAQs

In discussing our approach to lean product design, we've come across a few common questions, some of which are documented below. It is important to note that what follows is not necessarily meant to be an exhaustive list, but simply may be of use in generalizing some of the principles espoused above.

How can I determine which aspects of the business model should be held out of scope while others are explored?

For our purposes within FIBR, we've created a Lean FIBR Canvas, based on aspects of the original Lean Canvas⁴ and the Lean Impact Canvas⁵ frameworks. We incorporate this canvas into interactive workshops, in which we work to fill this chart out either on a physical printout with sticky notes, or digitally in a collaborative Google Doc. This has consistently helped us to narrow down the scope of our engagement, while starting to explicitly and transparently lay out our assumptions and other information that feeds into hypothesis building.



⁴ Lean Canvas is Ash Maurya's adaptation of Business Model Canvas by Alexander Osterwalder, but created in the Lean Startup spirit (Fast, Concise and Effective startup). Lean Canvas promises an actionable and entrepreneur-focused business plan by focusing on problems, solutions, key metrics and competitive advantages. See more at: https://leanstack.com/lean-canvas/ ⁵ Lean Change Canvas is Michel Gelobter's adaptation of Ash Maurya's Lean Canvas, presented in his book Lean Startups for Social Change. This approach notably differs from the Lean Canvas in that it focuses on affecting change for targets, as opposed to a primary focus of acquiring paying customers. See more at: http://www.leanchange.net/coresite/#principles



How can I identify which assumptions are the most critical to the outcome?

The types of assumptions that should typically be tested first can be bucketed into 3 broad categories: core assumptions, which must be true for your product to address the need; unknown assumptions, which are those that the team is most uncertain about; and risky assumptions, which would lead to the failure of the product if proven wrong.

For each of these, one can run an expected value exercise: take the likelihood that each assumption is incorrect times the quantitative impact of that assumption being wrong; the most negative value poses the highest risk and should be tested first.

That said, it can often be difficult to assess post ante the probability of being wrong, in which case we can fall back to the Lean FIBR Canvas mentioned above. Typically, items listed in the Targets, Problems, Channels, and Revenue Streams boxes are the highest risk to a product's success. If the customer/problem matching, monetization strategy, or communication channels are off, the product may never even get off the ground. After these are addressed, the next wave is converting these assumptions to solution hypotheses.

How long should I spend looking at data sets and research results before formulating hypotheses?

The important thing to keep in mind during the initial exploratory analysis of the data is that the primary goal is to produce a list of testable hypotheses, and nothing more than that. It's important to keep in mind that we are not necessarily looking to produce formal conclusions or even to ensure our results are statistically significant. At this point, any potential hypotheses should be noted as they arise, and the formal testing of these hypotheses should be left to the hypothesis-testing (not this generation) step.

Another important point, which is woven through the preceding discussion, is that there is no "cutoff" for stopping the generation of assumptions and hypotheses and moving into product development and testing. These two should be ongoing, often in parallel, and should feed off of each other continuously. There are always new data to be synthesized into learnings, improvements to the product, and ultimately to the quality of target users' experience.

How might I decide which hypotheses should be tested?

Ideally, if you have a domain expert available, they are typically a great option for your first pass. Experts know their industry and have directly experienced what has worked and what has not in the past. By simply consulting with an expert – or better yet multiple experts – some hypotheses from the list may be confirmed or disproven right out of the gate.

Another technique, called dot voting⁶, can be employed to produce a group consensus. In this technique, the team gathers and reviews the intended outcomes of the product before reviewing the list of hypotheses to be prioritized. Hypotheses are then listed out on a whiteboard or flipchart, and team members can place stickers or post-its next to the top 3 or 5 they think are most important. The items with the most "votes" should be addressed first. Rather than relying on one individual who likely has his own biases, this method leverages the team's collective expertise and experience to make the decision.

⁶ See <u>https://methods.18f.gov/discover/feature-dot-voting/</u> for details



How many target users must be counted to confirm hypotheses?

The answer to this question is highly dependent on the type of hypothesis being tested, the type of product being used, and the target audience's demographics to be tested, among other factors. Generally, lean testing experts suggest sizes on the order of 5 users for a usability study⁷ to 20 users for a quantitative study.⁸ But even among these experts, those figures are considered to be guidelines⁹ more than hard numbers to be stuck to at all costs.

The important thing to remember is that – even with properly instrumented digital products – there are diminishing returns with each incremental user tested in the early stages of development. As illustrated in the chart below¹⁰, even five users can typically account for 80% of identified usability problems in a given round of testing. Thus, it is often a more effective practice to address those issues first, and introduce the improved product in a new round, than it is to add five more users to confirm the results of the current round.



⁷ See, for example, "Why You Only Need to Test with 5 Users" here:

https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/

⁸ See, for example, "Quantitative Studies: How Many Users to Test?" here:

https://www.nngroup.com/articles/guantitative-studies-how-many-users/

⁹ See, for example, "Lean User Testing" here: <u>http://www.slideshare.net/jduverneay/lean-usertestingintro-ga</u> and "How Many Test Users in a Usability Study?" here: <u>https://www.nngroup.com/articles/how-many-test-users/</u>

¹⁰ See "A mathematical model of the finding of usability problems" here: <u>http://dl.acm.org/citation.cfm?id=169166</u>