Chapter 3: Approaches to Creating Personas

By **Steve Mulder** from *The User Is Always Right: A Practical Guide to Creating and Using Personas for the Web*

First things first: There is no one right way to create personas. Like almost everything you do, a big "It Depends" cloud hangs over your efforts:

- It depends on who your audience for the personas is and what they need in order to agree to use personas.

- It depends on how you plan to use personas and what types of decisions you'll make using them.

- It depends on how much time and money you want to (or can) invest in this user-centered design activity.

Although the persona creation process should be tailored for each situation, there are three primary approaches, based on the type of research and analysis performed:

- Qualitative personas

- Qualitative personas with quantitative validation

- Quantitative personas

Since the approach to creating personas largely depends on the type(s) of research you conduct, let's first explore the user research methodologies you can choose from.

The User Research Landscape

You do research to better understand your users, but exactly what is it that you want to find out about them? That's the first question you need to ask, and its answer dictates which research methods you should use, since specific methods are tailored to finding specific types of information.

One way to explore the landscape of user research methodologies is to classify them as qualitative or quantitative:

- **Qualitative research** is about discovering new things with a small sample size. User interviews and usability testing fall into this category, because they consist of interacting with a small number of users (10–20) to get new ideas or uncover previously unknown issues. Qualitative research doesn't *prove* anything, since you're talking with a limited number of people, but it's very valuable at uncovering insights that you can then test or prove. It's open-ended and often reveals things you didn't know.

- **Quantitative research** is about testing or proving something with a large sample size. Surveys and site traffic analysis (looking at your log files to see how users are moving around your site) are good examples. With hundreds or thousands of data points to analyze, you can look for statistically significant trends and be much more certain that your findings accurately reflect reality for all users. Quantitative research can help you test a hypothesis you uncover with qualitative research.

Here's another way to look at it: Quantitative research is better at telling you *what* is happening (e.g., your log files tell you that 35 percent of site visitors never reach a product page), and qualitative research is better at telling you *why* it's happening (e.g., usability testing shows that some users don't understand the terminology used in the site's navigation).

As you can imagine, both types of research can be critical. Qualitative research is an inexpensive way to gain new insights quickly and give some direction to your site. Quantitative research helps you determine if those insights are, in fact, real—and it gives you the evidence to prove it to your boss.

If qualitative versus quantitative is one axis for exploring user research methodologies, another is what people say versus what they actually do.

- **What people say** is important because it reveals their *goals* and their *attitudes*. Goals begin the conversation between your users and you, so you need to fully understand why they come to the site and what they're trying to do. Looking at site traffic logs may reveal problem spots, but if you don't know what users' goals are in the first place, you won't know what to improve or how to improve it. Attitudes, on the other hand, reveal how people perceive themselves and these online experiences, and it's important to understand these perceptions as well. Whether Francis the First-Time Home Buyer thinks of herself as ignorant or knowledgeable is going to affect how you communicate with her on the site. Interviews and surveys are very common methods for researching what people say, and for learning about goals and attitudes.

- **What people do** is just as important, since actual *behavior* can reveal more about people than what they say. Behavior reveals not only where they might be having problems (as in a usability test), but also how they tend to use Web sites in general, which can influence how you shape your site. For example, knowing how users scan a search results page is valuable information for ensuring that your search results page is useful and usable.

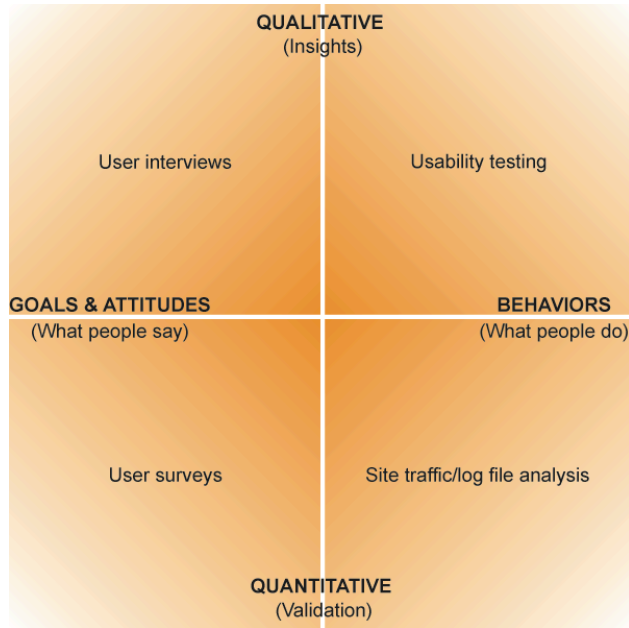The key point is this: *What people say isn't necessarily what they do*.

I've seen this again and again in usability tests. A user will clearly struggle trying to complete a task, and then claim the task was easy. Sometimes users are trying not to look stupid (guys often fall into this category), and sometimes users are trying too hard to tell me what they think I want to hear. But most often, users don't tell me the truth because they're not even aware of the truth of their behavior. It turns out that many people aren't very good at analyzing their own behavior or at paying attention to their actions. Basically, our image of ourselves doesn't always reflect reality. That's why it's important to watch what people really do, and not simply trust what they tell you they do.

My favorite story about this fundamental truth is somewhat famous in marketing circles. When Sony was introducing the boom box, the company gathered a group of potential consumers and held a focus group on what color the new product should be: black or yellow. After some discussion among the group of likely buyers, everyone agreed that consumers would better respond to yellow. After the session, the facilitator thanked the group, and then mentioned that, as a bonus, they were welcome to take a free boom box on the way out. There were two piles of boom boxes: yellow and black. (I bet you know where this is heading.) Every person took a black boom box.

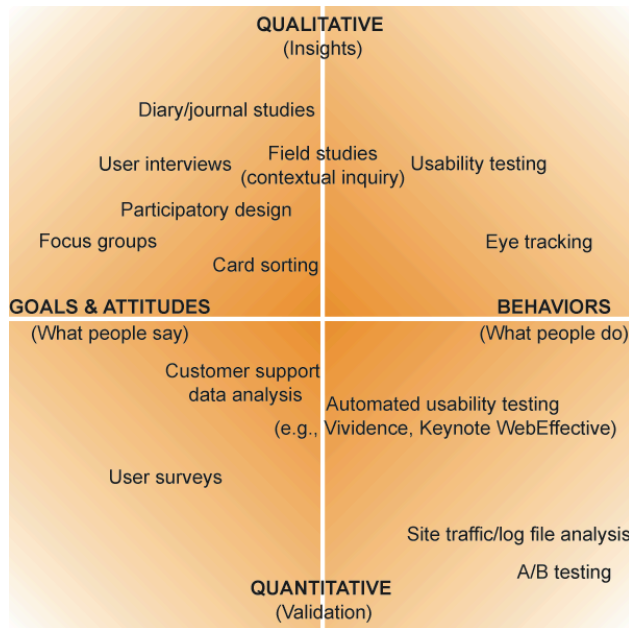What people say isn't necessarily what they do.

You need to make sure to understand both aspects, or you won't fully understand your users.

When you consider these two dimensions (qualitative versus quantitative research and what people say versus what they do), it becomes apparent how a few common user research methods support different research goals and complement one another. (I know I've spent too much time in the consulting business because I cannot resist creating a 2x2 matrix. It's like the siren's call.)

QUALITATIVE
(Insights)

User interviews

Usability testing

GOALS & ATTITUDES
(What people say)

BEHAVIORS
(What people do)

User surveys

Site traffic/log file analysis

QUANTITATIVE
(Validation)

User interviews uncover qualitative insights into users' goals and attitudes, and surveys are useful for testing and validating those insights. Usability testing reveals qualitative insights on user behavior, and site traffic analysis ensures that these behavior patterns are statistically real across the broader population.

I could add many other research methods to this framework, and of course, there is much crossover in what each method can be used for. The next diagram shows how various common methods might be positioned in this matrix. The next time you're planning user research, stop for a moment and make sure you know exactly what you're trying to find out and which research method is the best match.

QUALITATIVE
(Insights)

Diary/journal studies

User interviews

Field studies
(contextual inquiry)

Usability testing

Participatory design

Focus groups

Eye tracking

Card sorting

GOALS & ATTITUDES
(What people say)

BEHAVIORS
(What people do)

Customer support
data analysis

Automated usability testing
(e.g., Vividence, Keynote WebEffective)

User surveys

Site traffic/log file analysis

A/B testing

QUANTITATIVE
(Validation)

Now it's time to return to our regularly scheduled program: What are the three core approaches to

creating personas?

Qualitative Personas
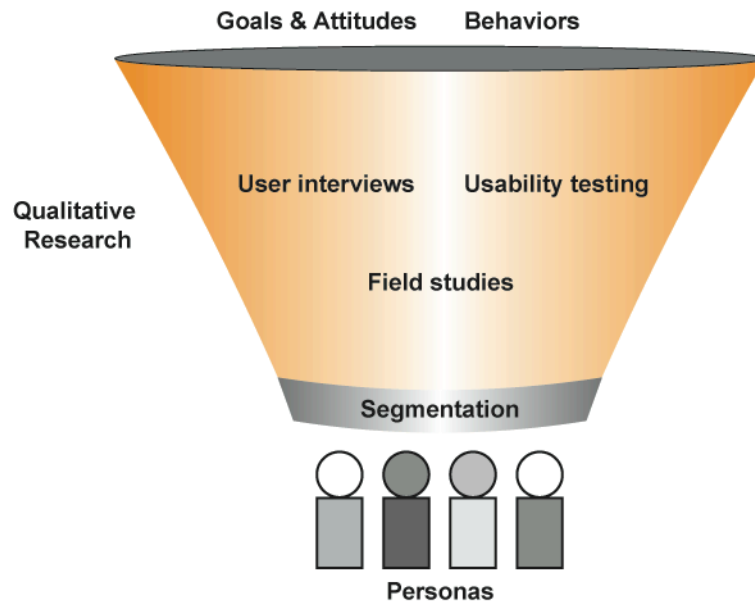
Let's start with the most traditional approach. Many organizations create personas following these simple steps:

1. Conduct qualitative research.

User interviews are the most common form of qualitative research, because talking one on one to 10 to 20 users is relatively easy for most companies. Some companies conduct field studies instead, where they observe users in their native environment (their office or their home), and can thus observe behavior while also asking about goals and attitudes. In addition, you can use usability testing to observe behavior, though this is less common for generating personas.

2. Segment users based on the qualitative research.

Segmentation is the art of taking many data points and creating groupings that can be described based on commonalities among each group's members. For personas, the goal is to find patterns that enable you to group similar people together into types of users. This segmentation is typically based on their goals, attitudes, and/or behaviors.

For qualitative personas, segmentation is, not surprisingly, a qualitative process. It's less about science and more about sitting in a room reviewing your notes and listening to your gut. For example, for a real estate site, you might interview users and then segment them based on overall goals: buy a house, find an apartment, sell a house, refinance a mortgage, and so on.

3. Create a persona for each segment.

Each type of user evolves into a persona as you add more detail to their goals, behaviors, and attitudes. Each one becomes realistic when you have supplied a name, a photo, demographic information, scenarios, and more.

These steps represent the most common approach to creating personas, and following them works successfully for a variety of organizations. It's a relatively quick process and an excellent way for companies to put their toes in the water and try out personas.

One of the first times I used personas was with a company that performed automated analyses of users' experience with a Web site. (I was and remain skeptical of the effectiveness of such a product, but maybe I'm just afraid of being replaced by a Hal 9000 computer.) As a young company testing new product ideas, this team interacted significantly with end users and already had a segmentation model in place. As part of redesigning their site and the reports they delivered, my team and I conducted interviews with existing and likely customers and found a new type of customer role that the company hadn't even considered. This new fifth segment, which soon became a fifth persona, resulted in important changes to the site and reports that increased overall customer satisfaction. Some quick qualitative research, made actionable through personas, made a difference to this business.

However, there are some drawbacks to this qualitative approach to building personas; so for each approach, let's look at the pros and cons.

Pros

- Relatively low effort is required. A basic process with, say, 15 user interviews, segmentation, and persona creation could take just three to four weeks. If an internal team creates the personas, costs are minimal, assuming you can find users to interview.
- Simpler persona stories increase understanding and buy-in. Because you're creating personas based on interviews and the stories you've heard, chances are your personas will be defined by just one or two attributes, making them relatively easy to understand at a glance. Simplicity breeds clarity, which stakeholders can more easily grasp and act upon.
- Fewer specialized skill sets are necessary compared to the other approaches. You still need people with interviewing skills and people who can find patterns and create personas out of interview findings, but you don't need any statistical analysis skills because you're not doing quantitative research.
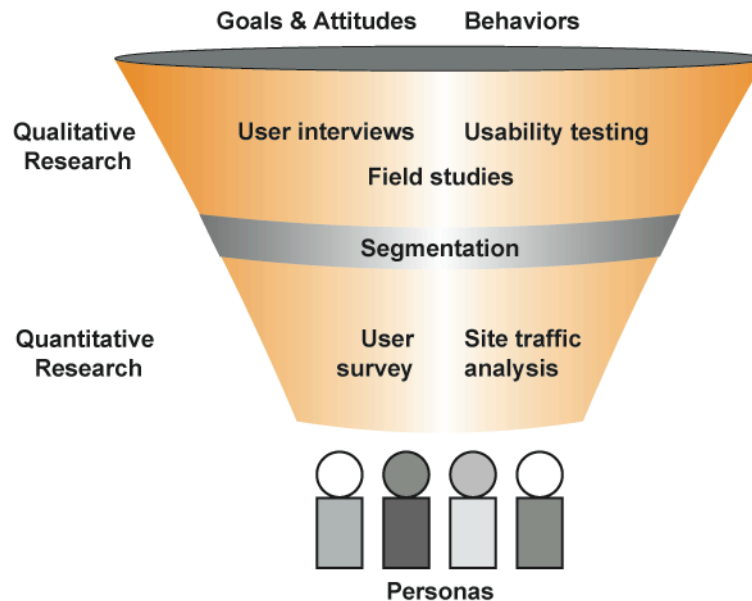
Cons

- There's no quantitative evidence. The most common pitfall in persona creation occurs when someone asks, "How can you be sure all our users are like the few you talked to?" Qualitative personas are based on the idea that you can talk to a small number of users and see patterns that apply to all of your users. In other words, the risk of being wrong is higher compared to having a large sample size to back you up. If your stakeholders need quantitative evidence to buy into your process, they'll disregard your personas as a creative but ultimately unreliable tool. Some people simply need the "proof" of hard data. And who can blame them? If you're going to be making critical business decisions based on these personas, you better be as certain as possible of their accuracy—and be able to persuade others.
- Existing assumptions don't tend to be questioned. You know your business, and you have assumptions about who the users are and what they need. When any person interviews users, he or she inevitably brings those assumptions to the research. The result: People find what they're looking for. Subconsciously, people look for the things that backup their own assumptions, so that instead of discovering surprises, they simply validate an existing worldview. Too often, their segmentation will look exactly like their original assumptions instead of being affected by the research. It won't always happen this way, but it's a serious risk when doing qualitative segmentation.

Use this approach if

- You can't invest much time and money in personas.
- Your stakeholders don't need to see quantitative data to believe and use your personas.
- The stakes aren't too high in terms of how you'll use the personas, so not having quantitative evidence is fine.
- You want to try using personas on a smaller project and see how they work before applying them to the larger business.

Qualitative Personas with Quantitative Validation

If you can invest a little more time and want to bring more quantitative objectivity to your personas, this approach is worth considering. Here are the basic steps:

1. Conduct qualitative research.

Just like the first approach, you begin by conducting qualitative research that reveals insights into user goals, behaviors, and attitudes.

2. Segment users based on the qualitative research.

You perform the same type of qualitative segmentation and end up with a number of segments based on particular user goals, behaviors, and/or attitudes.

3. Test the segmentation through quantitative research.

Here's the new step: Through a survey or other form of quantitative research, test your segmentation model using a larger sample size to be more certain it accurately reflects reality. The goals are to confirm that these segments are in fact different, and to have evidence to back up your personas in front of stakeholders. Survey data is best for testing goals and attitudes, and site traffic analysis (less commonly used for generating personas) gives you hard data on user behaviors. The analysis you do could involve simple cross-tabs, or you could use complex statistical analysis techniques.

For example, if you wanted to test a real estate segmentation model based on goals, you'd run a survey in which one of the questions covered the reason(s) users visit the site. You could perform a simple cross-tab analysis in Excel to examine how users' answers to this question affected their other answers (for example, did first-time home buyers rate certain features or content as more important?). If you see differences that support your plan to make these segments into separate personas, the segmentation model is successful. If you don't, you can try other ways of segmenting users and test those.

4. Create a persona for each segment.

When you have quantitative research for creating realistic personas, you can be more certain that decisions you make have statistical significance. Personas are no longer simply fictional creations,

but amalgams of research findings with evidence to back them up. Now when you say that Francis does something or wants something, you can back up that statement with hard numbers. Quantitative data doesn't necessarily *prove* your points, but it reduces the chance that you're wrong. It's also valuable for providing evidence of the magnitude and importance of a particular issue.

With this approach, you get a little more science and a little less art for your personas. The segmentation is still based on qualitative research, but you use quantitative research to obtain evidence to back up your decisions.

A couple of years ago I worked with a specialty retailer that wanted to better understand how its customers were using different channels to shop and make purchases. Given that the goal was channel usage, my team and I knew from the start that this was how we would segment our users into personas. So we ended up with four segments: store-only customers, Web site–only customers, combination Website and store customers, and combination Web and store and catalog customers. We conducted a survey and then tied in existing CRM data, so we knew which channels each survey respondent used. Then we analyzed the survey data by looking at all the questions' answers separated by these four segments. In this way, we could see at a glance how channel usage correlated to other aspects such as how much they spent, how they perceived the company, what they liked or disliked about shopping, how important certain factors like style were to each segment, and so on. The survey data confirmed that our channel-based segmentation was real and useful, so we created personas based on each of these four segments.

What follows are the advantages and disadvantages of this approach.

Pros

*   Quantitative evidence defends the personas. To the skeptical stakeholder, you can now answer that, yes, you are reasonably certain that these personas accurately represent the user base. Furthermore, you can rattle off data that defends your case, noting for example, that first-time home buyers make up 27 percent of site visitors, and that 82 percent of this segment believes that multimedia tutorials are important or very important.
*   Simpler persona stories increase understanding and buy-in. As with qualitative personas, you'll likely define them using just one or two attributes, rather than what advanced statistical analysis might reveal, which could be many factors working in combination. Simpler stories help people digest and use personas easier.
*   Depending on how you analyze the data, fewer specialized skill sets are necessary. You can perform simple cross-tab analysis of the survey results on your own. You can also dig deeper to validate the statistical significance of the data, which requires statistical analysis skills.


Cons

*   Additional effort is required. Due to the extra step, the overall process could take six to eight weeks. Surveys always add time, because you need to create the survey, deploy it, leave it active long enough to get sufficient completions, and then analyze the data. In addition, if your segmentation is based on multiple factors (e.g., the goal of home buying plus the attitude of anxiety plus the behavior of being ultra-organized), a simple cross-tab analysis of survey results won't be possible, and the level of effort (and skill set) could increase rapidly.
*   Existing assumptions don't tend to be questioned. A risk still exists in this approach, because you're still basing the initial segmentation on your own experiences and perspective. In other words, you could still be trying to prove what you already think you know, rather than opening up to new possibilities.
*   What if the data doesn't validate your theory? If the results of the survey or site traffic analysis don't support your segmentation model, then you're looking at additional work. In the best case, you can try different segmentation models one at a time to see if the data supports one of them. In the worst case, you have to start over and conduct a new survey or more site traffic analysis because you didn't ask all

the questions that are now needed to explore other options. Either way, it adds up to more time and money.

Use this approach if

- You have a bit more time and money to invest.
- Your stakeholders need to see quantitative data to believe and use your personas.
- You are very sure that your qualitative segmentation model is the right one.

Quantitative Personas
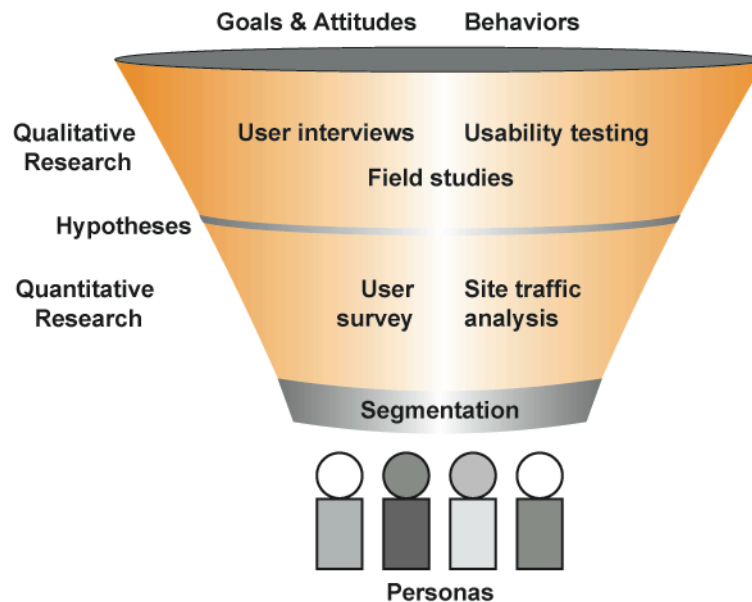
What if the process of creating personas could rely even a bit more on science? A third approach is emerging that I believe will become more and more common for creating personas. Instead of testing your qualitative assumption about one segmentation model, you use statistical analysis to test a variety of segmentation models at once, in order to find the model that is the most useful for creating personas.

1. Conduct qualitative research.

Once again, qualitative research reveals insights into user goals, behaviors, and attitudes.

2. Form hypotheses about segmentation options.

Rather than deciding on a final segmentation model right away, you use the qualitative research to come up with various potential ways you might segment the users. The goal is a list of a variety of candidates to analyze.

3. Gather data on segmentation options through quantitative research.

For each potential segmentation option, there are particular questions you need to ask in a survey or particular questions you need to answer using site traffic analysis. For example, if you think users' history with the site could be a way to segment, then a survey question about how long and how often they use the site would be appropriate. The quantitative research in this approach isn't trying to prove something; instead, its goal is to assemble more data for the next step.

4. Segment users based on statistical cluster analysis.

In this approach, statistical algorithms take a more active role in guiding you to a segmentation model, rather than just testing your existing assumptions. To oversimplify, you feed a set of variables into the machine, and it looks for naturally occurring clusters based on some set of commonalities. It tries many different ways of segmenting users, and through an iterative process, it finds a segmentation model that mathematically describes commonalities and differences. You could end up with any number of clusters and any number of attributes as key differentiators between the clusters. This process is a bit more complex, iterative, and still very much influenced by how you run it. But it's significantly different from the other approaches because the segmentation is data driven as well as human driven.

5. Create a persona for each segment.

When the cluster analysis spits out the segments, you take the data and make it real through the same process as before: by adding names, photos, and stories to turn these spreadsheets into real people.

As businesses rely on personas more and more for overall strategic decisions and marketing planning, quantitative personas will grow in popularity due to the scientific rigor they bring. The increased objectivity that the quantitative approach brings to the process of creating personas aligns much more closely with the data-driven decision making that occurs in so many organizations. Use of quantitative personas will also rise because the number of variables businesses have about users will only increase as research techniques continue to evolve. Machines are simply better than people at managing many variables at once.

One client I worked with (R.H. Donnelley, a yellow pages publisher) wanted to create a community-driven Web site in which consumer could provide ratings and reviews of local businesses. I knew from the start there were many different attributes that could drive segmentation of the consumers who would visit the site to read and post reviews. The user interviews conducted with a range of potential users led to a long list of potential segmentation options, which in turn led to an extensive survey for data gathering. An iterative process of cluster analysis on the survey results revealed that the best way to segment these consumers was based on a combination of attributes, including their usage of certain types of Web sites, their reliance on expert and consumer ratings and reviews, and their history contributing such ratings and reviews. Armed with this quantitative data, I created personas out of these segments, which made the users come to life for the company. The personas and the quantitative data behind them revealed many insights that guided decisions about the overall strategy, as well as feature prioritization and marketing plans.

Here are the pros and cons for the quantitative approach.

Pros

- Quantitative techniques, in combination with human smarts, generate the personas. Human bias influences this approach less than the others, since you have quantitative data not only to defend the personas you've created, but to test multiple segmentation models in the first place. To a skeptical stakeholder, this sophisticated process is likely to dispel any doubts he or she might have about personas as a valid decision-making tool.
- The iterative approach finds the best solution. Where the other approaches test one segmentation option, this approach iteratively finds the one segmentation model, among a great many possibilities, that you can best use to create personas. (You're still guiding the decisions about which segmentation model to ultimately use, but you have more data and more options for doing so.) It can reveal surprising patterns in the data that no one else would otherwise have considered, and it can lead to better insights about users and better actions to satisfy those users. It helps teams think in new ways.
- More variables can be examined. Machines are better at juggling a ton of different factors at once and looking for patterns and differences that the human eye can't discover. Cluster analysis can show you things you didn't know to look for.

Cons

- This approach requires significant effort. This type of statistical analysis takes time and is iterative, so total project length can vary from seven to ten weeks. Like all the approaches, it's mostly a serial process, since each step uses the previous step as an input (interviews to hypotheses to survey to cluster analysis).
- You need more specialized skill sets. There's a good chance you'll need to bring in a statistical analyst for this approach—someone who can apply various analysis techniques to the data. It's an advanced set of skills that you'll learn to admire deeply!
- The results could introduce new ways of thinking that could be more complicated or uncomfortable to the business stakeholders. The segmentation model revealed through cluster analysis may not be what anyone expected. It might involve a combination of many factors and be difficult to summarize. If the new segmentation goes against existing assumptions and ways of doing business, it can be a tough sell, but at least you'll have evidence to back up the major changes you'll be proposing.

Use this approach if

- You have time and money to invest.
- Your stakeholders need to see quantitative data to believe and use your personas.
- You want to explore multiple segmentation models to find the right one.
- You believe your personas will be driven by multiple variables, but you're not sure which ones are most important.

Other chapters in *The User Is Always Right* delve into complete detail on these three approaches to creating personas. There are chapters on conducting qualitative user research, conducting quantitative research, generating persona segmentation, and making personas real, as well as chapters on using personas once you have them. Learn more about the book and check out the companion blog at www.PracticalPersonas.com.