

Using K-Mode Clustering to Identify Personas for Technology on the Trail

Navyaram V Kondur

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science
in
Computer Science

Scott McCrickard, Chair
Steve Harrison
Michael Horning

April 19, 2018
Blacksburg, Virginia

Keywords: Technology, Hiking, Trail, Technology on the Trail, Persona, K-Mode
Clustering, Descriptive Analysis, Human Computer Interaction

Copyright 2018, Navyaram V Kondur

Using K-Mode Clustering to Identify Personas for Technology on the Trail

Navyaram V Kondur

(ABSTRACT)

Hiking is a widely-used term used differently by different people, and technology has an increasing role in the hikes that people take. Given the tremendous growth in technology capabilities for fitness, navigation, and communication, the breadth of devices and applications has expanded. Use of technology differs based on not only individuals but also the kinds of tasks performed. This research seeks to understand the different perspectives of the hikers and the technology they carry with them on the trail through a survey, analysis, and persona creation. 40 self-described hikers participated in a survey that asked about demographic data, hiking preferences, and would-you-rather preferences asking participants to choose between envisioned hiking and technology scenarios. The collected data were clustered into 6 clusters using K-Mode clustering, and descriptive analysis identified unique characteristics for each cluster. Five personas were crafted using the identified characteristics based on the analysis, taking into account the correlation between hiking preference and age, preferred activities and motivation, and maximum expenses they spend on the tech they carry. These personas seek to act as a guide to those who purchase, design and use hiking related products for both commercial and research purposes.

Using K-Mode Clustering to Identify Personas for Technology on the Trail

Navyaram V Kondur

(GENERAL AUDIENCE ABSTRACT)

Hiking is a widely-used term used differently by different people, and technology has an increasing role in the hikes that people take. Given the tremendous growth in technology capabilities for fitness, navigation, and communication, the breadth of devices and applications has expanded. Use of technology differs based on not only individuals but also the kinds of tasks performed. This research seeks to understand the different perspectives of the hikers and the technology they carry with them on the trail through a survey, analysis, and persona creation. 40 self-described hikers participated in a survey that asked about demographic data, hiking preferences, and would-you-rather preferences asking participants to choose between envisioned hiking and technology scenarios. The collected data was clustered into 6 similar clusters, and analysis identified unique characteristics for each cluster. Five personas were crafted to represent different types of hikers. The personas were created using the identified characteristics based on the analysis, taking into account the correlation between hiking preference and age, preferred activities and motivation, and maximum expenses they spend on the tech they carry. These personas seek to act as a guide to those who purchase, design, and use hiking related products.

Dedication

I would like to dedicate my work to my family(especially my dad) who helped me through the most difficult time in my life and made it possible for me to stand here today, and to my God who gave me the strength to fight the toughest battles.

Acknowledgments

I'd like to thank my advisor, Dr. Scott McCrickard , for being my academic father and guiding me through my research. When I started working with Scott I had no clue about how to do research or write a thesis. He helped me come out of my shell and gain confidence to express my ideas and thoughts with my fellow researchers. Without Scott's help I wouldn't be here today. He is a great mentor, very approachable, and hard to catch for a meeting. He is very kind and helped me through my masters journey in both good and bad times. I would like to thank my friends in the lab: Lindah, Tim, and Shuo for showing me my potential when I couldn't see it myself. I would also like to thank my committee for giving me opportunity to defend my thesis. Additionally I would like to thank the following departments for providing me with the resources to complete my thesis: the Department of Computer Science, the Institute for Creativity, Arts, and Technology (ICAT), the Center for Human-Computer Interaction (CHCI), the Social Informatics Group (SI), and the Notification Systems Research Group. Finally, I would like to thank my friends and roommates who were kind and helpful, and made Virginia Tech my home.

Contents

List of Figures	ix
1 Introduction	1
2 Related Work	4
2.1 Motivation	4
2.2 Technology on the Trail Initiative	5
2.3 Defining technology and trail	10
2.3.1 Technology	10
2.3.2 Trails	10
2.4 Social Media and Technology on the Trail	14
3 Approach	18
3.1 Statistical analysis methods	22
3.1.1 Survey analysis methods	22
3.1.2 K-Mode clustering of categorical data	24
3.2 Persona creation	25
3.3 IRB Approval	28

4 Analysis	29
4.1 K-Mode Clustering	31
4.2 Individual Cluster Analysis	34
4.2.1 Cluster 1 Analysis	36
4.2.2 Cluster 2 Analysis	39
4.2.3 Cluster 3 Analysis	42
4.2.4 Cluster 4 Analysis	43
4.2.5 Cluster 5 Analysis	44
4.2.6 Cluster 6 Analysis	47
5 Personas	49
5.1 Persona 1	50
5.2 Persona 2	53
5.3 Persona 3	56
5.4 Persona 4	59
5.5 Persona 5	61
6 Conclusion & Future Work	64
6.1 Conclusion	64
6.2 Future Work	65
Bibliography	67

Appendices	72
Appendix A Pre Cluster Data Analysis	73
Appendix B Cluster Visualizations:	78
B.1 Individual Analysis Cluster 1:	78
B.2 Individual Analysis Cluster 2:	85
B.3 Individual Analysis Cluster 3:	90
B.4 Individual Analysis Cluster 4:	95
B.5 Individual Analysis Cluster 5:	101
Appendix C Questionnaire	105

List of Figures

2.1	Facebook post for fellow hikers, Used with Permission from Erick Medrano	15
2.2	Facebook post to the fellow hikers, Used with Permission from Cheryl French Whitmer	16
3.1	Technology on the Trail workshop poster	20
3.2	2017 ICAT Day: Data display	21
3.3	2017 ICAT Day: Participants taking the survey and discussing with each other	22
4.1	Plot showing 6 clusters	33
4.2	Plot showing data distribution between 6 clusters	34
4.3	User's Age vs Activity preference in cluster 1	36
4.4	Users Activities and Gadget preference in cluster 1	37
4.5	Users Age and Activities Correlation in cluster 2	39
4.6	Users Age and Activities correlation in cluster 2	40
4.7	Users Age and Motivation correlation in cluster 5	45
4.8	Users Age and Gadget correlation in cluster 5	46
5.1	Persona 1. The left column describes key persona characteristics. The middle and right columns show cluster data used to craft the persona, Images used in this persona is used from Free image website[6].	52

5.2	Persona 2. The left column describes key persona characteristics. The middle and right columns show cluster data used to craft the persona, Images used in this persona is used from Free image website[6].	55
5.3	Persona 3. The left column describes key persona characteristics. The middle and right columns show cluster data used to craft the persona, Images used in this persona is used from Free image website[6].	58
5.4	Persona 4. The left column describes key persona characteristics. The middle and right columns show cluster data used to craft the persona, Images used in this persona is used from Free image website[6].	60
5.5	Persona 5. The left column describes key persona characteristics. The middle and right columns show cluster data used to craft the persona, Images used in this persona is used from Free image website[6].	63
A.1	Gadgets Preference	73
A.2	Age vs Expense	74
A.3	Age vs Gadgets	74
A.4	Age vs Activity	75
A.5	Correlation between Gadgets vs Would you rather Question 8	75
A.6	Correlation between Gadgets vs Would you rather Question 9	76
A.7	Correlation between Gadgets vs Would you rather Question 10	76
A.8	Correlation between Gadget vs Would you rather Question 11	77
A.9	Correlation between Gadgets vs Would you rather Question 12	77

B.1	User Age and Expense correlation in cluster 1	79
B.2	User Age and Gadget representation in Cluster 1	79
B.3	Users Activities and Motivation preference in cluster 1	80
B.4	Users Activities and Expense representation in cluster 1	81
B.5	Users Activities and Hike frequency preference in cluster 1	82
B.6	Users Activities and Motivation preference in cluster 1	82
B.7	Users Activities and Gadget preference in cluster 1	83
B.8	Cluster 1 Would you rather Preference	84
B.9	Cluster 1 Would you rather Preference	85
B.10	User Age and Expense correlation in cluster 2	86
B.11	Users Activities & expense Preference in cluster 2	86
B.12	Users Activities and Frequency Preference in cluster 2	87
B.13	Users Activities and Motivation Preference in cluster 2	88
B.14	Users Activities and Gadgets Preference in cluster 2	88
B.15	Users preferred gadget compared with what gadget they buy weather to look fit in someone else photo or take a picture that goes viral	89
B.16	Users Age and Maximum Expenditure in Cluster 3	90
B.17	Users Age and Activities Preference in cluster 3	91
B.18	Users Age and Gadget Preference in cluster 3	92
B.19	Users Activities and Maximum Hike Expenditure in cluster 3	92

B.20 Users Activities and Hike Frequency Preference in cluster 3	93
B.21 Users Activities and Motivation to hike Preference in cluster 3	94
B.22 Users Activities and Gadget Preference in cluster 3	94
B.23 Users Age and Maximum Expenditure in Cluster 4	96
B.24 Users Age and Activities Preference in cluster 4	97
B.25 Users Age and Gadget Preference in cluster 4	98
B.26 Users Activities and Maximum Hike Expenditure in cluster 4	98
B.27 Users Activities and Hike Frequency Preference in cluster 4	99
B.28 Users Activities and Motivation to hike Preference in cluster 4	100
B.29 Users Activities and Gadget Preference in cluster 4	100
B.30 Users Age and Motivation to hike in cluster 5	102
B.31 Users Age and Gadget Preference in cluster 5	103
B.32 Would you rather preference in cluster 5	104

Chapter 1

Introduction

Hiking is a widely-used term used differently by different people, and technology has an increasing role in the hikes that people take. Given the tremendous growth in technology capabilities for fitness, navigation, and communication, the breadth of devices and applications has expanded. Use of technology differs based on not only individuals but also the kinds of tasks performed. Understanding the usefulness of technology on the trail is complicated by the varying perspectives of different hikers. There are hikers who think technology enhances the hiking experience, but at the same time, some of them think it ruins the purpose of hiking.

A blog posting by a long-distance thru-hiker posed the following query: “*I thru-hiked back in 1999, and cell phones were not as omnipresent in society back then, and nobody had one on the trail for sure. One guy did have this thing he would type his journals into, then hook it to a telephone mouthpiece to send it home, but that’s it. Nobody had digital cameras back then either. I’m wondering what it’s like on the trail now. I would be really upset if I saw people talking on cell phones or texting while on the trail, but I guess that people are doing it. What’s the scene like? Is there a backlash against texters, if there are any?*”^[3] It was a thought-provoking query, and the replies to this post were equally thoughtful. Some of them replied with comments like “*Yes, technologies are used by hikers, and they dont care as long as others are not disturbed*”^[3]. In contrast, others noted an unwelcome atmosphere through comments like “*here was an overnighter on the trail in the Smokies in a shelter*

talking on a cell phone, and it left a pretty bad taste in the thru-hikers mouths.”[3] These sorts of comments raise questions about differences in approaches and perspectives of hikers, and how they are aligned with hiker characteristics like long distance multi-week hikers vs weekend hikers, or older hikers vs younger hikers, or walkers vs runners.

The goal of this thesis is to understand differing approaches and perspectives of hikers, with a focus on technology use. To address this goal, this thesis seeks to identify hiker personas based on personal characteristics collected through surveys of self-described hikers, including their reactions to technology on the trail. Personas are a widely used tool in human-computer interaction that capture archetypical users. Personas can then be used in analysis, design, and evaluation to help target interface development toward realistic users. This thesis provides a starting point toward understanding what sort of technology uses exist on hiking trails for different personas.

What is technology? Any human-made object can be viewed as technology, so we cannot avoid technology when we are hiking. For the purposes of this thesis, we are particularly interested in digital technologies, which we refer to as gadgets, including cameras, smartphones and their apps, smartwatches, location devices, and fitness trackers.

What do we consider to be hikers? This research takes a broad perspective on hiking, allowing people to self-identify as hikers whether they walk paved trails for a hour or two, take long walks in the woods, go on weekend backpacking adventures, or seek to hike the entire Appalachian Trail. It is these differences in hiking goals that proved important in crafting personas.

To understand what kind of technology hikers carry, and why, and whether their choice of gadget correlates with factors like age, culture, geographical, or financial considerations, we conducted a pair of surveys taken by 40 participants at a pair of hiking-focused events.

K-Mode clustering was used to identify clusters of hikers with common traits among these 40 users, with further analysis revealing pairwise correlations in the data. This analysis was used to identify 5 personas that capture archetypical hikers that can be found on the trail.

The motivation for this thesis in part came from a cultural probe conducted by Sarah Grace Fields in her thesis [18] on technologies on the trail. The cultural probe sought to encourage participants to think deeply about their hiking and outdoor experiences. A cultural probe is an excellent design technique in which human perspectives are understood in their real world context. Cultural probe activities encourage people to perform tasks in their natural environment through artifacts like maps, camera or diary[22][42][21]. One technique used by Fields is questions in the “Would you rather” form that ask people to choose between two options that focus on hiking challenges often with a technology aspect. This work uses such questions as a focus point to craft hiker personas.

This thesis is organized as follows. Chapter 2 is Related Work, elaborating on the motivations and discussing the foundational origins of this thesis, the Technology on the Trail initiative at Virginia Tech, the Spring 2017 class projects from the Technology on the Trail Initiative, social media and its impact on technology on the trail, and other related work in the Technology on the Trail research area. Chapter 3 describes the approach of the thesis, where we discuss how data was collected, the motivation for choosing the statistical analysis, and the approach for creating the personas. Chapter 4 discusses results of the K-Mode Clustering algorithm, which divided the clustered data in to 6 clusters with Individual Cluster Analysis conducted to define cluster characteristics. Chapter 5 discuss the designing of personas, based on individual cluster characteristics derived from the previous chapter. Chapter 6 outlines the conclusions and future work.

Chapter 2

Related Work

This section is divided into four major parts. The motivation behind the current research will be discussed first, followed by the initiation of technology on the trail in Virginia Tech. Details on the current research surrounding technology and hiking, interactive technology to enhance the hiking experience, and the technology designed for hiking are discussed next. Last comes the influence of social media and how it effects technology on the trail.

2.1 Motivation

When the *Technology on the Trail* initiative was started as a research focus within Virgina Tech's Center for Human Computer Interaction, one of the first research efforts was conducted by Sarah Grace Fields, who served as Graduate Research Asssistant in support of the initiative. In her graduate Master's thesis, Fields noted that "Technology as it is made its way to every part of human life, and now people started to carry the Smartphones and GPS systems in to remote, natural locations, recreational hiking or national parks." However, according to her research many hikers go out on the trail to get away or limit themselves from the use of technology for the limited time frame[18].

The main goal of Fields' thesis is not to design or build new technologies; rather it is more about understanding how people seek to make technology and nature more harmonious in the context of hiking. According to her research, hikers have a wide range of different perspectives

about how technology should and shouldnt fit into outdoor settings, and she thinks designers should listen to these perspectives before creating new technologies. Her main aim is to create a dialogue with hikers that explores their usage of and opinions on technology for hiking. She created and conducted several activities such as scrapbook activity or scavenger hunt, with multiple participants where they can creatively engage with the researchers. In her methods, she followed a type of research called cultural probes in which researchers use creative prompts and activities to explore the culture of a target audience[22][42][21].

As she mentions, her research does not intend to design new technology and her goal is to find a way where nature and technology go hand in hand with respect to hiking. Her motive is to shift the perspective of designers so that they can design products that can cater to all sorts of hikers. Current research was initiated as a continuation of Fields' work, but the main objective is to create a small set of hiker characteristics based on the gadgets they prefer, how frequently they hike, what a hike is according to them, and their motivation to hike. From those characteristics, six personas were created so that designers working for commercial or research institutes producing or designing hiking gears can use these personas as a guideline to design hiking products based on the target audience.

2.2 Technology on the Trail Initiative

“Technology on the Trail” was initiated with the support from Center for Human Computer Interaction (CHCI), Virginia Tech’s Institute for Creativity, Arts and Technology (ICAT), and Virginia Tech’s Department of Computer Science. This initiative seek to explore the influences, both positive and negative, of technology when used on extended trail hikes and similar activities. With this initiative, they seeks to understand and develop ways that technology is used on trails and in trail-like settings, such as extended and multi-day hikes,

where different user goals and desires affect our behaviors and interaction with others.

This initiative coincided with a course of the same name, that included grad projects that engaged with different aspects of the theme, including:

- Project 1: Augmented Reality on the Trail Wallace Santos Lages worked on this research. The main idea behind this project is to design and see how it is like to use augmented reality on the trail. Global Positioning System (GPS) is the main technology used today to support navigation on the trail. However, GPS-based routes are either offline (based on a previous track log) or built upon sparse data available on the field (such as topographic or satellite images). If the hiker decides to modify the current route to include or avoid a new unmarked waypoint, they often rely on visual information to do so. In his project Wallace investigated how augmented reality can be used to support route creation on the trail. By combining absolute GPS location with egocentric distance, it can allow users to create routes and mark physically distant landmarks.
- Project 2: Collaborative Exploration of Hiking Blogs with Surface Technologies Shuo Niu worked on this project. The main idea behind this project is to see how people commonly blog or take notes when on extended hikes, resulting in large collections of minimally-organized notes for which it is difficult to generate comprehensive and in-depth understanding of the blogs. For example, a journalist, who wants to write about someone's interesting journey, needs to read the journey-related blogs, figure out interesting information and summarize it into different sections of the article. Due to the huge corpus size, identifying and connecting useful information is time-consuming and very hard to process.

His project uses surface technologies to support dynamic explorations of blog data:

a tabletop computer and a vertical large display. The multi-touch tabletop provides a large interactive space to read, annotate and group blogs, and the vertical display shows a connected word-cloud which reflects users interaction and helps draw hints on possible interesting knowledge. The surface connects the intuitive touch interaction with the machine learning technologies to help readers explore self-defined topics in the blog data. The system hides the complex machine learning algorithm from the user and provides an easy way to define, express and explore user-defined semantics by interaction and visualization.

- Project 3: How fast can I hike? Phillip Summers worked on this project. The main idea behind this project is “When you want to go hiking, it can be hard to tell from guidebooks or maps just how long it might take to hike a trail”. Trail ratings are subjective, like “difficult” or “moderate”. If a hiker can estimate how long it would take to hike a trail, it could be easier to plan the supplies that the hiker needs to bring along the way. For multi-day ventures, this could save hikers from going hungry or carrying an extra 3 pounds of oatmeal he’ll never eat. With technology being more prevalent in our daily lives, there is a vast amount of data with which we can understand our environment.

It is not uncommon for users to have their GPS location monitored while out on trails, or to even upload this information to public sites such as WikiLoc. This project goal is to use this timestamped GPS data and other data such as heart rate readings to create a visualization of the trail that is representative of its subjective experience. This undertaking involves a non-trivial amount of data cleaning and normalization of data between separate users. When larger datasets become available, these visualizations will seek to provide information for both the novice and experienced hikers about what to expect from a trail that proves to be more useful than a traditional difficulty rating.

- Project 4: Finding identities of the Triple Crown of Hiking: Abigail Bartolome worked on this project and the objective of this project is to understand the Triple Crown trail AT, CDT and PCT, which all are part of the long-distance hiking, but each trail has its own mission statement, and therefore, its own identity. Do the microblogs (tweets) of the members of each community reflect the trail's mission statement? What can hiker learn about the communities who associate with these communities? Over the last decade, social networking has played a major role in archiving global trends. Now that more people are contributing to archiving documents via microblogging, there is an extensive library of documents that reflect the languages of different communities and their sentiments on new trends.

She conducted topic analysis using Latent Dirichlet Allocation (LDA) to collections of tweets pertaining to the three trail communities of interest, so that we can see the difference in topic models between the three distinct communities (AT hikers, PCT hikers and CDT hikers). She presumed that the topics discussed within each community reflect the values of that community. She compared the most popularly tweeted topics of each community with their respective mission statements and values. This ethnographic study will demonstrate the role of computer science in sociolinguistic studies, and provide an example for finding sub-communities in larger communities, extracting their topic models from micro-blogs (e.g., tweets) and relating those topics to their expected values.

- Project 5: Hiking the Appalachian Trail with Technology: Timothy Stelter worked on this project and this project started more of an exploratory venture to grasp the relationship between technology and nature. Tim planned to hike 100 miles of the Appalachian Trail (AT) that will span between 8 - 11 days during Spring break. Tim took this challenge up along with his father. His main aim was to understand two

things. One, to understand the use of technology while being on the trail for an extended period of time. This may offer ideas for how we can design technology to be transparent enough to not interfere with nature but still have the benefit of technological advancements. Two, to collect various amounts of data such as timestamped GPS location, audio recordings, pictures and written accounts to help find interesting pieces of information about the 100 miles hiked on the AT. The technology used to collect data will also serve as an opportunity to gain valuable insight on how technology may or may not interfere with the experience of the hike in relation to nature.

- Project 6: “Would you rather” - Probing Tradeoffs with Technology in Hiking and Outdoors Settings: Navyaram Kondur and Jagathshree Iyer worked on this project. The main objective for this project was to collect opinions using Would you rather cultural probe questions to understand perceptions of humans towards technology in hiking and outdoor settings. This project was the first step for this thesis. A set of would-you-rather questions under various occasions were presented and from the data gathered from these events, correlations and patterns in the way people from different background answered these questions were identified. Although significant correlations were not obtained, this project acted as a pilot study to conduct a fully controlled probe in future. Statics of the survey conducted while the participants registered for Technology on the trail workshop 2017 were presented as well. The below poster was presented at the reception of the workshop. A very encouraging feedback was obtained from the workshop audiences. This was a huge motivation for the team to conduct further survey to understand the hikers’ perspectives on Technology on the Trail.

Ongoing research projects in a follow-up class focus on different gadgets such as GPS systems, smartphone, smartwatch, and also the impact of social media such as using Twitter while on the hike or writing a detailed blog in Facebook after the hike.

All these projects addresses the different topics related to the hikers community. This motivated us to pursue research in the field and understand that there are various types of technologies that are already infused in the hikers hiking experiences and there is still a delicate argument happening between people who use technology on the trail and bystanders. Even though the current research was borne from a class project, it has deeper analysis, and doesn't concentrate on any single technology, attempting to identify common tech issues across people that are found in the trail.

2.3 Defining technology and trail

2.3.1 Technology

Over the decades, technology became ubiquitous, mobile and compact which users can carry easily and it has blended so deep into human life to a point where we cannot see anything else around us. From the moment they are awake, people are enslaved by the technology. But it has made human life easy in some ways. It makes the world smaller connecting people all around the world. But the same connectivity makes it impossible for humans to be off-work. Lot of people tried to escape this constant messiness by moving into nature by backpacking.

2.3.2 Trails

Trails are the scapegoat for most Americans from their daily routine. Lots of people go to a trail in search of peace and to stay disconnected from the digital world for a short period. Earlier, hikers use to make it a point that they will not carry any kind of technology with them in order to keep that distance from the world. As the time passed, technology had made its way into trails. Lot of hikers are contemplating whether they should let it happen

or find a balance between tech and the trail. Lot of hikers like to take a smartphone with them which is multi-purpose. Technology has reached the remote locations in the nature and people who are going for the adventures are taking professional cameras and capturing the real beauty of the locations and sharing it with fellow-hikers to encourage to get into the nature. There are some contra-indicative posts as well, where people are questioning if social media is ruining hiking[5].

In “Stories of the Smartphone in Everyday Discourse: Conflict, Tension, and Instability”, Harmon and Melissa Mazmanian discuss the ways in which smartphones have become the manifestation of work and social ties in an average person’s life. Through media and general culture, a pervasive image exists of a multitasking master who is also distracted by and addicted to the smartphone. The opposite image also exists, depicting an authentic human who is truly present when interacting with others, but who also winds up being an out-of-touch luddite. Harmon’s work also addresses the balance between use and non-use. Although participants in her ethnographic studies reported feeling constantly connected to their devices, her actual observations were of sporadic periods of use and non-use[28]. This suggests that even while not physically using or paying attention to their smartphones, participants feel them weighing on their minds.

In an attempt to bring harmony between the human-computer interaction and human-nature relationships, Jones from Brigham Young University mentions that technology has the ability to enhance the outdoor activities such as hiking, without disturbing the hiking experience[9]. To be able to study that, his team conducted a survey around the country using Amazon Mechanical Turk and collected 1002 responses. With this data they identified clusters with the participant’s attitudes towards the technology on the trail. They used K-Means clustering algorithm for categorizing the data in two parts, 5 clusters of hikers and 5 clusters of technology preferences when hiking. Later they identified the correlations between prefer-

ences regarding hiking and technology. Finally they gave 5 hiking clusters and their potential characterizations.

The idea and the desired results are very similar from our research. However, our approaches and data that was collected are different.

Observed differences:

1. Questionnaires are designed differently. Jones' questions are structured into two types, Likert-scale and dichotomous, whereas in current research, they are sectioned into three types, single-answer multiple-choice questions, multiple-answer multiple-choice questions, and dichotomous. There were no open-ended questions in the current research.
2. Data collection process: We both conducted surveys. While the survey conducted for the current research was taken personally from the 40 participants who attended the 2017 ICAT Creativity and Innovation Day using digital forms (i-Pads), he used crowd-sourcing as a tool to conduct his survey.
3. Analysis: Unlike him, we used K-Mode clustering Algorithm with $K = 6$ for the complete data. Once the data is clustered, we analyzed each cluster using descriptive analysis to define the characteristics.
4. Results: By using 6 Cluster characteristics we created 5 Personas. These personas are built based on the participant age, hiking as activity preference, frequency, motivation, maximum expenditure, and technology they carry while on hike—comparing these individual components to their Would you rather answers.

Jones also discussed his philosophy on Technology on Trails in a position paper ‘What are the Principles that Guide the Use of Tech on Trails?’[4]. He divided it into two sections -

‘recreational time spent on trails is a good thing’, and ‘interactive computing has a role in encouraging, enabling and enhancing time on trails’. Time on trails in nature temporarily quiets the demands of modern life which can lead to feelings of gladness, humility and gratitude - gladness, in the sense of a positive sense of well-being, humility, in the sense of feeling awe in the presence of nature, and gratitude for the natural world in which we live. These feelings may increase a sense of stewardship for the environment.

Couple of other interesting studies were focused on novel computing applications for outdoor activities (sports). In ClimbAware[34], Kosmalla explored wearables for use in rock climbing. In HOBBIT[35], Posti enables hikers to maintain solitude, and in [17], the researcher explored the cycling applications. Cheon [13] conducted a study through interviews to determine the use of health tracking habits among college-aged students and surprisingly they held back to old school methods such as manually logging their diet and activities. Thus, it is not always true that young people use technology a lot. Cheon[39] gave a very different perspective towards designing technology.

Current research is inspired by attendees of the Technology on the Trail workshop: Ellie Harmon, Alan Dix, Norman Su and Allison Druin. Harmon is an HCI researcher who has completed two of the three triple crown hikes. Her dissertation describes what it means for a human to be disconnected or constantly connected in a world where connectivity is inevitable [28]. An important finding of this research was the notion of feeling constantly connected even if, apparently, an individual did not use internet-enabled devices so frequently. In fact, in one of her papers [27], Harmon described how individuals experience conflicts in their efforts to be constantly connected or disconnected. The reason we believe her work impacts our study is that she is a computer scientist who is also a self proclaimed hiker who has completed thru-hikes. Her motivation to hike has been different in both the occasions. She made use of a lot of technology during her Pacific Crest Trail hike, and her motivation for

her research comes from the question 'Do people go on hikes to disconnect themselves from real world?'.

Dix walked a thousand miles along the complete periphery of Wales to understand and experience the technology needs in the areas where even phone connectivity is rough. He also explored how the local communities perceived technology and how it will improve their standards of living. He has written a blog about this on his website [2]. Druin addresses children and their perception towards technology. Her research group explore the design choices of technology for children to both learn and play. Su interviewed hunters regarding fair chase and how the various hunting devices affect their notion of fair chase. Su also suggested that when designing technology, developers need to consider the cultural impact such as the fair chase practice. Su has also explored various different communities such as musicians, nomads to understand how technologies impact these communities [40] [41]. Current research is similar to this line of work, but this is more about the influence of technology in hiking communities.

2.4 Social Media and Technology on the Trail

The world is depending on social media for everything day in day out. People started to believe what they see in the social media as a news. Even though hikers are facing challenges due to other people taking technology on the trail for personal or entertainment purposes, the usefulness of social media during the trail makes the cons of social media on trail looks smaller. On the contrast, hikers' pre-hiking knowledge is also coming from the blogs they read or the Facebook hiking groups they follow.

The smartphone that hikers carry is a passive data collector. Posting photos and tweets while on a hike, discussing in online communities about the dangers like shut-downs or wild

animals walking in a particular section of the trail helps other hikers so they can avoid going there as shown in Figures (2.1 and 2.2). In facebook there is online community called SoCal Hikers which is open to hikers, outdoor enthusiasts, backpackers and weekend warriors. Their goal is to connect with others in the community and discover new places to hike and explore. It is sponsored by SoCalHiker.net with participation from many hiking and outdoor bloggers throughout the area. There are lot of online communities which are helping each other using smartphones and th internet.

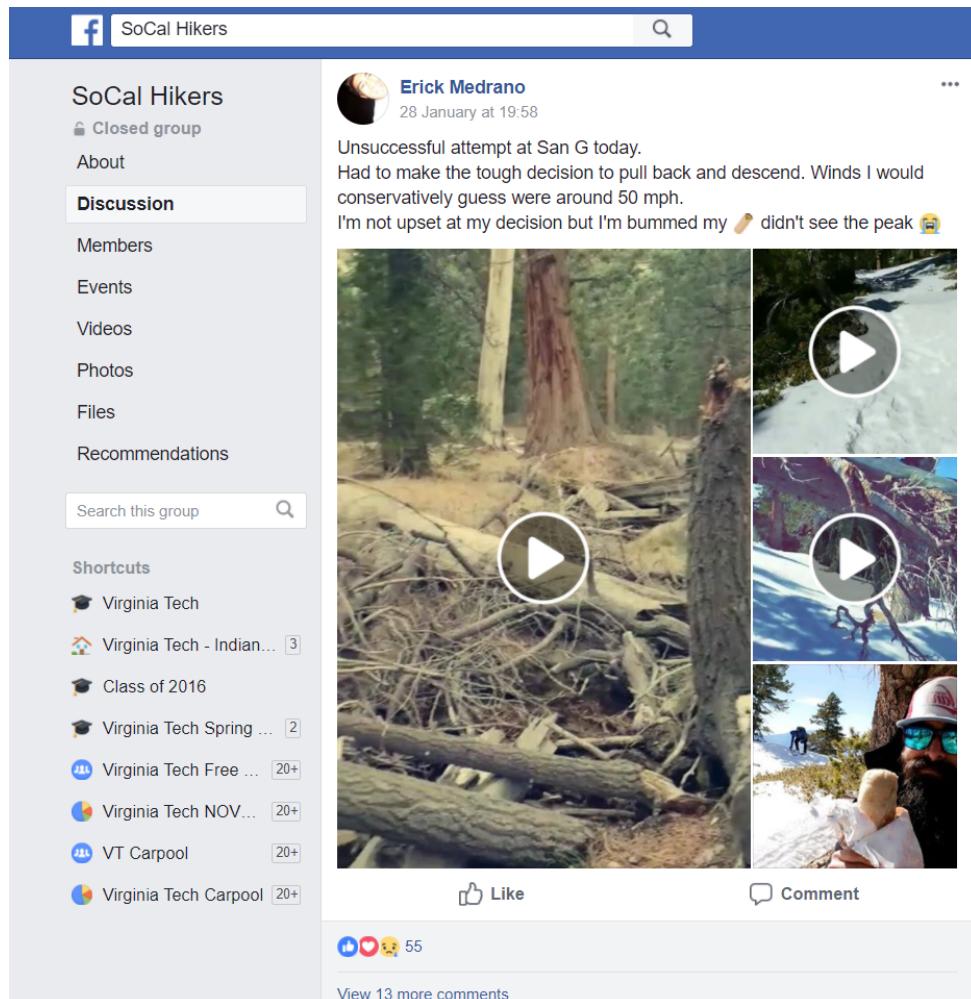


Figure 2.1: Facebook post for fellow hikers, Used with Permission from Erick Medrano



Figure 2.2: Facebook post to the fellow hikers, Used with Permission from Cheryl French Whitmer

In the recent past, Twitter is gaining momentum on the trail, and hikers have been actively sharing their hike experiences and pictures in short blogs. There are couple of ongoing research projects on tweets from trails. Bartolome[1] wrote a position paper in 2018 Group Conference “Exploring the cultural differences in Triple crown” in which she tries to address the questions - ‘What are the cultural differences between the trails?’,’Do the general community of hikers on each of these trails have different values?’. To answer these questions, she started analyzing the tweets that were collected in the Digital Library Research Laboratory of Virginia Tech from across the country relating to the Triple Crown trails, i.e.,

Appalachian Trail, the Pacific Crest Trail and the Continental Divide Trail, as well as tweets written by the conservancy organizations. She believes that tweets from the Appalachian Trail Conservancy would reflect the values of the Appalachian Trail community. She used Natural Language Processing techniques to understand what these communities care about from the tweets that are posted. She was able to identify multiple topics from the Triple Crown Trails that gave an overall understanding of cultural differences between different community preservers.

This current research identified 5 personas with specific characteristics. Of those 5, the majority of the personas preferred smartphone technology. The smartphones are very easy to carry and they are multi-purpose. Generally when hikers go on a hike, regardless of duration they prefer to travel light. So most of them prefer to take technology that does not weigh so much and which is multi-purpose. The smartphone is optimal choice for such hikers as evidenced in our observation.

Chapter 3

Approach

The goal of this thesis is to understand differing approaches and perspectives of hikers, with a focus on technology use. An important step in accomplishing this goal is to collect data about people's approaches and opinions regarding hiking. This chapter describes the methods used for collecting data to support understanding users. The first section describes the process of developing a questionnaire, initial data collection method. The second section describes the motivation for our statistical approach, compared to other methods. The survey method and different types of questions used, the categorical data, the importance of using K-Mode algorithms and the origin of our persona process are all described as well.

Initial data was collected from the registration event for 2017 Technology on the Trail workshop[4]. Workshop participants were asked to take a survey while registering. With the collected survey data, we generated some simple statistics and presented a poster to the guests at the workshop reception. The goal was to create awareness among the participants about *how other people answered to the same questions related to hiking* (poster shown in Figure 3.1). The second survey was conducted at the 2017 ICAT Day, a ‘showcase arts and technology’ event at Virginia Tech that included a Technology on the Trail section. At this venue all of our class members displayed their projects. Our team displayed the correlation graphs between attributes in large display screens. The people who were interested in the project topic and the results approached us for details, and after a brief discussion they were asked to take part in the survey. Participant participation is completely voluntary, and none

of the participants were asked to reveal their identifying details.

Surprisingly, a lot of participants were enthusiastic to see their answers and compared them with other responses for ‘would you rather..’ and ‘gadget preference’ questions. Some of them suggested to include a couple of questions if we are conducting a research further on this topic. Based on all the inputs gathered, we decided to continue the research. Hence we created the second survey questions with the combinations of some general hiking related questions and some would you rather questions.

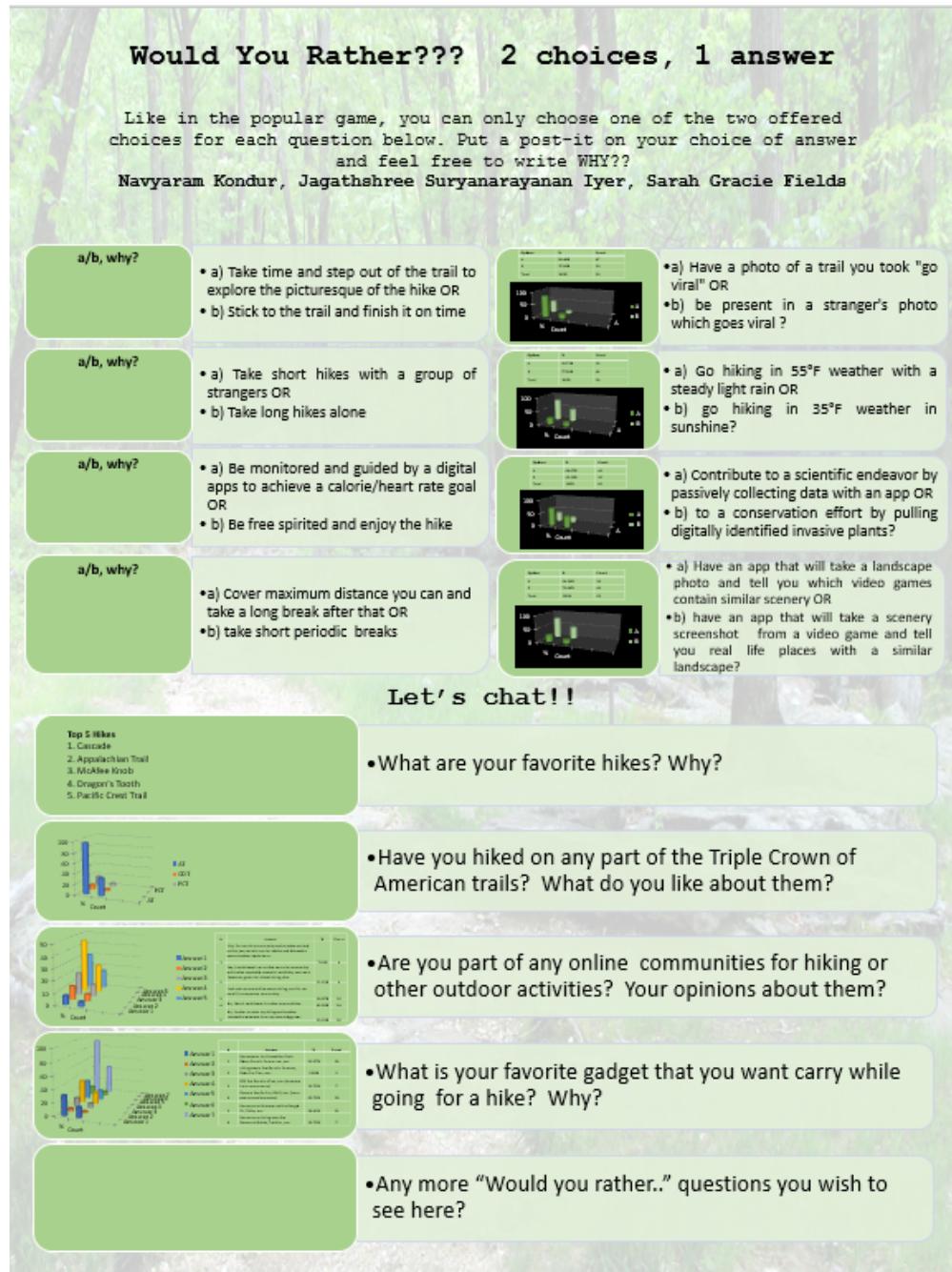


Figure 3.1: Technology on the Trail workshop poster

The questionnaire that was crafted was targeted to capture key, controversial concerns that we felt would allow us to understand distinctions between groups of hikers. The full ques-

tionnaire is in [Appendix C]. The first part of the questionnaire captures demographic information including age, experience, community involvement and technology use. It consists of 7 generic questions like user age, hiking habits and motivations, hike frequency, typical amount of money spent on a hike and preferred hiking gadgets.

The second part of the questionnaire asks people to make commitments between two exclusive ‘would you rather..’ options, as first featured in the thesis work of Gracie Fields [18]. Sometimes those are both somewhat undesirable, such as hiking in sunny but cold weather vs mild but rainy weather. Some address the tradeoffs in carrying (or leaving behind) technology on the trail. It consists of 5 would-you-rather questions; some focused on devices like ‘Buy a DSLR so that a picture you take on a hike goes viral? Or buy a smartwatch for fitness and look fit in a strangers image in a hike that goes viral?’; others focused on activities like ‘Tweet about an exciting spot in real time during your hike? Or write a detailed story in a hiking group on Facebook after finishing the hike?’.



Figure 3.2: 2017 ICAT Day: Data display



Figure 3.3: 2017 ICAT Day: Participants taking the survey and discussing with each other

3.1 Statistical analysis methods

This section describes the methods we employed for analyzing the results of our surveys. We describe our statistical analysis on the data, the need for categorical data, the selection of K-Mode clustering and the persona creation approaches.

3.1.1 Survey analysis methods

Surveys are information collection methods used to describe, compare or explain individual and societal knowledge, feelings, values, preferences and behavior. A survey can be a self-administered questionnaire that someone fills out alone or with assistance, or a survey can be an interview done in person or on the telephone. Some surveys are on paper or online, and the respondent can complete them privately at home or in a central location [19]. There

are different types of questions that can be asked on a survey questionnaire irrespective of whether it is online survey or in person.

A survey asks the respondents to answer questions, so the starting point in designing the survey instrument should always be the survey's purpose and objectives. These define the context of the survey and tell us what areas to cover and what types of information we need. In general, the purpose of a survey is wider in scope than the survey's objectives [33]. There are two types questions: open-ended and closed-ended. Open-ended questions give users the opportunity to write their own answers. Meanwhile, closed-ended questions usually consist of preset answers that the user has to choose from, which provide the researcher an easier way to analyze the data.

Types of closed-ended questions include:

1. Dichotomous: These types of closed-ended questions are binary questions that can be answered either in (yes or no) or (true or false).
2. Multiple Choice (single selection and multiple selection): These types of closed-ended questions are flexible and help the researcher obtain clean and easy-to-analyze data. These questions typically consist of multiple related options, allowing the user to indicate their preferences.
3. Rating: These types of closed-ended questions require the respondents to assign a fixed value in response, usually numeric. The number of scale points depends on what sort of questions a researcher is asking.
4. Likert Scale: These types of closed-ended questions are ideally five pointer or above scale questions where the respondent is required to complete the questionnaire that needs them to indicate the extent to which they agree or disagree.

5. Ranking: These types of closed-ended questions come with multiple options from which the respondent can choose based on their preference - from most-preferred to least-preferred (usually in bullet points).

In the current research, we used two types of closed-ended questions: multiple choice questions (both single selection and multiple selection) and dichotomous questions ('would you rather..'). Additionally, we asked the participants to write some comments about the survey at the end.

3.1.2 K-Mode clustering of categorical data

Most of the data that is collected in the real world is categorical data. A categorical variable has a measurement scale consisting of a set of unique categories [24]. For example ('Male', 'Female' for gender attributes) or ('house', 'condominium', 'apartment' for accommodation attributes). Each categorical attribute is represented with a set of unique categorical values like the examples given above. Categorical values are discrete and unordered, unlike numeric data, so the common clustering algorithms used for numeric data cannot be used for categorical data. The data collected through our survey is categorical – different age groups, preferred type of hike, gadget preference, etc. Huang from the University of Hong Kong discussed the origins of clustering algorithms for categorical data [43]. He explains how the K-Mode clustering algorithm is the most optimal method to use for categorical data. In the past decade there has been lot of research developing new algorithms for clustering categorical data [30] [31] [32]. The K-Mode algorithm is one of the first algorithms developed for clustering categorical data. K-Mode is an extension of the K-Means algorithm. K-Means is the most frequently used algorithm for clustering large datasets but it cannot be used for categorical data because of the Euclidean distance function and use of means to represent

cluster centers.

To understand whether K-Means clustering algorithm can be used on categorical data, Ralambondrainy converted each unique category to a dummy binary attribute and used 0 or 1 to indicate the categorical value either absent or present in a data record. For high dimensional categorical data this approach might not be ideal [37]. Huang described K-Modes as: “K-Modes modifies the standard K-Means process for clustering categorical data by replacing the Euclidean distance function with the simple matching dissimilarity measure, using modes to represent cluster centers and updating modes with the most frequent categorical values in each of iterations of the clustering process. These modifications guarantee that the clustering process converges to a local minimal result. Since the K-Means clustering process is essentially not changed, the efficiency of the clustering process is maintained.” [43]. There have been other research methods like the dynamic system approach[23], Cactus[20], ROCK[26] , Coolcat[11] and LIMBO[10]. However, these methods have largely stayed in research stage and not been widely applied to real world applications.

3.2 Persona creation

The designing process starts with understanding the customer: what he thinks, what he feels, and what he wants to do. To understand these questions, designers should first have a clear idea of their business goals. Based on the business goals, designers structure the market surveys. From the market survey, they collect data, initiate contextual analysis, and convert the raw data points into the different user work roles and requirements. Then they go through design process (alternatively called the UX life cycle): creation of a design perspective, ideation, identification of a persona, development of scenarios for the storyboard. In the past decade, personas have been a key aspect of the design process. Hartson of Virginia

Tech's Department of Computer Science gives an extensive design process in *The UX Book* [29].

Alan Cooper wrote *About Face* [15] and *Inmates Are Running the Asylum* [16] that touted the successes of personas in design. He created the goal-directed design methodology and pioneered the use of personas in it as practical interaction design tools to create high-tech products. He explained personas as an approach to “*develop a precise description of our user and what he wishes to accomplish. The sophistication comes from how we determine and use that precise description. We make up pretend users and design for them. We call these pretend users personas*” [16]. He describes how a simple concept like persona can be a powerful tool in the interaction design process. He believes understanding the end user and knowing what he really wants is the key factor for a designer to be able to design a product which can enhance user experience. The method may appear very trivial but it is very powerful and effective in every case.

Personas are created from a designer’s imagination but their characteristics are derived from actual users. As Cooper says, “They are hypothetical archetypes of actual users. Although they are imaginary, they are defined with significant rigor and precision” [14]. After the market research, the design team decides on the number of persona’s they are planning to create (persona goal). Those goals are based on the business goals and they may vary from 3 to 6 personas. In general, a persona creation team consists of product planners, usability engineers, interaction designers, market researchers and technical writers.

John Pruitt and Jonathan Grudin, researchers from Microsoft, examined the notion of personas with scientific rigor [36] [25]. As Pruitt and Grudin mentioned in their papers “*Personas as used by Cooper can be valuable, but they can be more powerful if used to complement, not replace, a full range of quantitative and qualitative usability methods. Personas amplify the effectiveness of other methods. Personas might be used by one designer to help focus.*

However, their greatest value is in providing a shared basis for communication. Cooper emphasizes communicating the design and its rationale among designers and their clients: It's easy to explain and justify design decisions when they're based on Persona goals....". They extended this using Personas to communicate a broader range of information to more people like designers, developers, testers, writers, managers, marketers and others. They claim that any data collected from the market research, ethnographic studies, instrumented prototypes, usability tests or any other source that relates to target users represented by the personas can be conveyed rapidly to all project participants. There are studies similar to Pruitt's and Grudin's research [12] [8] [38].

The primary objective of this research is to be able to distinguish the multiple personalities of hikers. These personalities can act as a base guideline to design a persona. Once the data is clustered into six groups and each group was analyzed based on the participants' age groups, preferences in terms of hiking as an activity, how frequently they hike, what motivates them to hike, the preferred gadget they carry and the hiking budgets. All of this analysis derived a set of characteristics for each cluster. By examining all these characteristics, five personas are presented. The bio information, quotations, and other information in the personas is imaginary. However, their unique characteristics are drawn from the analysis. For example, Persona 1 is middle aged man from Roanoke, working for a bank and is married with three kids. Due to his tight schedule and stressful work life he is not able to go for long distance hikes. So he prefers short or weekend backpacking where he can take his family for fishing, camping, etc. We created these bios like a scenario involving the user's emotional back-story, connecting the main characteristics identified in the process. For better understanding on primary personas, please refer chapter 5.

3.3 IRB Approval

This research leveraged a publicly available data set, generated as part of a Virginia Tech class. Since human subjects were involved, IRB approval was necessary. The IRB approval information can be found in Support documents, titled “Would you Rather” and “Data Analysis”. The approval number 18-203. The approval certificate is included in the Kon-dur_NV_T_Support2.

Chapter 4

Analysis

40 people successfully completed the survey and activities. The collected data was a combination of quantitative response to questions and qualitative feedback from discussions. This chapter highlights the analysis of the data, highlighting a few significant results but with a focus on a K-Mode clustering towards identifying key personas for technology on the trail.

The collected data is categorical in nature, the aim was to test the association of two categorical variables with a Chi-square test to find p-values (0.05). A Chi-square test is used to determine whether an relationship between two categorical variables in a sample is likely to reflect a real association between these 2 variables in the population. However, there was no significant correlation, perhaps because of a small sample size and the diversity of hiker types. Descriptive analysis was performed as a next step. By observing the data correlation where made on overall data using R programming, ggplot, and Bar-graphs. With this analysis we where able to define only 3 personalities, predominantly influenced by age groups. Here are few of those observations drawn from the first cycle of analysis. For more details look at Appendix A.

1. People who choose to carry 1 to 3 gadgets are in the age group <18 to 30 yrs; this shows that people in these age groups carry more devices while on a hike than the people who are more than 31. People in other groups carry gadgets far less frequently, suggesting that gadgets are mainly used by young people. This age group also tends

to spend more on the hiking gadgets.

2. Comparing the effect of age group on the mobile phone usage patterns for social networking while on a hike, it was evident that people who are less than 30 years old are more interested in tweeting while on a hike instead of going back and writing a blog about it. An interesting note is that people who are above 51 are not much interested in carrying a smartphone and tweeting real time on the spot.
3. People who are interested to collecting data have opted to carry multiple gadgets than the passive data collectors who want take a smartphone. Overall 67% of people are interested in bringing smartphone while collecting data passively.

With these analyses we were not able to obtain significant results. Hence, to identify multiple personas, first we decided to cluster the data, based on the cluster formations the data is further divided in to 6 groups and analyzed to identify the characteristics. The most common clustering algorithm is K-Means algorithm, generally K-Means algorithm is very efficient in clustering large datasets [7]. K-means works by selecting a small set of number called centroids, which are in the data space but not naturally in data points, and to work out which of the centroids each data points is closest then, it replaces each centroid with the new centroid in the crowd of the data points that were associated to it. We repeat this process multiple times until the centroids converge and forms the new clusters. In the process K-Means converts the data points in to vectors and start the clustering process. However, working only on numeric values constrain K-Means from being used to cluster real world data containing categorical values. The alternate solution for clustering categorical data is K-Mode clustering algorithm.

4.1 K-Mode Clustering

K-Mode clustering algorithm is built on the similar concept as K-Means. However, instead of dividing by the number of options each user selected in each question in the questionnaire and convert those selected options into vectors like in K-means, the K-modes algorithm simply records each question and then which answer got the most votes. This is the form of the responses to the most common answer, which is where the name K-modes comes from. So each centroid is in the same form as the original questionnaire data, like a set of responses to the different questions instead of multi-dimensional vectors[7]. Initially, to decide how many clusters we needed, we wanted to have more than three characteristics but at the same time we did not wanted to go over 10 clusters. As data set was small by creating the 10 clusters the data distribution was very unequal and there where many single data point clusters. Hence, we decided to try 6 clusters to check the decent amount of cluster distribution. We looked at 3 but there was too much variance within each cluster, so we looked at 6, which was a bit sparse for meaningful personas, so we cut back to 5 for the personas.

After multiple iterations in the current research we gave the maximum number of centroids as $K = 6$ and iterated the process 10 times. After the convergence the final clusters were formed and each time the cluster segregations where checked and iterated the program until we got the decent group segregations. Once the optimal clusters are formed, using descriptive analysis unique characteristics where identified.

For clustering the data, questions from 2 to 5 were used because they reflect participant interest in different outdoor activities including walking in a marked trails at urban areas or backwood unmarked trails, in rural areas, or jogging or biking Also it consists of demographical data like age. As mentioned previously the questionnaire was separated into 2 sections; first part is from 1 to 7 and second is 8 to 12. To run the algorithm we gave a range from

Q2 to Q5 in the data file. For further analysis, data was manually analyzed by observing each users choice of answer from Q1 to Q7 and conducted correlations for the clusters that were identified, thus leading to the creation of the personas.

Plot showing data clustered in to 6 clusters

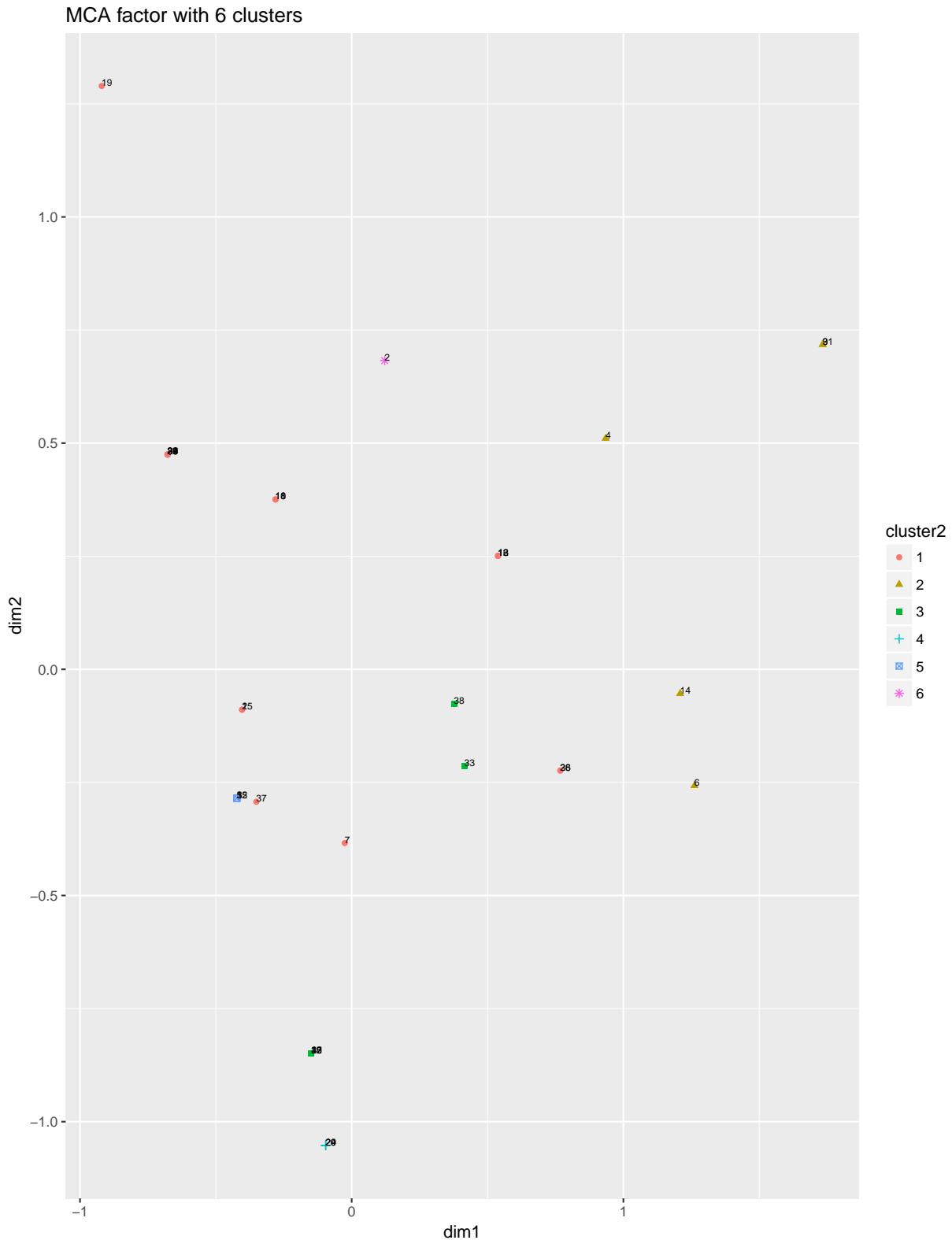


Figure 4.1: Plot showing 6 clusters

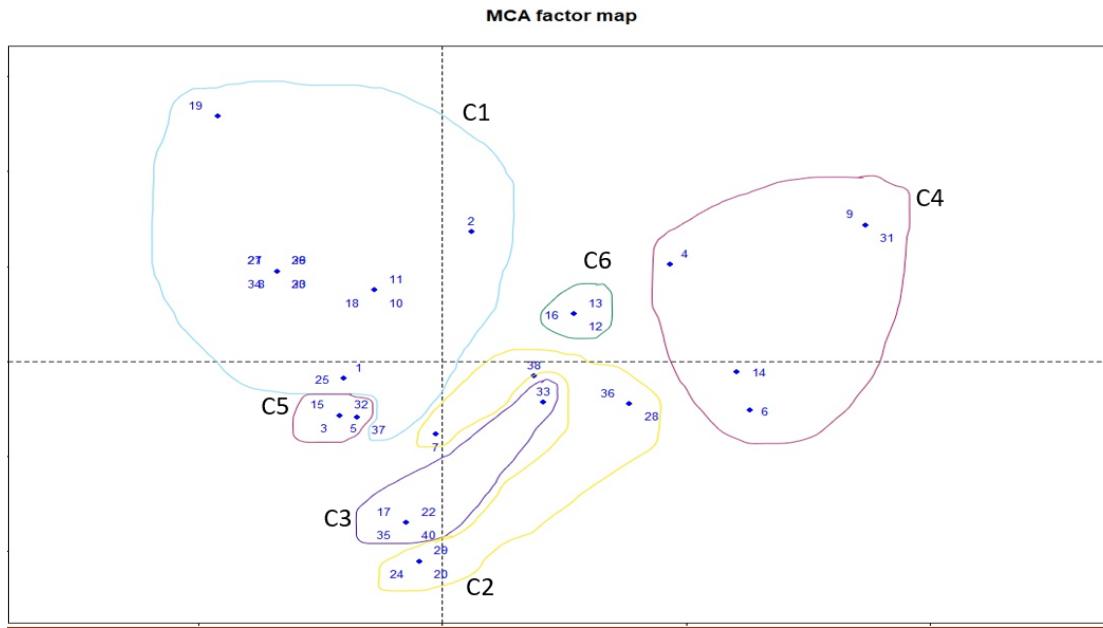


Figure 4.2: Plot showing data distribution between 6 clusters

4.2 Individual Cluster Analysis

In this section, we are discussing the analysis of individual clusters on each category, and at the end of each section, we discussed individual cluster characteristics. With the derived unique characteristics we created Persona in Chapter 5. Each cluster is analyzed into two parts: first, to give the background to the survey data we discussed raw data statistics. Second, correlation between categories are discussed. Detailed visualizations with brief descriptions are mentioned in [Appendix B]. All the correlations performed in the current research are briefed below.

1. Age with Expenses
2. Age with Activities

3. Age with Gadgets
4. Activities with Maximum expenses
5. Activities with Frequency
6. Activities with Motivation
7. Activities with Gadgets
8. comparing Gadget preference with Would you Rather question sections

We are only showing few of those correlations in the main document and the rest of the analysis and visualizations are in [Appendix B].

4.2.1 Cluster 1 Analysis

Age Vs Activities:

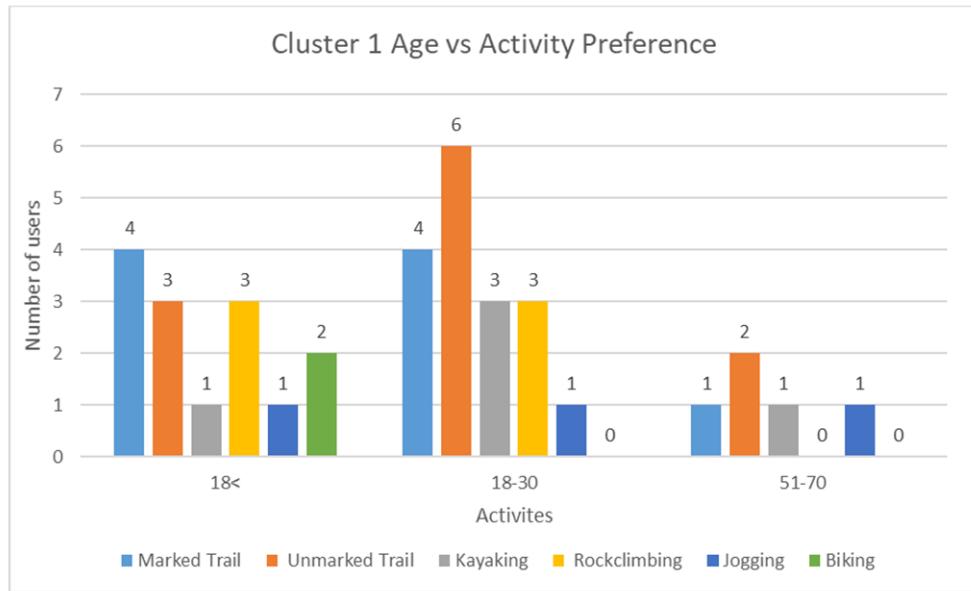


Figure 4.3: User's Age vs Activity preference in cluster 1

In this graph, we are projecting each user age and preferred activity in cluster 1. The X-axis represents the different age groups and each bar represents the activities users choose. As displayed in the graph in each age group we observed the most preferred activity in age group < 18 yrs are walking in the marked trail in urban area. Whereas, in older aged groups the preferred activity is walking in the backwoods in a rural area. Also as we can see younger generation liked to participant in most of the activities. This analysis is conducted to understand if the user activities are influenced based on the age group they belong. These are a very active group of people, irrespective of age differences all of them showed the eagerness in outdoor activities like kayaking, rock climbing, jogging, and biking. The uniqueness of

this graph is activity preference of the participant based in the age group they belong to.

Activities Vs Gadgets:

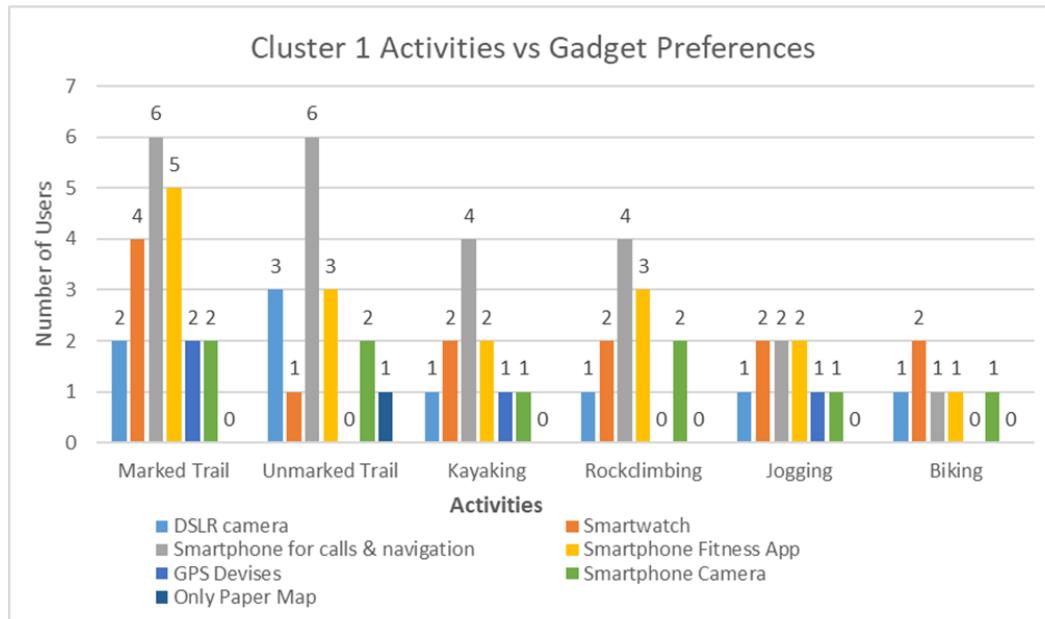


Figure 4.4: Users Activities and Gadget preference in cluster 1

In the above graph, we tried to understand the what are the participants choice of tech for each activity. The x-axis displays the different activities. In all the activities the most preferred gadget is smartphone for calls and navigations. This cluster participants are more active in the outdoor activities and the chose lot of fitness related gadgets like smartwatch and smartphone fitness app. People who likes to walk in backwood unmarked trails doesn't like to wear a smartwatch instead they like to carry multi-purpose smartphone. People who chose biking like to wear a smartwatch.

Cluster1 Characteristics

This is by far the largest group. 16 participants are in this cluster. The maximum number of participants are from age groups <18 and 18 to 30 yrs old majority is in between 18 to 30.

However, most participants who are between 50 to 70 years of age also fall in this group. The most preferred hike for this group is walking in backwoods, unmarked trails through rural areas and marked trails. Some of them choose activities like rock-climbing and Kayaking. This group likes to go for short hikes which takes 2 to 3 days at maximum, most of them said they like to hike once in a month or once in a week. To them the only motivation to hike is to stay way from the daily routine and spend time with family and friends.

As most of them preferred short hikes their maximum expenses are between <50 to 100 \$. Through they use multiple gadgets most of them choose smartphone for calling, navigation and fitness monitor apps. However, lot of them said they chose a DSLR camera over smartphone cameras for taking pictures. They choose to buy a DSLR camera to take quality pictures that may go viral. They hike as a recreational activity so they are not keen on collecting active data but they can use smartphone to takes pictures through apps as a passive data collection. As they all are smartphone users they like to tweet on the spot about a beautiful locations and use a off-line map in a smartphone over a paper map.

Based on these characteristics, we crafted Persona 1 as a middle aged man who is 58 yrs old who enjoys shorter hikes in the weekends. This cluster contained the oldest participants. As with other clusters, this persona enjoys urban hikes. Because most of they the users preferred to take smartphone as a hiking companion persona 1 takes only smartphone for calling and navigation purpose and DSLR camera for pictures. See Chapter 5 for the complete persona.

4.2.2 Cluster 2 Analysis

Cluster two is rather smaller group compared to cluster 1. This group consists of 7 participants.

Age Vs Activities

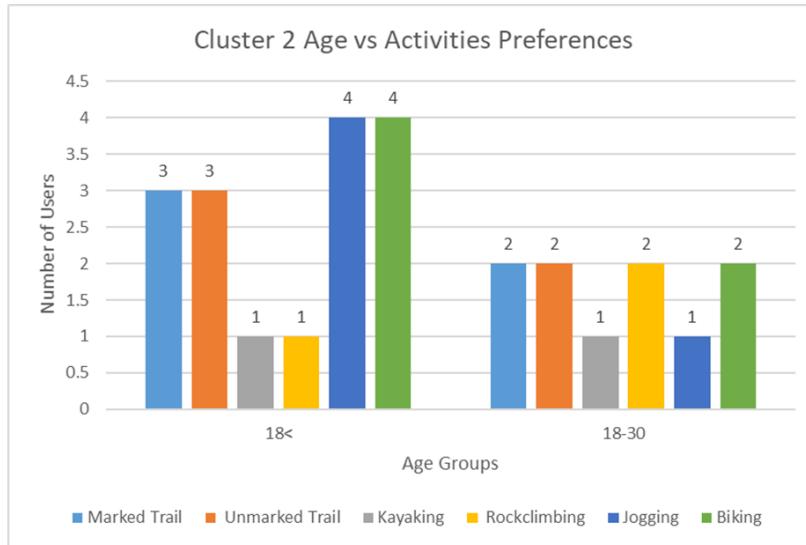


Figure 4.5: Users Age and Activities Correlation in cluster 2

In this graph, we are projecting the correlation between participants preferred activity with age in cluster 2. X-axis refers the age groups. This analysis is conducted to understand if the user activities are influenced based on the age group they belong to. Even though the age is split into 60/40 in this cluster most of them choose to go for biking through rural areas and the next most preferred activity is jogging outdoors through a rural area. The significant difference between two age groups that we observed is most of the users <18 havent opted rock climbing and kayaking. People who are above 18yrs have opted for multiple activities evenly. However, participants who are <18 have predominantly chose jogging and biking.

Age Vs Gadgets

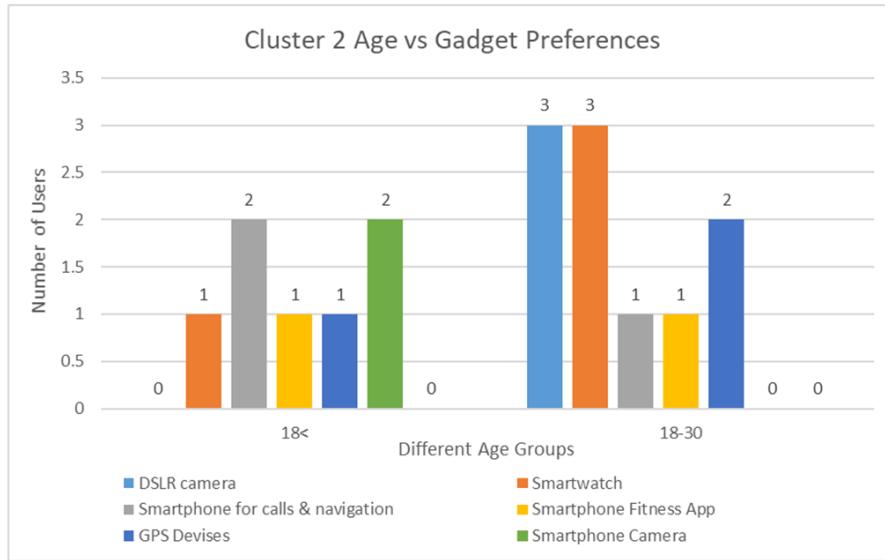


Figure 4.6: Users Age and Activities correlation in cluster 2

This graph represents the user ages and the tech they carry with them while on a hike. As mentioned before the first layer is age groups of the users and upper blocks are preferred gadget. We can identify a definite character here age has a powerful impact on the tech they carry. As we can see the participants who are < 18 yrs old prefer to take smartphone for calls, navigations and to take pictures. Participants who are older than 18yrs prefer to take DSLR camera and smart watch and a GPS devises for navigation instead of single smartphone. This group participants like to carry multiple gadgets while on hike not just the multipurpose gadget.

Cluster 2 Characteristics

7 participants are clustered in this cluster 2. This group of people is only between <18 and 18 to 30 years of age; they choose multiple activities as their preferred activities. 6 out of 7 people selected biking through the rural area as the favorite activity, 5 of them choose the following activities- walking in marked trails in urban areas, walking through backwoods,

unmarked trails through rural areas, jogging outdoors through rural areas. Few of them choose rockclimbing and Kayaking, The uniqueness of this group is most of them choose more than 3 options as their preferred activities. 2 out of 7 participants who are older than 18 years choose their hiking habit as thru-hikers, and 1 person said that he likes to go for a hike which takes 2 to 3 days. 3 of them said they are going like to go monthly once to the hike and couple of them noted 6 months once.

The reason to hike according to them is to stay away from the daily routine and explore nature. Also, they like to spend time with their friends and family. As a group they are spend a decent amount. Lot of them said they spend 50 to 100\$ while on a hike, specifically the person who said he like to go for a thru-hike will spend more than 1000\$ for his hike. The tech they like to carry with them is DSLR camera to take pictures and smartwatch to monitor their fitness and smartphone to call and navigation only.

They like to buy a smartwatch and get fit to look good in someone else's picture which goes viral. None of them like to be active data collectors. Hence they use a smartphone to take pictures and be passive data collectors. All of them like to tweet on the spot about an exciting location and let their friends know where they have been. This group of people is very active in physical activities, and they are young, and they dont mind spending money for the hike, and they like to enjoy the nature with their family and friends and share it in social media whenever they find.

Based on these characteristics, Persona 2 is a young high-school athlete who enjoys biking every day to his work and likes to collect data actively. He stays fit with inspiration from his smartwatch.

4.2.3 Cluster 3 Analysis

Cluster 3 is a rather small group compared to cluster 2. This group consists of 5 participants. Below we are only discussing about the characteristics of the cluster 3; for detailed analysis see [Appendix B].

Cluster 3 Characteristics:

5 participants are in cluster 3. 3 participants are in 18 to 30 yr group, and 2 are in 31 to 50 yr group. They tend to prefer to go for walking through marked trails through urban areas; few of them said they like to go on unmarked trails in rural areas. They not very keen on any other activities like rock climbing, kayaking, biking, or jogging. Cluster 3 participants hiking frequency is mostly once every 6 months. All of them are motivated to hike to stay away from the daily routine and explore nature and while doing that they want to spend time with their family and friends. As short-term hikers, they are fine to spend a lot of money; if they go for longer hikes they dont mind spending 100 to 500\$ on the tech they like to carry. The tech they want to carry with them includes smartphone for navigation and calls, DSLR camera for the taking pictures, and smartwatch to monitor their body fitness.

With these observations, we can understand that they dont mind carrying multiple gadgets, not just smartphone. 60% of participants prefer to buy a smartwatch for fitness and look fit in a stranger's picture in a hike that goes viral. They are not active data collectors, and they like to use the smartphone to take pictures with an app while on a hike as passive data collectors. 3 of them said they like to tweet about an exciting location on the spot, whereas the older people said they like to write entries in a Facebook hiking group. As they are multi-gadget people, they like to use a water-resistant offline map on a smartphone than a paper map. One of the participants commented on the paper map “**As a scout, and gadget freak, I'd love to have my gadgets, plus the backup of paper maps.**”

Based on these characteristics, Persona 3 was formed to be a 45 year old woman who lives in a big city and is only interested to go on hikes closer to home. She uses multiple gadgets for different purposes not only a smartphone user. See Chapter 5 for details about this persona.

4.2.4 Cluster 4 Analysis

Cluster 4 is a rather small group compared to the previous clusters. This group consists of 5 participants as well. Below we are only discussing about the characteristics of the cluster 4 and the detailed analysis see [Appendix B].

Cluster 4 is more demographically distributed regarding age groups from <18 yrs, 18 to 30 yrs and 31 to 50 yrs old. Irrespective of age difference all of them said their most preferred activity is to go for walking through backwoods, unmarked trails through rural areas and the second most preferred activity is jogging outdoors through rural areas. The most unique aspect of this cluster is that all of them said they would like to do thru-hiking that takes 3 to 6 months. A couple of them do weekly hikes currently, and a couple of them monthly, one person once every 6 months. The second unique aspect of this cluster is their motivation to hike; all of them chose the options: to stay away from the daily routine and explore nature, to spend time with their family and friends, and as part of fitness regime.

They do not lavishly spend money on hikes, but most of them said they spend 50 to 100\$ and one user said he spends more than 1000\$. Regarding gadgets also this cluster is unique because they all are very specific in choosing a smartphone as a most preferred tech on trial for both “calls and navigation” and “to take pictures no separate camera.” These observations show they are thru-hikers like to spend time on their own away from the world, but they don't like to carry multiple gadgets the Smartphone. Though they chose to take the only smartphone to take pictures while on a hike, in the would you rather question they

choose to buy a DSLR camera so that a picture they take can go viral. We think they select this option because in the would you rather section they were not given a choice of a smartphone to take pictures.

This group of people is very similar with each other in most of the categories none of the other clusters choose rainy weather and carry batteries for charging devices; this would you rather question was always a one-sided in all the cluster other than this cluster 4. As we said before if given a choice they are choosing a smartphone as the favorite gadget because in the question 10 all of them said they use a smartphone to collect data passively rather take multiple gadgets to locate poisonous plants. For question 11 as well all of them enormously said they like to write a detailed story in a hiking group on Facebook after finishing the hike instead of Tweeting while on a hike. This shows that this group is more on serious hiking and older than the previous clusters. Finally, they all said they like to use an offline map offered by a water-resistant smartphone app but has charging issues instead of a paper map.

Based on these characteristics, Persona 4 was formed to be a young graduate student who is very fitness oriented and primarily likes to go thru-hiking. The motivations for the choices is discussed in Chapter 5.

4.2.5 Cluster 5 Analysis

Figure 4.6 shows the correlation between participant age and motivation to hike. None of the participants are motivated to hike to collect the data. The participant <18 years old says they want to go for a hike as a fitness regime and to spend time with his family and friends but not to stay away from the daily routine. However, an older group member said they like to stay away from daily routine along with fitness. The interesting observation is that participant 15 the oldest member of the group said he only like to hike to stay way on

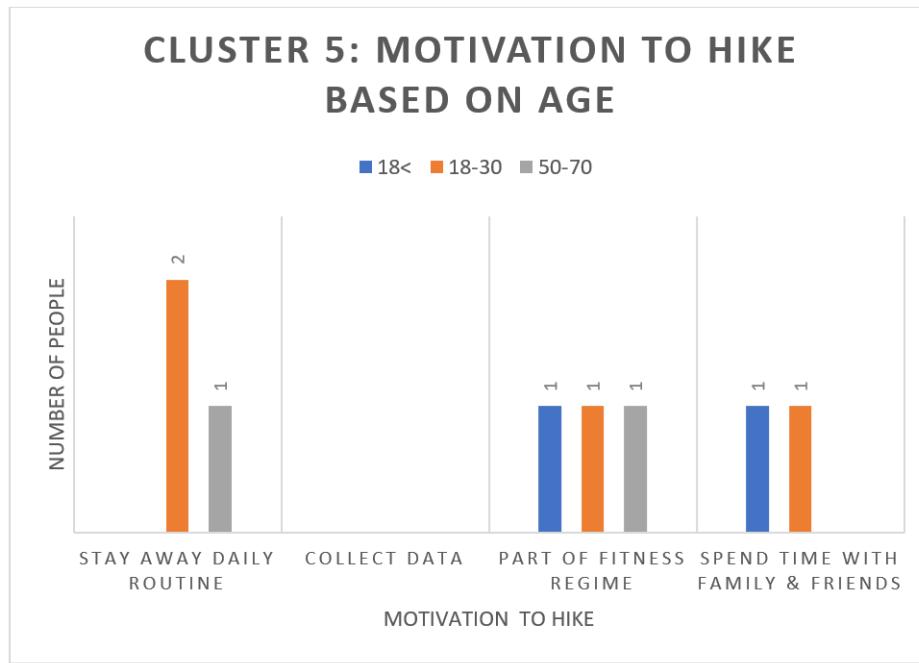


Figure 4.7: Users Age and Motivation correlation in cluster 5

a daily routine and fitness but not spending time with family and friends.

Figure 4.7 shows the correlation between the Age and Gadget preference in Cluster 5. The youngest member of this group most preferred gadget is smartwatch to monitor his fitness and smartphone to navigate and make calls. On the other hand, the older member in this group likes to take only DSLR camera as a tech with him nothing else. However, this shows the vast differences between different age groups and their gadget preference on this cluster.

Cluster 5 is demographically distributed regarding age groups from <18 yrs - 18 to 30 yrs and 51 to 70 yrs old but the 2 out of 4 participants are 18 to 30 yrs old. Irrespective of age difference all of them said their most preferred activity is to go for walking through backwoods, unmarked trails through rural areas. None of them gave their hiking preferences. 3 out of 4 chose once monthly and only 1 participant said they would hike once in a year. The uniqueness of this cluster is that they said their motivation to hike is to stay away from the daily routine and explore nature and a part of fitness regime.

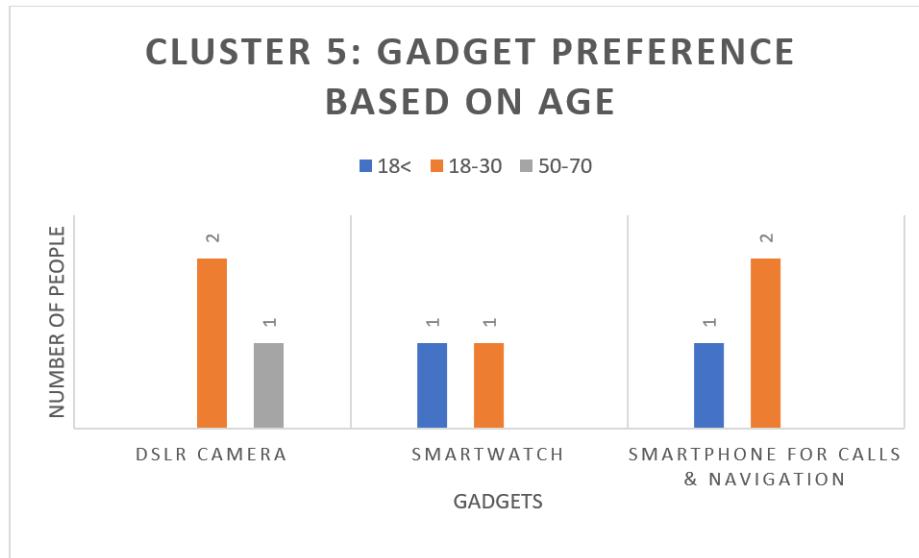


Figure 4.8: Users Age and Gadget correlation in cluster 5

This group of people is not high spending hikers, but most of them said they spend 50 to 100 \$ and one user said he spends more than 1000\$. Unlike the previous cluster, these participants do not want the only smartphone as a tech on the trail, their preference DSLR camera, and smartwatch and they like to use a smartphone only for navigation and calling. These observations show they are thru-hikers who like to spend time on their own away from the world, but they dont mind to carry multiple gadgets not only smartphone. In the previous clusters the people who choose the only smartphone to take pictures while on a hike choose to buy a DSLR so that a picture they take a hike goes viral, but this group selected separate camera which is not part of smartphone as tech on trail preferred option 1(buy a DSLR camera so that a picture they take can go viral).

Regarding the data collection, there was 50/50 preference: 2 of them said they are passive data collectors with a smartphone and 2 prefer active data collection using multiple gadgets. Surprisingly, participant 5 and 15, the youngest and the oldest, choose the smartphone passive data collection. When it comes to the social media and expressing the hiking experience, 2 of them said they like to tweet on the spot and 2 of them said they like to go home and

write a detailed blog about their hiking experience.

Clusters 5 and 6 were both small; they were combined into a single persona.

4.2.6 Cluster 6 Analysis

Cluster 6 is a group of 3 participants; all of them are in different age groups, but older group members like one are 18 to 30 yrs, second is 31 to 50 yrs, and third is 51 to 70 yrs. The most preferred activity is option 2 & 6 that is walking through backwoods, unmarked trails through rural areas and biking through rural areas. Other than the most preferred activities there are some other options they choose - one participant age 51 to 70 years old said he likes to go on rock climbing and another one said who is 31 to 50 yrs old like to go for jogging through a rural area.

Most of the participants said they prefer to go on a hike weekly once and one of them said once in 6 months. Their motivation to hike is for them to hike is to stay away from the daily routine and explore nature and fitness regime couple of them said they like to spend time with their family and friends. Their maximum amount of money to spend on the hike is between <50 \$, but one participant whose age is 51 to 70 yrs said he spends 1000\$. The gadgets preference is more diverse in this group; one participant said he likes to take only a paper map while on a hike and others are a mixture of DSLR camera, smartwatch and navigation and calls.

However, when it comes to the would you rather prefer their intention of tech on trails gets more clear. Most of them prefer to buy a DSLR camera so that a picture they take can go viral and for data collection, most of them are active data collectors with multiple gadgets like DSLR to take pictures and GPS device to locate poisonous plants and contribute data actively. After the hike, they like to go home and write a detailed story in hiking group on

facebook instead of tweeting on the spot. Other the person who choose only paper map rest the people decide to use the offline smartphone to navigate even though there is a charging issue.

Cluster 5 and 6 are smaller groups: cluster 5 consists of four members and cluster 6 consists of 3 members. We combined these two clusters to form persona 5, a 30 year old graduate student who is interested to go on a hike to collect data actively for her research. This persona is described in detail in Chapter 5.

Chapter 5

Personas

This chapter describes the five personas crafted for this thesis. The personas correspond to data from each of the six clusters described in the previous chapter, with the two smallest clusters combined into a single persona. As described in the related work, a persona does not always consist of characteristics most frequently found in a cluster (though it sometimes does). Rather, it seeks to leverage the survey data in a creative way toward crafting an interesting, compelling, and distinct set of representations, to be used in tasks like purchasing, development, and further research. Each persona in our set seeks to include unique characteristics that distinguish it from other personas. This allows a reader to understand differences between the personas.

The remainder of this chapter describes each persona in its own section. Each section begins with a narrative about the persona, followed by an explanation for why key characteristics were chosen for each persona. The explanations tend to be rooted in the statistics for the corresponding cluster; i.e., Persona 1 roughly corresponds to Cluster 1.

5.1 Persona 1

John Smith is 58 yrs old and lives in Roanoke VA. He works for a bank and, like many Americans, John works hard and spends a lot of time on his job works on the weekend to reach his goals. He lives with his wife and 3 kids. His parents live in Virginia countryside, being raised in the country John likes to hike but does not have enough time to go regularly. He is not fitness enthusiast, but his childhood experiences with his father going on a hike to hunt or to fish in the nearby trails make him go for a shorter hike whenever he gets time. Like his father, he likes to take his kids for walks through backwood areas and unmarked trails through a rural area. He has no desire to go for long-term thru-hikes, but he likes to go with his family on short hikes of 2 to 3 days. He acknowledges that perhaps in the distant future he will go on a thru-hike for 3 to 6 months.

For his short hikes he doesn't like to spend a lot of money, so he always keeps a budget of 100\$ for the family for a 3-day trip. He already owns hiking and camping gear. Whenever they all go on a hike, he likes to spend some quality time on the trail with his family and to stay away from daily pressure and routines. As a seasoned hiker he always makes it a point not to carry unnecessary weight in his backpack, but he prefers to take a smartphone for navigation, emergency calls, and fitness tracking apps.

However, he is keen on taking good pictures with DSLR camera of his family so that the photographs he takes in the most exotic locations goes viral. He goes hiking to enjoy nature; he may use some apps to collect data passively for research but not an active data collector. Being father of young boys, he is also interested in the current social media trends, and he tweets his experience on the spot, once he goes back to the daily routine he doesn't have much time to sit and write a detailed blog about the hike.

Key in the John Smith persona is that he is the oldest of the personas. Even though this

group consists younger participants we decided to create John to be 58 yrs because most of the participants who are in the age group from 50 to 70 belong to cluster 1. Hence giving priority to that we decided to go with 58. Most of the participants in this group said they like to go only for 2 to 3 day hike. However, they like to go fairly often, like once in a week or once in a month. In the analysis, it was clear that the motivation for all of them is to stay away from the daily routine and spend time with the family. That is the reason we thought John should be little older: instead of hiking with his friends he like to spend time with his family.

The below figure is sectioned into 2 parts. The left corner consists of persona descriptions like profile, personnel information, background, frustrations. Goals are formed based on characteristics and persona back story. From the activities section, the rest of them are derived from the unique characteristics observed in the previous chapter.

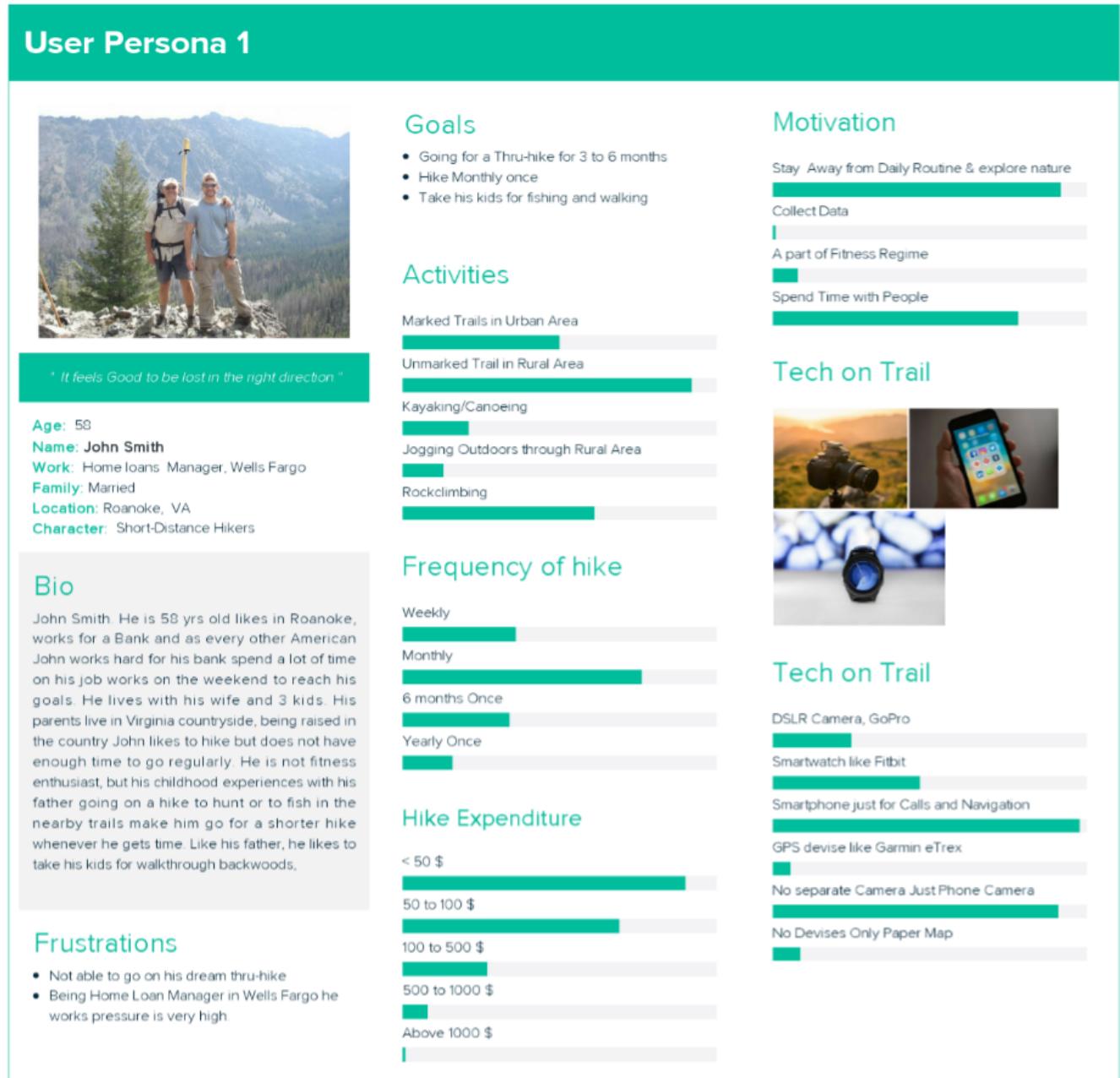


Figure 5.1: Persona 1. The left column describes key persona characteristics. The middle and right columns show cluster data used to craft the persona, Images used in this persona is used from Free image website[6].

5.2 Persona 2

William Knot is 16 yrs old, lives in Radford VA, and works in a restaurant part-time to save money for his college. He is in high school as a three-sport athlete. He likes to stay fit, so he likes to jog everyday morning, go rock-climbing once in a week, and bike to his job every day. He tries to stay fit and his major mission is to get into college with a recreational sports focus. Being an athlete, he likes to keep healthy, whenever he gets the time, he wants to go on short hikes frequently, at least monthly.

His motivation to hike is to stay away from daily routine and explore nature and spend time with his friends and family. For fitness, he likes to go for other activities like jogging, biking. The maximum amount he spends while on a hike is 50 to 100\$. As a fitness enthusiast William always wears a smartwatch and look fit in pictures when his friends take a picture. He likes to carry multiple gadgets while on a hike, he doesn't mind carrying weights while on a hike. Whenever he goes to very specific hiking locations he doesn't mind helping his friends in college to collect data passively with the pictures he takes but not active data collector.

William is very active in social media and likes to tweet on the spot whenever he finds some interesting locations instead of taking pictures and go home and write a detailed blog on Facebook hiking groups. He likes to spend some time collecting data while on a hike to help researchers for some extra money. Whenever he gets chance, he likes to tweet about his experience about the hike and doesn't want to go home and write a detailed blog about it.

Key in the William Knot persona is that he is the youngest. Even though the cluster is a mix of <18 and 18 to 30 groups we decided to create William to be 16 because most of the participants in this group are under 18. Being a younger crowd most of them said they participate in most of the activities from walking on marked and unmarked trails, Rock-climbing, biking and jogging. Because all of them are active in all of the activities in the

survey, we gave a athlete profile to this persona. Everyone in the cluster preferred to carry multiple gadgets and collect data actively for research purposes, so we gave a back story to William as a high-school athlete who is trying to save money for college tuition, so he will be interested to go that extra mile to earn some money.

The below figure is sectioned into 2 parts. The left corner consists of persona descriptions like profile, personnel information, background, frustrations. Goal are formed based on few characteristics and persona back story. From activities section rest of them are derived from the unique characteristics observed in the previous chapter.

User Persona 2



"It feels Good to be lost in the right direction."

Age: 16
Name: William Knot
Work: Part-time employee, Wendy's
Family: Single
Location: Radford, VA
Character: Fitness Enthusiast

Bio

William Knot. He is 16 yrs old lives in Radford, works in a restaurant part-time to save money for his college. He is high school student. He is part of high school athlete, and he likes to stay fit, so he likes to jog everyday morning. Rock-climbing once in a week and bike to his job every day. He tries to stay fit and his major mission is to get into college with sports quota. Being an athlete, he likes to keep healthy, whenever he gets the time, he wants to go for Rock climbing and Kayaking and short hikes preferably monthly once.

Frustrations

- Not able to go on his dream thru-hike
- Being Home Loan Manager in Wells Fargo he works pressure is very high.

Goals

- Stay fit
- Hike Monthly once
- Save money for College
- Get into College with sports Quota

Activities

Activity	Percentage
Marked Trails in Urban Area	Low
Rock Climbing	Very Low
Unmarked Trail in Rural Area	Medium
Jogging Outdoors through Rural Area	Medium
Biking Through Rural Area	High

Motivation

Motivation	Percentage
Stay Away from Daily Routine & explore nature	High
Collect Data	Medium
A part of Fitness Regime	Medium
Spend Time with People	Medium

Frequency of hike

Frequency	Percentage
Weekly	Very Low
Monthly	High
6 months Once	Medium
Yearly Once	Very Low

Hike Expenditure

Expenditure	Percentage
< 50 \$	High
50 to 100 \$	Medium
100 to 500 \$	Medium
500 to 1000 \$	Very Low
Above 1000 \$	Medium

Tech on Trail



Tech on Trail

Tech	Percentage
DSLR Camera, GoPro	Medium
Smartwatch like Fitbit	High
Smartphone just for Calls and Navigation	Medium
GPS devise like Garmin eTrex	Medium
No separate Camera Just Phone Camera	Very Low
Smartphone for Fitness	Medium

Figure 5.2: Persona 2. The left column describes key persona characteristics. The middle and right columns show cluster data used to craft the persona, Images used in this persona is used from Free image website[6].

5.3 Persona 3

Jane Undergound is 45 yrs old lives in DC, works with a government organization in Washington DC. She is married with 2 kids; the eldest son is in high school and younger one in middle school. She is born and raised in Blowing Rock NC. Her father is a retired Park Ranger on the Blue Ridge Parkway. Since her childhood, living close to nature she always loved to go beautiful locations and take great pictures. Though she is not in extreme outdoor activities Jane likes to wear a smartwatch to monitor her fitness and health.

Even after having kids she never stopped going for short hikes for 2 to 3 days, but because of her job and personal responsibilities she is not able to go for longer hikes. So being in the city, she can only go for marked hikes which are close to home, and she tries to take her family with her at least once in 6 months and leave behind all her work pressures and daily routines. For short hikes, she plans a budget of 100\$ for each person in the family.

She likes to carry a DSLR camera with her to take beautiful pictures of natures, she is a very active blogger, and she is part of a couple of hiking groups on Facebook. Before every hike, she talks to her fellow hikes in the group for suggestions and tips. She always keeps her phone on hand for emergency calls and navigation. As a seasoned hiker, she always keeps a paper map with her just in case all the tech dies out of battery for some reason, and she can safely bring back her family to the civilization.

Key in the Jane's persona is that she is older persona and not keen on going for other outdoor activities. She likes to hike only on the marked trails in the urban area. When we gave a back-story that she lives in Washington DC and likes to do short hikes closer to the city. However, as derived in the cluster characteristics all of these group members are interested in buying multiple gadgets and they can spend money for it, so we said the maximum amount of money she spend for short-hikes are 100\$ per person in the family.

The below figure is sectioned into 2 parts. The left corner consists of persona descriptions like profile, personnel information, background, frustrations. Goal are formed based on few characteristics and persona back story. From activities section rest of them are derived from the unique characteristics observed in the previous chapter.

User Persona 3



"It feels Good to be lost in the right direction."

Age: 45
Name: Jane Underground
Work: Government Organization
Family: Married
Location: Washington DC, The District of Columbia
Character: Thru-Hiker

Bio

Jane Austin is 45 yrs old lives in DC, works with a Government Organization. She is married with 2 kids, the eldest son is in high school and younger one in middle school. She is born and brought up from Blowing Rock North Carolina; her father is retired Park Ranger in Blue Ridge Parkway. Since her childhood, she has been an active hiker, and she is thru-hiked 2000 mile Appalachian crust trail when she was a college graduate.

Frustrations

- No time to go for hikes regularly
- Work Pressure

Goals

- Stay close to nature
- Go back to Thru-hike once again
- Take her kids for short hikes often

Activities

Activity	Frequency
Marked Trails in Urban Area	High
Unmarked Trail in Rural Area	Medium
Rock Climbing	Low
Jogging Outdoors through Rural Area	Low
Biking Through Rural Area	Low

Motivation

Motivation	Strength
Stay Away from Daily Routine & explore nature	Very High
Collect Data	Low
A part of Fitness Regime	Medium
Spend Time with People	Medium

Brands & Influencers



Frequency of hike

Frequency	Strength
Weekly	Low
Monthly	Medium
6 months Once	High
Yearly Once	Low

Hike Expenditure

Expenditure Range	Strength
< 50 \$	Very High
50 to 100 \$	Low
100 to 500 \$	Medium
500 to 1000 \$	Medium
Above 1000 \$	Low

Tech on Trail

Device	Strength
DSLR Camera, GoPro	Medium
Smartwatch like Fitbit	Low
Smartphone just for Calls and Navigation	Very High
GPS devise like Garmin eTrex	Medium
No separate Camera Just Phone Camera	Medium
No Devise Only Paper Map	Low

Figure 5.3: Persona 3. The left column describes key persona characteristics. The middle and right columns show cluster data used to craft the persona, Images used in this persona is used from Free image website[6].

5.4 Persona 4

Henry Young is 24 yrs old and lives in Blacksburg VA. He is a 1st-year Ph.D. student at Virginia Tech. He is an international student from South Africa. He is born and raised in Cape Town. Living in a city all through his life he is a fitness enthusiast, but from a lack of exposure he doesn't like outdoor activities like rock climbing, trail running, and kayaking. When he came to Blacksburg, he always wanted to go on thru-hike for the sense of accomplishment it would provide, and last summer he went for the thru-hike on the Appalachian Trail with friends he met in the Venture Out outdoors organization for 4 months. Since his first thru-hike, he decided to hike once in a 6 months for at least ten days long trips, including on isolated trails and even unmarked areas. For all short hikes he spends 50 to 100\$, and for the thru-hike, he spends 1000\$ maximum. Being a student, he can't afford to spend more than 1000\$ on an extended hike.

The primary motivation is for him to hike to stay away from his class work and research pressures through the semester. While on a hike he doesn't like to carry multiple gadgets, his only preferred Tech is Smartphone trial for both "calls and navigation" and "to take pictures no separate camera." Henry likes to write, and he maintains a blog about his travels and about his research and whenever he goes for a hike he likes to write a detailed blog and share it on facebook hiking communities.

Key in the Henry's persona is that he is a thru-hiker. Unlike other personas, Henry is very keen on big adventures, including bushwacking in unmarked trails. He might not lavishly spend a lot of money on the hikes but for the thru-hikes he can spend up to 1000\$. That is why we gave a back-story that he is an international PhD Student and he likes to take only smartphone with him for a hike because being a student he won't be able to afford it.

The below figure is sectioned into 2 parts, the left corner consists of persona descriptions

like profile, personnel information, background, frustrations. Goal are formed based on few characteristics and persona back story. From activities section rest of them are derived from the unique characteristics observed in the previous chapter.

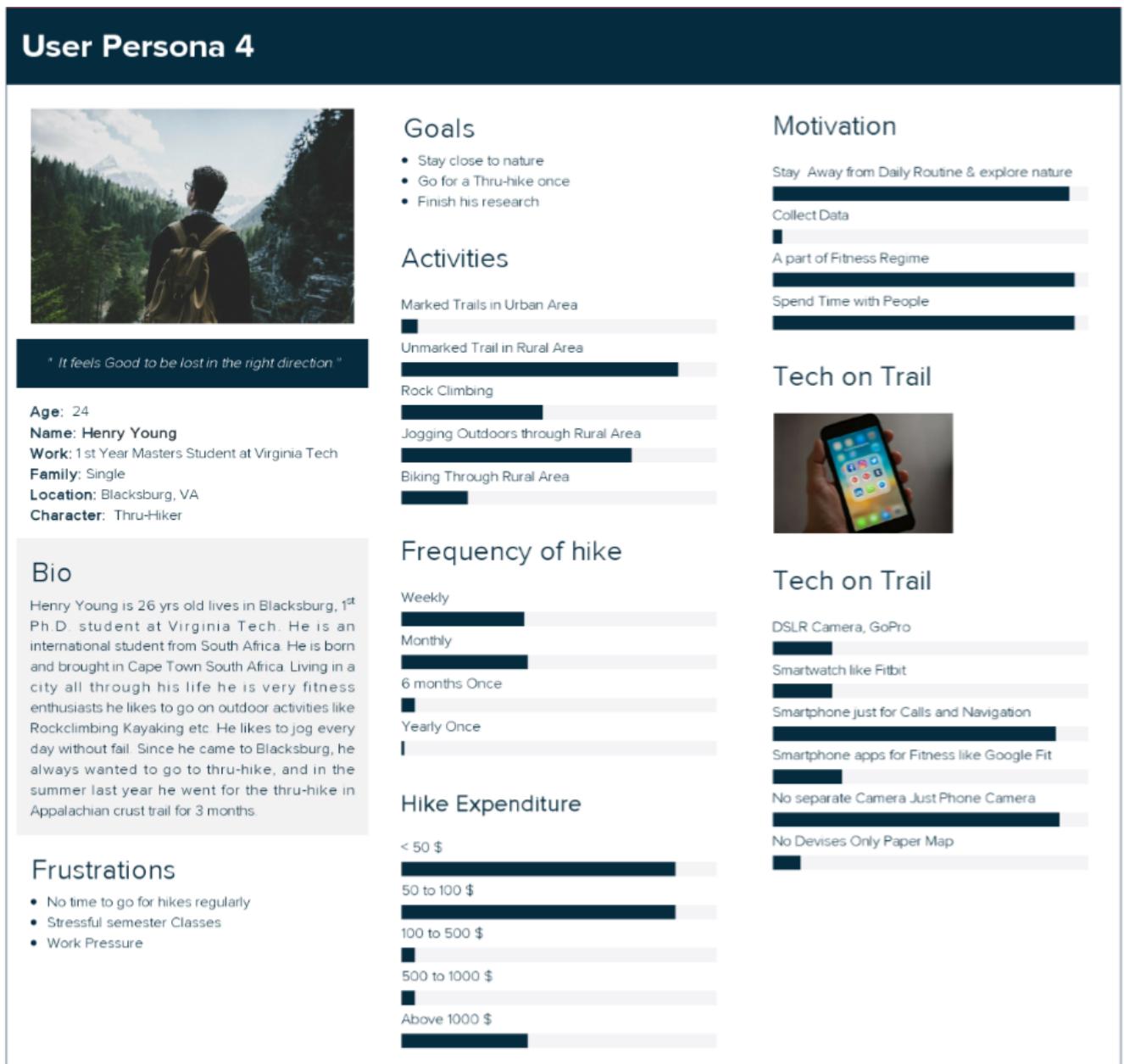


Figure 5.4: Persona 4. The left column describes key persona characteristics. The middle and right columns show cluster data used to craft the persona, Images used in this persona is used from Free image website[6].

5.5 Persona 5

Cluster 5 and 6 are smaller groups, cluster 5 consists of four members and cluster 6 consists of 3 members. So we combined these two clusters to form persona 5.

Andrea Madison 30 year old graduate student at Appalachian State University, North Carolina. Andrea is a botany major studying poisonous plants and their effect on mushroom hunters in marked and unmarked trails. Living very close to nature she is an avid hiker; whenever she gets the time she will go for a hike near her campus, preferably once in a month or more. Like others, Andrea also likes to hike just to enjoy nature and avoid daily routine, and also as a fitness regime.

As an grad student, she doesnt earn so much, and her daily allowance from her parents is limited, so for extracurricular activities like hiking she spends less than 70\$ at max. As a botany student, Andrea always carries multiple gadgets like a DSLR camera to take pictures and smartwatch to monitor her health and a smartphone with a navigation apps and for emergency calls. Whenever she finds an exotic plant or a poisonous plant she takes pictures and notes for her research and saves the location of the plant so that she can tweet on the spot to her group members to be aware of the poisonous plan location or to visit those fantastic places she just found.

Both the clusters are long distance hikers and like to hike once in a month and not very casual spenders who carry multiple gadgets and active data collectors. When we see them individually as a different clusters they may look different however when we combined the data and reanalyze them they are almost similar. Both the clusters are mixed age group people, but mostly above 30 yrs old so we created the persona 5 to be between 30 to 50 yrs based on the their most favorite activity, frequency, motivation and gadget preference we decided to create this persona 30 yrs old and a research student who is interested to collect

data for her research using multiple gadgets.

The below figure is sectioned into 2 parts. The left column consists of persona descriptions like profile, personnel information, background, frustrations. Goals were formed based on characteristics and persona back story. From activities section, the rest of them are derived from the unique characteristics observed in the previous chapter.

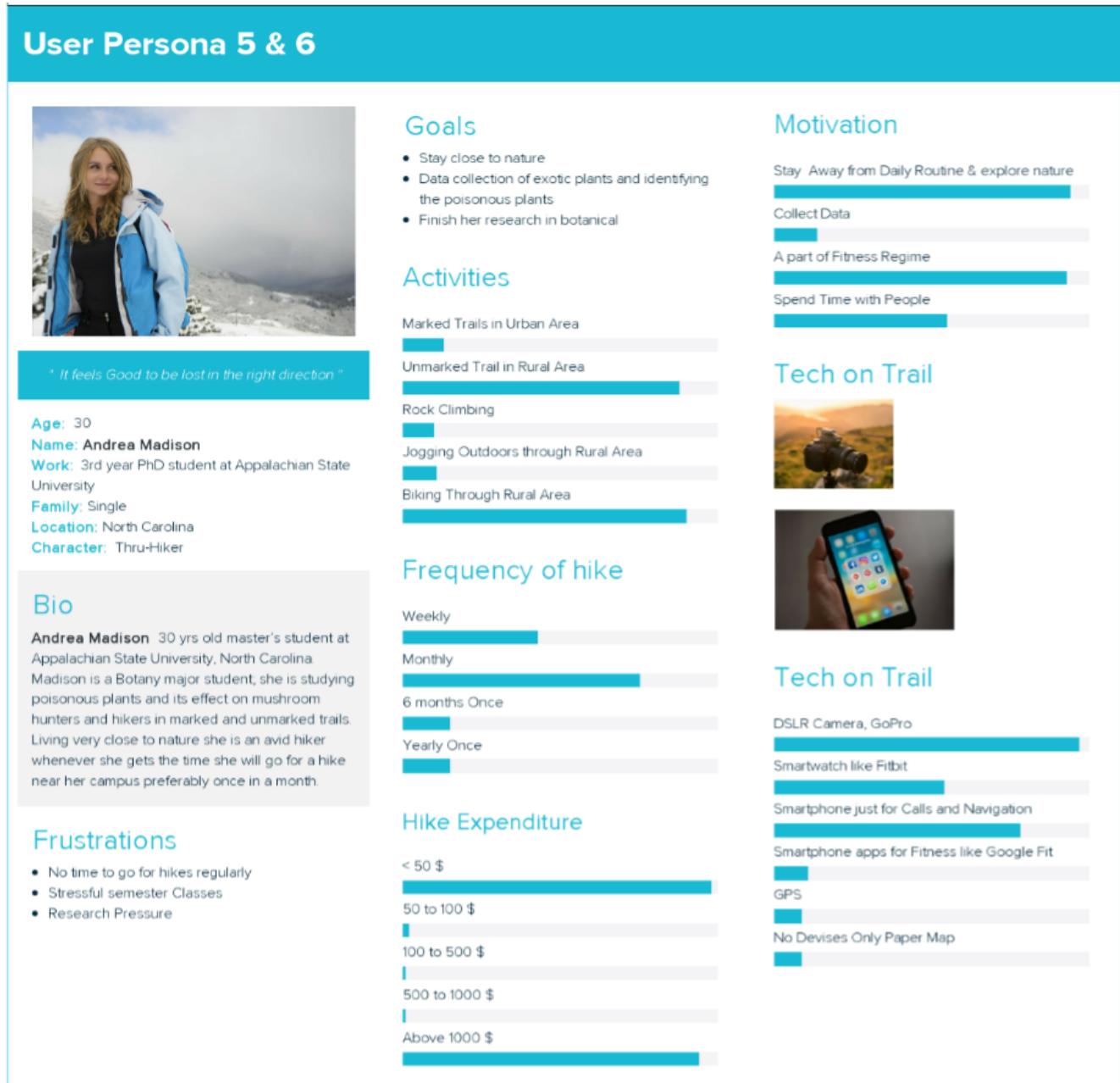


Figure 5.5: Persona 5. The left column describes key persona characteristics. The middle and right columns show cluster data used to craft the persona, Images used in this persona is used from Free image website[6].

Chapter 6

Conclusion & Future Work

6.1 Conclusion

There are different types of people that we meet on a trail, and each one of them has a different personality. Everyone has a reason for being there at that moment. Some of them want to spend time with friends and family, some may want to stay away from the daily routine and tensions, some may go for rehabilitation. These hikers carry all sorts of technology with them just to make the journey safe, comfortable, or memorable. Everyone has a different choice and perspective towards the tech they like to carry with them. Some may feel they need their iPhone to take pictures of beautiful scenery they come across in the trail, while some may like to use a high definition Canon's 50.6-megapixel camera to take the same picture. This research seeks to understand the different perspectives of the hikers and the technology they carry with them on the trail.

40 self-described hikers participated in a survey that consisted of 12 questions in two parts. The first part consists of questions asking about the participants' demographic data, hiking habits, frequency, expenditures, and technology they carry. The second half is "would you rather" questions asking participants to choose between envisioned hiking scenarios. The collected data was clustered into 6 clusters using K-Mode clustering, and descriptive analysis identified unique characteristics for each cluster.

Five personas were crafted using the identified characteristics based on the analysis, taking into account the correlation between hiking preference and age, preferred activities and motivation, and maximum expenses they spend on the tech they carry. These personas are designed to act as a guide to those who purchase, design, and use hiking related products for both commercial and research purposes. Each persona represents a set of unique set of characteristics for a group of people who share similar likes and dislikes for hiking gear and hiking goals.

6.2 Future Work

The future scope of research in this area lies in conducting deeper analysis to create the design guidelines for each technology we have discussed in this research. The current study is a first step towards the technology on the trail product development. This research should be conducted on wider range of hiking gadgets, but we can narrow down the focus to each gadget for example conducting survey on only smartphones or only smartwatches or only navigations devices or only high definition cameras.

This research is limited to 40 participants who are in and around Blacksburg, but for future studies a survey can be done through venues such as Facebook hiking groups, hiking-related web sites, groups like Virginia Tech's Venture Out, and product web sites. That will give more perspectives of hikers in different trails, and thus pursue breadth in what hikers across the country think about technology reactions. We expect that the results of this thesis, and of follow-up survey and analysis work, can support projects such as: smartphone multi-purpose hiking app creation, crafting augmented reality on the smartphone for people to understand the plants and animals they come across in the hike, smartphone camera apps that very specific to the trail-related goalss, smartwatch apps which aid in wellness and

navigation in planning hike and camping distances, and understanding social media effects on the adventure of hiking.

Bibliography

- [1] Technology on the trail position papers. URL <https://technologyonthetrail.wordpress.com/workshop-group2018/workshop-position-papers/>.
- [2] Alan walks wales. URL <http://alanwalks.wales/>.
- [3] Reddit online discussion: Technology on the trail. URL https://www.reddit.com/r/AppalachianTrail/comments/1h2z1o/technology_on_the_trail.
- [4] Technology on the trail 2017 workshop. URL <https://technologyonthetrail.wordpress.com/workshop-2017/>.
- [5] Is social media ruining hiking in b.c.? URL <http://www.cbc.ca/news/canada/british-columbia/social-media-hiking-1.3755738>.
- [6] Free photos of hikers and gadgets. URL <https://www.pexels.com/photo-license/>.
- [7] The shape of data: K-modes. URL <https://shapeofdata.wordpress.com/2014/03/04/k-modes>.
- [8] Tamara Adlin, John Pruitt, Kim Goodwin, Colin Hynes, Karen McGrane, Aviva Rosenstein, and Michael J Muller. Putting personas to work. In *CHI'06 Extended Abstracts on Human Factors in Computing Systems*, pages 13–16. ACM, 2006.
- [9] Zann Anderson, Candice Lusk, and Michael D Jones. Towards understanding hikers’ technology preferences. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*, pages 1–4. ACM, 2017.

- [10] Periklis Andritsos, Panayiotis Tsaparas, Renée J Miller, and Kenneth C Sevcik. Limbo: Scalable clustering of categorical data. In *International Conference on Extending Database Technology*, pages 123–146. Springer, 2004.
- [11] Daniel Barbará, Yi Li, and Julia Couto. Coolcat: an entropy-based algorithm for categorical clustering. In *Proceedings of the eleventh international conference on Information and knowledge management*, pages 582–589. ACM, 2002.
- [12] Daniel G Cabrero. User-created persona: Namibian rural otjiherero speakers. In *Proceedings of the 33rd Annual International Conference on the Design of Communication*, page 28. ACM, 2015.
- [13] EunJeong Cheon, Mohammad Hossein Jarrahi, and Norman Makoto Su. technology isnt always the best: The intersection of health tracking technologies and information practices of digital natives. In *Healthcare Informatics (ICHI), 2016 IEEE International Conference on*, pages 207–215. IEEE, 2016.
- [14] Alan Cooper, Robert Reimann, and Dave Cronin. Modeling users: Personas and goals. *Erschienen in About Face*, 2:55–74, 2003.
- [15] Alan Cooper, Robert Reimann, David Cronin, and Christopher Noessel. *About face: the essentials of interaction design*. John Wiley & Sons, 2014.
- [16] Alan Cooper et al. *The inmates are running the asylum:[Why high-tech products drive us crazy and how to restore the sanity]*. Sams Indianapolis, 2004.
- [17] Alexandru Dancu, Velko Vechev, Adviye Ayça Ünlüer, Simon Nilson, Oscar Nygren, Simon Eliasson, Jean-Elie Barjonet, Joe Marshall, and Morten Fjeld. Gesture bike: examining projection surfaces and turn signal systems for urban cycling. In *Proceedings*

- of the 2015 International Conference on Interactive Tabletops & Surfaces*, pages 151–159. ACM, 2015.
- [18] Sarah Grace Fields. *Technology on the Trail: Using Cultural Probes to Understand Hikers*. PhD thesis, Virginia Tech, 2017.
- [19] Arlene Fink. *How to conduct surveys: A step-by-step guide: A step-by-step guide*. Sage Publications, 2012.
- [20] Venkatesh Ganti, Johannes Gehrke, and Raghu Ramakrishnan. Cactusclustering categorical data using summaries. In *Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 73–83. ACM, 1999.
- [21] Bill Gaver, Tony Dunne, and Elena Pacenti. Design: cultural probes. *interactions*, 6(1):21–29, 1999.
- [22] William W Gaver, Andrew Boucher, Sarah Pennington, and Brendan Walker. Cultural probes and the value of uncertainty. *interactions*, 11(5):53–56, 2004.
- [23] David Gibson, Jon Kleinberg, and Prabhakar Raghavan. Clustering categorical data: An approach based on dynamical systems. *Databases*, 1:75, 1998.
- [24] James E Grizzle, C Frank Starmer, and Gary G Koch. Analysis of categorical data by linear models. *Biometrics*, pages 489–504, 1969.
- [25] Jonathan Grudin and John Pruitt. Personas, participatory design and product development: An infrastructure for engagement. In *PDC*, pages 144–152, 2002.
- [26] Sudipto Guha, Rajeev Rastogi, and Kyuseok Shim. Rock: A robust clustering algorithm for categorical attributes. *Information systems*, 25(5):345–366, 2000.

- [27] Ellie Harmon and Melissa Mazmanian. Stories of the smartphone in everyday discourse. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*, 2013. doi: 10.1145/2470654.2466134.
- [28] Mary E. Harmon. *Computing as Context: Experiences of Dis/Connection Beyond the Moment of Non/Use*. PhD thesis, 2015.
- [29] Rex Hartson and Pardha S Pyla. *The UX Book: Process and guidelines for ensuring a quality user experience*. Elsevier, 2012.
- [30] Zhexue Huang. Clustering large data sets with mixed numeric and categorical values. In *Proceedings of the 1st pacific-asia conference on knowledge discovery and data mining, (PAKDD)*, pages 21–34. Singapore, 1997.
- [31] Zhexue Huang. A fast clustering algorithm to cluster very large categorical data sets in data mining. *DMKD*, 3(8):34–39, 1997.
- [32] Zhexue Huang. Extensions to the k-means algorithm for clustering large data sets with categorical values. *Data mining and knowledge discovery*, 2(3):283–304, 1998.
- [33] Barbara A Kitchenham and Shari Lawrence Pfleeger. Principles of survey research: part 3: constructing a survey instrument. *ACM SIGSOFT Software Engineering Notes*, 27(2):20–24, 2002.
- [34] Felix Kosmalla, Frederik Wiehr, Florian Daiber, Antonio Krüger, and Markus Löchtefeld. Climbaware: Investigating perception and acceptance of wearables in rock climbing. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 1097–1108. ACM, 2016.
- [35] Maaret Posti, Johannes Schöning, and Jonna Häkkilä. Unexpected journeys with the

- hobbit: the design and evaluation of an asocial hiking app. In *Proceedings of the 2014 conference on Designing interactive systems*, pages 637–646. ACM, 2014.
- [36] John Pruitt and Jonathan Grudin. Personas: practice and theory. In *Proceedings of the 2003 conference on Designing for user experiences*, pages 1–15. ACM, 2003.
- [37] Henri Ralambondrainy. A conceptual version of the k-means algorithm. *Pattern Recognition Letters*, 16(11):1147–1157, 1995.
- [38] Emma Siddall, Corelia Baibarac, Aimee Byrne, Niall Byrne, Ashling Deasy, Nuala Flood, Camille Goulding, Shane O’Driscoll, Niamh Rabbitt, Angharad Sweeney, et al. Personas as a user-centred design tool for the built environment. 2011.
- [39] Norman Makoto Su and Eunjeong Cheon. Reconsidering nature. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI ’17*, 2017. doi: 10.1145/3025453.3025966.
- [40] Norman Makoto Su and Bryan Duggan. Tunetracker: tensions in the surveillance of traditional music. In *Proceedings of the 2014 conference on Designing interactive systems*, pages 845–854. ACM, 2014.
- [41] Norman Makoto Su and Gloria Mark. Designing for nomadic work. In *Proceedings of the 7th ACM conference on Designing interactive systems*, pages 305–314. ACM, 2008.
- [42] Jayne Wallace, John McCarthy, Peter C Wright, and Patrick Olivier. Making design probes work. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 3441–3450. ACM, 2013.
- [43] John Wang. *Encyclopedia of data warehousing and mining*, pp 246 - 248. IGI Global, 2005.

Appendices

Appendix A

Pre Cluster Data Analysis

Pre-Cluster Data Analysis: 1. Gadget preference: In the below-mentioned graph we can observe that gadget 3 (Smartphone for calls and Navigation apps) is the most preferred gadget. Age Groups from 18 to 30yr are more interested in the smartphone navigations than using

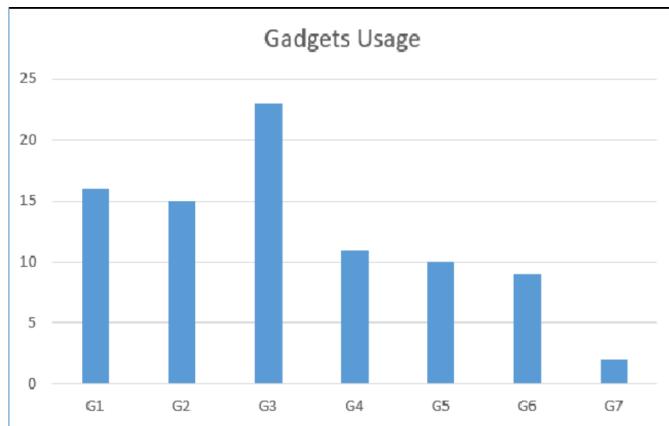


Figure A.1: Gadgets Preference

2.Age vs. Expenses: This is bar Plot from R software. we used this plot to see the clear fluctuation between two data points age and expense. As we can see in the graph some participants ages from 18 to 30 tend to spend more money on the gadgets for a hike than the people who are between 31 to 50 age group. However, its is clear that people who are in age group 18 to 30 are the high spenders and they could be the target age group for consumer product companies.

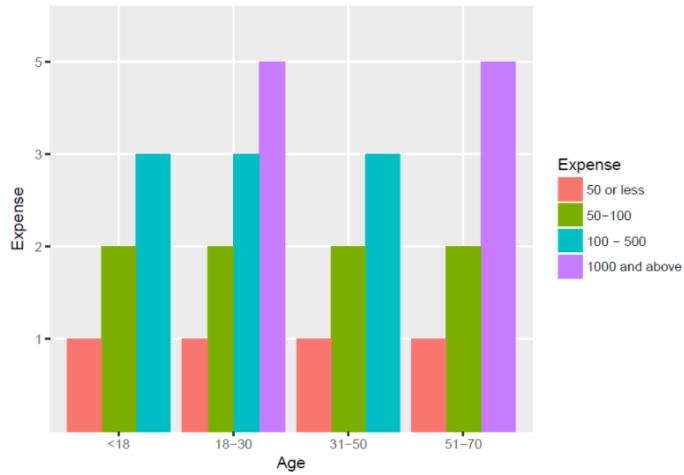


Figure A.2: Age vs Expense

3. Age vs. gadgets: The below graph is a bar graph created using R software and In the graph people who choose to carry more than three gadgets are in the age groups if <18 to 30 yrs. However, If we observe the older people are more likely to carry a DSLR camera and a GPS system over a smartphone.same time they are not interested to use a smartwatch to monitor their health.

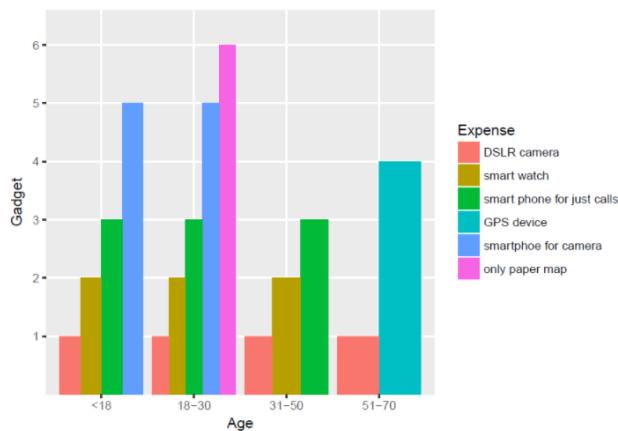


Figure A.3: Age vs Gadgets

4. Age vs. Activity: This is parallel graph, it's created using JMP software and density if

the links connecting between two lines shows that more people have selected those options, In the below graph activities A, C, and E are the most preferred activity, and most of the age groups have opted the A and C activities.

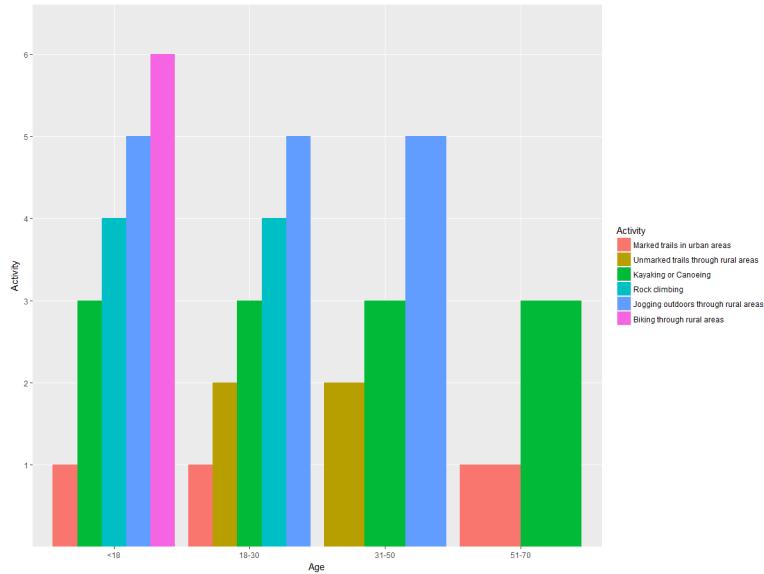


Figure A.4: Age vs Activity

5. Comparing would you rather with gadgets In this graph we can see 62.7 of hikers choose option A Over B. People who choose option gadgets 1 and 3. This analysis says that more people are inclined to take a DSLR camera and Fitbit over a smartphone with multiple Apps.

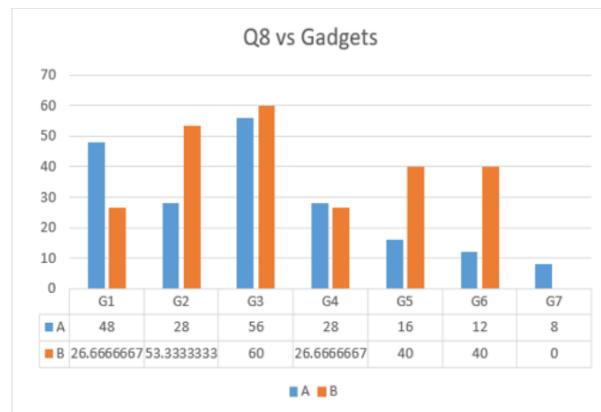


Figure A.5: Correlation between Gadgets vs Would you rather Question 8

6. In the below graph its very evident that hikers would like to go on a hike when it is sunny but cold weather with a solar energy charger over the heavyweight batteries. In the graph, 80% of the participants choose to go on a hike in sunny weather with a solar energy charger.

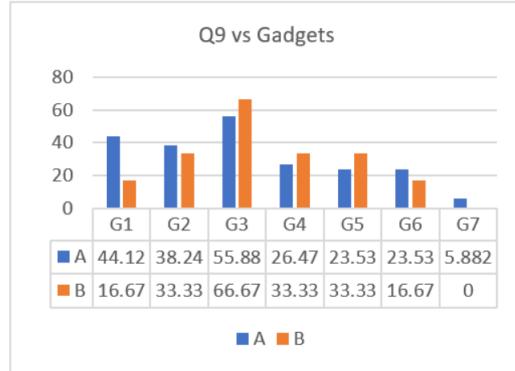


Figure A.6: Correlation between Gadgets vs Would you rather Question 9

7. In the graph, we can see that people who are interested to collect data have opted to carry multiple gadgets than the passive data collectors who want take a smartphone. Overall 67% of people are interested in bringing the smartphone while collecting data passively.

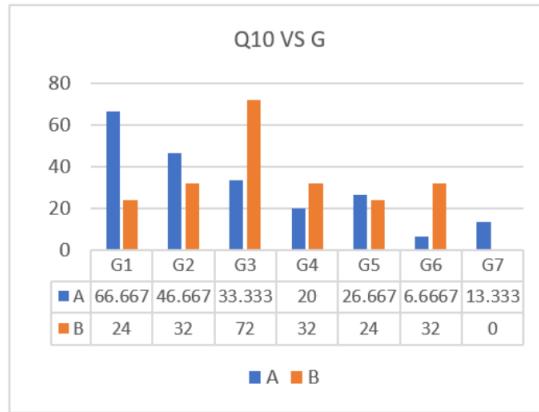


Figure A.7: Correlation between Gadgets vs Would you rather Question 10

8. Comparing Age of the hikers effect on the mobile usage patterns for social networking while on a hike? In the graph, we can see people who are less than 30 years are more interested

in tweeting while on a hike instead go back and write a blog about it. An interesting note is that people who are above 51 are not much interested in carrying a smartphone and tweet the real time on the spot.

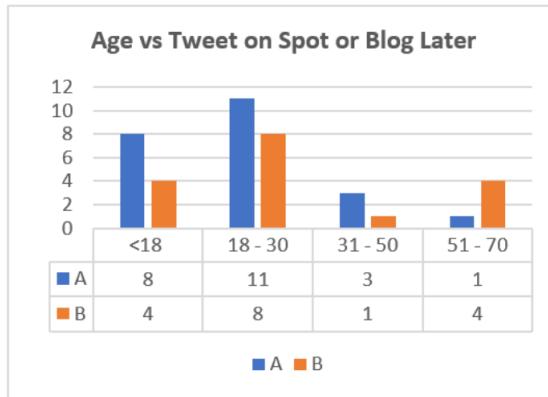


Figure A.8: Correlation between Gadget vs Would you rather Question 11

9. Comparing smartphone off-line maps vs Correlation between Gadgets vs Would you rather Question 8old school paper maps.In the graph as we observed 60% of people choose to use smart phone with a water-resistant on line map over paper map.

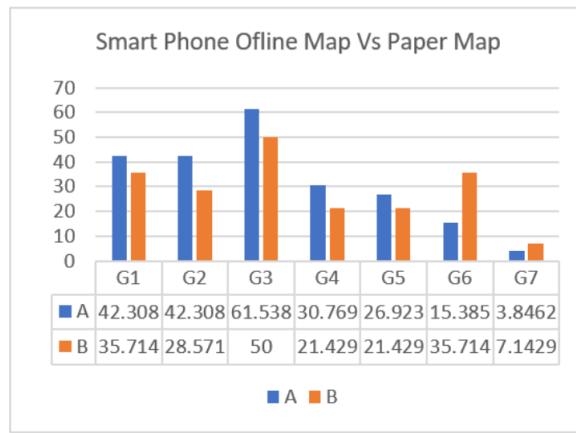


Figure A.9: Correlation between Gadgets vs Would you rather Question 12

Appendix B

Cluster Visualizations:

B.1 Individual Analysis Cluster 1:

Age, Activities and Gadgets:

In this cluster 16 participants are clustered together. Out of 16 participants 5 of them belong to < 18 yrs, and 7 of them belong to 18 to 30 yrs, and 3 of them belong to 51 to 70 yrs. As we can see the major population is in 18 to 30 age group. For preferred activities: 11 participants selected option 2(Walking through backwoods, unmarked trails through rural areas), 9 selected option 1(Walking through backwoods, unmarked trails through rural areas), 6 selected option 4(Rock climbing), 5 selected option 3(Kayaking/canoeing) and 3 selected option 5(Jogging outdoors through rural areas).

Gadget preferences for cluster 1: 10 participants choose option 3(Smartphone for just calls and navigation apps like Guthook), 6 choose option 4(Smartphone apps for fitness like google fit), 5 choose option 5(GPS devise like Garmin eTrex, etc. (separate from smartphone)), 4 choose option 1(DSLR camera, Go Pro, etc. any camera which is not part of smartphone), 3 choose option 6(No separate camera, just a smartphone camera). The majority of the people are in this cluster is smartphone users.

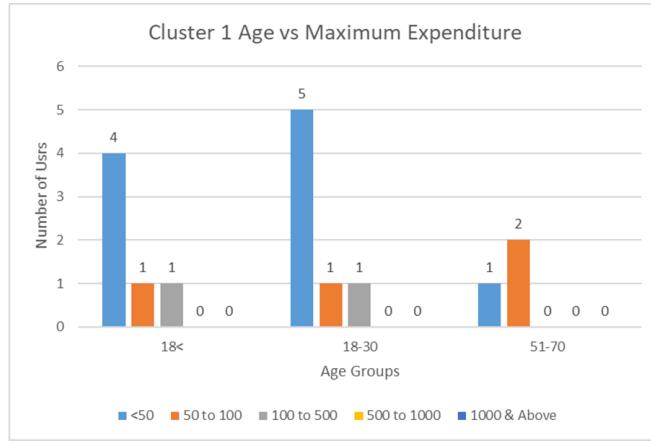
Age vs Expense:

Figure B.1: User Age and Expense correlation in cluster 1

The above graph displays the correlation between the age and expenses in cluster 1. The X-axis represents the participants Expense categories and Y-axis the number of users choose these options. Most of the participants in this cluster are between 18 to 30 yrs old and the maximum money they spend is <50\$. This graph represents that most of the participants are a young crowd and their maximum amount is between <50 and 50 to 100\$. However, these young people are the ones who are willing to spend 100 to 500\$.

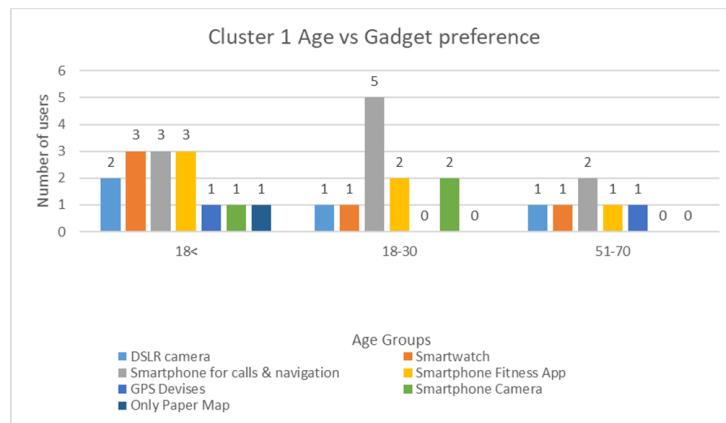
Age Vs Gadgets:

Figure B.2: User Age and Gadget representation in Cluster 1

This graph represents the user ages and the tech they carry with them while on a hike. The most preferred gadget is the smartphone for navigation and the calling, irrespective of different age groups all of them choose the same gadget. The second most preferred gadget is smartphone apps for fitness like Google Fit. We can identify a definite characteristic here participant 1 is below 18 yrs old, and he likes to take multiple devices like a smartwatch for fitness and smartphone for calls and navigation and also smartphone apps for fitness and a smartphone camera.

Activities Vs Motivation:

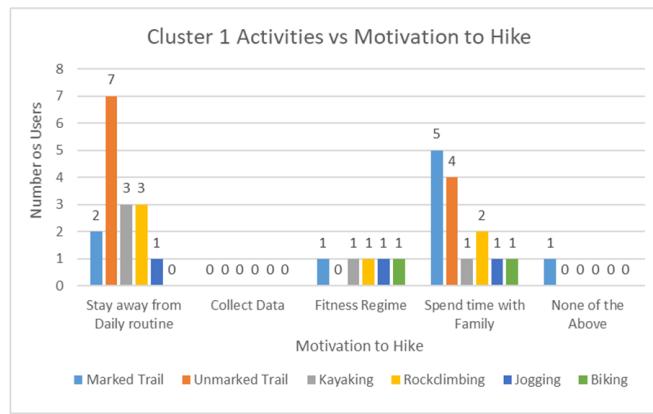


Figure B.3: Users Activities and Motivation preference in cluster 1

Why do people like to go on a hike? What motivates them to walk for straight 8 to 11 hrs. in a day or go for thru-hike for a 2000-mile hike for 3 months?. As Dr. Ellie Harmon in her dissertation, she shared her experience with other hikers while she went for thru-hike in PCT. While on a hike she met a lot of hikers who are true-hikers as well as short-distance hikers and discussed technology and hiking motivation. Lot of them said they are hiking to get away from other people, but more than that, they were attempting to get away from society and its never-ending obligations.

In my survey, we have similar questions, and most the participants also said that they like

to go on a hike to stay away from the daily routine and explore nature. Few of them said they like to go on a hike as a fitness regime. Some of them said they want to spend time with family and friends we consider these later group of people as recreational hikers they like to go for short hikes and like to spend time with friends. For example, user 8 like to go on an unmarked hike in a rural area, and the motivation for the hike is to stay away from daily routine and explore nature. Options wef gave the participants in the survey is given below.

Activities Vs Expense:

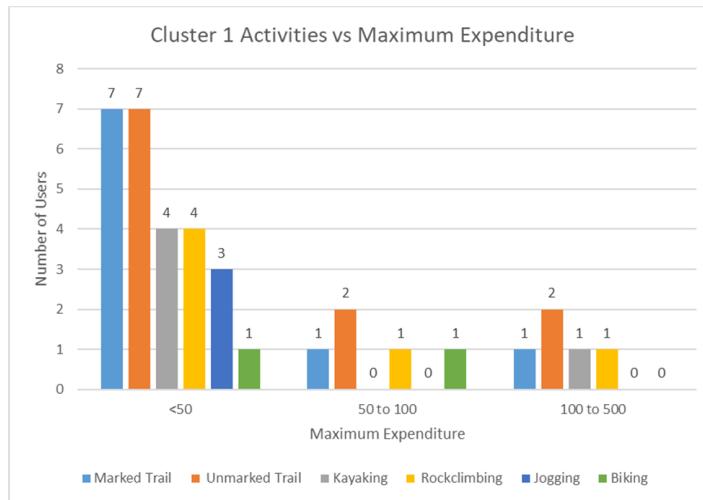


Figure B.4: Users Activities and Expense representation in cluster 1

In this graph, we are representation each user activities and their maximum expenses while on a hike. User 8 likes to go on an unmarked trail through a rural area and walk through backwoods and likes to spend 100 to 500 \$. There are some participants who just randomly selected all the activities like participant 19. User 37 likes to walk on the marked trail and unmarked trail and kayaking and rock-climbing and the max amount he wants to spend on 100 to 500\$. This shows that user 37 is an active person and he likes to spend money on his tech.

Activities Vs Frequency:

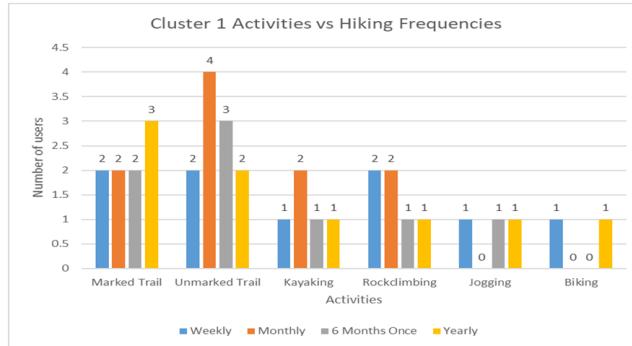


Figure B.5: Users Activities and Hike frequency preference in cluster 1

In the above graph, we are displaying the user preferred activities and hike frequency. The x-axis represents the users, and y-axis user chose option blocks. This show how often these participants are going on a hike, for example, user 27 said he likes to go on an unmarked trail in the rural area, and he would like to go monthly once. If we compare this with the expense, we can identify concrete characteristics of Cluster 1. As you can see most of the users are preferring to go on marked and unmarked trails but being a young crowd these group of people is active in a lot of outdoor activities.

Activities Vs Motivation:

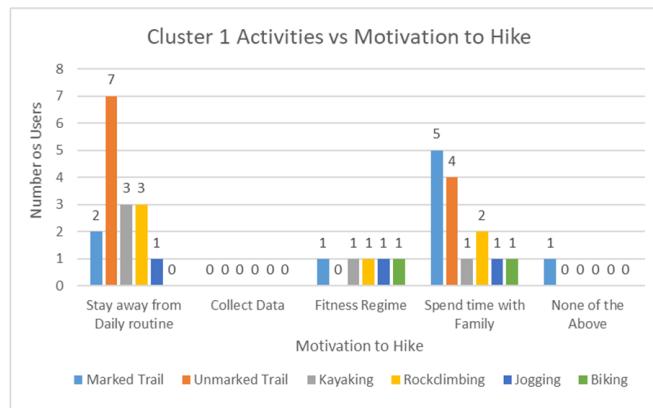


Figure B.6: Users Activities and Motivation preference in cluster 1

Why do people like to go on a hike? What motivates them to walk for straight 8 to 11 hrs. in a day or go for thru-hike for a 2000-mile hike for 3 months?. As Dr. Ellie Harmon in her dissertation, she shared her experience with other hikers while she went for thru-hike in PCT. While on a hike she met a lot of hikers who are true-hikers as well as short-distance hikers and discussed technology and hiking motivation. Lot of them said they are hiking to get away from other people, but more than that, they were attempting to get away from society and its never-ending obligations.

In my survey, we have similar questions, and most the participants also said that they like to go on a hike to stay away from the daily routine and explore nature. Few of them said they like to go on a hike as a fitness regime. Some of them said they want to spend time with family and friends we consider these later group of people as recreational hikers they like to go for short hikes and like to spend time with friends. For example, user 8 like to go on an unmarked hike in a rural area, and the motivation for the hike is to stay away from daily routine and explore nature. Options we gave the participants in the survey is given below.

Activities Vs Gadgets:

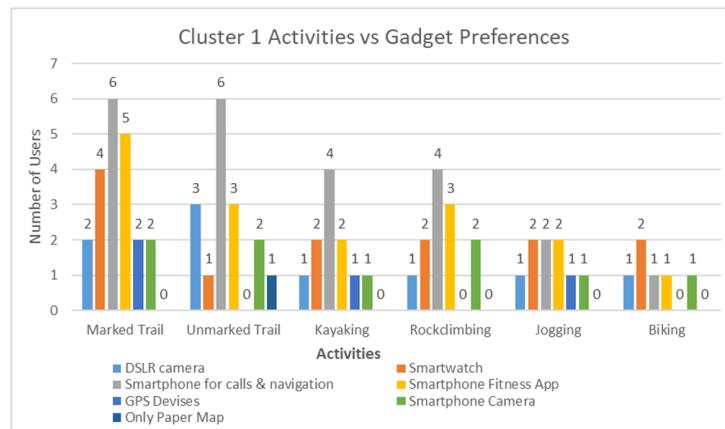


Figure B.7: Users Activities and Gadget preference in cluster 1

In the above graph, we tried to understand the what is the users choice of gadget for each activity. The first half blocks are activities, and upper layers blocks are gadgets. One interesting observation the participant 19 selected all the options in the activities, but in the gadgets, he just picked Smartwatch as a preferred tech on trial. This gives him distinguishing characteristics, i.e., he is very active outdoors person, and he is keen on fitness. One more example participant 25 like to go on marked and unmarked trails in urban and rural areas. His choice of tech is the smartphone for navigation, calls and camera purpose. He doesn't like to take any other gadget with him other than a smartphone multipurpose.

Discussions over Would you rather:

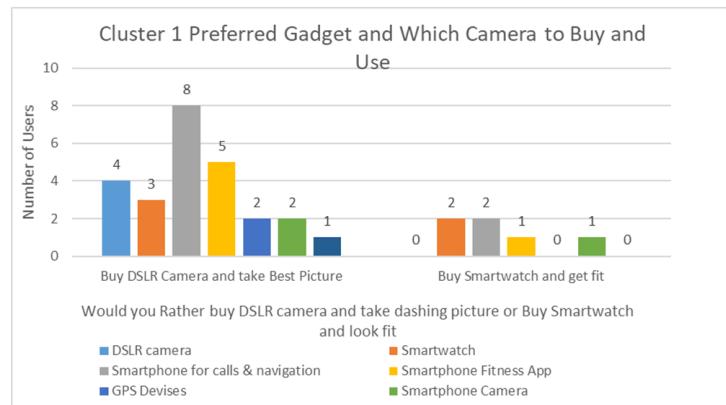


Figure B.8: Cluster 1 Would you rather Preference

In the above graph we are projecting cluster 1 participants would you rather question preferences. This projection completes the connection between the previous questions. With this, graph we are combining previous graphs with this we are identifying the characteristics to create the persona for cluster 1.

Discussions over Would you rather:

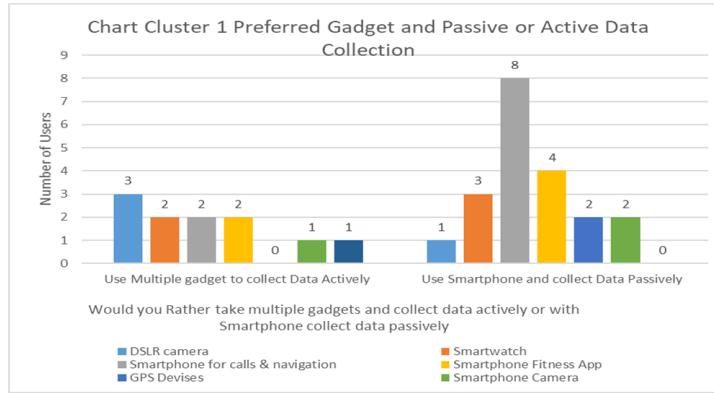


Figure B.9: Cluster 1 Would you rather Preference

B.2 Individual Analysis Cluster 2:

Cluster two is rather smaller group compared to the cluster 1. This group consists of 7 participants and below is the analysis for this cluster and like the previous analysis here also we tried to address both the Raw data representation and the correlation between the categories.

Age, Activities, Gadgets:

In this group we have two categories one <18 yrs and the other one is 18 to 30 group. But the significant chunk of people is from former <18 yrs old. Most of the users choose multiple activities; as we observed most of them preferred to go biking through rural areas and almost all them said they like to go for jogging outdoors through rural areas. Very few of them selected the outdoor activities like rock climbing and kayaking. This group of people selected multiple gadgets, but the most preferred tech on the trail is smartwatch like Fitbit for fitness. The second most preferred tech is DSLR camera. Here people who like a smartwatch doesn't want to carry a smartphone for the taking photos they prefer GoPro or

DSLR cameras.

Age vs Expense:

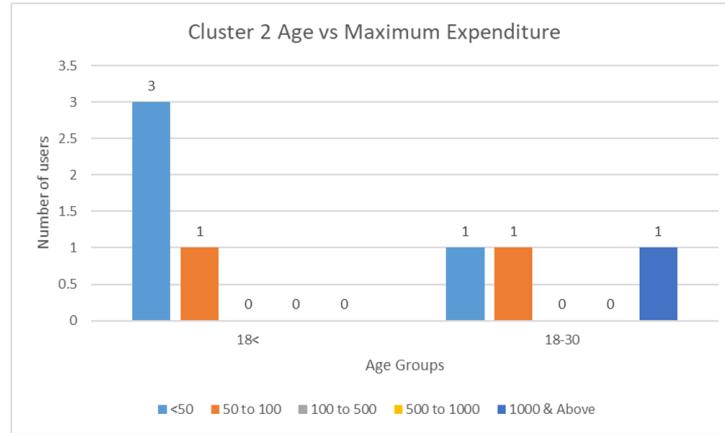


Figure B.10: User Age and Expense correlation in cluster 2

The above graph displays the correlation between the age and expenses in cluster 2. Most of the users in this cluster are between <18 yrs old. This graph represents that most of the participants are a young crowd and their maximum amount is between <50 and 50 to 100 \$.

Activities Vs Expense:

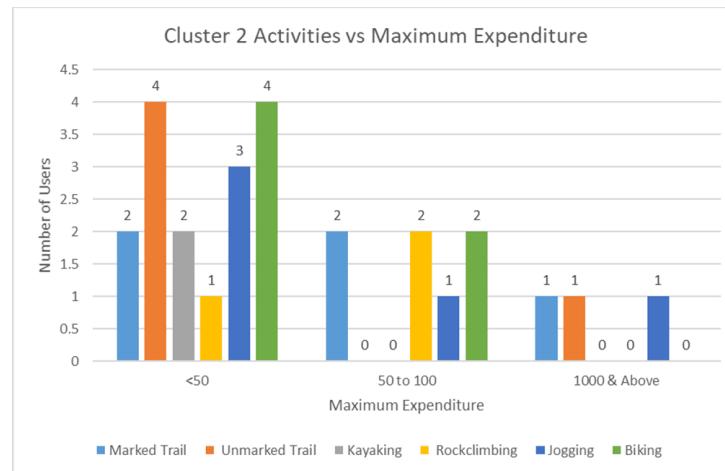


Figure B.11: Users Activities & expense Preference in cluster 2

In this graph, we are representation each user activities and their maximum expenses while on a hike. As observed most users who choose to spend <50\$ choose to take part in all the activities.

Activities Vs Frequency:

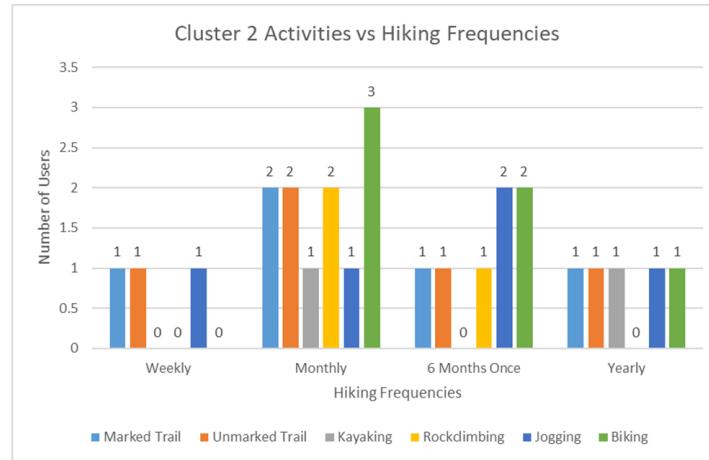


Figure B.12: Users Activities and Frequency Preference in cluster 2

In the above graph, we are displaying the user preferred activities and hike frequency. This show how often these participants are going on a hike. As observed most of the users are preferring to go on marked and unmarked trails but being a young crowd these group of people is very interested in biking through rural areas and go for a hike monthly once.

Activities Vs Motivation:

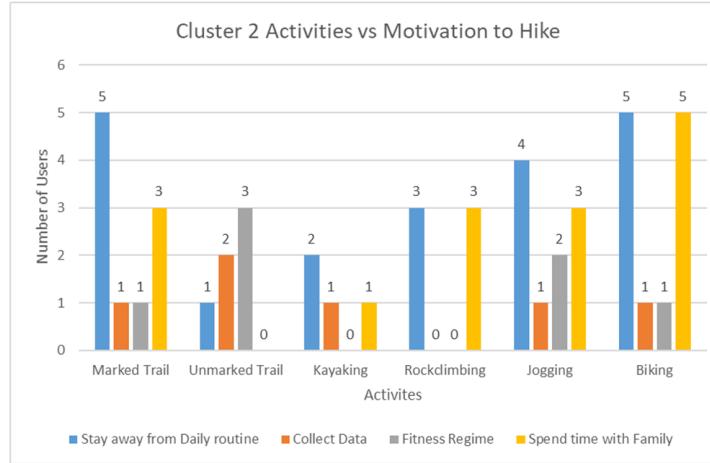


Figure B.13: Users Activities and Motivation Preference in cluster 2

In the above graph, we are connecting the hiker's activities and their motivation to go on a hike. Most of the participants said that they like to go on a hike to spend time with family and friends and almost same amount of people said they like to stay away from the daily routine and explore nature. This group of people is not serious hikers they like to hike but with family and friends.

Activities Vs Gadgets:

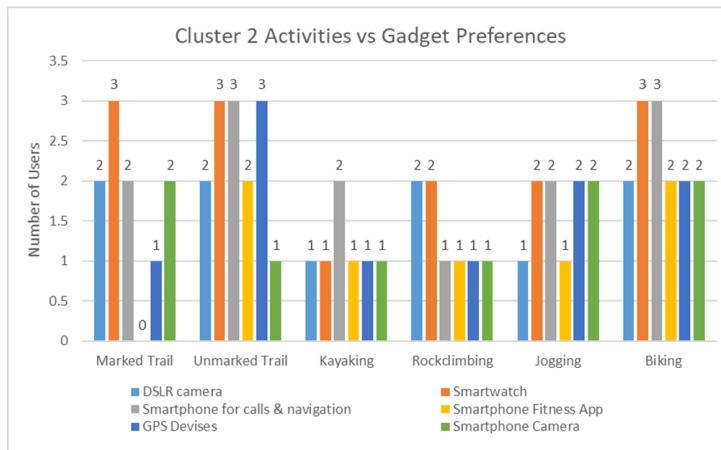


Figure B.14: Users Activities and Gadgets Preference in cluster 2

In the above graph, we tried to understand the what is the users choice of gadget and their preferred activity. One interesting observation the participant all of them selected all of the options in the activities, and gadgets. The most preferred gadget is Smartphone as a preferred tech on trial. But a lot of them preferred option 2 that smartwatch and GPS devise which is separate from a smartphone.

Cluster 2 Preferred Gadget and Which Camera to Buy and Use:

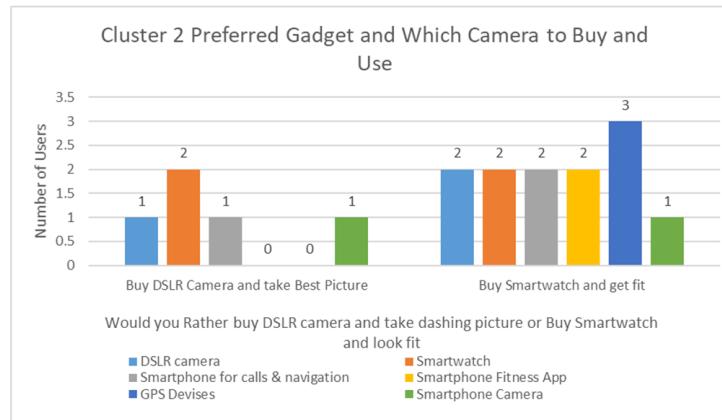


Figure B.15: Users preferred gadget compared with what gadget they buy weather to look fit in someone else photo or take a picture that goes viral

In the above graph we are comparing the Users preferred gadget compared with gadget they wish to buy. This is the correlation between Question 7 Preferred gadgets with Question 8 would you rather buy a DLSR camera and take a stunning photo which goes viral or buy a smartwatch and get fit and be part in strangers picture that goes viral. As observed people who choose multiple gadgets like to buy a smartwatch and get fit for strangers photos.

B.3 Individual Analysis Cluster 3:

Cluster three is rather smaller group compared to the cluster 2. This group consists of 5 participants and below is the analysis for this cluster and like the previous analysis here also we tried to address both the Raw data representation and the correlation between the categories.

Age, Activity, Gadget:

In this cluster as well we have only two categories but different groups 1. 18 to 30 group, 2. 30 to 50 group. But the significant chunk of people are between 18 to 30 yrs old. Most of the users choose multiple activities; as observed most of them like walking in marked trails in the urban area and the second most preferred activity is walking through backwoods, unmarked trails through rural areas. Again this group selected multiple combinations of devices, but the most preferred tech on the trail is the smartphone for just calls and navigation apps like Guthook. The second most preferred tech is evenly distributed between smartphone for taking pictures, GPS devise separate from a smartphone, smartwatch, and DSLR camera.

Age vs Expense:

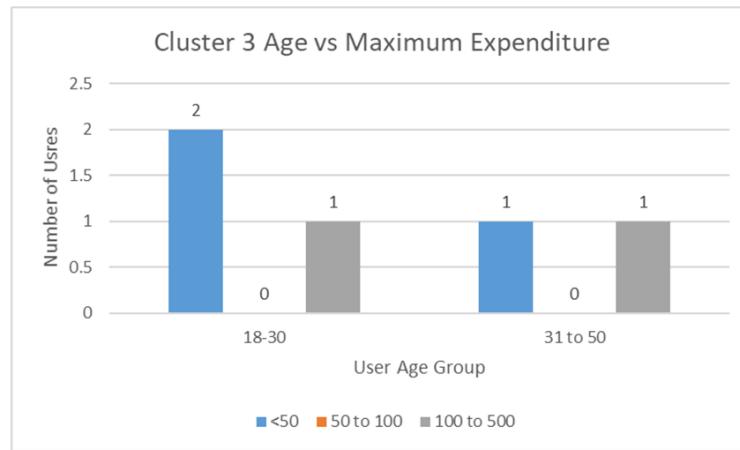


Figure B.16: Users Age and Maximum Expenditure in Cluster 3

The above graph displays the correlation between the age and maximum expenses in cluster 3. This graph represents that most of the participants are a little older crowd and the maximum amount they spend while on a hike is also more. Even though three of the participants said they spend <50, but the two of them said straight to 100 to 500\$. Comparing with previous clusters, this is a unique observation.

Age Vs Activities:

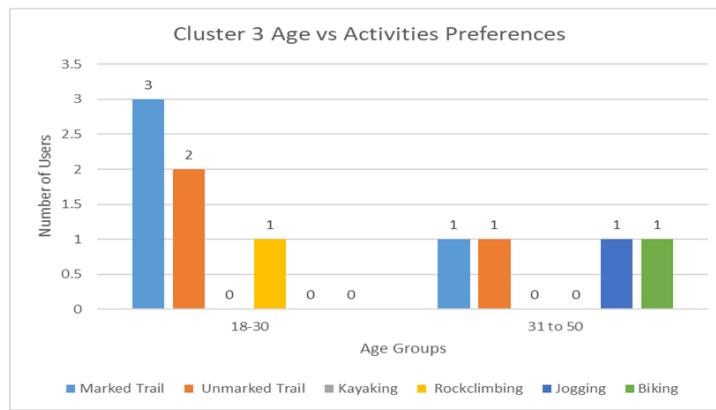


Figure B.17: Users Age and Activities Preference in cluster 3

In this graph, we are projecting each user preferred activity correlation with age in cluster3. This analysis is conducted to understand if the user activities are influenced based on the age group they fall. Irrespective of their age this group of people most preferred activity is walking through in marked trails in urban areas and walking through backwoods, unmarked trails through rural areas.

Age Vs Gadgets:

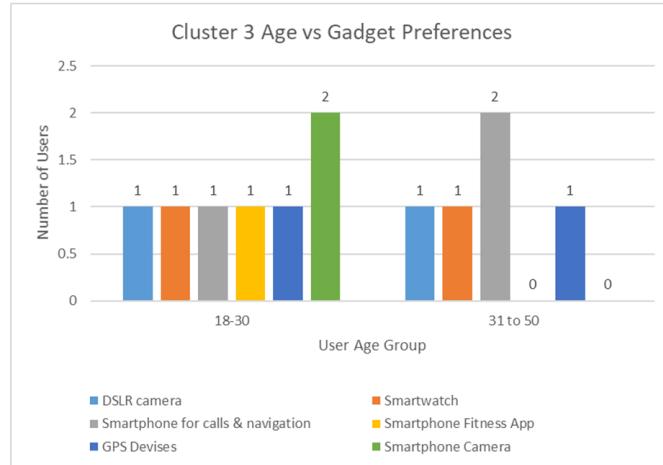


Figure B.18: Users Age and Gadget Preference in cluster 3

This graph represents the user ages and the tech they carry with them while on a hike. As we can see, there is no differentiation in gadget preference based on the user age group. Irrespective of their age group cluster 3 most preferred tech on the trail is Smartphone for just calls and navigation apps.

Activities Vs Expense:

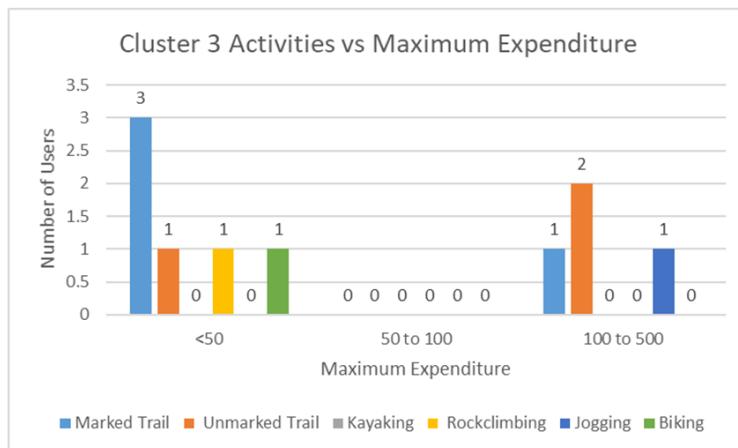


Figure B.19: Users Activities and Maximum Hike Expenditure in cluster 3

In this graph, we are representation each user activities and their maximum expenses while on a hike. As discussed before the second most preferred activity in this cluster is option two, i.e., walking through backwoods, unmarked trails through rural areas. Out these three users, 2 of them said their maximum expenditure is between 100 to 500\$.

Activities Vs Frequency:

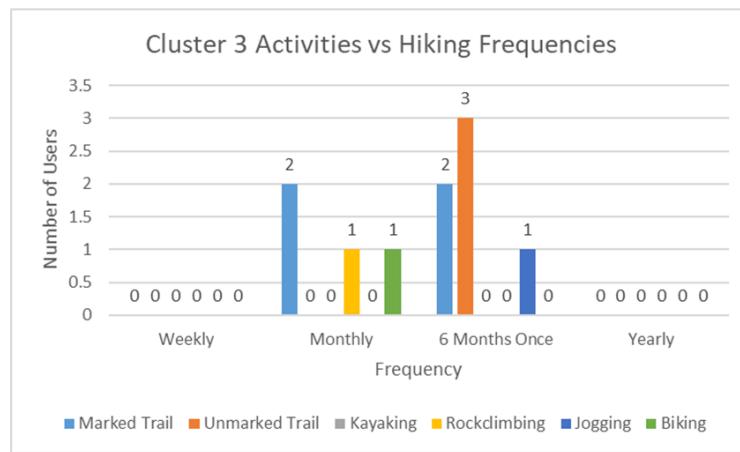


Figure B.20: Users Activities and Hike Frequency Preference in cluster 3

In the above graph, we are displaying the user preferred activities and hike frequency. This show how often these participants are going on a hike. As this group of people likes to hike more than doing other outdoor activities their frequency gives cluster characteristics. All the of the users said either they go on a hike once in month or 6months once.

Activities Vs Motivation:

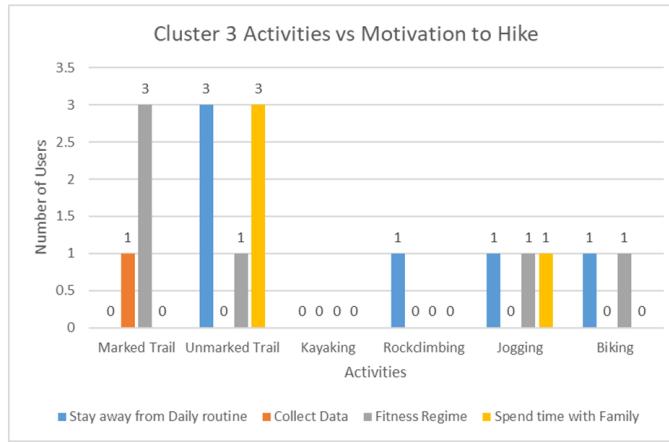


Figure B.21: Users Activities and Motivation to hike Preference in cluster 3

In the above graph, we are connecting the hiker's activities and their motivation to go on a hike. All of them said they like to go hiking to stay away from the daily routine and explore nature and spend time with family and friends. Out of 5 people, two of them said they also like to see the hike as a fitness regime. This group of people is serious hikers though they said they want to spend time with family and friends.

Activities Vs Gadgets:

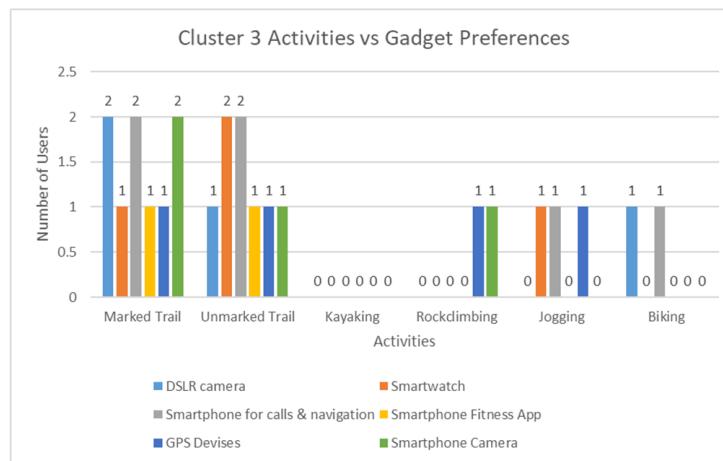


Figure B.22: Users Activities and Gadget Preference in cluster 3

In the above graph, we tried to understand the what is the users choice of gadget and their preferred activity. Most of them like to hike in a marked trail in urban areas like to use a smartphone for navigation and for calling purpose but use DSLR camera to take pictures and use the smartwatch to monitor their fitness.

B.4 Individual Analysis Cluster 4:

Cluster four is similar to cluster 3 strength wise. This group consists of 5 participants as well and below is the analysis and as defined in the previous clusters here also we tried to address both the Raw data representation and the correlation between the categories.

In this cluster as well we have only two categories but different groups 1. 18 to 30 group, 2. 30 to 50 group. But the significant chunk of people are between 18 to 30 yrs old. Most of the users choose multiple activities; as observed most of them like walking in marked trails in the urban area and the second most preferred activity is walking through backwoods, unmarked trails through rural areas. Again this group selected multiple combinations of devices, but the most preferred tech on the trail is the smartphone for just calls and navigation apps like Guthook. The second most preferred tech is evenly distributed between smartphone for taking pictures, GPS devise separate from a smartphone, smartwatch, and DSLR camera.

Age, Activity, Gadget:

In this cluster, we have only three categories. 2 of them are in <18 yrs and 2 of them in 18 to 30 yrs and 1 in 31 to 50 yrs old group. Most of the users selected multiple activities; as we can observe most of them like walking through backwoods, unmarked trails through rural areas and second most preferred activity is jogging outdoors through rural areas. This group of people selected multiple devices, but the most preferred tech on the trail is the

smartphone for just calls and navigation apps like Guthook. The second most preferred tech is option 6, ie. No separate camera, Just smartphone for taking photos.

Age vs Expense:

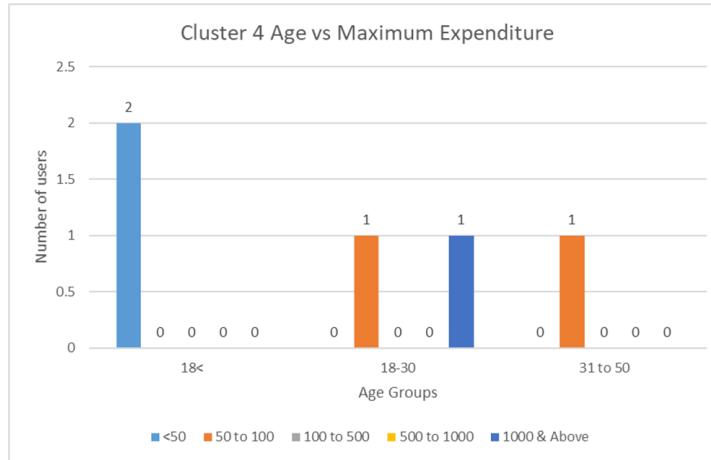


Figure B.23: Users Age and Maximum Expenditure in Cluster 4

The above graph displays the correlation between the age and maximum expenses in cluster 4. Most of the users in this cluster are between 18 to 30 yrs and 31 to 50 yrs old. Blue bar displays the age and orange is for the expenses. This graph represents that most of the participants are an older crowd and the average amount they spend while on a hike is 100 to 500\$.

Age Vs Activities:

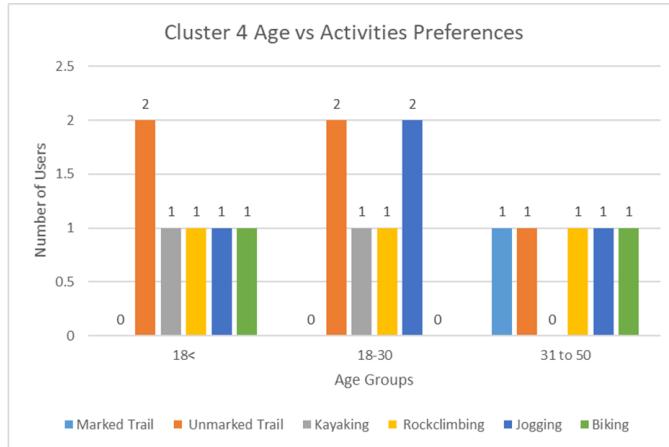


Figure B.24: Users Age and Activities Preference in cluster 4

In this graph, we are projecting each user preferred activity correlation with age in cluster 4. This analysis is conducted to understand if the user activities are influenced based on the age group they fall. Irrespective of their age this group of people most preferred activity is walking through backwoods, unmarked trails in rural areas and Jogging outdoors through a rural area. As per the graph, this group of people is more of outdoor activist because most of them selected all the activities. There is no age influence on the number of activities they choose because in cluster consist of 3 different age groups from 18 to 50 yrs.

Age Vs Gadgets:

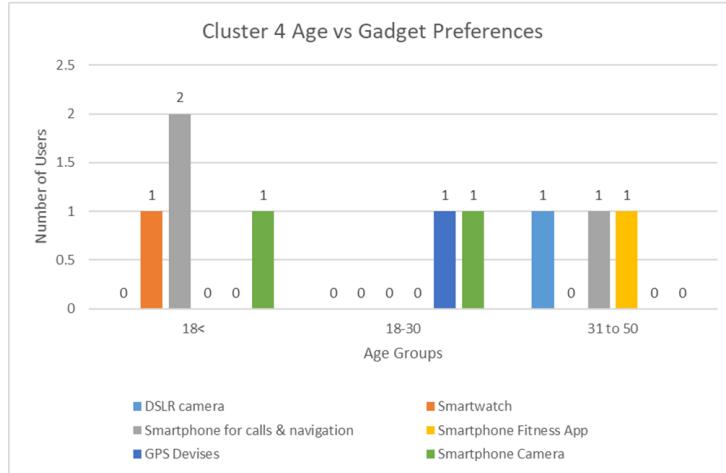


Figure B.25: Users Age and Gadget Preference in cluster 4

This graph represents the user ages and the tech they carry with them while on a hike. As observed gadget preference varies from the each of the age groups. Participants who are 18 to 30 yrs old did not choose smartphone for calls and navigation, instead they choose GPS system.

Activities Vs Expense:

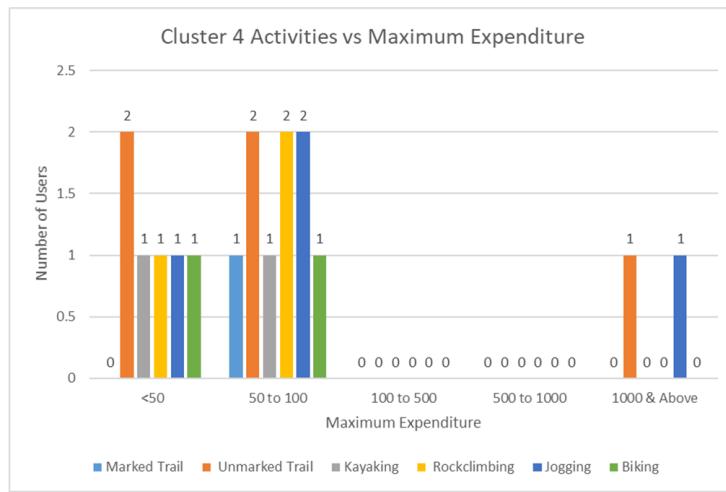


Figure B.26: Users Activities and Maximum Hike Expenditure in cluster 4

In this graph, we are representing each user activities and their maximum expenses while on a hike. As observed most of the participants who choose multiple activites choose to spend <50\$ and 50 to 100\$. People who like to go on hike in a unmarked trail said he will spend more than 1000\$.

Activities Vs Frequency:

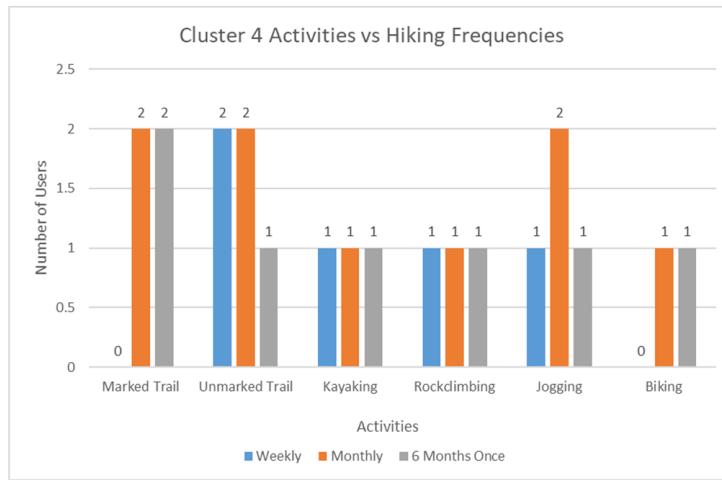


Figure B.27: Users Activities and Hike Frequency Preference in cluster 4

In the above graph, we are displaying the user preferred activities and hike frequency. This show how often these participants are going on a hike. All of the participants prefers to hike once in a week or monthly or 6 months once.

Activities Vs Motivation:

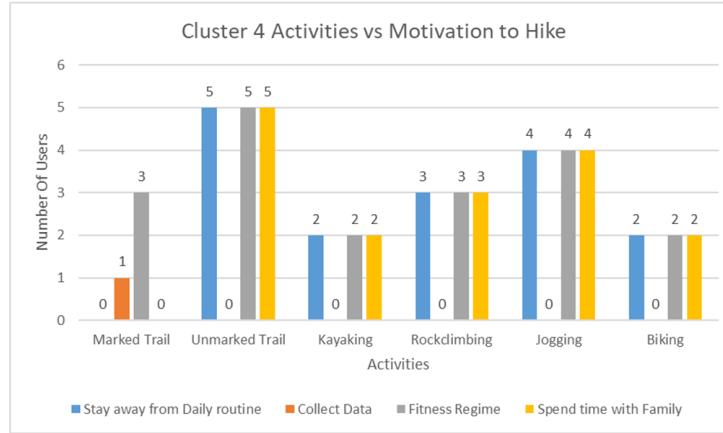


Figure B.28: Users Activities and Motivation to hike Preference in cluster 4

In the above graph, we are connecting the hiker's activities and their motivation to go on a hike. All of them said they like to go hiking to stay away from the daily routine and explore nature and spend time with family and friends, also as a fitness regime. This group of people are serious hikers who are fitness oriented and likes to spend some quality time in nature.

Activities Vs Gadgets:

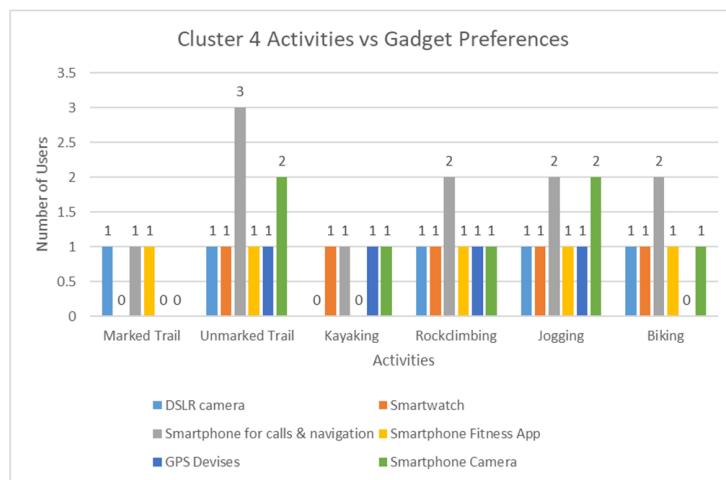


Figure B.29: Users Activities and Gadget Preference in cluster 4

In the above graph, we tried to understand the what is the users choice of gadget and their preferred activity. Most of them like to hike in a unmarked trail in rural areas and prefers like to use a smartphone for navigation and calls and to take pictures. Though they said they hike as a fitness regime they didnt choose smartwatch to monitor their fitness.

B.5 Individual Analysis Cluster 5:

Age, Activity and Gadgets:

This cluster consists of 4 users, 2 of them belong to age group 2 that is 18 to 30 yrs and 1 in 51 to 70 yrs old and 1 in <18 yrs old. This is one group where most participants are above 18 yrs older people. All of the users in cluster 5 preferred to go for a walk through backwoods, and unmarked trails through rural areas. 3 of the participants choose option 1 (DSLR camera or GoPro(or any camera not part of a smartphone)) as the preferred tech on the trail, and 2 of them choose option 2 (Smartwatch, Fitbit, etc. for Fitness), and 2 of them choose option 3 (Smartphone for just calls and Navigation apps like Guthook). These Gadget, Activites preferences of participants and the Age group they fall-in in cluster 5 shows that they are hikers who dont like smartphone much other than calls and navigation. Pictures are the reflection of our experiences and memories we make when we are traveling, especially when we are in nature walking on an unmarked trail in a rural area nature could be extraordinary to capture that these group of people prefer a more professional camera like DSLR and GoPros, etc.

Age vs Expense:

In this section an discussing the correlation between the Age and Maximum Expenditure in cluster 5. Irrespective of age difference 3 out of 4 participants choose option 1 (<50\$) and

participant 15 whose age is 50 to 70 yrs old and maximum money he likes to spend while on a hike is more than 1000\$. Other than him rest of the group is between 18 to 30 yrs old as we observed previous most the participants who are in this age group spend a less, i.e., maximum of 500\$. Very few of them said they spend like 500 to 1000\$.

Age Vs Motivation:

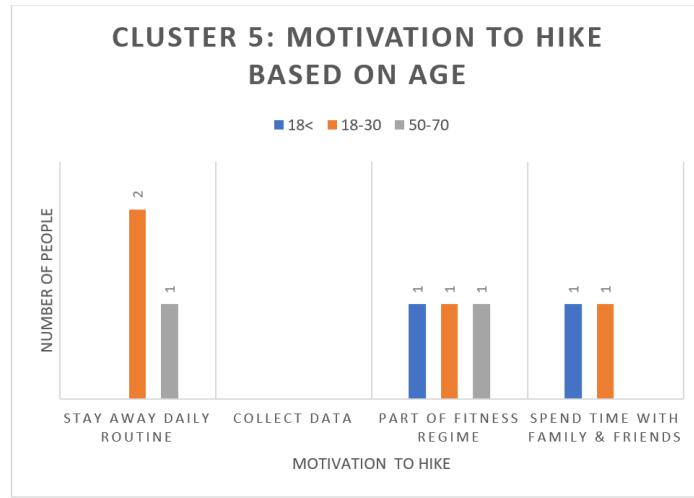


Figure B.30: Users Age and Motivation to hike in cluster 5

In this section am discussing the correlation between participants Age and the Motivation to hike. As we can see that none of the participants are motivated to hike to collect the data. The participant <18 years old says he want to go for a hike as a fitness regime and to spend time with his family and friends but not to stay away from the daily routine. However, an older group member said they like to stay away from daily routine along with fitness. The interesting observation is that participant 15 the oldest member of the group said he only like to hike to stay way on a daily routine and fitness but not spending time with family and friends.

Age Vs Gadgets:

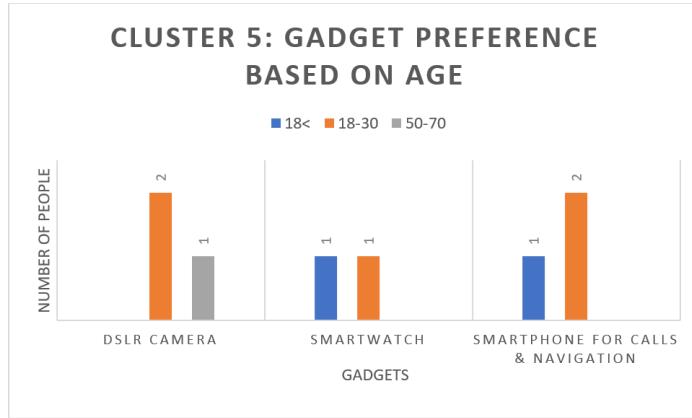


Figure B.31: Users Age and Gadget Preference in cluster 5

In this section, I am discussing the correlation between the Age and Gadget preference in Cluster5. The youngest member of this group is number 5, and his gadget preference is smartwatch to monitor his fitness and smartphone to navigate and make calls. On the other hand, the older member in this is number 15, and he likes to take only DSLR camera tech with him nothing else. However, this shows the vast differences between different age groups on this cluster.

Activities Vs Frequency: In this section am discussing the correlation between the Participants Activities and Frequency to hike. In Cluster 5 the only activity all of the participants choose is to go walking through backwoods, unmarked trails through rural areas and regarding frequency 3 of them preferred to hike monthly once, and one user said he likes to go for a hike once in a year. This characteristic states that this group goes for long walking hiking and once in a month like 3 to 4 days.

Discussions over Would you rather:



Figure B.32: Would you rather preference in cluster 5

In this section I am discussing the would you rather questions preference in cluster 5. As seen before 3 out of 4 participants choose to buy a DSLR so that a picture they take a hike goes viral instead of purchasing a smartwatch for fitness and look fit in a strangers picture in a hike that goes viral. However, user 5 the youngest of all choose the otherwise. When it comes to the data collection, 2 of them said they are passive data collectors with a smartphone and 2 for the active data collection using multiple gadgets, surprisingly participant 5 and 15 the youngest and the oldest choose the smartphone passive data collection. When it comes to the social media and expressing the hiking experience, 2 of them said they like to tweet on the spot and 2 of them said they like to go home and write a detailed blog about their hiking experience.

Appendix C

Questionnaire

1. Age:

- (a) <18
- (b) 18 to 30
- (c) 31 to 50
- (d) 51 to 70
- (e) 71 and Above

2. What is a hike to you?

- (a) Walking in marked trails in urban areas
- (b) Walking through backwoods, unmarked trails through rural areas
- (c) Kayaking / Canoeing
- (d) Rock climbing
- (e) Jogging outdoors through rural areas
- (f) Biking through rural areas
- (g) None of the above

3. Which of these define your hiking habits?

- (a) I like to go on hikes that take 2 to 3 days

- (b) I like to do thru-hiking that takes 3 to 6 months
- (c) I like to go on short hikes that take 4-5 hours
- (d) None of the above

4. How frequently you hike?

- (a) Weekly
- (b) Monthly
- (c) Once in 6 months
- (d) Yearly

5. What is the motivation for you to hike?

- (a) Stay away from daily routine and explore nature
- (b) Collect data about plants
- (c) A part of your fitness regime
- (d) Spend time with people (family, friends or strangers)
- (e) None of the above

6. What is the maximum amount of money you have spent for a hike?

- (a) <50
- (b) 50 to 100
- (c) 100 to 500
- (d) 500 to 1000
- (e) 1000 and above

7. Which of the following gadgets would you prefer to carry for a hike?

- (a) DSLR camera, GoPro (or any camera not part of a smartphone)
- (b) Smartwatch, Fitbit, etc. for fitness
- (c) Smartphone for just calls and navigation apps like Guthook
- (d) Smartphone for just calls and navigation apps like Guthook
- (e) GPS device like Garmin eTrex, etc. (separate from smartphone)
- (f) No separate camera, just a smartphone camera G. None of the above, only paper map

8. Would you rather.

- (a) Buy a DSLR so that a picture you take a hike goes viral?
- (b) Buy a smartwatch for fitness and look fit in a strangers picture in a hike that goes viral?

9. Would You rather.

- (a) Hike in 35 F but sunny weather and use solar energy for charging devices?
- (b) Hike in 55 F but rainy weather and carry heavy batteries for charging devices?

10. Would you rather.

- (a) Make use of multiple gadgets like DSLR to take pictures or GPS device to locate poisonous plants and contribute to data actively?
- (b) Just use a smartphone and collect data passively through an app while you hike?

11. Would you rather.

- (a) Tweet about an interesting spot in real time during your hike?
- (b) Write a detailed story in a hiking group on Facebook after finishing the hike?

12. Would you rather.

- (a) Use an offline map offered by a water-resistant smartphone app but has charging issues?
- (b) Prefer to carry a paper map that is more prone to damage?