2020

Exposome Factors: Exploratory Study Approach and the Role of Persuasive Technology to Raise Awareness About the Exposome Concept

Abdulrahman Alzahrani
Claremont Graduate University

Exposome Factors: Exploratory Study Approach and the Role of Persuasive Technology to Raise Awareness about the Exposome Concept

By

Abdulrahman Alzahrani

Claremont Graduate University

2020

APPROVAL OF THE DISSERTATION COMMITTEE

This dissertation has been duly read, reviewed, and critiqued by the Committee listed below, which hereby approves the manuscript of Abdulrahman Alzahrani as fulfilling the scope and quality requirements for meriting the degree of Doctor of Philosophy (Ph.D.) in Information Systems and Technology.

Lorne Olfman, Co-Chair

Claremont Graduate University

Professor

Brian Hilton, Co-Chair

Claremont Graduate University

Professor

Yan Li, Member

Claremont Graduate University

Professor

Abstract

Exposome Factors: Exploratory Study Approach and the Role of Persuasive Technology to Raise Awareness About the Exposome Concept

by

Abdulrahman Alzahrani

Claremont Graduate University, 2020

Exposome is a new concept that emerged at the beginning of the 21st century to complement genome research. The exposome encompasses the entirety of a person's environmental exposures from birth to death. This study focused on environmental rather than genetic factors that are related to chronic diseases. It had two phases. Phase 1 involved building a regression model aimed at investigating the influence of various indicators on cancer cases in Los Angeles County. The investigation of the potential correlation between the cancer cases based on census tract data and the exposome factors follows an exploratory approach. Multiple regression stepwise analysis (using SPSS) are reported and discussed. Phase 2 aimed to design, build, and evaluate a mobile-based and text message awareness campaign artifact based on the results of the regression model built in Phase 1. I utilized the Geller proenvironmental model to govern the design requirements of the proposed artifacts to educate people on the exposome factors resulting from the regression model that correlated with cancer cases in Los Angeles County. Novel persuasive technology techniques were utilized within the system artifact. This technology uses direct and indirect persuasion routes for delivering interventions.

First, the mobile application indirectly persuades users through providing educational videos and an e-fotonovela. Second, text messages directly persuade users by providing supportive tips. A one-way between subjects ANOVA was conducted to compare the effect of three methods of persuasion on presenting information to increase awareness. A significant effect was present at the $p < .05$ level for the three conditions [$F(2)= 6.056$, $p = 0.007$]. A post hoc Tukey test indicated that Group A and Group B differed significantly at $p < .05$; and Group A and Group C differed significantly at $p < .05$; Group C was not significantly different from GroupB. This study contributed to the body of knowledge through providing a solution that aims to raise the awareness level about the concept of exposome.

DEDICATION

To my dear parents, Ahmed and Soda

&

My lovely wife, Rehan, and wonderful daughters, Rossol and Rateel

&

My sister and brothers

Acknowledgment

My PhD journey was full of challenges, experience, lessons, and achievements. I would not have achieved my goal without help and support from faculty members of CISAT. First, I would like to acknowledge my previous committee chair, Dr. Samir Chatterjee. I thank you for all the guidance, insights, and feedback. You helped me to develop the idea of this study. I really appreciate that I worked with you to defend my proposal and I am thankful for your supervision. Second, I am thankful to my committee member cochairs, Dr. Lorne Olfman and Dr. Brian Hilton. Dr. Olfman, you supported me in my worst time and stand by my side and embraced me with love and kindness. You were always available to answer my questions. I am thankful that you were there for me and I acknowledge your work. Dr. Hilton, you dedicated your time listening to me and you guided me throughout this journey. I acknowledge your help and support. I am also thankful to the committee member Dr. Yan Li for working with me in a short period of time to finish this journey. I really appreciate your valuable feedback in developing the first artifact of my dissertation. Your inspiration guided me to learn more and achieve this moment and I acknowledge your work. Third, I am thankful to Dr. June Hilton for her help and support. You helped me with your insight and analysis of the model developed in this study. I acknowledge your help and support. Fourth, I acknowledge the work of Carolina Fautsch and Chris Hunter. Thank you, Carolina, for helping me develop the script for e-fotonovela. Thanks, Chris, for offering your help to develop the script into comic pictures. Fifth, I acknowledge the work of Homud Alhugbani. Thanks, Homud, for helping me design and develop the application. I am thankful for your patience and help in testing the application functionality and fixing bugs.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

A substantial amount of literature addresses chronic disease causes and preventions. A summary of the current state of chronic diseases in the United States can be found in Appendix A. Most causality research has focused on genetic factors as primary causes and environmental exposure as secondary. Rappaport (2016) conducted a study on the number of citations from PubMed using the keywords *disease causes* and *genetics* and compared it to the number of citations with the keywords *disease causes* and *exposure.* Articles with the keywords *disease causes* and *genetics* were cited 566,685 times compared to 71,922 articles with the keywords *disease causes* and *exposure.*

The role of genomics is relevant to the prevention of chronic diseases (Johnson et al., 2005; Bodzin et al., 2005); however, genetic variations do not fully explain potential chronic disease risk, which leaves a high possibility that environmental exposures play a large role in cause. Hence, the *exposome* concept plays an important role in the assessment of lifelong exposure history (Siroux et al., 2016a): “At its most complete, the exposome encompasses life-course environmental exposures (including lifestyle factors), from the prenatal period onwards” (Wild, 2005, p. 2). The human exposome or environmental (non-genetic) exposures have three overlapping domains: internal, specific external, and general external. It is widely known that a small part of chronic disease can be explained by genetic factors alone; however, environmental exposure information is important in broadly evaluating chronic diseases. The exposome must be measured at repeated points during a lifetime, such as during pregnancy, infancy, adulthood, and aging.

Most environmental health studies have limitations. Prior studies have primarily focused on a single time point to assess environmental exposure, thus minimizing the effect of the change of exposure risks over time. Furthermore, regulation agencies such as the Occupational Safety and Health Administration and the Environmental Protection Agency have shifted the focus of exposure research toward chemical assessment and compliance and ignored the human health side of exposure research. Between 2006 and 2010, 491 papers were accepted for publication in the Exposure Science and Environmental Epidemiology journal; only 40 papers, or 8%, involved health effects, while the rest focused on chemical regulations (Rappaport, 2011). Moreover, exposure scientists have shifted their focus away from empirical measurements toward probability models based on observational data.

Despite these limitations concerning causes and prevention of chronic disease, one significant opportunity to impact health-related behaviors is made possible by the increasing pervasiveness of smartphones (Chen et al., 2018). The percentage of Americans who own a smartphone increased from 60% in 2014 to 77% in 2016 (Smith, 2017), with mobile applications (apps) projected in 2017 to comprise 57% of digital media usage (Lella & Lipsman, 2017). In March 2018, Google Play offered 3.8 million apps for Android devices and the Apple Store offered about 2 million apps for IOS devices (Statista, 2018). Health apps such as mHealth were estimated to reach 325,000 apps in 2017 (Larson, 2018); therefore, smartphones play an important role in patient education, disease self-management, and the remote monitoring of patients (Mosa et al., 2012), with both patients and healthcare experts increasingly relying on mobile apps to raise health awareness and promote a healthy lifestyle.

In this research study, I used the design science research (DSR) approach developed (Hevner & Chatterjee, 2010) to design, build, and evaluate a regression model that highlights the most significant exposome factors that correlate with cancer incidents in Los Angeles County. The results were used to develop a mobile application that intended to increase awareness of exposome factors within the common population of Los Angeles County. I developed a text-messaging intervention that provided tips and shocking facts about exposome factors in the environment to persuade users to take advantage of the application. The DSR method has three iterative and interrelated cycles: relevance, design, and rigor. The developed information technology (IT) artifacts went through all three DSR cycles to produce a novel regression model and an application that can contribute to the scientific knowledge base.

1.1 Problem Statement

The U.S. National Center for Health Statistics defined chronic disease as “a disease that persists for a long time. A chronic disease is one lasting 3 months or more and cannot be prevented by vaccines or cured by medication, nor do they just disappear” (Schiel, n.d.). Chronic diseases—including heart disease, stroke, cancer, enduring respiratory diseases, and diabetes—are responsible for 60% of all deaths globally (World Health Organization [WHO], n.d.). Eighty percent of chronic disease deaths occur in low-income countries and the remaining 20% of chronic disease deaths occur in high-income countries (WHO, n.d.). Twenty percent is considered high for developed countries where individuals have access to health care providers; this high percentage will lead to greater costs if nothing is done to slow the growth rate of chronic diseases (Nugent, 2008). The United States has a higher rate compared to other high-income countries, with 50% of

Americans living with at least one chronic disease (Raghupathi & Raghupathi, 2018). Chronic diseases are associated with age; thus, 88% of Americans over 65 years of age have a chronic disease.

Chronic diseases are caused by genetic factors, environmental factors, or both. The majority of chronic disease deaths are caused by environmental factors rather than genetic factors (Sainani, 2016); however, genetic variability seems to be even less important when it comes to the cause of chronic disease. Genetic differences account for only 10% to 20% of chronic diseases (Rappaport, 2011; Rappaport & Smith, 2010).

Wild (2005) coined the term exposome to address the problem of environmental exposure and develop a new perception about environmental exposure: “At its most complete, the exposome encompasses life-course environmental exposures (including lifestyle factors), from the prenatal period onwards” (p. 2). The term exposome was coined to match the term “genome;” however, raising people’s awareness about the exposome concept is still in its infancy. This study focused on the factors that are significantly correlated with the incidence of cancer in Los Angeles County to raise the awareness level among the Los Angeles County population.

Another significant problem is the lack of information regarding the development of mobile applications to raise awareness of environmental exposure. Major studies have focused on providing mobile applications that are dedicated to increasing awareness about the diseases themselves; however, these studies do not address the causes of the disease. For example, Quinn et al. (2008) conducted an experiment on 30 patients with Type 2 diabetes—which is caused by environmental factors—to assess the impact of using a mobile phone system for managing diabetes. Quinn et al. found a statistically significant improvement in patients who used the mobile phone system to manage the disease; however, the system did not raise awareness about the core environmental causes or root causes of diabetes. In addition to this, Bender et al. (2013) conducted a review to categorize the purpose and content of 295 cancer-focused smartphone apps. Bender et al. found that just 32% of apps were dedicated to promoting awareness about cancer or providing cancer educational content. However, Bender et al.'s review did not address the source of the diseases.

The use of mobile applications plays an important role in promoting a healthy lifestyle and improving patients' health. With the widespread adoption of smartphones worldwide, more people are on their phones than before. Smartphones have become an ideal platform to reach the population and raise awareness; however, no applications are designed to address the exposome as a concept or to increase people's awareness of it. The goal of the current study is (a) to fill the gap in literature on the lack of exposome awareness and (b) to design an application that uses the design science principles developed by Hevner and Chatterjee (2010) to increase awareness about the exposome. This study also examined the effectiveness of a persuasive technology using text messages as a direct persuasive approach in motivating and facilitating healthy behavioral change. I aimed to use my application along with text messages to raise awareness about exposomes, the causes of chronic diseases, and the harmful effects of our environment and processed food.

The structure of the dissertation is as follows. First, I review the current literature that relates to the exposome, and provide its definition, domains, pathways, and measurements. Second, I review the literature on the role of Information Technology

(IT)—specifically mobile applications—in the health care sector and examine the role of persuasive technologies in designing processes to change human behavior. Third, I present the theoretical grounding drawn from the human-computer interaction field, which informed the design and evaluation of the exposome mobile application and text message campaign. Fourth, I discuss the research methodology and the DSR approach and explain how DSR guided the development of the regression model and exposome mobile application. Fifth, I explain the methods used to design, build, and evaluate the proposed artifacts. Sixth, I discuss research limitations and provide a conclusion. The research contributions to scientific knowledge and society are presented in the last section.

1.2 Objectives and Research Questions

One goal of the current study is to aid others in understanding the exposome and to design, develop, and evaluate a regression model using an exposome data set to examine the exposome factors that are significantly correlated with the incidence of cancer in Los Angeles County. The second goal is to design, build, and evaluate a mobile application to raise awareness about the significant factors of the exposome that resulted from the first goal. The third goal is to use persuasive technologies to educate users through video and graphic stories using an e-fotonovela and to conduct a text message campaign to send tips and shocking facts about exposome factors. The following objectives were defined to help me achieve these goals.

- Give the exposome definition, domains, pathways, and measurements.
- Give persuasion a definition, levels, strategies, and technology.

- Design, build, and evaluate a regression model to investigate the potential correlation between cancer cases in Los Angeles County as the dependent variable and geosocial data that includes environmental measurements and Esri's consumer data as the independent variables.
- Design, build and evaluate an awareness artifact that educates the people of Los Angeles County about the exposome.

The research questions that guide this research are as follows.

RQ1. What set of the following factors best correlates with cancer cases in Los Angeles County: ozone diesel, Tox_release, traffic, Avg_fast_food, Avg_Alcoholic_Beverages, Avg_Canned_Fruit, Avg_Canned_Vegetable, Avg_Canned_Beans, Avg_Processed_Fruit, Avg_Processed_Vegetables, Hispanic, White, Asian_American, Other_population, education, poverty, and Avg_Income?

RQ2. Can a mobile application that is designed and developed to raise people's awareness level of the exposome be effective?

Chapter 2: Background and Literature Review

2.1 The Exposome

The following sections focus on the exposome.

2.1.1 What is The Exposome?

The exposome, coined to match the concept of genome, is a new paradigm for studying the effect of the environment on human health. A genome is defined as follows. The complete set of genetic information in an organism. It provides all of the information the organism requires to function. In living organisms, the genome is stored in long molecules of DNA called chromosomes. Small sections of DNA, called genes, code for the RNA and protein molecules required by the organism ("Genome," n.d.).

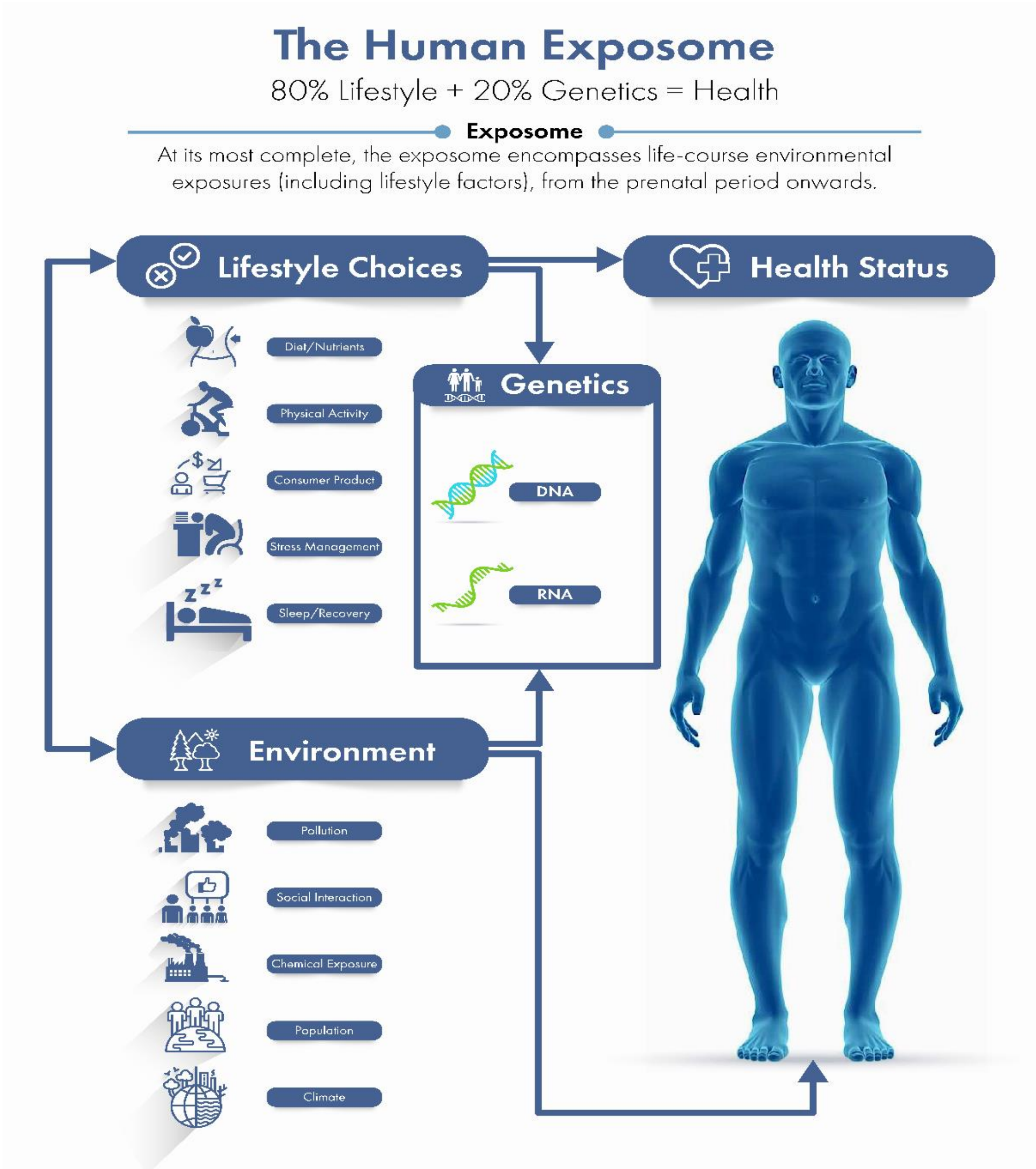
The exposome complements the genome by providing a comprehensive description of lifelong exposure history.

In 2005, Christopher Wild published a paper titled "Complementing the Genome with an 'Exposome:' The Outstanding Challenge of Environmental Exposure Measurement in Molecular Epidemiology." Wild coined and defined the term exposome: "At its most complete, the exposome encompasses life-course environmental exposures (including lifestyle factors), from the prenatal period onwards" (Wild, 2005, p. 2; see Figure 1). Rappaport and Smith (2010), who called for a comprehensive quantitative approach to investigate environmental exposure and discover the causes of chronic disease, refer to an exposome as "the totality of environmental exposures from conception onwards" (Rappaport & Smith, 2010, p. 2). The Human Exposome Project (n.d.) added that "the exposome, conceptually and practically, provides a holistic view of human health and disease. It includes exposures from our diets, our lifestyle, and our behaviors. It also includes how our bodies respond to these challenges" (p. 1). For example, food generates several chemicals in our bodies from the processing and digesting of food by gut bacteria. Alcohol consumption or smoking tobacco products also produce chemical changes that are associated with different cancer types. Other lifestyle factors, such as stress and lack of physical activity, can increase the risk of chronic disease.

The exposome can be defined as the measure of all the exposures of an individual in a lifetime and how those exposures relate to health (National Institute for Occupational Safety and Health [NIOSH], 2014). An individual's exposure begins before birth and includes health risks from environmental and occupational sources. The exposome is defined as the understanding of how exposures from our environment, diet, and lifestyle interact with our own unique characteristics—such as genetics, physiology, and epigenetics—to ultimately impact our health (see Figure 1).

Figure 1

The Human Exposome

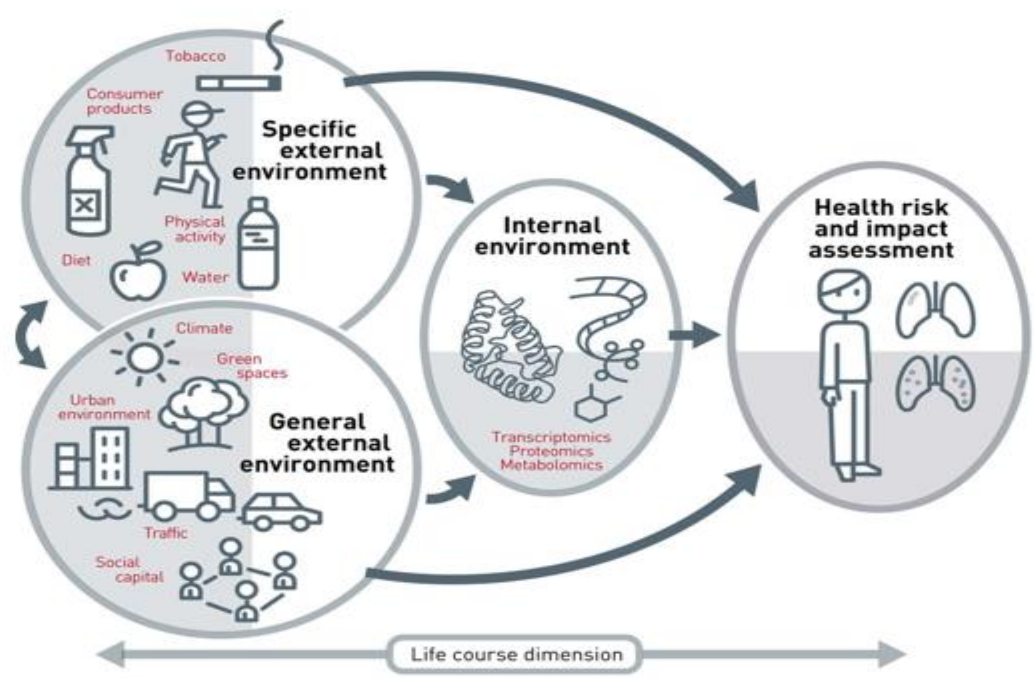


2.1.2 Exposome Domains

The human exposome or environmental (nongenetic) exposures have three overlapping domains: internal, specific external, and general external (Wild, 2012; see Figure 2). The internal environment of the human body refers to biological factors "such as metabolism, hormones, body morphology, physical activity, gut bacteria, inflammation, lipid peroxidation, age and stress" (Wild, 2012, p. 1). Specific external exposures on the individual level include radiation, viral or bacterial infection, chemical contaminants and pollutants, diet, lifestyle, and drugs. General external exposure at the population level includes a wider range of factors such as social capital, education, financial status, psychological and mental stress, urban environment, and climate. It is also important to highlight the environment of early life, namely the exposure of a fetus inside the mother's body.

Figure 2

Exposome Domains



Note. Adapted from M. Vrijheid (2014). The exposome: A new paradigm to study the impact of environment on health. *Thorax, 69*(9), 876–878 (https://doi.org/10.1136/thoraxjnl-2013-204949).

2.1.3 Exposome Sub-domains

The human exposome concept has been extended and developed to include sub-domains, which include the following:

Eco-exposome. According to the National Research Council of the National Academies (2012), "exposure science extends from the point of contact between stressor and receptor inward into the organism and outward to the general environment, including the ecosphere" (p. 12). Eco-exposome includes all environmental exposures to humans, such as air, food, water, and consumer products (Lioy & Smith, 2013).

Adverse outcome pathways (AOP) exposome. The AOP concept is a sub entity of the exposome domain that emphasizes the understanding the chemical toxicity mechanism of interactions when our bodies are exposed to chemicals. The AOP and exposome concept can be integrated to map a biological pathway from an initial chemical interaction to an adverse outcome (Escher et al., 2017).

Indoor exposome. Dai et al. (2017) introduced *indoor exposome* as a sub-domain of the human exposome. Dai et al. advocated for a holistic understanding of the exposome in the man-made environment—which includes residences, workplaces, and public buildings—by considering three important components: indoor air, water, and surfaces.

Pollutome exposome: Landrigan et al. (2018) introduced the term *pollutome,* defined as "the totality of all forms of pollution that have the potential to harm human health" (Landrigan et al., 2018, p. 7). The pollutome concept can be viewed as a subset of the exposome domain with three zones. Zone 1 includes known health effects of known pollutants, like the association between airborne pollutants and noncommunicable diseases. Zone 2 includes the unknown health effects of known pollutants, like pollution from industrial sites and its relationship with diseases. Zone 3 includes unknown health effects from new pollutants.

2.1.4 Characterizing the Exposome

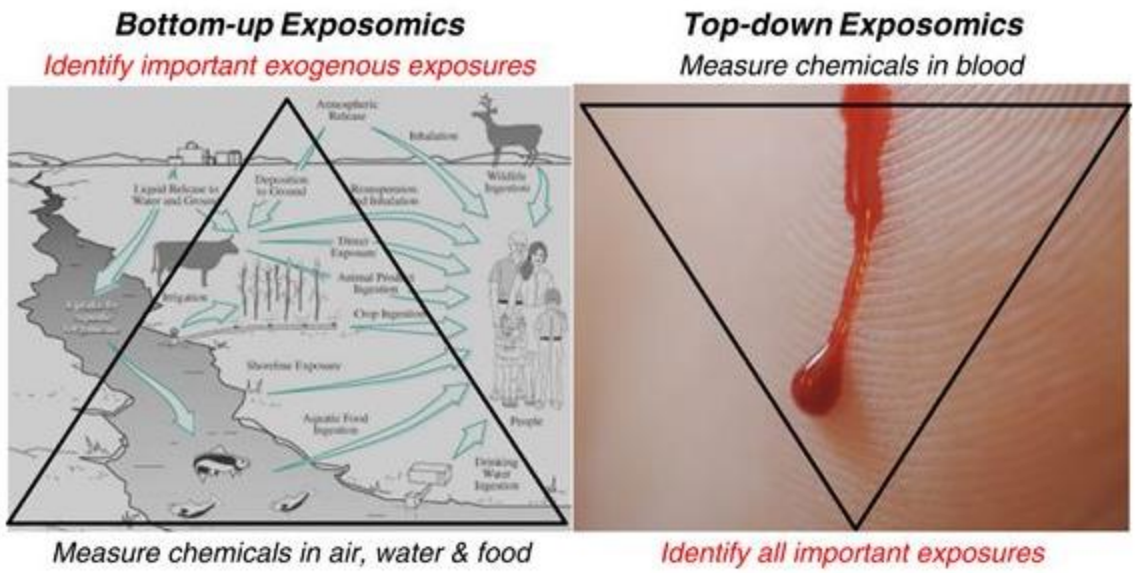
Two generic strategies can be used to characterize the exposome: bottom-up and top-down (Rappaport, 2011; see Figure 3). In the bottom-up approach, the focus is on identifying external exposure—such as air, water, and lifestyle—and assessing the level of exposure. This approach is effective in providing information about sources and environmental exposure levels; however, this approach does not address internal exposures and requires a huge effort to evaluate and analyze different external media and sensors.

The top-down approach focuses on biological monitoring through repeated sampling of blood, urine, or other tissues over a lifespan. This approach covers both external and internal chemical exposures to the human body. The top-down approach is a more efficient approach compared with the bottom-up approach; however, the top-down approach does not provide information about the source or level of exposure.

van Tongeren and Cherrie (2012) suggested a third approach for characterizing the exposome. They advocated for an integrated approach that measures the exposome by considering all available data on both internal and external exposures, as well as data on behavior patterns collected routinely from newly developed sensors. However, the ability to fully characterize the human exposome remains a challenge (Escher et al., 2017).

Figure 3

Characterizing the Exposome



Note. Adapted from S. M. Rappaport (2011). Implications of the exposome for exposure science. *Journal of Exposure Science and Environmental Epidemiology*, *21*(1), 5–9 (https://doi.org/10.1038/jes.2010.50).

2.1.5 Exposome Pathways

The Agency for Toxic Substances and Disease Registry (ATSDR; 2005) has defined an exposure pathway as “the link between environmental releases and local populations that might come into contact with, or be exposed to, environmental contaminants” (ATSDR, 2005). A complete exposure pathway has five elements:

1. Source: A source is the origin of environmental contamination, such as chemical spills or leaks, and open burn areas.
2. Mechanism: A mechanism is the means of carrying chemicals from the source to different media such as air, water, and soil.

3. Exposure point: The exposure point is the point of human contact with the hazard or chemicals.

4. Exposure route: The exposure route is the means of chemical entry into the human body.

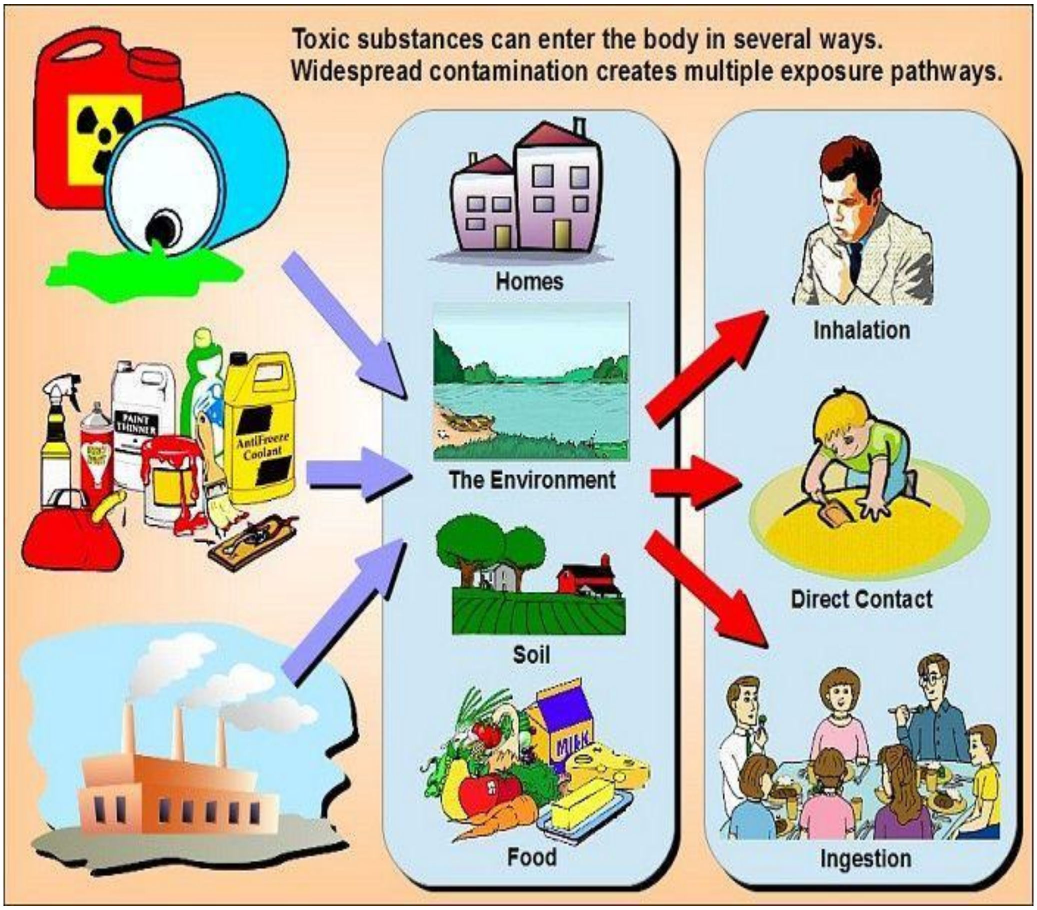
5. Exposed population: The exposed population is the human population that may potentially come into contact with contaminants (ATSDR, 2005).

The ATSDR (2005) listed common routes for human exposure to contaminants (see Figure 4). The common routes are as follows.

- Inhalation: breathing in of chemicals.
- Ingestion: consumption of contaminated food or water.
- Dermal: skin exposure to a contaminant.
- Direct external exposure**:** exposure to outside stressors, such as radiation, heat and noise.

Figure 4

Exposure Pathways



Note. Adapted from W. Sawyer (2010). *Toxic exposures.* Toxicology consultants and assessment specialists (http://experttoxicologist.com/toxicology-toxic-exposures.aspx?cln=1).

2.1.6 Measuring the Exposome

The assessment of the human exposome depends on a variety of tools and technologies. Measuring the exposome requires adapting conventional approaches and methods and combining these approaches with advanced technologies such as the biomarker approach (see Figure 5). The National Institutes of Health Biomarkers Definitions Working Group (2001) defined biomarker as "a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention" (p. 2). Measuring the internal exposome depends on the field of study; these fields of study include genomics, metabonomics, lipidomics, transcriptomics, and proteomics (NIOSH, 2014). The external exposome can be measured using conventional devices such as portable or personal monitors for measuring air pollution through environmental sensors, physical activities, and vital signs. New technologies and devices such as smartphones or wearable sensors that track location through GPS, monitor stress, count steps, and determine dietary intake can be used to measure the external exposome (Siroux et al., 2016a; Turner et al., 2017; Wild, 2012). See Table 1.

Table 1

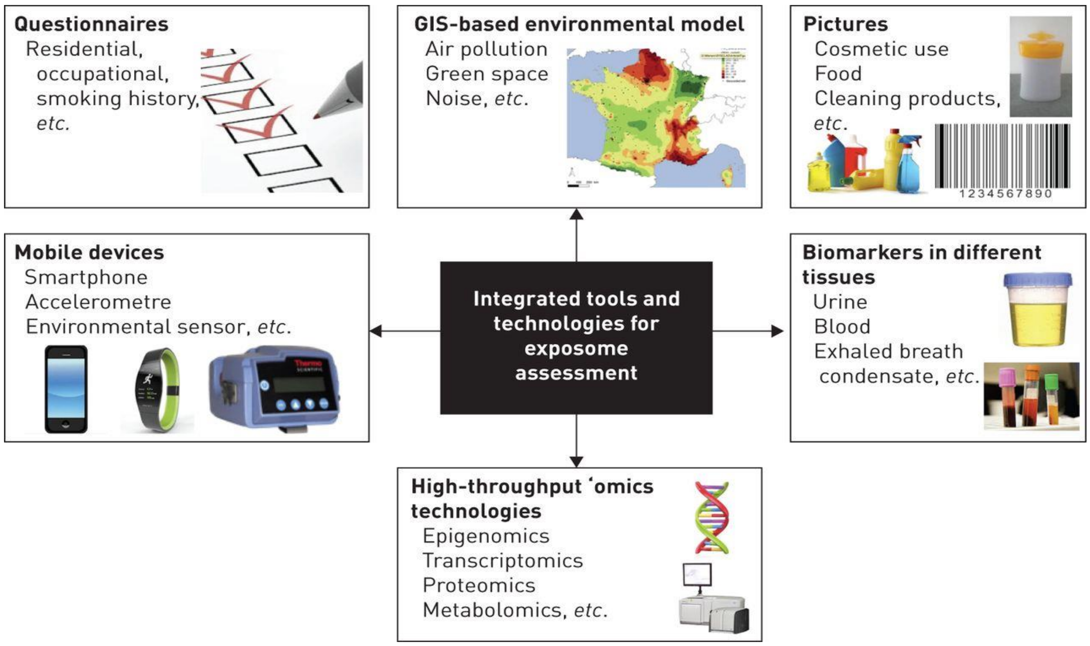
Approaches and Tools for Measuring the Exposome

Exposome Domain	Approach	Tools
Internal	Biomarkers	Omics technologies
General External & Specific External	Sensor technologies (including mobile phones)	Environmental pollutants, physical activity, stress, circadian rhythms, location [global positioning systems (GPS)]
General External & Specific External	Imaging (including mobile phones, video cameras)	Diet, environment, social interactions
General External & Specific External	Portable computerized devices (including palmtop computers)	Behavior and experiences (ecological momentary assessment), stress, diet, physical activity
General External & Specific External	Improved conventional measurements (combined with environmental measures)	Job-exposure matrices; dietary recall (e.g., EPIC-Soft)

Note. Adapted from C. P Wild (2012). The exposome: From concept to utility. *International Journal of Epidemiology, 41*(1), 24–32 (https://doi.org/10.1093/ije/dyr236).

Figure 5

Measuring the Exposome



Note. Adapted from V. Siroux, L. Agier, & R. Slama (2016). The exposome concept: A challenge and a potential driver for environmental health research. *European Respiratory Review, 25*(140), 124–129 (https://doi.org/10.1183/16000617.0034-2016).

The current study focused on using external exposome domains and Los Angeles County general exposome and specific general exposome data to build a regression model that highlighted the factors that correlated significantly with incidents of cancer cases and raise the awareness level of people of Los Angeles County.

2.2 Persuasion

Philosophers and scholars have studied persuasion intensively throughout history. Aristotle, an ancient Greek philosopher, is well known in the field of rhetoric. Rhetoricians used to give public speeches aimed at influencing and persuading their audiences. Rigorous studies on persuasion continued to evolve during the 1900s, especially in the field of social science; however, scholars and researchers have not agreed on a single definition of persuasion. Jones and Simons (2017) defined persuasion as “human communication designed to influence the autonomous judgments and actions of others” (p. 7). Reardon (1991) includes a change concept in his definition and defined persuasion as “the activity of attempting to change the behavior of at least one person through symbolic interaction” (p. 3). Other scholars have viewed persuasion as a means of changing, shaping, or reinforcing behavior (Stiff & Mongeau, 2016). I have adapted Fogg’s (2003) definition of persuasion for this study; Fogg stated that persuasion is “an attempt to change attitude or behavior or both (without coercion or deceptions)” (p. 1). It is important to differentiate persuasion from coercion and deception. Coercion implies force and deception indicates some form of misleading behavior, whereas persuasion suggests a voluntary change in behavior or attitude (Fogg, 2003). Fogg’s persuasive systems can be applied to other fields, such as education, safety, environmental preservation, occupational effectiveness, and health care.

Persuasion is considered a key factor in changing attitudes and behaviors. Scholars have studied the persuader, persuasive message, and persuadee. The way in which information systems influence people’s attitudes and motivate changes in behavior has received increased attention in recent literature. Theories and models from social psychology and other fields have been adapted to predict user behavior. Teng et al. (2015) critically reviewed five major behavior change theories to identify trends of applying behavior change to social media marketing. The review assessed cognitive dissonance theory (CDT), social judgment theory (SJT), the heuristic-systematic model (HSM), the elaboration likelihood model (ELM), and the theory of planned behavior (TPB).

CDT. Cognitive Dissonance Theory was developed by (Festinger, 1957). Festinger argued that people who experience dissonance between belief and behavior are motivated to reduce their discomfort level by altering beliefs, actions, or perceptions of actions. CDT offers an interpretation and prediction of human behavior according to "the moral, legal and social values of society" (Freedman, 1965, p. 146). However, Festinger acknowledged that his theory does not provide accurate predictions. Other scholars have criticized CDT because it cannot provide either a valid measurement of dissonance or the level of dissonance (Teng et al., 2015).

SJT. Social Judgment Theory was articulated by M. Sherif, C. Sherif, and Hovland in the 1960s. SJT proposes that understanding a person's attitude is the key to persuasion (Teng et al., 2015). SJT explains when persuasive messages are most likely to succeed and how people make judgments about persuasive messages. SJT suggests that changing attitudes is a 2-step process. A person who receives a persuasive message will either accept or reject the message based on their knowledge and judgment by comparing the persuasive message with their preexisting attitudes. Second, this person will make a judgment or assessment about the message. The Ordered Alternatives Questionnaire was developed to determine people's positions on certain issues by listing statements that represent different perspectives on a subject. The questionnaire represented three measurements of judgments on possible positions: latitude of acceptance, latitude of rejection, and latitude of non-commitment. SJT offers a framework for understanding attitudes and behavior; however, SJT has been criticized for its inability to measure correlations between ego-involvement variables (O'Keefe, 2002) that explain how the issue is important to the individual (Teng et al., 2015).

HSM. The Heuristic-Systematic Model was developed by Chaiken in the 1980s to explain how people process persuasive messages. HSM suggests that our thoughts about the world are shaped based on two factors: the motives being pursued and the way information is processed. The basic assumption regarding HSM is that “when not sufficiently motivated to engage in effortful processing, the default will be to use less effort, to follow an information processing strategy based on simple rules, schemas, and prior knowledge–what is known as heuristic processing” (Bohner et al., 1995, p. 34). Heuristics are socially shared, result from people’s experience, and have some empirical validity (Oinas-Kukkonen & Harjumaa, 2008b). HSM suggests that information can be processed in two ways: (a) through a systematic processing mode that requires deep analysis, assessment, and understanding of available information; and (b) through a heuristic processing mode that involves applying heuristic cues (Oinas-Kukkonen & Harjumaa, 2008a).

ELM. The Elaboration Likelihood Model was developed by Petty and Cacioppo in the 1980s. ELM addresses the process of persuasion. ELM is a general theory about changing attitudes that has a dual-process model in which people are processing information via a central route or a peripheral route (Teng et al., 2015). People who use the central route think critically about the persuasive messages before judging the messages. People who use the peripheral route do not think critically about the persuasive messages but may be persuaded by simple cues or rules of thumb. Ability and motivation are important factors in determining the route of persuasion. People are more likely to be persuaded through the central route if the message is given to an audience whose members are highly able and motivated to process the information. People are less likely to be persuaded if the audience is less able or unwilling to process the persuasive message; however, these individuals may be persuaded through the peripheral route. ELM has been criticized by scholars for misalliance between theory and data (Teng et al., 2015) and inability to describe how humans process information (Stiff, 1986).

TPB. The Theory of Planned Behavior was articulated by Fishbein and Ajzen in the 1980s. TPB mainly focuses on predicting and understanding the influences of motivation on behavior (Teng et al., 2015). According to TPB, human behavior can be predicted through examining intention and perceived behavioral control. People who are strongly motivated are more likely to change their behavior (Teng et al., 2015); the TPB refers to this strong motivation as intention. Perceived behavioral control refers to how difficult one thinks it is to perform the behavior. TPB provides a useful framework for predicting and understanding behaviors (Teng et al., 2015); however, TPB has been criticized for its weakness in predicting the gap between intention and actual behavior (Teng et al., 2015).

2.2.1 Levels of Persuasion

One of two forms of persuasion can be used to change attitudes and behavior: macrosuasion and microsuasion (Fogg, 2003).

Macrosuasion. Technology products at the macrosuasion level of persuasion are designed with the intent of persuading and influencing users. For example, Baby Think It Over is a high-tech doll designed for school parenting programs to influence teenagers not to become parents. The doll has an imbedded computer chip that is programmed to have the doll cry at different times for different periods of time. The caregiver—in this case, a teenage girl or boy—must carry the doll everywhere and pay attention to it by holding it and inserting a key when the baby cries. A report will be issued if the caregiver neglects the doll. This report will be shown to the teacher at the end of the learning experience, which lasts for a few days. The core purpose of this doll is to increase students' awareness by showing them how their lives would be impacted by having a baby at their age. The doll is designed to motivate the students to avoid becoming teen parents.

Microsuasion. Computing products at the microsuasion level of persuasion do not have an overall intent to persuade. Instead, these computing products include minor persuasive elements to achieve overall goals. Microsuasion elements can be designed as dialogue boxes, icons, or interaction patterns between the computer and the user, such as praising users and sending reminders or visualizations. The eBay rating system feedback is an example of microsuasion, where the seller and buyer review each other after a transaction is completed. This is designed to motivate people to be honest, responsive, and courteous in their interactions.

2.2.2 Persuasion strategies

Many approaches can be used in the process of persuasion; persuasion processes can be classified into two groups: (a) processes that focus on the content of the persuasive messages and (b) processes that focus on the characteristics of persuasive situations (Stiff, 1986). Two main strategies can be used for persuasion: direct and indirect. Direct persuasion is defined as "persuasion that has clear and apparent intentions" (Aleahmad et al., 2008, p. 2) and indirect persuasion is defined as persuasion that "does not clearly expose its own position, confront or condemn users' existing attitudes, or adopt an identity typical of people who already agree with the message. Indirect persuasion should incur less resistance from users" (Aleahmad et al., 2008, p. 2).

As described earlier, the ELM is a general theory of attitude change in which the fundamental idea is that two routes to persuasion exist: a central (direct) and a peripheral (indirect) route. In the HSM, the direct route is called a systematic route and the indirect route is called a heuristic route. The difference between HSM and ELM concerns the simultaneity of the direct and indirect processes; the direct process in the ELM excludes the indirect process, whereas the direct and indirect processes in HSM can exist simultaneously (Oinas-Kukkonen & Harjumaa, 2008b).

This study used a video clip and an e-fotonovela as an indirect persuasion route and text messages as a direct persuasion route; both routes delivered instructional intervention. The video clip and e-fotonovela provided educational material as a way to indirectly communicate with the application's user. The explanation of the exposome in a short video provided viewers with insights toward the origin of the phrase. The compelling graphical story adapted in the e-fotonovela explained the five factors that significantly correlated with the incidence of cancer in Los Angeles County. I also used text message campaign, which used a direct persuasion route that included medical tips and shocking facts about the danger of exposome factors and the risk of having a chronic disease to raise the awareness level surround health and chronic disease.

2.2.3 Persuasion Types

Three different types of persuasion exist: interpersonal, computer-mediated, and human computer (Harjumaa & Oinas-Kukkonen, 2007).

Interpersonal persuasion is a traditional form of persuasion that occurs when a person interacts with another person or group of people with the intention to persuade them. Interpersonal persuasion depends on verbal and nonverbal symbols to change a persuadee’s attitude or behavior (Oinas-Kukkonen & Harjumaa, 2008b). Interpersonal persuasion should not include coercion, deception, or money incentives (Simons, 1970). In interpersonal persuasion situations, the persuader has the capability to respond instantly according to the receiver’s response; however, the persuader might not be fully prepared if the conversation takes a different direction (Reardon, 1991).

Computer-mediated persuasion is a form of computer-mediated communication (Oinas-Kukkonen & Harjumaa, 2008b) such as e-mail and SMS messages. In the current study, the computer was the communication channel that facilitated interactions between people for persuasion purposes.

Human-computer persuasion occurs when a user interacts with a computer and consequently is persuaded to change, shape, or reinforce their behaviors or attitudes. Computer products do not have intentions; however—in the current study—the product designer’s intention to change behaviors or attitudes resulted in persuasion (Fogg, 1998). Fogg (1998) proposed three types of persuasive intent: endogenous, exogenous, and autogenous. Endogenous intent is persuasion intent that results from creating or designing computing products with the purpose to persuade users. Miranda et al. (2013) used endogenous intent to design a product and conduct a study aimed at increasing user awareness of the dangers of texting while driving. Miranda et al. increased user awareness by showing an emotional video narrative followed by text reminders to reduce texting while driving. Miranda et al.’s findings provided preliminary evidence for the efficacy of systems that use endogenous intent. Computer technology inherits exogenous intent when a person or firm uses computing products in an attempt to change others' attitudes or behaviors. Persuasion in this approach does not come from the designer or the users; computing products inherit persuasion from the product distributor. It is important to note that computing products that use exogenous intent are not designed for persuasion purposes. For example, many organizations have adapted Google Calendar for persuading their employees to be more organized and productive.

A computing product inherits autogenous intent when a person picks a technology to change their own attitudes or behaviors. In this case, persuasion power does not reside in the hands of either the designer or the distributor; persuasion resides in the hands of the user. For example, a person might use a step-counting application to encourage themselves be more active and maintain a healthy lifestyle.

2.2.4 Persuasive Technology

Persuasive technology, also known as Computer As a Persuasive Technology (CAPTOLOGY), is "the study of computers as persuasive technology focuses on human-computer interaction" (Fogg, 2003, p. 16). CAPTOLOGY examines people's motivation when they interact with computing technology products. Computing products that are included in the interaction process can be the source of persuasion. Persuasion requires intentionality; therefore, CAPTOLOGY focuses on planned persuasive effects of computer technologies rather than side effects of a technology.

2.2.5 Using Mobile Applications in Health Care (mHealth)

Smartphones are used intensively in health care as a form of delivering health-related information and services using mobile computing capabilities with communication technologies (Carter et al., 2015). Mobile health services include applications that use technologies such as voice, video, text, and sensors to deliver multimedia messages (Marcolino et al., 2018). mHealth is a method for improving health outcomes and reducing costs. mHealth is used intensively to monitor, educate, and communicate with patients, and help patients manage chronic diseases (Marcolino et al., 2018). Moreover, mobile health applications for health care users have been designed to promote a healthy lifestyle and to change behavior(Mahmood et al., 2019).

2.2.6 Use of Smartphones in Promoting a Healthy Lifestyle

Smartphones are being used to promote a healthy lifestyle. Scholars have extensively studied the adaptation of mHealth to persuade individuals and populations to change their behavior. Martin et al. (2015) studied how mobile applications for tracking and texting could be used to motivate cardiac patients to increase their physical activity. Similarly, investigated the effectiveness of using a smartphone display to raise awareness for self-monitoring physical activity. They found that participants with the awareness display maintained their physical activity and participants without the awareness display reported a significant drop in their physical activity.

2.2.7 Use of Smartphone in Managing Chronic Diseases

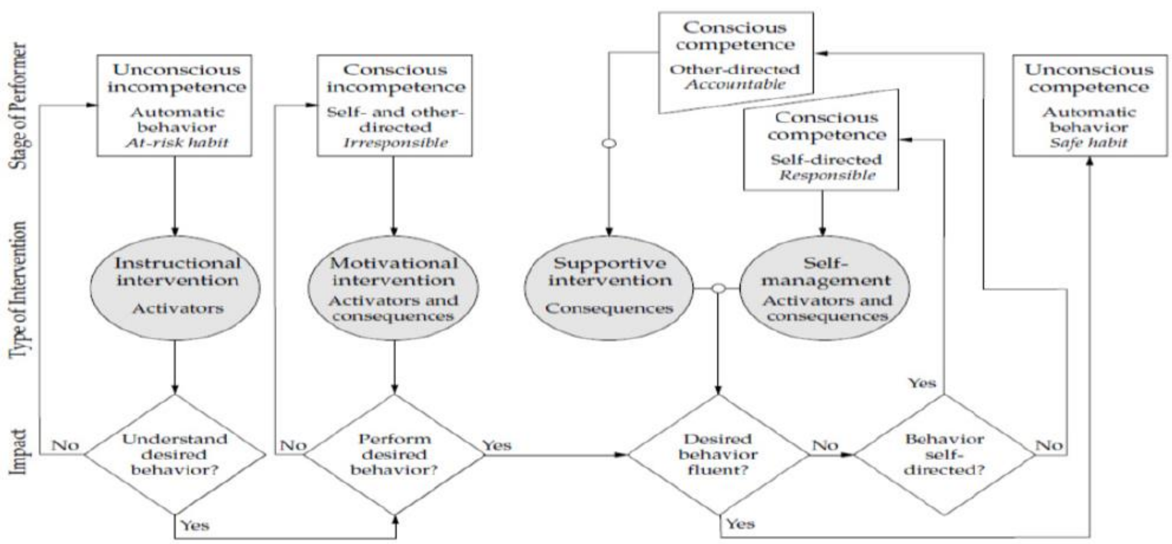
Addressing chronic disease is a major challenge for health care providers because chronic disease patients require continued supervision and care. Patient-level outcomes can be improved with the help of self-management interventions such as smartphone applications (Reynolds et al., 2018). For example, Marshall et al. (2008) designed a mobile application to help chronic obstructive pulmonary disease patients improve self-management and reduce healthcare costs. Azevedo et al. (2015) conducted a literature review on rheumatic disease self-management. They argued that a rheumatic disease self-management intervention using a smartphone application could help patients, thus increasing patient satisfaction. conducted a survey for a diet application used in five countries to evaluate the use and perceptions of users. Their survey findings indicated that one-third of participants used diet apps due to the apps' usefulness in tracking food intake.

2.3 Theoretical Grounding

I adopted Geller's (2002) proenvironmental behavioral model as a theoretical framework to design an exposome solution for behavioral change (see Figure 6). The Geller model has four behavioral change phases: unconscious incompetence, conscious incompetence, conscious competence, and unconscious competence. People who are in the unconscious incompetence stage are not aware of the problem; thus, the aim at this stage is—through instructional intervention—to get people's attention and educate them on how to shift their behavior from unconscious incompetence to conscious competence. The educational videos and e-fotonovela of the exposome application used in the current study aimed to increase individuals' awareness level; however, changing or modifying people's behavior requires more than information or advice alone (Skinner, 1989). Therefore, motivation and support interventions are required to achieve desired behaviors in people who are conscious and willing to change their behavior. Motivation intervention requires external influences—such as incentives or rewards—to encourage change. The exposome application provided a compelling story to motivate users to explore aspects of the exposome. The objective of this stage was to promote a transition from conscious incompetence to conscious competence. Targeted behavior may not last in the absence of motivational intervention; hence, supportive intervention is needed to help individuals turn their behaviors into a habit. The intervention used in the current study focused on achieving positive consequences by providing reminders, feedback, or recognition. The exposome solution provided a text message campaign designed to provide hints and reminders for users about living a healthy lifestyle. No further educational or motivational interventions were required at this stage.

Figure 6

Adopted Model: Geller Behavioral Change Model



Note. From E. S. Geller (2002). The challenge of increasing proenvironment behavior. In R. B. Bechtel & A. Churchman (Eds.), *Handbook of Environmental Psychology* (pp. 525–540). Wiley.

Chapter 3: Research Methodology

3.1 Design Science Research

This study follows the DSR approach developed by Hevner and Chatterjee (2010). DSR is a practical research approach to design, develop, and evaluate IT artifacts. Hevner and Chatterjee defined DSR as follows.

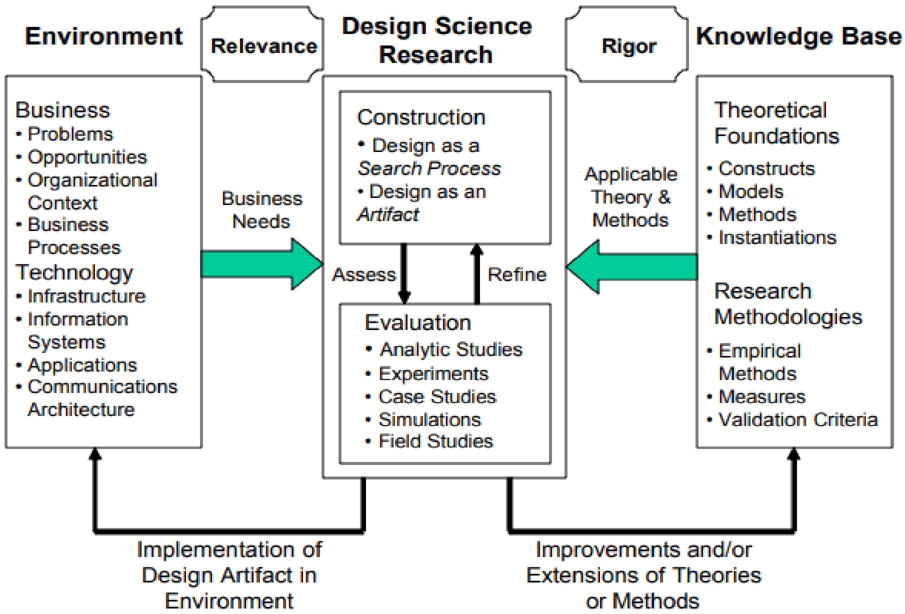
> A research paradigm in which a designer answers questions relevant to human problems via the creation of innovative artifacts, thereby contributing new knowledge to the body of scientific evidence. The designed artifacts are both useful and fundamental in understanding the problem. (p. 5)

DSR is an iterative approach with three main cycles (see Figure 7).

- Relevance cycle: This cycle ensures that the developed IT artifact is designed with relevance to the problem's contextual environment. The relevance cycle provides both requirements and field testing in an iterative way.
- Rigor cycle: This cycle ensures that the designed IT artifacts are developed and evaluated with relevance to the knowledge base of theories, models, and experience. Moreover, the research design should contribute to the body of knowledge through IT artifacts resulting from the design research.
- Design cycle: This central cycle involves both building and evaluating IT artifacts and iterates between relevance and rigor to make sure that IT artifacts have reached a satisfactory level.

Figure 7

Design Science Research Cycles



Note. Adapted from A. Hevner & S. Chatterjee (2010). Design science research in information systems. In A. Hevner & S. Chatterjee (Eds), *Design research in information systems* (Vol. 22, pp. 9–22). Springer.

In the current study, I developed two IT artifacts using the 3-component DSR cycle. First, I developed a regression model for exposome factors that significantly correlated with the incidence of cancer in Los Angeles County. The factors were selected based on the prior literature and studies related to the relationship between those factors and cancer incidents. I used multiple statistical analyses to build the regression model, which was evaluated quantitatively. The second artifact was an awareness application based on the result of the regression model to raise the level of awareness among residents of Los Angeles County about the significantly correlated exposome factors. I evaluated the artifact effectiveness by conducting a pre- and post-survey to measure the difference in awareness level. I also evaluated the usability of the artifact.

Chapter 4: Proposed Artifacts

4.1 Exposome Regression Model (Artifact 1)

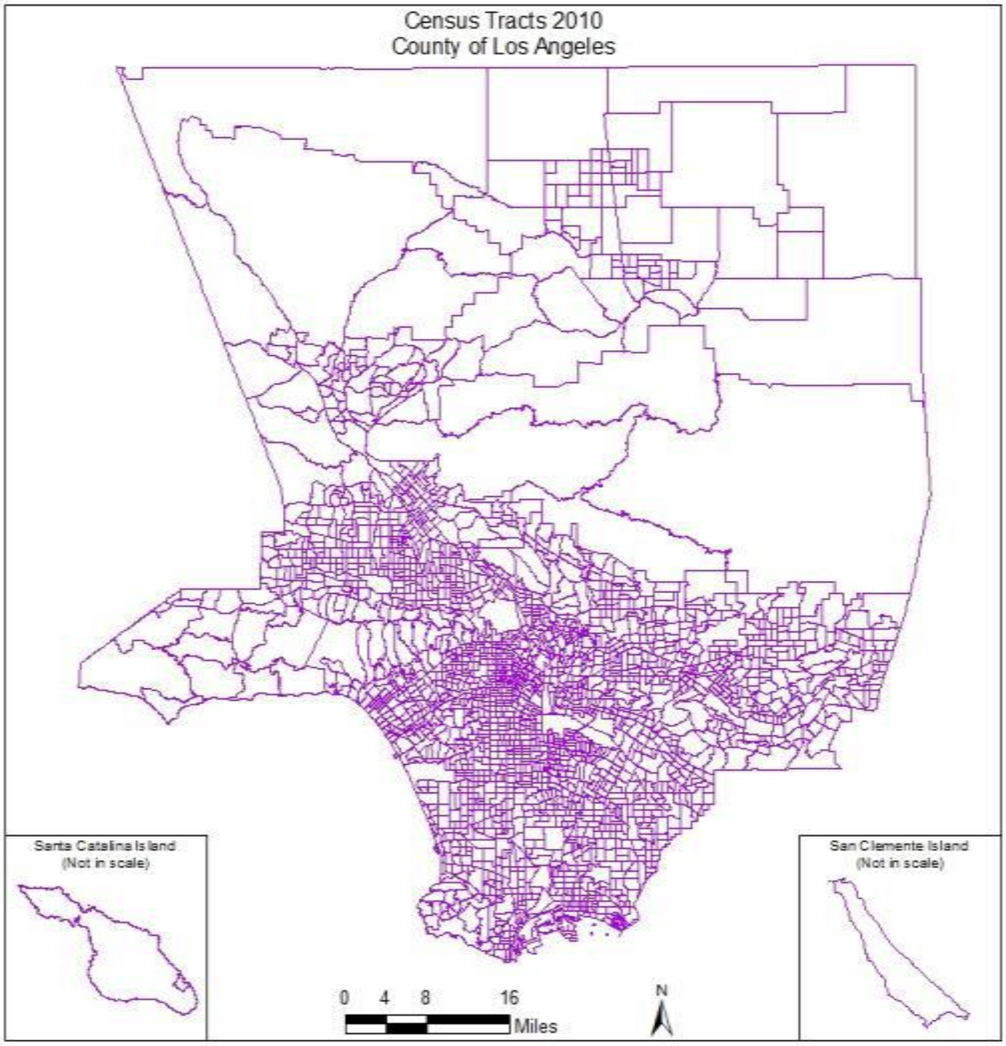
The goal was to build a regression algorithm that explains the influence of the exposome factors on the number of cancer incidents in Los Angeles County per census tract on the population level. Regression analysis was used to understand how the exposome factors were significantly correlated with cancer incidents.

4.1.1 Data Sources

The exposome data set was comprised of geospatial data that were collected from three different sources: California Office of Environmental Health Hazard Assessment, Think Health LA, and Esri consumer spending data. The data were constructed based on a location on the surface of the Earth or census tracts. A census tract can be defined as “an area roughly equivalent to a neighborhood established by the Bureau of Census for analyzing populations”(US Census Bureau, 2019). All three data sets were combined based on census tracts. Figure 8 illustrates the census tract map of Los Angeles County.

Figure 8

Census Tract Map of Los Angeles County



Note. Adapted from Los Angeles County GIS Data Portal (https://egis3.lacounty.gov/dataportal/).

4.1.1.1 California Office of Environmental Health Hazard Assessment Dataset

The first data set was obtained from the California Office of Environmental Health Hazard Assessment and contained 2,343 census tracts for Los Angeles County with 11 attributes (see Table 2). The data set had some attributes that needed to be transformed and some records had missing values. Preprocessing techniques were applied to improve the quality of the model.

Table 2

List of Data set Indicators Adapted from The California Office of Environmental Health Hazard Assessment

Variable Name	Variable Description
Census Tract	Census Tract ID from 2010 Census
Tract Population	2010 population in census tracts
Ozone concentration	Amount of daily maximum 8-hour concentration
Particulate matter 2.5	Annual mean PM 2.5 concentrations
Diesel	Diesel PM emissions from on-road and non-road sources
Toxic release	Toxicity-weighted concentrations of modeled chemical releases to air from facility emissions and off-site incineration
Traffic	Traffic density, in vehicle-kilometers per hour per road length, within 150 meters of the census tract boundary
Education	Percent of population over 25 with less than a high school education
Poverty	Percent of population living below two times the federal poverty level
Population	Race or ethnicity from 2010 Census (White, Hispanic, Asian American, Other population)
Average income	Average income per census tract

Note. Retrieved from https://oehha.ca.gov/.

4.1.1.1.1 Air Quality

Ozone. Long-term exposure to ozone is associated with chronic disease and mortality. Jerrett et al. (2009) conducted a study on the association between exposure to ozone and the risk of death. Their findings demonstrated that an increase in ozone concentration was associated with a significant increase in the risk of death from respiratory causes. In the current study, ozone concentration was measured using the daily maximum 8-hour ozone concentration (ppm) during the summer months (May through October) averaged over three years (2012 to 2014).

Particulate matter. Pun et al. (2017) studied the connection between long-term exposure to particulate matter ($PM_{2.5}$) and chronic disease. Their findings indicated that long-term exposure to $PM_{2.5}$ is associated with increased chronic ailments such as cardiopulmonary disease.

In the current study, $PM_{2.5}$ exposure was measured using the annual mean concentration of $PM_{2.5}$ (average of quarterly means) over 3 years (2012 to 2014).

Diesel PM: Exposure to diesel emissions is associated with pulmonary diseases such as COPD (Vaughan et al., 2019). Exposure to diesel emissions has a severe impact and can affect human health. In the current study, $PM_{2.5}$ data were collected based on spatial distribution of gridded diesel PM emissions from on-road and nonroad sources for a summer day in July 2012 (kg/day).

4.1.1.1.2 Pollution Burden: Environmental Effects Indicators

Pesticide. Alavanja et al. (2004) reviewed the impact of chronic pesticides exposure on cancer. A substantial number of studies indicated a strong relationship between pesticide exposure and different cancer types. In the current study, pesticides were measured by the total pounds of selected active pesticide ingredients (filtered for hazard and volatility) used in production-agriculture per square mile, averaged over 3 years (2012 to 2014).

Toxic releases from facilities. The association between various toxic exposures from facilities and cancer has been discussed extensively (Dutzik et al., 2003). In the current study, toxic releases were measured by toxicity-weighted concentrations of modeled chemical releases to air from facility emissions and off-site incineration (averaged over 2011 to 2013).

Traffic density. Living near dense traffic areas is associated with multiple pulmonary diseases, such as asthma and COPD (Lindgren et al., 2009). In the current study, traffic density was measured using the sum of traffic volumes adjusted by road segment length (vehicle-kilometers per hour) divided by total road length (kilometers) within 150 meters of the census tract boundary.

Cleanup sites. Hou et al. (2012) studied the impact of exposure to hazardous chemicals in dump sites and provided a summary of the change of human genome that is related to exposure to chemicals, and linked the diseases associated with chemical exposure. In the current study, the cleanup sites attribute was scored on a weight scale of 0 to 12 and measured by taking the sum of weighted score for within each census tract. The data were downloaded in December 2016.

Groundwater threats. Toxic chemicals such as excessive use of pesticides in agricultural areas can result in chemical exposure to groundwater and increase health risks (Zhao & Pei, 2012). In the current study, groundwater threats were scored on a weighted scale of 1 to 15 and were measured by taking the sum of weighted scores for sites within each census tract. Data were downloaded in December 2016.

Hazardous waste generators and facilities. Exposure to toxic chemicals is associated with health risk. The US Centers for Disease Control and Prevention conducted blood and urine screening tests and found 148 toxic chemicals (Hou et al., 2012). In the current study, the hazardous waste generators and facilities attribute was calculated by taking the sum of weighted permitted hazardous waste facilities and hazardous waste generators within each census tract. Permitted hazardous waste facilities data were downloaded in December 2016 and hazardous waste data was collected from 2012 to 2014.

Impaired water bodies. Exposure to impaired water bodies has been linked to chronic ailments such as kidney and thyroid disease (Iglesias & Díez, 2009). In the current study, the impaired water bodies attribute was measured by adding the number of pollutants across all water bodies designated as impaired within the area.

Solid waste and facilities. Exposure to chemicals in solid waste sites and facilities has potential health hazards for people living nearby (Mattiello et al., 2013). In the current study, the solid waste and facilities attribute was scored on a weighted scale in consideration of CalRecycle's prioritization categories and measured by taking the sum of weighted solid waste sites and facilities as of December 2016.

4.1.1.1.3 Drinking Water Contaminants

Drinking water that contains dangerous levels of contaminants can cause chronic diseases such as cancer (United States Environmental Protection Agency [EPA], 2017). In the current study, average contaminant concentrations over one compliance cycle (2005–2013) were obtained.

4.1.1.1.4 Population Characteristics: Socioeconomic Factor Indicators

Population. Race or ethnicity from 2010 Census (White, Hispanic, Asian American, Other population).

Low Education. The education attribute was measured using population data for individuals over age 25 with less than a high school education (5-year estimate, 2011-2015).

Poverty. This attribute was measured using population data on the number of people living two times below the federal poverty level (5-year estimate, 2011-2015).

4.1.1.2 Think Health LA! Dataset

The second data source was obtained from the Think Health LA! website. Think Health LA! is a web platform supported by the Los Angeles County Department of Public Health and provides acceding health indicator data related to the county. The Think Health LA! dataset used in the current study provided the percentage of adults aged 18 and over who were diagnosed with any type of cancer (excluding skin cancer) for the 1,760 census tracts in Los Angeles County (see Table 3). This data represent the percentage of people with cancer from 2014 to 2016.

Table 3

List of Variables Adapted from Thinking Health LA!

Variable Name	Variable Description
Census Tract	Census Tract ID from 2010 Census
Adult with cancer	This indicator shows the percentage of adults aged 18 and over who have ever been told by a health professional that they have any type of cancer, except skin cancer

Note. Adapted from Think Health LA! (https://www.thinkhealthla.org/)

4.1.1.3 Esri Dataset

The third data source was obtained from the Esri website. Esri's consumer spending database provided data on the spending habits of the Los Angeles County population for 2016. The spending data set provided comprehensive information on how individuals from Los Angeles County spend their money. The current study used data on canned, processed, and fast food; and data on alcoholic beverage spending (see Table 4).

Table 4

List of Variables Adapted from Esri

Variable Name	Variable Description
Census Tract	Census Tract ID from 2010 Census
Avg_Canned_Fruit	Average spending on canned fruit
Avg_Canned_Beans	Average spending on canned beans
Avg_Vegetable_Canned	Average spending on canned vegetables
Avg_Processed_Fruit	Average spending on processed fruit
Avg_Processed_Vegetable	Average spending on processed vegetables
Avg_Alcoholic_Beverages	Average spending on alcoholic beverages
Avg_Fast_Food	Average spending on fast food restaurants

Note. Adapted from Esri (https://www.esri.com/en-us/arcgis/products/arcgis-business-analyst/data-reports)

Canned food. Canning is a method of processing food by adding sugar or salt to preserve it for a longer time. Foods that are rich in saturated fat, salt, and sugar are linked to multiple chronic diseases such as obesity, cardiovascular disease, and cancer (Lucan et al., 2010). The current study used data on individuals' average spending per census tract on canned fruit, vegetables, and Beans in 2016.

Processed food. Processed food is a type of food that includes artificial colors, flavors, and preservatives. High levels of processed food or drink consumption is linked to chronic diseases such as obesity and tooth decay (Canella et al., 2014). The current study used data on individuals' average spending per census tract on processed fruit and vegetables in 2016.

Alcoholic beverages. Poor lifestyle choices such as excessive alcohol consumption paired with poor diet and smoking is associated with a variety of chronic diseases (Fine et al., 2004). The current study used data on individuals' average spending per census tract on alcoholic beverages in 2016.

Fast food. Fast food can be defined as "convenience food purchased in self-service or carry-out eating venues without wait service" (Rosenheck, 2008, p. 535). Several studies have linked high calorie intake, which is associated with fast foods, with chronic diseases such as obesity (Rosenheck, 2008). The current study used data on individuals' average spending per census tract on fast food in 2016.

4.1.2 Data Aggregation

Data were downloaded and obtained in the form of an Excel sheet, and formatted as a comma-separated values. All three data sets were combined to form the exposome data set. The exposome data set had 1,760 records based on the census tracts of cancer cases data set where all the three data sets were joined using census tract as a unique key. RStudio software was used to bind all three Excel sheets to form the Exposome data set. RStudio is a free platform for statistical computing that compiles and runs on multiple operating systems.

4.1.3 Data Preprocessing

Data on the adult with cancer attribute were collected from 2014 to 2016 based on census tracts for Los Angeles County. This attribute represented the percentage of adults aged 18 and over who currently or previously had any type of cancer, excluding skin cancer. I used Excel software to calculate the average number of individuals who had been diagnosed with cancer per census tract. Then, I calculated the average cancer incidents by multiplying each census tract percentage by census tract population.

4.1.4 Missing Values

Abu-Bader (2010) provided a general guideline for dealing with missing values. Abu-Bader stated that "if only 5 percent (or less) of cases have missing values at random, then almost any procedure for handling missing values yields similar results" (p. 30). In the current study, only 69 records out of 1,760 (3.9%) had missing values. The decision was made to exclude the missing data from the analysis. RStudio were used to identify missing values and exclude these from the data set.

4.1.5 Dealing with Outliers

Univariate outlier scores. The exposome data set was uploaded to SPSS to identify and deal with univariate outliers. SPSS is statistical software widely used in social science and health research.

Univariate outliers occur when one variable has extreme scores at either end of the distribution (Abu-Bader, 2010). Several SPSS methods can be used to identify univariate outliers, such as calculating a z-score, and examining boxplots and stem-and-leaf plots (Abu-

Bader, 2010). I used z-scores in the current study. Raw scores were converted into z scores using the following equation:

$z = (x - \mu) / \sigma$

Where *z* is the new converted variable score, *x* is the old variable score, μ is the variable mean and σ is the variable standard deviation. Abu-Bader (2010) provided a general rule for identifying univariate outliers; raw scores that have a *z*-score greater than +3 and lower than -3 can be considered univariate outliers. Exposome data were converted into *z*-scores using SPSS (see Table 5). As a result, 282 scores were removed from the exposome data set and 1,409 records remained.

Table 5

Excluded Outliers Scores for Each Variable of Exposome Data Set.

Variable name	Univariate outlier number
Average cancer cases	20
Ozone	63
Diesel PM	15
Toxic release	16
Traffic	27
Fast food	23
Average alcohol beverages	27
Average canned fruit	0
Average canned vegetables	3
Average canned beans	0
Average processed fruit	0
Average processed vegetables	0
Hispanic population	5
White population	1
Asian American population	34
Other population	37
Education	8
Poverty	1
Average income	2

Multivariate outliers. Multivariate outliers occur when two or more variables have extreme scores. Multivariate outliers can be identified using two methods: scatterplots or a Maholanobis distance test (Abu-Bader, 2010). The Maholanobis distance test was used in the current study to identify multivariate outliers for the exposome data set. The Maholanobis distance test uses the chi-square distribution to measure the distance of each score from the centroid of the distribution. A score is considered a multivariate outlier when the Maholanobis value exceeds the chi-square value at alpha = .001 (Abu-Bader, 2010). SPSS was used to eliminate univariate outliers.

As a result of conducting the Maholanobis distance test, 91 scores were identified as outliers and were thus excluded from the exposome data set. The number of remaining records was 1,318.

4.1.7 Multiple Linear Regression Analysis

I used multiple regression to examine the effect of the independent variables (air quality, environment, consumer spending variables and demographics) on the dependent variable (average cancer cases). The following steps must be conducted before conducting a multiple regression analysis.

Step 1: State the research question. A null and alternative hypothesis are not necessary in multiple regression analysis because the goal is not to verify or falsify research hypotheses; rather, the goal is to develop a regression model that best correlates with the criterion (Abu-Bader, 2010). The current study's research question was as follows.

RQ1. What set of the factors shown in Table 6 best significantly correlated with average cancer incidents in Los Angeles County?

Table 6

Exposome Factors

	Variable
1	Avg_ cancer_cases
2	Ozone
3	Diesel
4	tox_release
5	Traffic
6	Avg_fast_food
7	Avg_Alcoholic_Beverages
8	Avg_Canned_Fruit
9	Avg_Canned_Vegetable
10	Avg_Canned_Beans
11	Avg_Processed_Fruit
12	Avg_Processed_Vegetables
13	Hispanic
14	White
15	Asian_American
16	Other_population
17	Education
18	Poverty
19	Avg_Income

Step 2: Choose alpha. I set alpha (α) at 0.05. The result would be statistically significant if $p \leq .05$.

Step 3: Select the appropriate statistical test. The goal of this study was to build a regression model that significantly correlated multiple factors with cancer incidents in Los Angeles County. Before running the model, data were evaluated and examined to meet multiple regression analysis assumptions. The multiple regression assumptions are as follows.

1. Level of measurement: In this assumption, the dependent variable must be measured at the interval level or higher. In the current study, the square root of average cancer incidents in Los Angeles County was the dependent variable or criterion and was measured at the interval level of measurement.
2. Linearity: This assumption emphasizes the linear relationship between the independent variable and dependent variables. To evaluate this assumption, I used SPSS to create scatterplots that showed the linear relationship between the square root of average cancer incidents in Los Angeles County and the exposome variables. A linear relationship between the independent variables and the dependent variables existed with minor deviation (see Appendix B for plots).
3. Normality of residuals: This assumption emphasizes the shape of the distribution of the residuals to form the shape of a normal distribution. To evaluate this assumption, I used SPSS by creating a histogram and normal plots (see Figure 9). Inspection of histogram A indicates that the residuals are normally distributed and the points in plot B fall on a straight diagonal line.

Figure 9

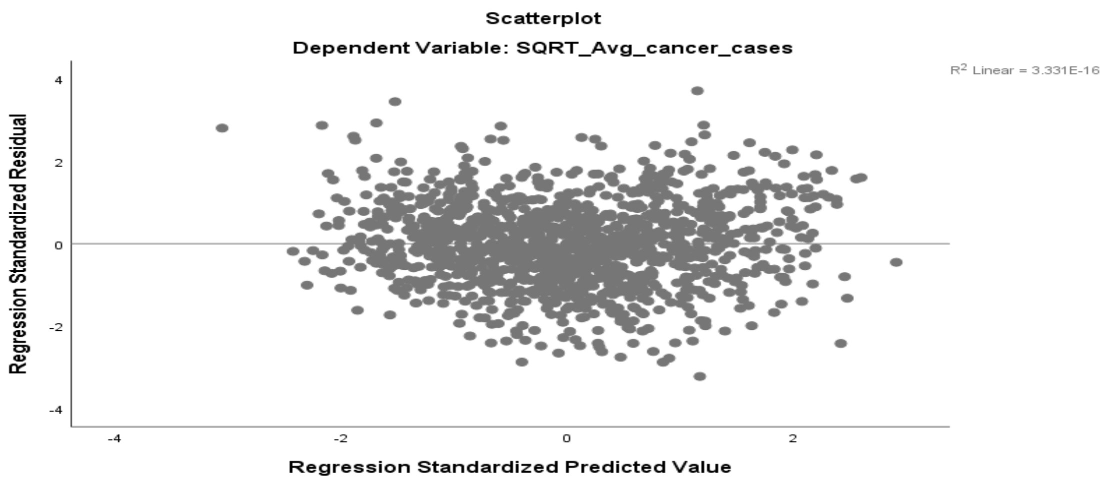
Histogram and Probability Plot for the Residuals

A B

4. Homoscedasticity: In this assumption, the variance around the regression line should be the same for all values of the independent variables. To evaluate this assumption, I created a scatterplot for the residual against the predicted values (see Figure 10). The points seem to be distributed equally along the horizontal line for each level of intendent variables, thus indicating that the exposome data are homoscedastic.

Figure 10

Scatterplot for Predicted Scores and Residual Scores



5. Multicollinearity: The assumption of multicollinearity emphasizes that the relationship between all pairs of the independent variables must not exceed .80 ($r <= .80$). To examine this assumption, I used SPSS to calculate Pearson correlation coefficients between all independent variables. The process required four iterations (see Appendix C for details of these iterations) to remove all multicollinearity. Also, the assumption of multicollinearity can be evaluated by examining Variance Inflation Factor (VIF). Table 7 displays these values and shows the VIF and tolerance value for each factor with all others. Abu-Bader (2010) provided a general rule for dealing with

multicollinearity using the VIF method. If the VIF exceeds 10 or the tolerance value is smaller than .1 then multicollinearity does exist. Table 7 shows that all VIF values are below 10 and the tolerance value is bigger than .1, hence no multicollinearity exists.

Table 7

VIF Values

	Collinearity Statistics	
	Tolerance	*VIF*
White	0.326	3.064
Hispanic	0.149	6.715
Other_population	0.942	1.061
Asian_American	0.842	1.188
Fast_Food	0.287	3.478
tox_release	0.589	1.696
Ozone	0.129	7.742
Diesel	0.784	1.275

6. Sample Size: There are 18 factors in this research question. A sample size of 194 or more was needed to use multiple regression analysis when applying the sample size (N) formula ($N \geq 50 + 8m$, where m = number of factors (Abu-Bader, 2010). In the current study, the exposome data set had 1,318 records which exceeded the minimum required sample size.

Step 4. Select factors that will be entered in the analysis. Step 4 was to select which factors or independent variables should be entered in the regression analysis. To determine this, I ran a Pearson correlation between the criterion (Avg_cancer_cases) and each factor (ozone, diesel, tox_release, traffic, Fast_Food, Hispanic, White, Asian_American, and Other_population; see Figure 11).

Figure 11

Results of Running Pearson Correlation Test

Correlations

		SQRT_Avg_cancer_cases	SQRT_ozone	SQRT_diesel	SQRT_tox_release	SQRT_traffic	SQRT_Fast_Food	SQRT_Hispanic	SQRT_White	SQRT_Asian_American	SQRT_Other_poulation
SQRT_Avg_cancer_cases	Pearson Correlation	1	.156**	-.361**	-.117**	-.110**	.472**	-.061*	.678**	.338**	.194**
	Sig. (2-tailed)		0.000	0.000	0.000	0.000	0.000	0.027	0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_ozone	Pearson Correlation	.156**	1	-.335**	-.729**	0.048	.190**	-.075**	.299**	.071*	-.285**
	Sig. (2-tailed)	0.000		0.000	0.000	0.084	0.000	0.006	0.000	0.010	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_diesel	Pearson Correlation	-.361**	-.335**	1	.237**	.227**	-.373**	.117**	-.335**	-0.038	-0.018
	Sig. (2-tailed)	0.000	0.000		0.000	0.000	0.000	0.000	0.000	0.171	0.505
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_tox_release	Pearson Correlation	-.117**	-.729**	.237**	1	-.126**	-.286**	.318**	-.403**	0.012	.099**
	Sig. (2-tailed)	0.000	0.000	0.000		0.000	0.000	0.000	0.000	0.673	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_traffic	Pearson Correlation	-.110**	0.048	.227**	-.126**	1	-0.052	-0.040	-.055*	-0.001	0.032
	Sig. (2-tailed)	0.000	0.084	0.000	0.000		0.058	0.151	0.045	0.980	0.243
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Fast_Food	Pearson Correlation	.472**	.190**	-.373**	-.286**	-0.052	1	-.652**	.766**	.232**	-.136**
	Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.058		0.000	0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Hispanic	Pearson Correlation	-.061*	-.075**	.117**	.318**	-0.040	-.652**	1	-.614**	-.184**	.062*
	Sig. (2-tailed)	0.027	0.006	0.000	0.000	0.151	0.000		0.000	0.000	0.023
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_White	Pearson Correlation	.678**	.299**	-.335**	-.403**	-.055*	.766**	-.614**	1	.329**	-.098**
	Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.045	0.000	0.000		0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Asian_American	Pearson Correlation	.338**	.071*	-0.038	0.012	-0.001	.232**	-.184**	.329**	1	-.142**
	Sig. (2-tailed)	0.000	0.010	0.171	0.673	0.980	0.000	0.000	0.000		0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Other_poulation	Pearson Correlation	.194**	-.285**	-0.018	.099**	0.032	-.136**	.062*	-.098**	-.142**	1
	Sig. (2-tailed)	0.000	0.000	0.505	0.000	0.243	0.000	0.023	0.000	0.000	
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Figure 11 illustrates that Avg_cancer_cases has a significant correlation with ozone ($r = .156$, $p < .05$), diesel ($r = -.361$, $p < .05$), tox_release ($r = -.117$, $p < .05$), traffic ($r = -.110$, $p < .05$), Fast_Food ($r = .472$, $p < .05$), Hispanic ($r = -.061$, $p < .05$), White ($r = .678$, $p < .05$), Asian_American ($r = .338$, $p < .05$), Other_population ($r = .194$, $p < .05$). All the factors have significant bivariate relationships with Avg_cancer_cases and were entered into the regression analysis.

Step 5: Run multiple regression analysis. A stepwise multiple regression analysis was conducted to estimate a regression model that significantly correlated exposome factors with average cancer incidents in Los Angeles County (Avg_Cancer_cases). In the stepwise method, factors are entered based on the size of their partial correlation coefficient; the factor with the largest correlation coefficient is entered first in the analysis (Abu-Bader, 2010). The model summary in Table 8 illustrates that the regression model has eight steps in the regression equation. The model summary indicates that the adjusted coefficient of determination or Adjusted $R^2 = 0.759$.

Table 8

Results of Multiple Regression Analysis – Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.678[a]	0.460	0.459	1.95975	0.460	1119.614	1	1316	0.000
2	.814[b]	0.663	0.662	1.54927	0.203	790.737	1	1315	0.000
3	.855[c]	0.731	0.730	1.38509	0.068	331.229	1	1314	0.000
4	.863[d]	0.745	0.744	1.34816	0.014	73.962	1	1313	0.000
5	.869[e]	0.755	0.754	1.32111	0.010	55.323	1	1312	0.000
6	.871[f]	0.759	0.758	1.31103	0.004	21.256	1	1311	0.000
7	.872[g]	0.760	0.759	1.30806	0.001	6.953	1	1310	0.008
a. Predictors: (Constant), White									
b. Predictors: (Constant), White, Hispanic									
c. Predictors: (Constant), White, Hispanic, Other_poulation									
d. Predictors: (Constant), White, Hispanic, Other_poulation, Fast_Food									
e. Predictors: (Constant), White, Hispanic, Other_poulation, Fast_Food, tox_release									
f. Predictors: (Constant), White, Hispanic, Other_poulation, Fast_Food, tox_release, diesel									
g. Predictors: (Constant), White, Hispanic, Other_poulation, Fast_Food, tox_release, diesel, ozone									

The results of ANOVA for the eighth model is presented in Table 9; the stepwise multiple model coefficients are shown in Table 10. This model included eight of the nine factors that significantly correlated with cancer incidents in Los Angeles County ($F_{1, 1309}$) = 593.866, p = .000). The only factor that was not included in the model was “Traffic”.

Table 9

Results of Multiple Regression Analysis-ANOVA.

	Model	Sum of Squares	df	Mean Square	F	Sig
7	Regression	7112.838	7	1016.120	593.866	.000[h]
	Residual	2241.441	1309	1.711		
	Total	9354.279	1317			

Table 10

Results of Multiple Regression Analysis –Coefficients

	Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
7	(Constant)	-1.694	1.292		-1.311	0.190		
	White	0.158	0.004	0.968	42.697	0.000	0.356	2.809
	Hispanic	0.101	0.003	0.578	29.476	0.000	0.475	2.105
	Other_poulation	0.082	0.004	0.274	18.629	0.000	0.845	1.183
	Fast_Food	0.088	0.013	0.158	6.689	0.000	0.326	3.065
	tox_release	0.017	0.002	0.166	7.528	0.000	0.376	2.658
	Diesel	-0.151	0.040	-0.059	-3.768	0.000	0.737	1.357
	Ozone	12.565	4.765	0.060	2.637	0.008	0.355	2.820

a. Dependent Variable: Avg_cancer_cases

With a beta of .97 ($p < .01$), the White race category was the most significantly correlated with average cancer incidents, accounting for 46% of the variance in average cancer incidents. The second strongest factor was the Hispanic race category ($\beta = .58$, $p < .01$), which accounted for an additional 20.3% of the variance in average cancer incidents. The third strongest factor was the other population race category ($\beta = .27$, $p < .01$), which accounted for an additional 6.8

percent of the variance in average cancer incidents. The other four factors accounted for an additional 5.5% of the variance in average cancer incidents.

These results indicate that higher average cancer incidents in Los Angeles County are a function of higher White race, higher Hispanic rate, higher other population, higher Asian American race, higher fast food spending, higher toxic release in the air, higher ozone concentration, and lower diesel emission into the air. Overall, the model explained almost 76% of the variance in average cancer cases (R = .872). Thus, about 24% of the variance in average cancer cases in Los Angeles County is still unaccounted for in this model. The regression equation for average cancer incidents in Los Angeles County is as follows:

Average cancer incidents = -1.694+ (.158 *White) + (.101*Hispanic) + (.082*Other_poulation) + (.088*Fast_Food) + (.017* tox_release) + (-.151 * diesel) + (12.56*ozone).

4.2 Exposome Awareness Application (second artifact)

The goal of building an exposome awareness app was to raise awareness of exposome factors through using persuasive technology techniques based on significantly correlated exposome factors with cancer incidents in Los Angeles County per census tract. As described in the previous section, the exposome factors that were statistically significant in the regression model were race (White, Hispanic and other population), fast food, toxic release, ozone concentration, and diesel emissions.

The persuasive application consisted of two components.

1. The application followed the Geller model and consisted of instructional and motivational interventions to raise users' awareness level of the exposome, thereby indirectly persuading individuals of Los Angeles County through educational video

and a compelling graphical story to shift their attitudes regarding their unhealthy behaviors and become more receptive to healthier behaviors.

2. A text message campaign represented the supportive intervention of the Geller model. The campaign used a direct persuasive strategy of tips and hints for making healthy lifestyle changes, such as quitting smoking and avoiding processed foods, thus minimizing the effects of harmful environmental exposures.

4.2.1 Application Contents

Exposome video explainer. I made a video using Vyond, an online paid platform. I recorded a voice-over of the script before using that the Vyond software. I began to create the video after the voice-over was recorded. Avatars, or characters, were designed according to the requirements and the context of the exposome. The early mapping included the addition of facial features, such as eyes, hair and face shape, then attire and facial expressions. After this, a timeline was created to determine when to begin the voice-over, followed by which specific characters and actions were to appear at what time in the video. The next stage was deciding when and for how long to bring in icons and animation. Next and most importantly, transitions were added along with text. Transitions were selected and customized for duration, color, and direction of movement on screen. They were then added to their respective position on the timeline. A text box was then generated and the text type was selected. At this point, the video was almost complete. The next task was to ensure that visuals were properly synchronized with the voice-over and, if not, that elements of the video were rearranged accordingly. Next, music from the Vyond software library was added to the timeline, with audio fades placed at the beginning and end of the music. At this point, the video was complete.

4.2.2 e-fotonovela

An e-fotonovela is a type of comic that tells a story through images and bubble dialogs. The e-fotonovela demonstrated desirable environmental and personal health behaviors to educate residents of Los Angeles County about environmental risks surrounding their communities. The scenario included multiple characters and scenes. The developed scenario was as follows.

Lawrence and Sarah are a couple in Los Angeles. Their child, Daniel, is often sick. Although they have been to doctors, Lawrence and Sarah do not know why Daniel is sick. Daniel was not sick nearly as often before the family moved to Los Angeles two years ago. The sick days mean that Daniel often misses school, and Lawrence and Sarah have to miss work to take care of him. One day, a school administrator calls and says that Daniel's sick days are concerning and that he may fall behind significantly if he is not able to improve his attendance Lawrence and Sarah begin to investigate the potential causes of his sickness. They visit Daniel's pediatrician and the doctor provides Lawrence and Sarah with information about the environmental exposome. The doctor informs Lawrence and Sarah about a new study conducted in Los Angeles County that links cancer to environmental exposure. Figures 12, 13, and 14 illustrate this portion of the e-fotonovela.

Figure 12

A Suburban Home in Los Angeles. It's the Morning, and It's A Normal Middle-Class Home: Small Yard, etc.

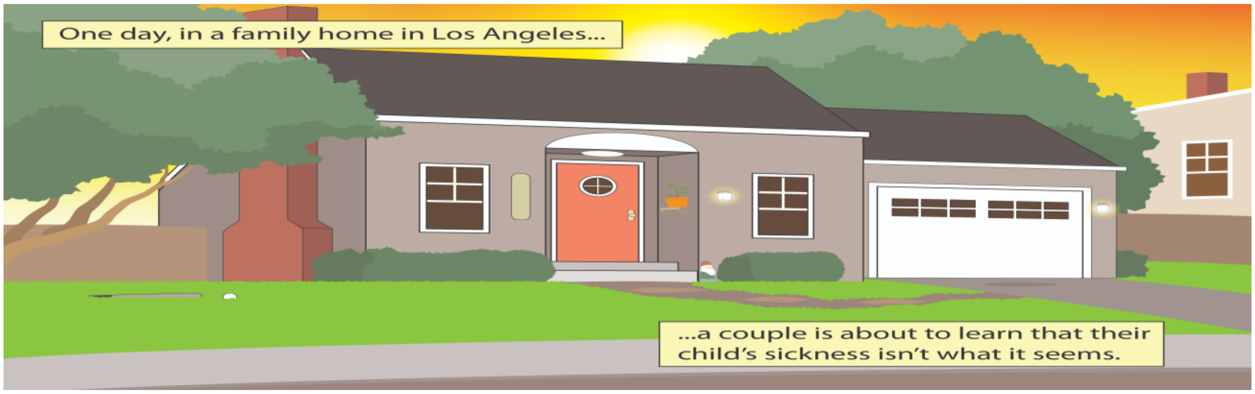


Figure 13

Lawrence and Sarah Look At Each Other With Concern Across the Table

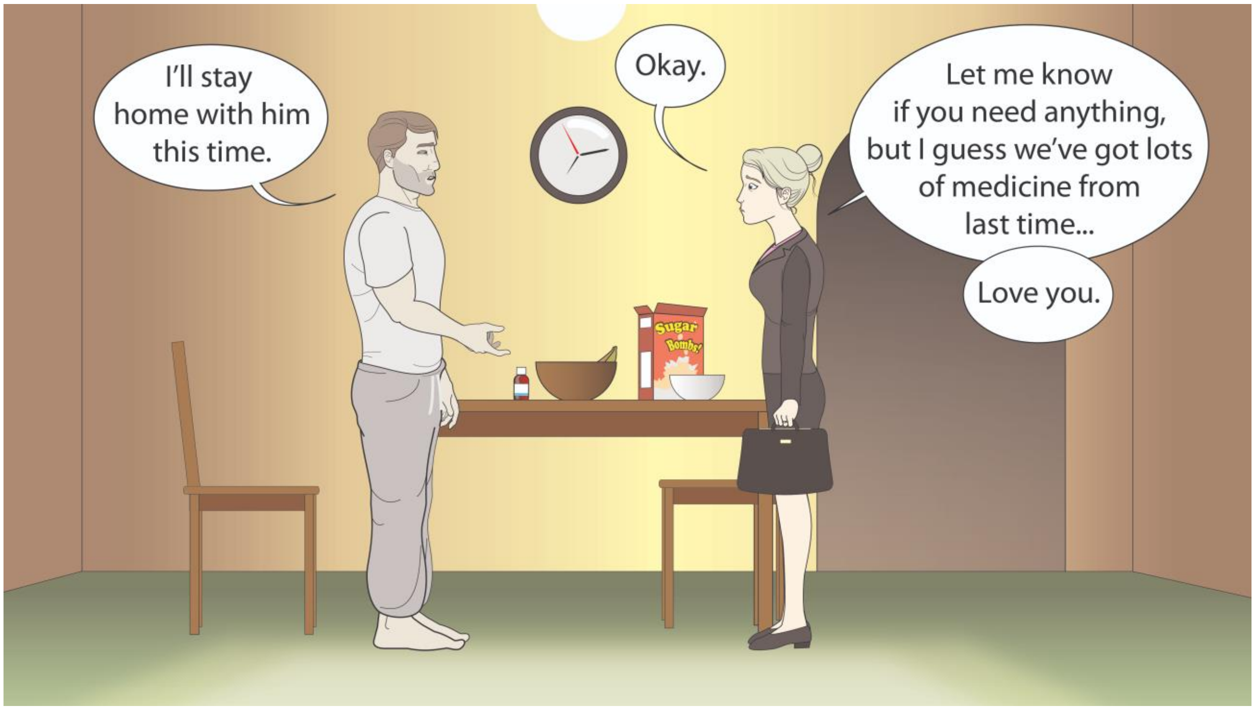


Figure 14

Shot of the Doctor

Eventually, Sarah and Lawrence learn that a possible source of Daniel's sickness is the environmental exposome; specifically, poor air quality. Their home is in an area with a high concentration of ozone, diesel emissions, and a nearby facility that produces toxic chemical and causes air pollution. Sarah and Lawrence discover that other parents are concerned as well. They started to take protective measures by using an indoor air purifier and reduce the time the windows are open. They also pay attention to the air quality index and reduce outdoor activities and wear masks when going outside in poor air quality. After 6 months, Daniel starts to feel better and most of his parents' concerns fade away. Figures 15, 16 and 17 illustrate this portion of the e-fotonovela. Appendix D provides all developed scenarios and screens.

Figure 15

Shot of Lawrence's Laptop Screen, Open to an Air Quality Index Page

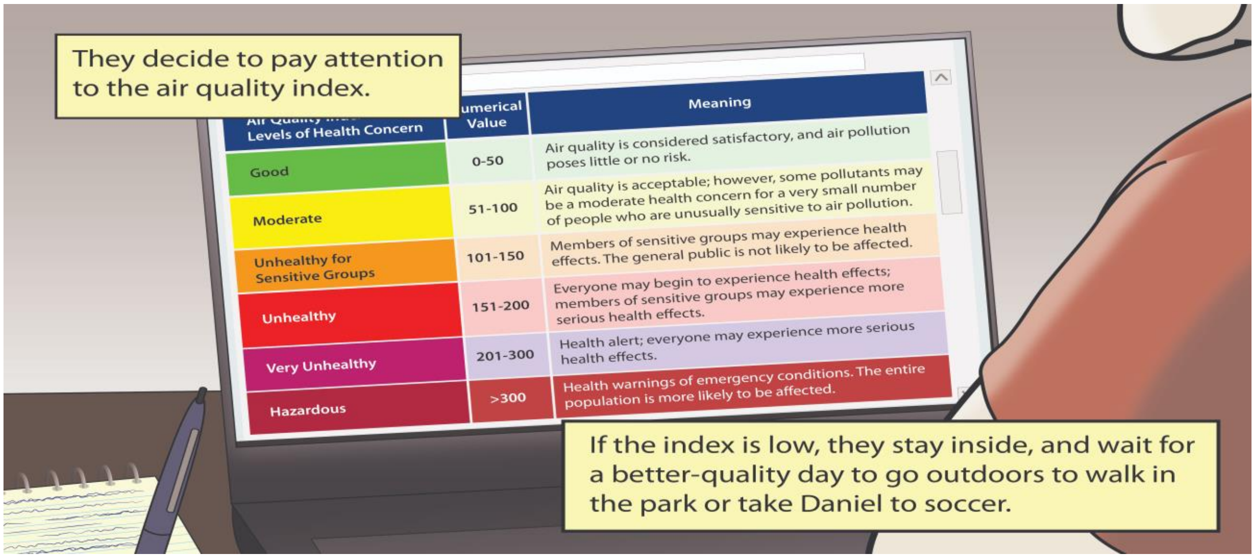
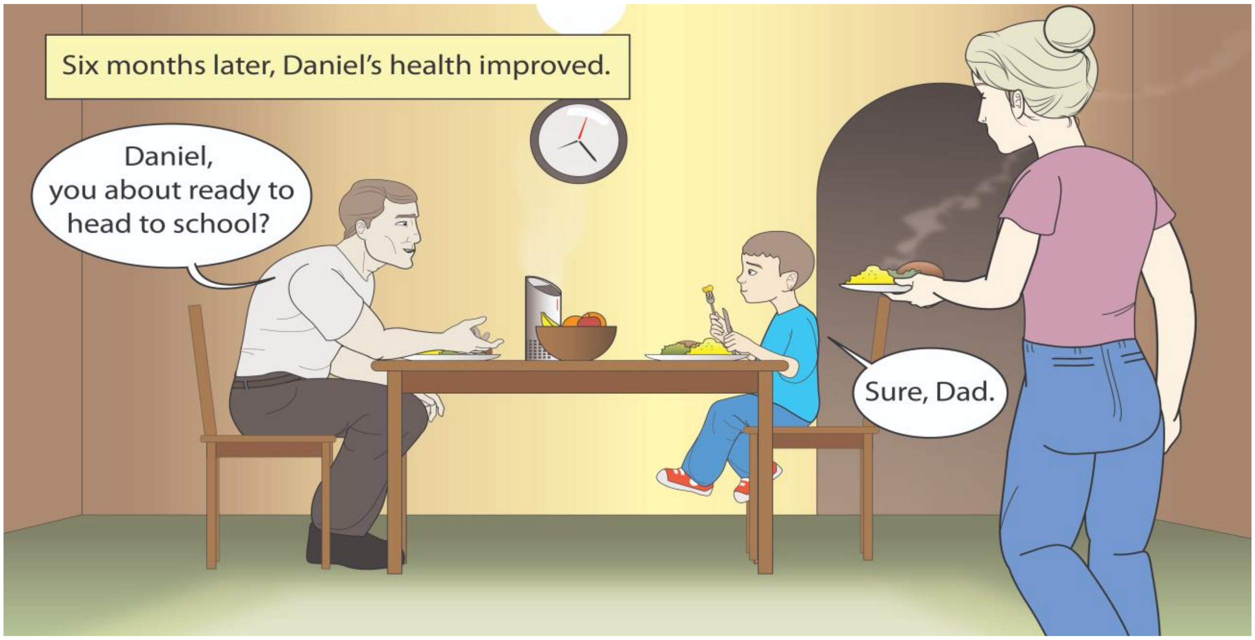


Figure 16

Sarah and Daniel Wearing an Air Mask

Figure 17

Shot of the Living Room, with the Air Purifier, with Daniel Eating a Healthy Breakfast (eggs, greens) on the Dining Room Table

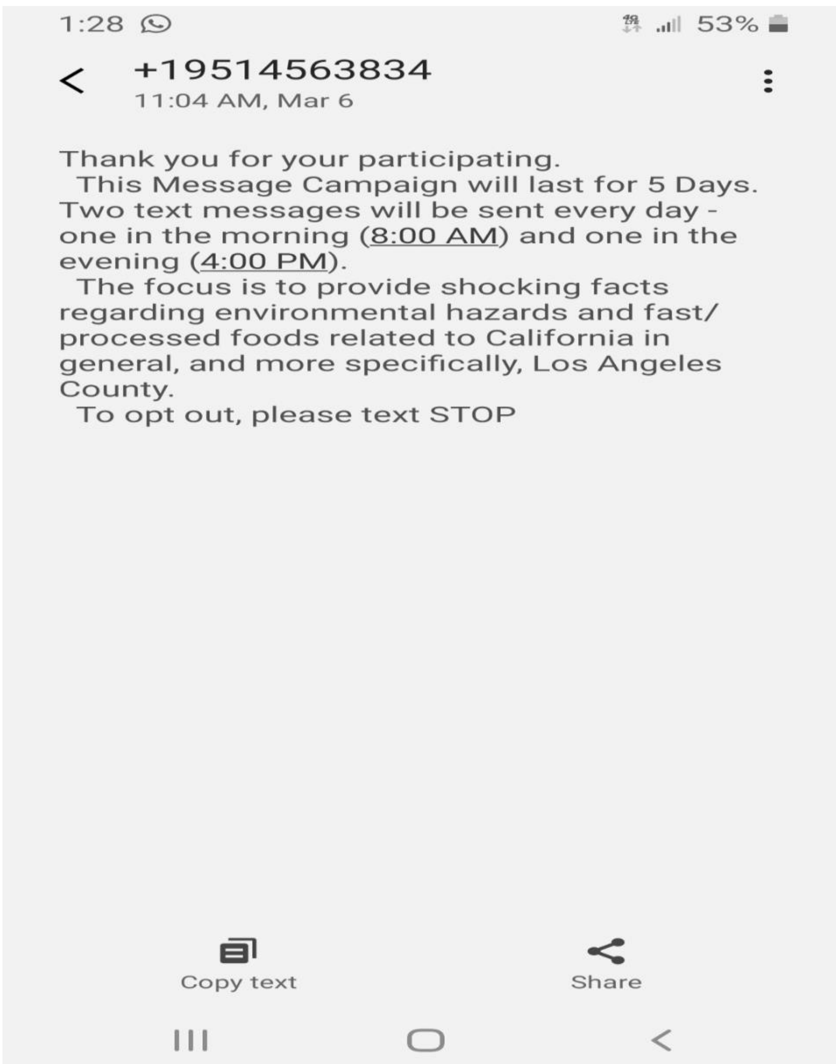


4.2.3 Text Message Campaign

The text messaging campaign included a series of shocking facts sent to the participants; a direct persuasion strategy was used in these messages. The campaign was an effective approach to raising the awareness level of the various risks to human health. Khurshid et al. (2015) developed the "text4health" campaign for individuals at risk of Type 2 Diabetes; the campaign was conducted through social media and text messages. Khurshid et al. found a positive trend in reaching their targets and educated individuals at risk for Type 2 Diabetes through a series of educational strategies. Participants who subscribe to the text4health campaign receive two texts a day for 5 days upon first subscribing. The last screen of the app encourages users to subscribe to the text messaging campaign. The campaign uses Twilio, a cloud-based text message platform that can be programmed to send and receive text messages. Twilio allows for text scheduling. The text4health facts are sent at a specific time every day; for example, texts are sent at 8:00 am and 4:00 pm. The first message greets the participants and thanks them for participating. This text explains the campaign duration and timing, and allows participants to opt out at any time by texting the word “STOP” (see Figure 18).

Figure 18

Thanking Message Sent to Participant After Subscription to the Campaign



The first environmental tip was articulated as “Inhaling air pollution can reduce your lifespan by 1–2 years.” This tip was inspired by an article titled “Air Pollution Reduces Global

Life Expectancy by More than One Year” (University of Texas at Austin, 2018). The second tip was articulated as “Pollutants that are released into the air, as opposed to land and water pollutants, are the most harmful to humans” (National Geographic Society, n.d.). The third tip was articulated as “Toxic air pollution poses a greater threat to children, due to their smaller physical size and lung capacity” (World Health Organization, 2018). The fourth tip was articulated as “5,000 premature deaths in Southern California are caused by exposure to pollution from diesel trucks” (Mpatino, 2015). The last environmental tip was articulated as “According to the 2019 ‘State of the Air’ report from the American Lung Association, California topped the list with the worst air pollution, and Los Angeles is the city with the worst ozone pollution. Los Angeles has topped that list for the last 19, out of 20, of these reports” (American Lung Association, 2018).

The first fast food and processed food tip was articulated as “Every 10% increase in the consumption of processed foods is associated with a 12% higher risk for cancer in general, and an 11% increase risk for breast cancer” (WebMD, 2018). The second tip was articulated as “Fast food has been associated with poor diet quality and higher fat, saturated fat, and sugar intake, which are known contributors to heart disease” (Summit Medical Group, 2018). The third tip was articulated as “From 1994 through 2006, caloric intake from fast food increased from 10% to 13% among children aged 2–18 years” (Vikraman et al., 2015). The fourth tip was articulated as “Living near a fast food restaurant is linked to a 5.2% greater risk of obesity” (Lee, 2018). The last tip was articulated as “Regularly eating fast food doubles your chance of developing insulin resistance, which heightens risk of developing Type 2 Diabetes” (Leonard, 2019). Screenshots of these sent messages are provided in Figures 19 and 20.

Figure 19

Screenshot of the Sent Messages.

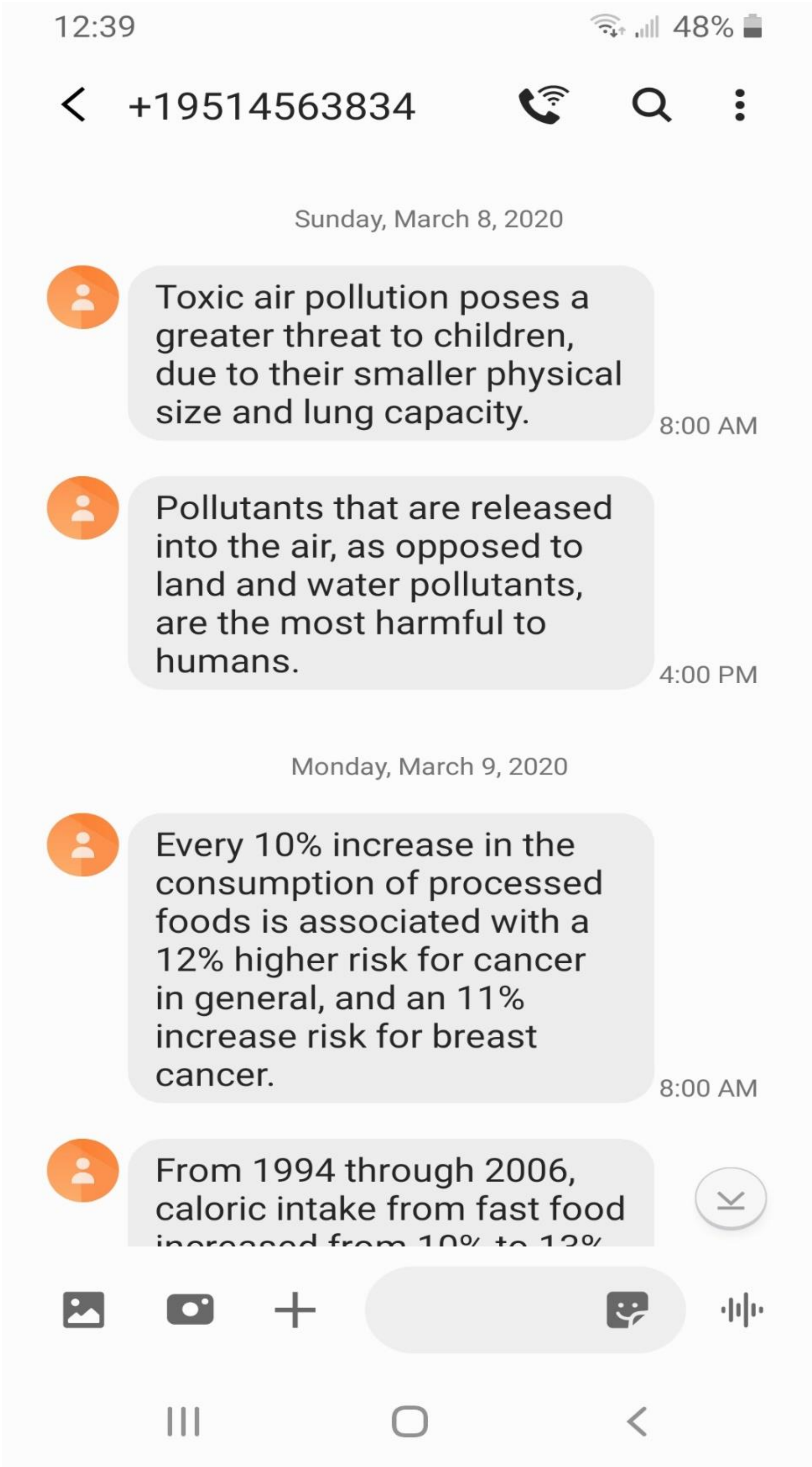
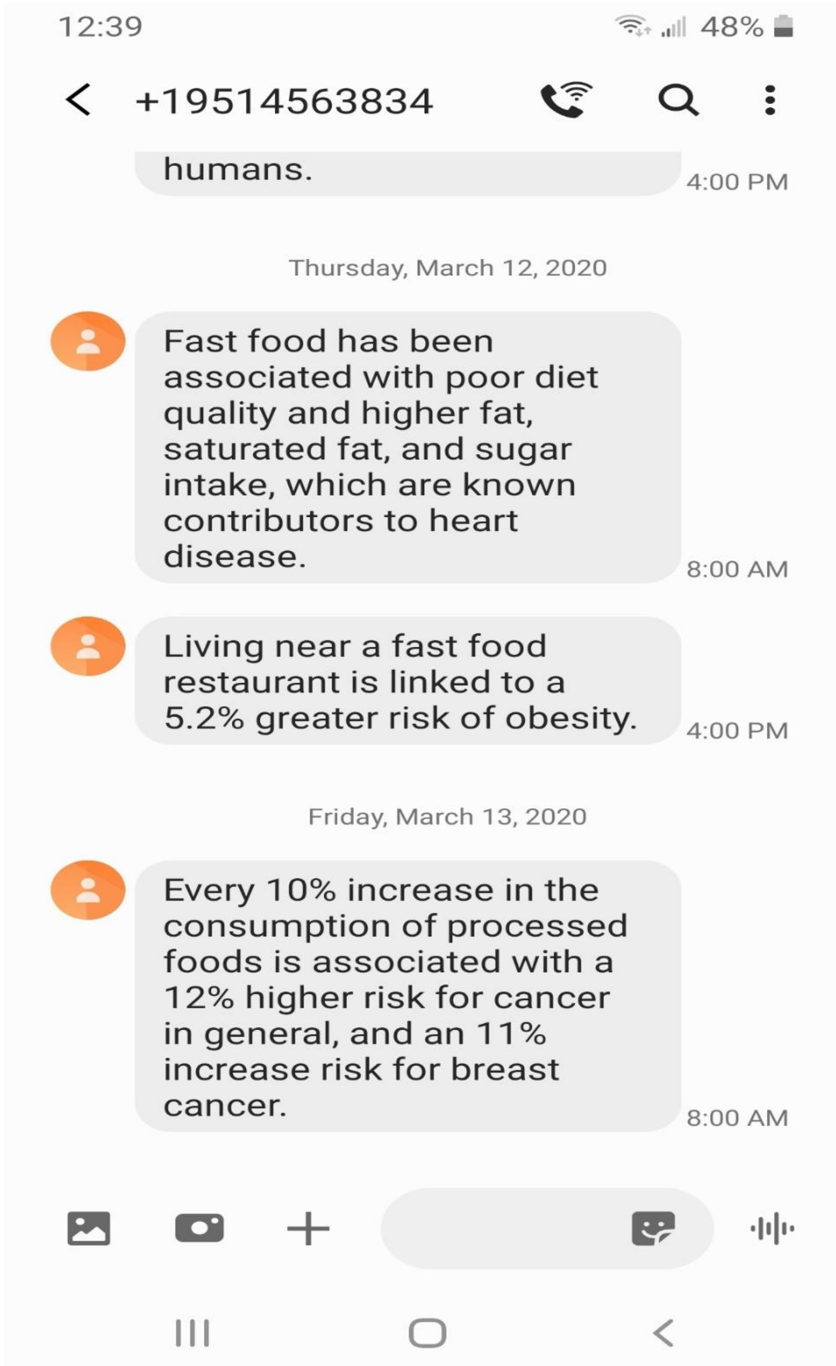


Figure 20

Screenshot of the Sent Messages



4.2.4 Participants

I applied inclusion criteria to select participants who were familiar with using smartphones and mobile apps, and who were aged between 18 and 65. The participants had to be willing to participate in the pre- and post-surveys and had to be willing to receive text message tips for 5 days. The target sample size was 30 participants. This study follows an exploratory approach and literature suggested that sample size between 10 and 30 would have practical advantages (Johanson & Brooks, 2010)

4.2.5 Exposome Mobile App Design

The exposome application adapted the PSD model (Oinas-Kukkonen & Harjumaa, 2008a) for designing a persuasive system. The system features design principles are built on the Fogg model that provides primary task, dialogue, system credibility, and social support (Oinas-Kukkonen & Harjumaa, 2008a).

Primary task support. The design principles for primary task support focus on the user's primary task and include reduction, tunneling, tailoring, personalization, self-monitoring, simulation, and rehearsal (Oinas-Kukkonen & Harjumaa, 2008a). The exposome application implemented the tunneling design principle, defined as "using the system to guide users through a process or experience provides opportunities to persuade along the way" intended to help users navigate the application contents. The application consisted of a 1-minute educational video, followed by infographics that provided further details about the exposome factors and text campaign. Upon downloading the application, an avatar welcomes the user and provides brief information about what the user is expected to learn.

Dialogue support. The design principles for dialogue support are associated with implementing computer-human dialogue support to assist users in achieving their target behavior. Dialogue support includes praise, rewards, reminders, suggestions, similarity, liking, and social role. The exposome application implemented rewards. The rewards design principle helps users be more open to persuasion via words, images, and symbols as a way to provide positive feedback (Oinas-Kukkonen & Harjumaa, 2008b). In the exposome design, the avatar character provided users with motivation to learn more in each part and badges were awarded after completing the assigned contents.

System credibility. The design principles for the credibility of a persuasion system include trustworthiness, expertise, surface credibility, real-world feel authority, third-party endorsements, and verifiability. The exposome application used (a) trustworthiness by providing truthful, fair, and unbiased information from peer-reviewed studies and (b) expertise by incorporating knowledge and experience of scholars like Christopher Wild, former director of the World Health Organization's International Agency for Research on Cancer; and Stephen Rappaport, Emeritus Professor in Environmental Health Sciences and Director of the Berkeley Center for Exposure Biology.

Social support. The design principles for social support are used to design a motivational system to exert social influence. These principles include social facilitation, social comparison, normative influence, social learning, cooperation, competition, and recognition. The exposome application used the recognition design principle by offering public recognition to increase the likelihood of users' adopting the targeted behavior. Users could share their achievements on social media by clicking a share link.

4.2.6 Workflow Design

The app workflow started with an introduction to the app's main goal. Then, the participants had to go through different parts of the app depending on their assigned group. The app had three different versions depending on the group assigned to the participants.

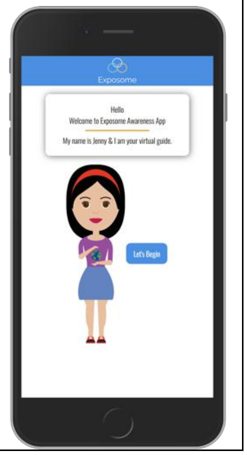
Version 1 (Group A) participants had to watch a 1-minute video that introduced the concept of exposome. Then, the user was asked to join a text message campaign. Version 2 (Group B) participants had to watch a -minute e-video that introduced the concept of the exposome. Then, a graphical story was presented. Finally, the user was asked to join a text message campaign. Version 3 (Group C) participants had to watch a 1-minute video that introduced the concept of the exposome. Then, a graphical story was presented. Users in groups A and B joined a text campaign. Once joined, the first text message was sent immediately followed by two text messages a day (8:00 am and 4:00 pm) for 5 days.

The app was composed of multiple pages that users navigated in a sequence depending on their assigned version or group. Table 11 illustrates the application design.

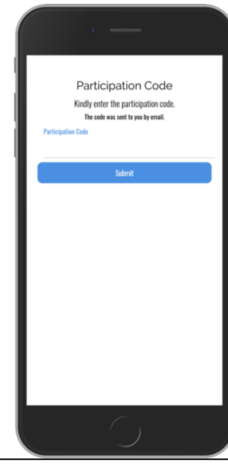
Table 11

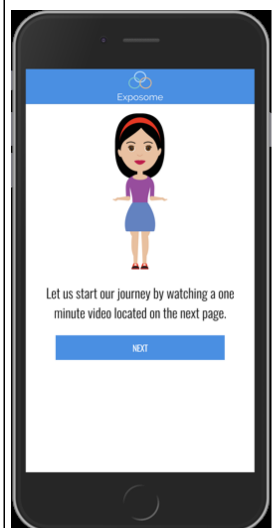
Application Design

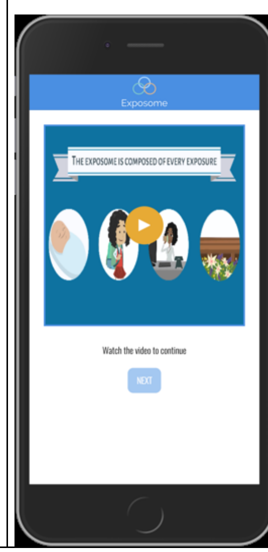
<table>
<tr>
<td>
Page 1: Landing:
Globe image with the app name.
• User click on “LETS GET STARTED” button to continue.
</td>
<td>


</td>
</tr>
<tr>
<td>
Page 2: Introduction:
Introducing Jenny, the virtual character.
• User clicks on “Let’s begin” button to continue
</td>
<td>


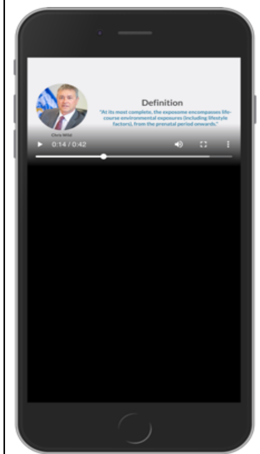
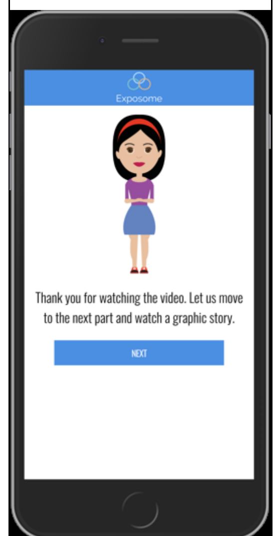
</td>
</tr>
</table>

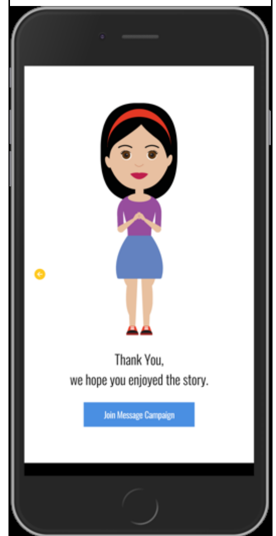
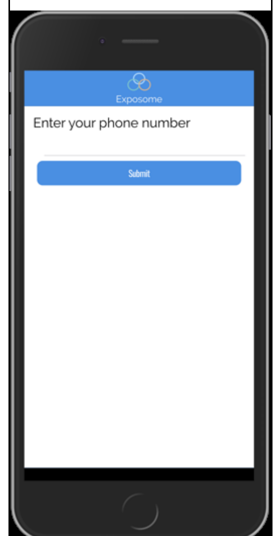
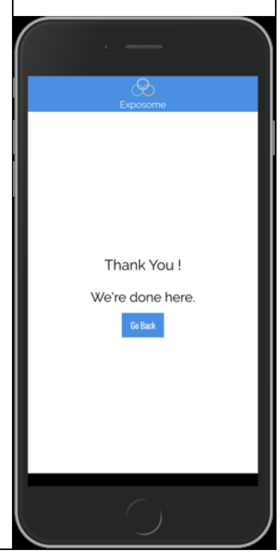
<table>
<tr>
<td>
Page 3: Code Assigned:
Asks the user for the participation code.
User must enter the provided participation code to continue. Codes are stored locally in the app, and user input is validated to match the stored codes.
</td>
<td>


</td>
</tr>
<tr>
<td>
Page 4: Video Introduction:
Jenny asks the user to watch the video.
User clicks on “NEXT” button to continue.
</td>
<td>


</td>
</tr>
<tr>
<td>
Page 5: Video Page:
A thumbnail with play button.
User must watch the video by clicking on the thumbnail, and then click “Next” button to continue.
</td>
<td>


</td>
</tr>
</table>

Playing the video in full screen.	
Page 6: Graphical story introduction: Jenny introduces the user to the graphic story. • User clicks on "Next" button to continue.	
Page 7: Graphical Story Slides: A slideshow of the graphic story (total of 30 images) • User swipes or click to watch all images.	

Last slide of graphic story: The last slide of the story asks the user to join the message campaign. • User clicks on the “Join Message Campaign” to continue.	
Page 8: Phone Number Form A form to ask the user for their phone number. • User enters his phone number and click “Submit” to continue.	
Page 9: Thank You A thank you message for finishing the app. • User can click the “Go Back” button to start the app again.	

4.2.7 Exposome Mobile App Build

The Exposome Awareness App is a multiplatform app built using the Ionic Framework, which is a cross-platform framework for building web and mobile apps. I chose to use Angular integration as the front-end framework for the app. I used Capacitor which is the underlying bridge to run the app on mobile devices. The app is available on three platforms: iOS, Android, and Web. The web version can be accessed using the link http://exposome-awareness.web.app/.

4.2.8 Services Used

I used Google Firebase as the backend of the mobile and web apps. Google Firebase provides many products for mobile and web app developers. I used the following Firebase products.

1. Cloud Firestore: Cloud Firestore stores users' phone numbers, text messages, and queued messages.
2. Cloud Functions: Cloud Functions handle the logic of joining message campaigns and sending the text messages.
3. Authentication: This product was used to authenticate the app to use Firebase services.
4. Hosting: This product was used to host the web app.
5. Google Analytics: Google analytics were used to collect app usage data (see Figure 21).

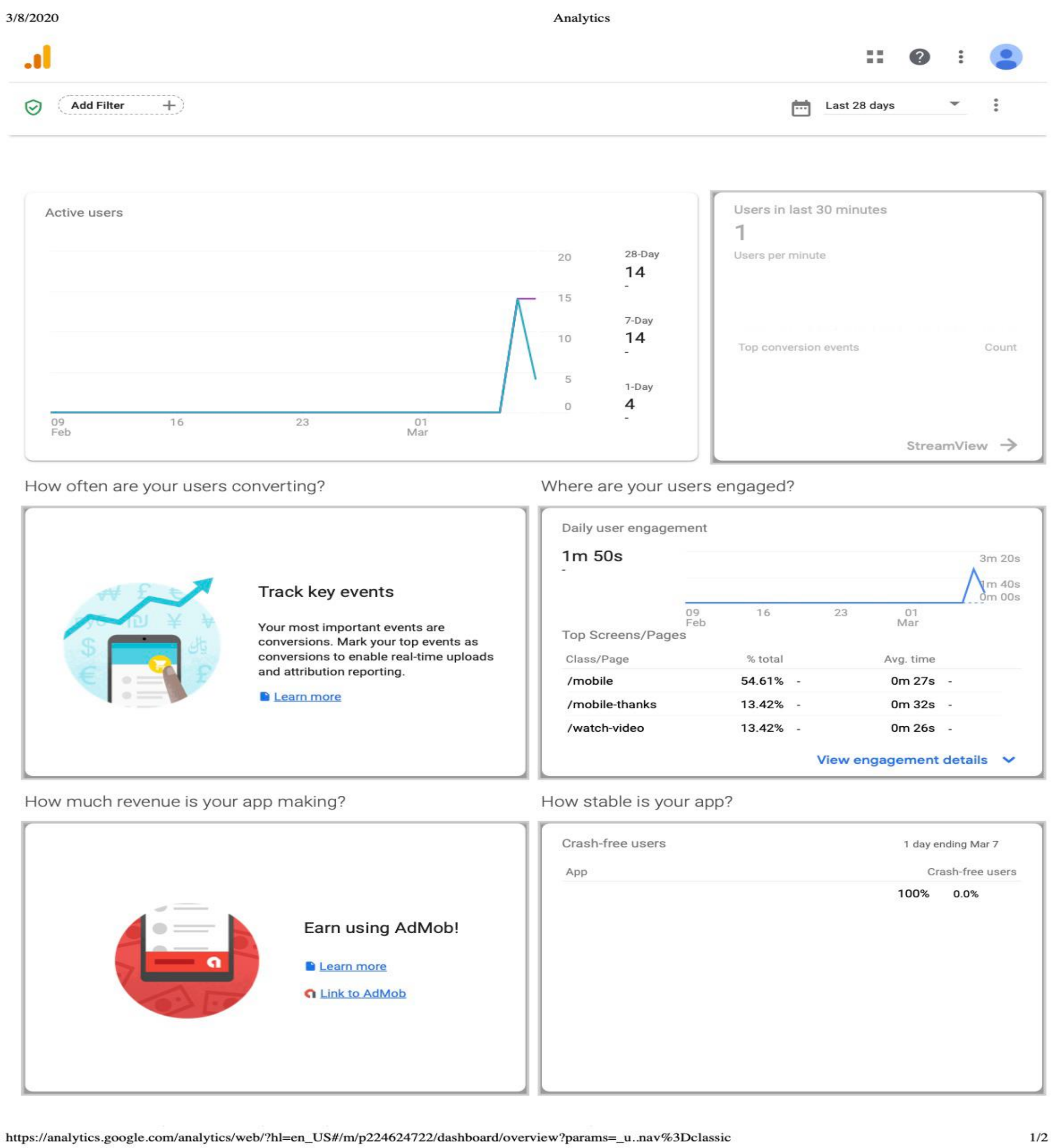
I also used some services provided by Google Cloud, such as Cloud Computing Services. Cloud Storage was used to store the video. Cloud Scheduler was used to send the messages on schedule. I used Twilio software for sending text messages, which is a business communications provider. Twilio's programmable SMS API was used to send text messages to users who subscribed to the text message campaign.

4.2.9 App Logic

- I signed up users anonymously using the Firebase authentication system when the user first opened the app. This allowed the app to communicate with Firebase services.
- I also activated Firebase analytics to collect data about the app usage.
- When the user was asked for the participation code, I compared the value entered by the user with the codes stored in the local memory of the app.
- The graphic story images are stored locally in the app.
- The video is stored in the Google Cloud Storage services and streamed to the user.
- When the user entered their phone number to join the message campaign, I called a Firebase function to validate the number stored the number in the Firebase Firestore database, and sent the user an initial message using Twilio. The function added 10 messages in the queued messages collection of the Firestore database to be sent to the user at a later time.
- A scheduled Firebase function ran twice a day and used Google Scheduler to send the messages from the queued messages collection.

Figure 21

Screenshot of Google Analytics



4.2.10 Evaluation

Evaluation is a vital part of design science research because evaluation provides valuable feedback and a better understanding of a problem, thus improving the design process, quality of

IT artifacts, and scientific rigor (Hevner & Chatterjee, 2010). The evaluation phase is strongly connected with relevance of the problem and scientific rigor. The evaluation of IT artifacts can be done through testing functionality, completeness, consistency, accuracy, performance, quality, reliability, usability, and efficacy (Hevner et al., 2004). The evaluation process started after I received Institutional Review Board (IRB) protocol approval from Claremont Graduate University (CGU).

This study proposed an exposome mobile application that aimed to increase awareness levels through education and proposed a text message campaign aimed at supporting desired behavior. The exposome mobile app was evaluated by conducting a field study. A presurvey and postsurvey were conducted to measure the effectiveness of the exposome awareness mobile app in educating people living in Los Angeles County about the exposome concept. The presurvey started by providing a list that contained 18 factors that are thought to be related to the incidence of cancer in Los Angeles County as described in Section 4.1.1. Participants were asked to select the five factors they believed to be significantly correlated with the incidence of cancer in Los Angeles County. The 18 factors are listed below.

- Ozone concentration
- Diesel emissions
- Toxic releases from facilities
- Traffic density
- Consuming alcoholic beverages
- Consuming fast food
- Consuming canned fruit (added sugar)
- Consuming canned vegetables (added salt)

- Consuming canned beans (added salt)
- Consuming processed fruit (includes artificial colors, flavors and preservatives)
- Consuming processed vegetables
- Being White
- Being Hispanic
- Being Asian American
- Being another race
- Income level
- Poverty
- Education level

The presurvey asked participants for their demographic information such as education level, race, gender, income level, marital status, and employment status.

Similar to the presurvey, participants were asked to complete the postsurvey by selecting the five factors that they believed to be significantly correlated with the incidence of cancer in Los Angeles County. Participants were also asked to fill out a system usability scale (SUS) survey. The SUS was created by John Brooke in 1996 as a reliable and validated measuring tool to evaluate the usability of a given product or service(Peres et al., 2013) The SUS is widely used and can be adapted to different contexts. International standard ISO 924-11 proposed that the measures of usability should cover three aspects: effectiveness, efficiency, and satisfaction (Jordan et al., 1996). SUS is a 10-statement survey that has a 5-point Likert scale where the strength of agreement ranges between strongly agree to strongly disagree. The survey questions are as follows.

1. I think that I would like to use this system frequently.

2. I found the system unnecessarily complex.

3. I thought the system was easy to use.

4. I think that I would need the support of a technical person to be able to use this system.

5. I found the various functions in this system were well integrated.

6. I thought there was too much inconsistency in this system.

7. I would imagine that most people would learn to use this system very quickly.

8. I found the system very cumbersome to use.

9. I felt very confident using the system.

10. I needed to learn a lot of things before I could get going with this system.

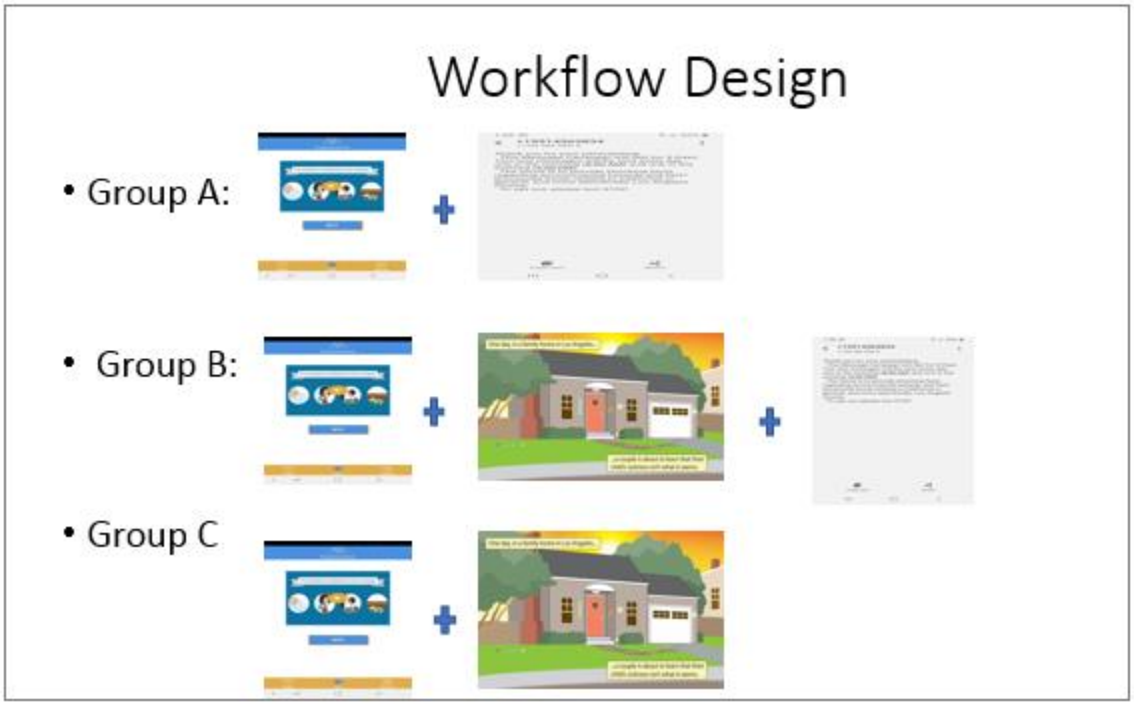
4.2.11 Exposome Awareness App Research Procedures

Invitations to participate in the study were sent using e-mail addresses from the CGU student e-mail list. I asked the individuals who agreed to participate in the study to share the study invitation with their families and friends who live within Los Angeles County. I could not distribute brochures or flyers around the CGU campus due to the COVID-19 “shelter in place” order and campus closures.

The study began on March 25th, 2020 after receiving IRB approval and lasted until April 19th, 2020. Thirty-nine individuals agreed to participate and signed the informed consent form. Each e-mail was assigned a unique identification number in order to track the same participant in the following phases of this study. A presurvey invitation was sent by e-mail using Qualtrics software. Participants were assigned to one of the following groups (see Figure 22).

Figure 22

Workflow Design of Three Groups



- Group A: 13 participants were assigned to watch a short video and then were persuaded directly through the text message campaign.
- Group B: 13 participants were assigned to watch a short video and then were persuaded indirectly through the e-fotonovela and asked to join the text message campaign as a direct route of persuasion.
- Group C: 13 participants were assigned to watch a short video and were persuaded indirectly through the e-fotonovela.

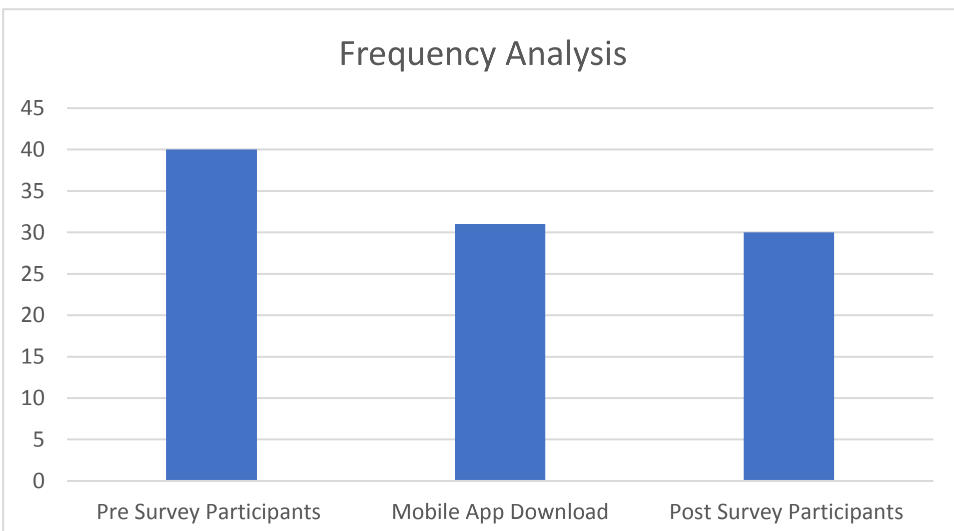
Once participants filled out the presurvey, an email was sent to them with instructions on how to download the app or access it through the web. A code number was provided to participants so that they could access the materials assigned for their group. Participants assigned to Groups A and B had to navigate the app and were then asked to join a text message campaign for 5 days. On the sixth day, an e-mail was sent to the participants asking them to fill out the postsurvey. Participants in Group C had to navigate through the app and an e-mail was sent to the participants to fill out the postsurvey the next day.

4.2.12 Results and Discussion

Figure 23 provides a frequency analysis that shows that 39 participants filled out the presurvey. The exposome awareness mobile app was downloaded 28 times and accessed through direct link to web page 3 times. Thirty participants who filled out the postsurvey received a $20 Amazon gift card. One participant in Group B opted out of the text messages and did not fill out the postsurvey; thus, this participant was excluded from the analysis. Ten participants in Group A, 12 participants in Group B, and 8 participants in Group C completed the postsurvey.

Figure 23

Frequency Analysis for Exposome Awareness Application Facts



Data preparation. I used Qualtrics software to collect pre- and postsurvey data. I used Excel software to compile all observations into one sheet. Participants who did not fill out the postsurvey were excluded from the analysis. Table 12 provides the number of participants in each group along with the total number of participants who completed each survey.

Table 12

Number of Participants in Each Group

Group ID	Presurvey total participants	Postsurvey total participants
A	13	10
B	13	12
C	13	8
Total	39	30

In the Excel sheet, each identification number had one row that included participants' presurvey and postsurvey data. I created a new column named "Group ID" for each participant and converted the group column into numeric values, which were coded as A = 1, B = 2, or C = 3. Participants were asked in the pre and postsurveys to select five factors out of a list of 18 that participants believed to be significantly correlated with the incidence of cancer in Los Angeles County. I numbered each factor from 1 to 18 based on the known effects of the factor, with 1 being the highest-ranked factor and 18 being the lowest. The list was ordered based on the results of the regression model where the strongest factor that correlates with cancer incidents ranked number 1, the second factor ranked number 2, and so on for the first nine factors. The rest of the factors were not part of the regression model (canned food, processed food, education and poverty), so they were ranked based on the following: canned food (beans, vegetables, fruits) were ranked next on the list based by the fact that they are minimally processed with only sugar and/or salt, processed food (vegetable, fruit, fast food) were ranked next based on the fact that they have preservatives, and finally education and poverty ranked last see table 13 . Next, I created 10 new columns: five columns for the presurvey factors and five columns for the posttest factors. I put the five numeric prefactors in each column followed by the five numeric postfactors. Next, I created two new columns; the first column was named “pre survey sum” which added each of the five presurvey scores and the second column was named “post survey sum” which added each of the five postsurvey scores. Next, I created a new column named “Difference” by subtracting presurvey scores from postsurvey scores to calculate the change in participant’s awareness. Note that the lower the sums the more aware the participant is of the leading causes of cancer (in Los Angeles County).

For demographic data, I coded all of the variable data into numeric values. Participants were asked to answer 10 questions as part of the SUS. For questions Q1, Q3, Q5, Q7, and Q9, I coded participants’ responses with a number between 1 and 5, where 1 = strongly disagree and 5= strongly agree. Questions Q2, Q4, Q6, Q8, and Q10 were negatively worded. To calculate each response, 5 were subtracted from the even number and the odd number were subtracted by 1. Then, the total was multiplied by 2.5. The SUS score for each response ranged between 0 and 100.

Table 13

Ranking Exposome Factors

Assumption	Factor	Ranking
Regression Model Results	White Race	1
	Hispanic Race	2
	Other Race	3
	Asian American Race	4
	Fast Food	5
	Toxic Release	6
	Ozone concentration	7
	Diesel Emission	8
	Traffic	9
Minimal Processed Food	Canned Fruit	10
	Canned Vegetable	11
	Canned Beans	12
Highly Processed Food	Fast Food	13
	Processed Fruit	14
	Processed Vegetable	15
	Alcoholic Beverages	16
	Poverty	17
	Education	18

4.2.13 Demographic Analysis

This study included 30 participants; 9 females (30%) and 21 males (70%) participated in the study (see Table 14). For most part, the groups had a similar profile of participants. Both genders participated in each group. Female participants tended to download the application while male participants downloaded the application on their devices and accessed it through the web link too. Thirteen participants were White (43.3%), 2 participants were African American (6.7%), 11 participants were Middle Eastern (36.7%), 1 participant was American Indian (3.3%), 2 participants were Asian (6.7%), and 1 participant was Native Hawaiian (see table 15). White, Black and Middle East ethnicities participated in group A. All ethnicities participated in group B except for the black ethnicity. Group C included only white and middle-eastern participants. Twenty-one participants were married (70%), 1 participant was a widow (3.3%), 1 participant was separated (3.3%), and 7 participants were never married (23.3%; see table 16). Married and never married participants participated in group A. For group B, married, never married, separated and divorced participants agreed to join the study while only married and never married participants were in group C. One participant had only a high school diploma (3.3%), 2 participants had an associate's degree (6.7%), 6 participants had a bachelor's degree (20%), 11 participants had a master's degree (36.7%), and 10 participants had doctoral degrees (33.3%; see table 17). In group A, participants have all levels of education while in group B all participants had education beyond high school. And, all group C participants had an education beyond an associate degree. Three participants preferred not to disclose their income (10%), 4 participants earned less than $10,000 annually (13.3%), 2 participants earned between $10,000 and $19,999 annually (6.7%), 1 participant earned between $20,000 and $29,000 annually (3.3%), 3 participants earned between $30,000 and $39,999 annually (16.7%), 5 participants earned between $40,000 and $49,000 annually (16,7%), 5 participants earned between $50,000 and $59,999 annually (6.7%), 5 participants earned between $60,000 and $69,999 annually (6.7%), 2 participants earned between $70,000 and $79,000 annually (10%), and 3 participants earned between $80,000 and $89,000 annually (see table 18). One participant in group A has the highest yearly income of $60,000 while all other participants in the three groups were below $50,000 a year.

Table 14

Gender

Gender	Group A	Group B	Group C	Total
Male	6	8	7	21
Female	4	4	1	9
Total	10	12	8	30

Table 15

Race

Race	Group A	Group B	Group C	Total
White	5	4	4	13
Black or African American	2	0	0	2
Middle Eastern	3	4	4	11
American Indian or Alaska Native	0	1	0	1
Asian American	0	2	0	2
Native Hawaiian or Pacific Islander	0	1	0	1
Total	10	12	8	30

Table 16

Marital Status

Marital Status	Group A	Group B	Group C	Total
Married	7	8	6	21
Widowed	0	0	0	0
Divorced	0	0	1	1
Separated	0	1	0	1
Never Married	3	3	1	7
Total	10	12	8	30

Table 17

Education

Educational Level	Group A	Group B	Group C	Total
High school	1	0	0	1
Associate degree in college (2-year)	1	1	0	2
Bachelor's degree in college (4-year)	2	3	1	6
Master's degree	3	3	5	11
Doctoral degree	3	5	2	10
Total	10	12	8	30

Table 18

Income

Income	Group A	Group B	Group C	Total
Less than $10,000	3	1	0	4
$10,000 to $19,999	1	1	0	2
$20,000 to $29,999	1	0	0	1
$30,000 to $39,999	0	1	2	3
$40,000 to $49,999	3	1	1	5
$50,000 to $59,999	2	1	2	5
$60,000 to $69,999	0	2	0	2
$70,000 to $79,999	0	1	1	2
$80,000 to $89,999	0	2	1	3
$90,000 to $99,999	0	0	0	0
$100,000 to $149,999	0	0	0	0
$150,000 or more	0	0	0	0
Prefer not to disclose	0	2	1	3
Total	10	12	8	30

4.2.14 Data Analysis

I used SPSS software and ran a simple *t*-test to observe the change in participants' awareness level. The pair sample descriptive statistics presented in Table 19 indicate that the presurvey sum's mean is 40.70 with a standard deviation of 7.7. The posttest sum's mean is 23.20 with a standard deviation of 9.8. As expected, the mean of the postsurvey was lower than the mean of the presurvey. Paired samples correlations are presented in Table 20. The result indicated that the pretest and posttest were positively correlated: Pearson's $r = .49$, $p < .001$.

Table 19

Paired Sample Statistics

		Mean	*N*	Std. deviation	Std. error mean
Pair 1	Presurvey Sum	40.70	30	7.724	1.410
	Postsurvey Sum	23.20	30	9.785	1.787

Table 20

Paired Samples Correlations.

		N	Correlation	*Sig.*
Pair 1	Presurvey sum & Postsurvey sum	30	.485	.007

Table 21 presents the result of running a paired sample *t*-test. The results were statistically significant at $p < .001$. This indicates that a difference exists between presurvey and postsurvey scores; however, running a *t*-test would not specify which persuasive strategy was more effective.

Table 21

Paired Samples Test of Presurvey vs. Postsurvey Scores

		Paired differences							
					Interval of the				
		Mean	Std. deviation	Std. error mean	Lower	Upper	*t*	*df*	*Sig*. (2-tailed)
Pair 1	Presurvey sum-postsurvey sum	17.500	9.058	1.654	14.118	20.882	10.582	29	0.000

Univariate analysis of variance for change in awareness. A 1-way between subjects ANOVA was conducted to compare the effect of the three methods of persuasion (independent variable) on increasing awareness (dependent variable). Table 22 shows the descriptive statistics. The Levene's Test for Equality of Variances is not statistically significant ($p > .05$, see Table 23). This indicates that the group variances are equal in the population. There is a significant effect of the three methods that were used to present the information on increasing awareness at the $p < .05$ level for the three conditions: F (2, 29) = 6.056, $p = .007$ (see Table 24). A post hoc Tukey test (see Table 25) showed that Group A and Group B differed significantly at $p = .003$. Group A and Group C differed significantly at $p = .011$; Group C was not significantly different from Group A ($p = .862$). Thus, the group (A) who watched a video and received a text reminder did not increase their awareness as much as the groups (B and C) who viewed the e-Fotonovela.

Table 22

Descriptive Statistics for Difference of Awareness

Dependent Variable: Difference			
group ID	Mean	Std. Deviation	N
A	10.50	10.763	10
B	21.25	6.210	12
C	20.63	5.012	8
Total	17.50	9.058	30

Table 23

Levene's Test of Equality of Error Variances[a, b]

		Levene statistic	*df1*	*df2*	*Sig.*
Awareness	Based on Mean	2.110	2	27	.141
	Based on Median	1.162	2	27	.328
	Based on Median and with adjusted df	1.162	2	18.062	.335
	Based on trimmed mean	1.861	2	27	.175

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Dependent variable: Difference

b. Design: Intercept + groupID

Table 24

Tests of Between-Subjects Effects: Awareness

Source	Type III sum of squares	*df*	Mean square	*F*	*Sig.*
Corrected model	736.875a	2	368.438	6.056	.007
Intercept	8896.672	1	8896.672	146.236	.000
groupID	736.875	2	368.438	6.056	.007
Error	1642.625	27	60.838		
Total	11567.000	30			
Corrected total	2379.500	29			

a. R Squared = .310 (Adjusted R Squared = .259)

Table 25

Pairwise Comparisons: Awareness

Dependent variable: difference						
					95% Confidence interval for difference, b	
(I) group ID	(J) group ID	Mean difference (I-J)	Std. Error	*Sig.b*	Lower bound	Upper bound
A	B	-10.750*	3.340	.003	-17.603	-3.897
	C	-10.125*	3.700	.011	-17.716	-2.534
B	A	10.750*	3.340	.003	3.897	17.603
	C	.625	3.560	.862	-6.680	7.930
C	A	10.125*	3.700	.011	2.534	17.716
	B	-.625	3.560	.862	-7.930	6.680

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

4.2.15 System Usability Scale Analysis

The SUS was adopted to measure the exposome awareness application usability. I used SPSS to test SUS reliability; the result indicated that the survey responses were reliable (.726 – see Table26). SUS scores for each response ranged between 0 and 100. Test scores above 68 are considered to be above average (Jordan et al., 1996). The exposome awareness application scored 80.3 (see Table 27) based on the responses of 30 participants. This suggests that participants found the app to be usable. The Levene's Test for Equality of Variances is not statistically significant ($p > .05$, see Table 28). This indicates that the group variances are equal in the population. An ANOVA for the SUS by Group was not significant, $p = .209$, see table 29. This indicates that there was no difference across groups in terms of their usability assessment of the exposome application.

Table 26

Reliability Statistics for the System Usability Scale

Reliability statistics	
Cronbach's Alpha	N of Items
.726	10

Univariate Analysis of Variance

Table 27

Descriptive Statistics for the System Usability Scale

Dependent variable: SUS total			
Group ID	Mean	Std. Deviation	*N*
A	78.5000	2.22111	10
B	80.8000	3.77793	12
C	81.5000	3.65474	8
Total	80.3000	3.37945	30

Table 28

Levene's Test of Equality of Error Variances[a,b]

		Levene Statistic	df1	df2	Sig.
SUS_Total	Based on Mean	1.471	2	27	.247
	Based on Median	.746	2	27	.484
	Based on Median and with adjusted df	.746	2	21.610	.486
	Based on trimmed mean	1.389	2	27	.267

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Dependent variable: SUS_Total

b. Design: Intercept + groupID

Table 29

Tests of Between-Subjects Effects

Dependent Variable: SUS_Total					
Source	Type III sum of squares	*df*	Mean square	*F*	*Sig.*
Corrected model	36.300[a]	2	18.150	1.662	.209
Intercept	59854.776	1	59854.776	5480.091	.000
groupID	36.300	2	18.150	1.662	.209
Error	294.900	27	10.922		
Total	62166.000	30			
Corrected total	331.200	29			

a. R Squared = .110 (Adjusted R Squared = .044)

4.2.16 Demographic Characteristics Analysis

Details for this analysis are provided in Appendix E. The sample has relatively more male participants in each group and fewer females. Group C had the highest number of male participants with seven males and only single female. The sample has relatively well educated participants. Group A is the only group that has a participant with high school degree. The sample has more participants from White and Middle Eastern races in each group. Group A has two Black Americans but other groups have none. Group B has one American Indian and Native Hawaiian, two Asian Americans and the other groups do not have those races. The majority of participants in each group are married. No widows participated in the study. One participant in group B and one in group C reported being separated. The rest of the participants reported had never been married. The sample income ranged between less than $10,000 and up to $89,000. Group A ranged between less than $10,000 and $59,999 while group B ranged between less than $10,000 and $89,999. Group C emerged as the highest income group as it ranged between $39,000 and $89,999.

Chi Square tests were run for each of the demographic categories with respect to group (A, B, and C), and none of them were significant, see Tables 30, 31,32,33,34. However, the sample size was too small to conclude that there are no significant differences within demographic categories. To achieve 80% power, with a medium effect size, the sample size would have to be about 90. Given the lack of sample size it is unclear whether the experimental groups were similar in terms of demographic variables.

Table 30

Chi Square Test for Gender

	Value	df	Asymp. sig. (2-sided)
Pearson chi-square	18.800[a]	18	.404
Likelihood ratio	23.558	18	.170
Linear-by-linear association	1.764	1	.184
N. of valid cases	30		

Note. [a.] 30 cells (100%) have expected count less than 5. The minimum expected count is .27.

Table 31

Chi Square Test for Race

	Value	df	Asymp. sig. (2-sided)
Pearson chi-square	11.007[a]	10	.357
Likelihood ratio	12.716	10	.240
Linear-by-linear association	.197	1	.657
N. of valid cases	30		

Note. [a.] 17 cells (94.4%) have expected count less than 5. The minimum expected count is .27.

Table 32

Chi Square Test for Marital Status

	Value	df	Asymp. sig. (2-sided)
Pearson chi-square	4.905[a]	6	.556
Likelihood ratio	5.197	6	.519
Linear-by-linear association	.264	1	.607
N. of valid cases	30		

Note. [a.] 9 cells (75%) have expected count less than 6. The minimum expected count is .27.

Table 33

Chi Square Test for Education

	Value	df	Asymp. sig. (2-sided)
Pearson chi-square	5.598[a]	8	.692
Likelihood ratio	6.133	8	.632
Linear-by-linear association	1.144	1	.285
N. of valid cases	30		

Note. [a.]15 cells (100%) have expected count less than 5. The minimum expected count is .27

Table 34

Chi Square Test for Income

	Value	df	Asymp. sig. (2-sided)
Pearson chi-square	18.800[a]	18	.404
Likelihood ratio	23.558	18	.170
Linear-by-linear association	1.764	1	.184
N. of valid cases	30		

Note. [a.] 30 cells (100%) have expected count less than 5. The minimum expected count is .27

This paragraph describes the relationship between the demographic characteristics and the change in awareness scores. The results indicated that there is no difference in awareness score based on the demographic characteristics. Gender is a two-level factor (male and female). I ran a t-test on gender and the result indicated that there is no significant difference (t = 1.51, p

= .142; see Table 35). Marital status is consolidated as a two-level factor: married or not. I ran a t-test on marital status and the results indicated that there is no significant difference (t = -.586, p = .562; see Table 36). Education is consolidated into two level factors: postgraduate or not. I ran a t-test on education and the result indicated that there is no significant difference (t = 1.037, p = .309; see Table 37). Income is sorted into two-levels: $39,999 or less and greater than $39,999. I ran a t-test on income and the result indicated that there is no significant difference (t = .837, p = .410; see Table 38). Race is consolidated into a three-level factor: White, Middle Eastern, and other. I ran an ANOVA on race and the result indicated there is no significant difference (F= 2.038, p = .150; see Table 39). The sample size required to achieve power of 0.80 is less than the sample size for the above tests, which indicates that the change in awareness is not related to demographic factors.

Table 35

t-test for Gender

	T-test for equality of means			
	t	df	Sig (2-tailed)	Mean difference
Post equal variances assumed	1.510	28	.142	6.36508
Post equal variances not assumed	1.361	12.329	.198	6.36508

Table 36

t-test for Marital Status

	T-test for equality of means			
	t	df	Sig (2-tailed)	Mean difference
Post equal variances assumed	-.586	28	.562	-2.55556
Post equal variances not assumed	-.592	15.528	.562	-2.55556

Table 37

t-test for Education

	T-test for equality of means			
	t	df	Sig (2-tailed)	Mean difference
Post equal variances assumed	1.037	28	.309	4.46032
Post equal variances not assumed	.879	11.117	.398	4.46032

Table 38

t-test for Income

	T-test for equality of means			
	t	df	Sig (2-tailed)	Mean difference
Post equal variances assumed	.837	28	.410	3.35294
Post equal variances not assumed	.814	22.787	.424	3.35294

Table 39

ANOVA Test for Race

	Sum of squares	df	Mean square	F	Sig.
Between groups	444.780	2	222.390	2.038	.150
Within groups	2945.920	27	109.108		
Total	33390.700	29			

Chapter 5: Conclusion, Limitations, and Future Work

The first study using an exposome approach was conducted in the context of Type 2 Diabetes and used a publicly available data set. This study will open the door for new discoveries; however, exposome as a concept is still in its infancy and faces challenges with different technical issues such as measuring tools, data availability and accessibility, advanced statistical methods and bioinformatics, as well as dealing with dig data. The use of new measuring tools, data management software, and advanced statistical approaches might help the research community to link chronic diseases with environment risk factors.

The purpose of this study was to (a) design, build, and evaluate two IT artifacts to better understand the implications associated with environmental exposure on human health and (b) raise the awareness level of individuals residing in Los Angeles County. I adopted the DSR approach methodology to guide this study. The DSR methodology was used to solve a relevant problem by designing, building, and evaluating two IT artifacts. The results of the evaluations illustrated the artifacts' utility and usability.

First, I developed a regression model algorithm to highlight the exposome factors that are significantly correlated with cancer incidents in Los Angeles County. The regression model equation that was developed:

Average cancer incidents = -1.694+ (.158 *White) + (.101*Hispanic) + (.082*Other_poulation) + (.088*Fast_Food) + (.017* tox_release) + (-.151 * diesel) + (12.56*ozone).

The regression model explained 78% of the variance in correlating exposome factors with cancer incidents in Los Angeles County. However, the study had certain limitations in terms of training the model with data sets using a single county in California. The model performance could be enhanced through adding multiple counties' environmental data and combining the biomarker data. Moreover, machine learning tools such as neural network can use sthe data to make generalizations based on selected labels. The cancer data included data on all types of cancer except for skin cancer. This made it difficult to correlate specific types of cancer to exposome factors. Dr. Lee, an assistant professor in Claremont Graduate University's Center for Information Systems and Technology, mentioned that this model would not halt when it comes to predicting on an individual level; more data are required to predict cancer incidents. Future scholars may seek to include different chronic diseases such as asthma, cardiovascular diseases, and obesity in future studies. Future studies could also use data from different counties and group the data into a cluster based on census tract. For example, the pesticide data indicated that some census tracts have values greater than zero and some locations do not have pesticides at all. Certain patterns may emerge from the model through (a) clustering census tracts based on areas that have values greater than zero and (b) studying the effect of those areas on human health. In addition, revisiting the exposome data for population after the 2020 census is complete will update the numbers and allow for further accuracy. The current population variables are based on the census data from 2010. Calculating the yearly increase in population from 2010 to 2020 would allow scholars to revisit the estimates of the 2016 population and rebuild the model to observe any changes in the results.

Second, I developed a mobile application based on the results of the regression model to raise the awareness level of the statistically significantly exposome factors that are correlated with cancer incidents in Los Angeles County. The mobile application provided both direct and indirect persuasion to influence users' attitude regarding the danger of exposure to environmental factors. To evaluate the application, I ran a simple *t*-test, and as expected, the results showed that the mean of the postsurvey was lower than the mean of the presurvey. I also ran an ANOVA, and the results indicated a significant effect indicating that the means of the three methods that were used to present the information on increasing awareness were not equal at the $p < .05$ for the three conditions: $F(2, 29) = 6.056$, $p = .007$. A post hoc Tukey test showed that Group A and Group B differed significantly at $p = .003$. Group A and Group C differed significantly at $p = .011$; Group C was not significantly different from Group A ($p = .862$). Thus, the group (A) who watched a video and received a text reminder did not increase their awareness as much as the groups (B and C) who viewed the e-Fotonovela.

The design of this application can be applied to raise awareness of similar problems. For example, the *Los Angeles County Department of Public Health* can use application design to raise awareness of Coronavirus (COVID-19) among the Los Angeles County population. The contents of the mobile application can be modified to persuade certain populations and change their attitude. The indirect approach can adapt the video and short graphical story with material related to COVID-19 to educate people on how to deal with the pandemic, apply the best practice to stay safe, and minimize the spread of the virus. A direct approach would involve sending text messages to the Los Angeles County population with facts about new cases, recovered cases, and terminal cases. Other text messages could be sent to encourage populations to practice social distancing and provide awareness messages.

The study was limited by the small sample size (30 participants) due to the spread of COVID-19 and campus closures; the small sample size made it more difficult to test the impact of the application on developed awareness. It was difficult to determine which group had statistically significant results when comparing variance between groups with a small sample size. The use of a larger sample size to test the exposome application effectiveness and usability would give more reliable results and better precision. To refine the research design, adding a control group would make it possible to set a benchmark to measure the results of other groups. Control group participants would have to fill out the presurvey and the postsurvey without any intervention. Adding a control group would provide accurate analysis and establish research validity.

Pre and postsurveys were distributed among participants to measure their responses, and I applied quantitative analysis (t-test and ANOVA) to evaluate the effectiveness and usefulness of the exposome awareness application. The research design can be modified to include qualitative research, which could involve an open-ended survey to further understand participants' opinions and thoughts of the application.

To encourage participants to use the application, the text message campaign could be redesigned to include built-in notifications that participants can receive for free. The original design of the survey asked participants to select 5 out of 18 factors that they believed to be significantly correlated with the incidence of cancer in Los Angeles County. The presurvey and postsurvey could also be redesigned to break down the large list into small related groups to encourage participants to make their choices less randomly. This can be done by providing participants with pre- and postsurveys that consist of three items: (a) an air quality group that includes air pollution variables and asks the participants to select only two choices, (b) a food group that includes food variables and asks the participants to make two choices, and (c) a demographic group with demographic variables and asks the participant to select two choices. Finally, the content of the text messages can be changed to relate to the subject, be more compelling, and allow for more participants to engage in to the experiment.

The goal of this study was to develop and evaluate two IT artifacts. The first artifact was a regression model that highlighted the most significant exposome factors that correlated with cancer incidents in Los Angeles County. The second artifact was an awareness application used to raise awareness of the exposome concept based on the result of the regression model. This study contributed to the body of knowledge of the IS&T field by (a) providing a regression model that highlighted the exposome factors that most significantly correlated with cancer incidents in Los Angeles County and (b) providing a health care application that uses persuasive technologies (exposome mobile application and text message campaign) to implement direct and indirect persuasion strategies.

The IT artifacts have the design, build, and evaluation phases of design science principles. Moreover, the artifacts contribute to society by offering a system to help users understand environmental effects on human health. Finally, this study contributed to the body of knowledge by assessing the effectiveness of the mobile app through measuring the effects of specific interventions.

References

Abu-Bader, S. H. (2010). *Advanced and Multivariate Statistical Methods for Social Science Research with a Complete SPSS Guide*. Lyceum Books.

Air pollution reduces global life expectancy by more than one year. (2018). ScienceDaily. https://www.sciencedaily.com/releases/2018/08/180822112406.htm

Alavanja, M. C. R., Hoppin, J. A., & Kamel, F. (2004). Health effects of chronic pesticide exposure: Cancer and neurotoxicity. *Annual Review of Public Health*, *25*, 155–197. https://doi.org/10.1146/annurev.publhealth.25.101802.123020

Alzheimer's Association. (2018). *Facts and Figures*. Alzheimer's Disease and Dementia. https://alz.org/alzheimers-dementia/facts-figures

American Heart Association, A. H. (2017). Heart Disease and Stroke Statistics 2017 At-a-Glance. *On-Line at: Http://Www. Heart. Org/Idc/Groups/Ahamahpublic/@ Wcm/@ Sop/@ Smd/Documents/Downloadable/Ucm_491265. Pdf*.

ATSDR. (2005). *Chapter 6: Exposure Evaluation: Evaluating Exposure Pathways | PHA Guidance Manual*. https://www.atsdr.cdc.gov/hac/phamanual/ch6.html

Azevedo, A. R. P., Sousa, H. M. L. de, Monteiro, J. A. F., & Lima, A. R. N. P. (2015). Future perspectives of Smartphone applications for rheumatic diseases self-management. *Rheumatology International*, *35*(3), 419–431. https://doi.org/10.1007/s00296-014-3117-9

Barbour, K. E., Helmick, C. G., Boring, M., & Brady, T. J. (2017). Vital Signs: Prevalence of Doctor-Diagnosed Arthritis and Arthritis-Attributable Activity Limitation — United States, 2013–2015. *MMWR. Morbidity and Mortality Weekly Report*, *66*(9), 246–253. https://doi.org/10.15585/mmwr.mm6609e1

Bohner, G., Moskowitz, G. B., & Chaiken, S. (1995). The Interplay of Heuristic and Systematic Processing of Social Information. *European Review of Social Psychology*, *6*(1), 33–68. https://doi.org/10.1080/14792779443000003

Cancer Statistics. (2015, April 2). [CgvArticle]. National Cancer Institute. https://www.cancer.gov/about-cancer/understanding/statistics

Canella, D. S., Levy, R. B., Martins, A. P. B., Claro, R. M., Moubarac, J.-C., Baraldi, L. G., Cannon, G., & Monteiro, C. A. (2014). Ultra-Processed Food Products and Obesity in Brazilian Households (2008–2009). *PLOS ONE*, *9*(3), e92752. https://doi.org/10.1371/journal.pone.0092752

Carter, A., Liddle, J., Hall, W., & Chenery, H. (2015). Mobile Phones in Research and Treatment: Ethical Guidelines and Future Directions. *JMIR MHealth and UHealth*, *3*(4), e95. https://doi.org/10.2196/mhealth.4538

CDC. (2018, October 3). *6 Tips for a Healthy Summer*. Centers for Disease Control and Prevention. https://www.cdc.gov/chronicdisease/index.htm

Chen, S., Gong, E., Kazi, D. S., Gates, A. B., Bai, R., Fu, H., Peng, W., De La Cruz, G., Chen, L., Liu, X., Su, Q., Girerd, N., Karaye, K. M., Alhabib, K. F., Yan, L. L., & Schwalm, J. (2018). Using Mobile Health Intervention to Improve Secondary Prevention of Coronary Heart Diseases in China: Mixed-Methods Feasibility Study. *JMIR MHealth and UHealth*, *6*(1). https://doi.org/10.2196/mhealth.7849

Dai, D., Prussin II, A. J., Marr, L., Vikesland, P. J., Edwards, M., & Pruden, A. (2017). *Factors Shaping the Human Exposome in the Built Environment: Opportunities for Engineering Control*. https://pubs-acs-org.ccl.idm.oclc.org/doi/full/10.1021/acs.est.7b01097

Definition of Chronic disease. (n.d.). MedicineNet. Retrieved February 6, 2019, from https://www.medicinenet.com/script/main/art.asp?articlekey=33490 diabetes.org. (2018). *The Cost of Diabetes*. American Diabetes Association. http://www.diabetes.org/advocacy/news-events/cost-of-diabetes.html

Dutzik, T., Baumann, J., & Purvis, M. (n.d.). *TOXIC RELEASES AND HEALTH*. 78.

Escher, B. I., Hackermüller, J., Polte, T., Scholz, S., Aigner, A., Altenburger, R., Böhme, A., Bopp, S. K., Brack, W., Busch, W., Chadeau-Hyam, M., Covaci, A., Eisenträger, A., Galligan, J. J., Garcia-Reyero, N., Hartung, T., Hein, M., Herberth, G., Jahnke, A., … Wambaugh, J. F. (2017). From the exposome to mechanistic understanding of chemical-induced adverse effects. *Environment International*, *99*, 97–106. https://doi.org/10.1016/j.envint.2016.11.029

Festinger, L. (1957). A theory of cognitive dissonance: Stanford Univ Pr. *Fornell, C., & Larcker, DF (1981). Evaluating Structural Equation Models With*.

Fine, L. J., Philogene, G. S., Gramling, R., Coups, E. J., & Sinha, S. (2004). Prevalence of multiple chronic disease risk factors: 2001 National Health Interview Survey. *American Journal of Preventive Medicine*, *27*(2, Supplement), 18–24. https://doi.org/10.1016/j.amepre.2004.04.017

Finkelstein, E. A., Trogdon, J. G., Cohen, J. W., & Dietz, W. (2009). Annual Medical Spending Attributable To Obesity: Payer-And Service-Specific Estimates. *Health Affairs*, *28*(5), w822–w831. https://doi.org/10.1377/hlthaff.28.5.w822

Fogg, B. (1998). Persuasive computers: Perspectives and research directions. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '98*, 225–232. https://doi.org/10.1145/274644.274677

Fogg, B. J. (2003). Overview of captology. In *Persuasive Technology* (pp. 15–22). Elsevier. https://doi.org/10.1016/B978-155860643-2/50003-2

Freedman, J. L. (1965). Long-term behavioral effects of cognitive dissonance. *Journal of Experimental Social Psychology*, *1*(2), 145–155. https://doi.org/10.1016/0022-1031(65)90042-9

Geller, E. S. (2002). The challenge of increasing proenvironment behavior. *Handbook of Environmental Psychology*, *2*, 525–540.

Genome | Learn Science at Scitable. (n.d.). Retrieved November 1, 2018, from https://www.nature.com/scitable/definition/genome-43

Harjumaa, M., & Oinas-Kukkonen, H. (2007). Persuasion Theories and IT Design. In Y. de Kort, W. IJsselsteijn, C. Midden, B. Eggen, & B. J. Fogg (Eds.), *Persuasive Technology* (Vol. 4744, pp. 311–314). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-77006-0_37

Hevner, A., & Chatterjee, S. (2010). Design Science Research in Information Systems. In A. Hevner & S. Chatterjee, *Design Research in Information Systems* (Vol. 22, p. 5). Springer US. https://doi.org/10.1007/978-1-4419-5653-8_2

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). *Design Science in Information Systems Research*. 32.

Hou, L., Zhang, X., Wang, D., & Baccarelli, A. (2012). Environmental chemical exposures and human epigenetics. *International Journal of Epidemiology*, *41*(1), 79–105. https://doi.org/10.1093/ije/dyr154

Hurd, M. D., Martorell, P., Delavande, A., Mullen, K. J., & Langa, K. M. (2013). Monetary Costs of Dementia in the United States. *New England Journal of Medicine*, *368*(14), 1326–1334. https://doi.org/10.1056/NEJMsa1204629

Iglesias, P., & Díez, J. J. (2009). Thyroid dysfunction and kidney disease. *European Journal of Endocrinology*, *160*(4), 503–515. https://doi.org/10.1530/EJE-08-0837

J Natl Cancer Inst. (2011). *Cancer Prevalence and Cost of Care Projections*. https://costprojections.cancer.gov/

Jerrett, M., Burnett, R. T., Pope, C. A., Ito, K., Thurston, G., Krewski, D., Shi, Y., Calle, E., & Thun, M. (2009). Long-Term Ozone Exposure and Mortality. *New England Journal of Medicine*, *360*(11), 1085–1095. https://doi.org/10.1056/NEJMoa0803894

Johanson, G. A., & Brooks, G. P. (2010). Initial Scale Development: Sample Size for Pilot Studies. *Educational and Psychological Measurement*, *70*(3), 394–400. https://doi.org/10.1177/0013164409355692

Johnson, J., Giles, R. T., Larsen, L., Ware, J., Adams, T., & Hunt, S. C. (2005). Utah's Family High Risk Program: Bridging the Gap Between Genomics and Public Health. *Preventing Chronic Disease*, *2*(2). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1327718/

Jones, J. G., & Simons, H. W. (2017). *Persuasion in society*. Routledge.

Jordan, P. W., Thomas, B., McClelland, I. L., & Weerdmeester, B. (1996). *Usability Evaluation In Industry*. CRC Press.

Kardia, S. L., Bodzin, J., Goldenberg, A., Citrin, T., Raup, S. F., & Bach, J. V. (2005). Genomics and Public Health: Development of Web-based Training Tools for Increasing Genomic Awareness. *Preventing Chronic Disease*, *2*(2). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1327719/

Khurshid, A., Brown, L., Mukherjee, S., Abebe, N., & Kulick, D. (2015). Texting for Health: An Evaluation of a Population Approach to Type 2 Diabetes Risk Reduction With a Personalized Message. *Diabetes Spectrum : A Publication of the American Diabetes Association*, *28*(4), 268–275. https://doi.org/10.2337/diaspect.28.4.268

Landrigan, P. J., Fuller, R., Acosta, N. J. R., Adeyi, O., Arnold, R., Basu, N. (Nil), Baldé, A. B., Bertollini, R., Bose-O'Reilly, S., Boufford, J. I., Breysse, P. N., Chiles, T., Mahidol, C., Coll-

Seck, A. M., Cropper, M. L., Fobil, J., Fuster, V., Greenstone, M., Haines, A., … Zhong, M. (2018). The Lancet Commission on pollution and health. *The Lancet*, *391*(10119), 462–512. https://doi.org/10.1016/S0140-6736(17)32345-0

Larson, R. S. (2018). A Path to Better-Quality mHealth Apps. *JMIR MHealth and UHealth*, *6*(7). https://doi.org/10.2196/10414

Lella, A., & Lipsman, A. (2017). *The 2017 U.S. Mobile App Report*. ComScore, Inc. http://www.comscore.com/Insights/Presentations-and-Whitepapers/2017/The-2017-US-Mobile-App-Report

Leonard. (2019). *Junk food and diabetes: The link, the effects, and tips for eating out*. https://www.medicalnewstoday.com/articles/317122

Lindgren, A., Stroh, E., Montnémery, P., Nihlén, U., Jakobsson, K., & Axmon, A. (2009). Traffic-related air pollution associated with prevalence of asthma and COPD/chronic bronchitis. A cross-sectional study in Southern Sweden. *International Journal of Health Geographics*, *8*(1), 2. https://doi.org/10.1186/1476-072X-8-2

Lioy, P. J., & Smith, K. R. (2013). A Discussion of Exposure Science in the 21st Century: A Vision and a Strategy. *Environmental Health Perspectives*, *121*(4), 405–409. https://doi.org/10.1289/ehp.1206170

Lucan, S. C., Karpyn, A., & Sherman, S. (2010). Storing Empty Calories and Chronic Disease Risk: Snack-Food Products, Nutritive Content, and Manufacturers in Philadelphia Corner Stores. *Journal of Urban Health*, *87*(3), 394–409. https://doi.org/10.1007/s11524-010-9453-5

Mahmood, A., Kedia, S., Wyant, D. K., Ahn, S., & Bhuyan, S. S. (2019). Use of mobile health applications for health-promoting behavior among individuals with chronic medical conditions. *Digital Health*, *5*. https://doi.org/10.1177/2055207619882181

Marcolino, M. S., Oliveira, J. A. Q., D'Agostino, M., Ribeiro, A. L., Alkmim, M. B. M., & Novillo-Ortiz, D. (2018). The Impact of mHealth Interventions: Systematic Review of Systematic Reviews. *JMIR MHealth and UHealth*, *6*(1), e23. https://doi.org/10.2196/mhealth.8873

Marshall, A., Medvedev, O., & Antonov, A. (2008). Use of a Smartphone for Improved Self-Management of Pulmonary Rehabilitation. *International Journal of Telemedicine and Applications*, *2008*, 1–5. https://doi.org/10.1155/2008/753064

Martin, S. S., Feldman, D. I., Blumenthal, R. S., Jones, S. R., Post, W. S., McKibben, R. A., Michos, E. D., Ndumele, C. E., Ratchford, E. V., Coresh, J., & Blaha, M. J. (2015). mActive: A Randomized Clinical Trial of an Automated mHealth Intervention for Physical Activity Promotion. *Journal of the American Heart Association: Cardiovascular and Cerebrovascular Disease*, *4*(11). https://doi.org/10.1161/JAHA.115.002239

Mattiello, A., Chiodini, P., Bianco, E., Forgione, N., Flammia, I., Gallo, C., Pizzuti, R., & Panico, S. (2013). Health effects associated with the disposal of solid waste in landfills and incinerators in populations living in surrounding areas: A systematic review. *International Journal of Public Health*, *58*(5), 725–735. https://doi.org/10.1007/s00038-013-0496-8

Miranda, B., Jere, C., Alharbi, O., Lakshmi, S., Khouja, Y., & Chatterjee, S. (2013). Examining the Efficacy of a Persuasive Technology Package in Reducing Texting and Driving Behavior. In S. Berkovsky & J. Freyne (Eds.), *Persuasive Technology* (pp. 137–148). Springer Berlin Heidelberg.

Mosa, A. S. M., Yoo, I., & Sheets, L. (2012). A Systematic Review of Healthcare Applications for Smartphones. *BMC Medical Informatics and Decision Making*, *12*, 67. https://doi.org/10.1186/1472-6947-12-67 mpatino. (2015, April 22). Lung Health Institute | 20 Air Pollution Fast Facts for Earth Day. *Lung Health Instititute*. https://lunginstitute.com/blog/earth-day-2015/

National Institutes of Health Biomarkers Definitions Working Group. (2001). Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework. *Clinical Pharmacology & Therapeutics*, *69*(3), 89–95. https://doi.org/10.1067/mcp.2001.113989

National Research Council. (2012). *Exposure Science in the 21st Century: A Vision and a Strategy*. National Academies Press. https://doi.org/10.17226/13507

NIOSH. (2014). *CDC - Exposome and Exposomics—NIOSH Workplace Safety and Health Topic*. https://www.cdc.gov/niosh/topics/exposome/default.html

Nugent, R. (2008). Chronic Diseases in Developing Countries: Health and Economic Burdens. *Annals of the New York Academy of Sciences*, *1136*(1), 70–79. https://doi.org/10.1196/annals.1425.027

Oinas-Kukkonen, H., & Harjumaa, M. (2008a). A Systematic Framework for Designing and Evaluating Persuasive Systems. In H. Oinas-Kukkonen, P. Hasle, M. Harjumaa, K. Segerståhl, & P. Øhrstrøm (Eds.), *Persuasive Technology* (Vol. 5033, pp. 164–176). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-68504-3_15

Oinas-Kukkonen, H., & Harjumaa, M. (2008b). Towards Deeper Understanding of Persuasion in Software and Information Systems. *First International Conference on Advances in Computer-Human Interaction*, 200–205. https://doi.org/10.1109/ACHI.2008.31

O'Keefe, D. J. (2002). *Persuasion: Theory and Research*. SAGE.

Peres, S., Pham, T., & Phillips, R. (2013). Validation of the System Usability Scale (SUS). *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *57*, 192–196. https://doi.org/10.1177/1541931213571043

Pun, V. C., Kazemiparkouhi, F., Manjourides, J., & Suh, H. H. (2017). Long-Term PM2.5 Exposure and Respiratory, Cancer, and Cardiovascular Mortality in Older US Adults. *American Journal of Epidemiology*, *186*(8), 961–969. https://doi.org/10.1093/aje/kwx166

Quinn, C. C., Clough, S. S., Minor, J. M., Lender, D., Okafor, M. C., & Gruber-Baldini, A. (2008). WellDoc™ Mobile Diabetes Management Randomized Controlled Trial: Change in Clinical and Behavioral Outcomes and Patient and Physician Satisfaction. *Diabetes Technology & Therapeutics*, *10*(3), 160–168. https://doi.org/10.1089/dia.2008.0283

Raghupathi, W., & Raghupathi, V. (2018). An Empirical Study of Chronic Diseases in the United States: A Visual Analytics Approach to Public Health. *International Journal of Environmental Research and Public Health*, *15*(3). https://doi.org/10.3390/ijerph15030431

Rappaport, S. M. (2011). Implications of the exposome for exposure science. *Journal of Exposure Science and Environmental Epidemiology*, *21*(1), 5–9. https://doi.org/10.1038/jes.2010.50

Rappaport, S. M. (2016). Genetic Factors Are Not the Major Causes of Chronic Diseases. *PLOS ONE*, *11*(4), e0154387. https://doi.org/10.1371/journal.pone.0154387

Rappaport, S. M., & Smith, M. T. (2010). Environment and Disease Risks. *Science (New York, N.Y.)*, *330*(6003), 460–461. https://doi.org/10.1126/science.1192603

Reardon, K. K. (1991). *Persuasion in practice*. Sage.

Reynolds, R., Dennis, S., Hasan, I., Slewa, J., Chen, W., Tian, D., Bobba, S., & Zwar, N. (2018). A systematic review of chronic disease management interventions in primary care. *BMC Family Practice*, *19*. https://doi.org/10.1186/s12875-017-0692-3

Rosenheck, R. (2008). Fast food consumption and increased caloric intake: A systematic review of a trajectory towards weight gain and obesity risk. *Obesity Reviews*, *9*(6), 535–547. https://doi.org/10.1111/j.1467-789X.2008.00477.x

Sainani, K. (2016). *Taking on the Exposome | Biomedical Computation Review*. http://biomedicalcomputationreview.org/content/taking-exposome

Simons, H. (1970). *Requirements, problems, and strategies: A theory of persuasion for social movements: Quarterly Journal of Speech: Vol 56, No 1*. https://nca.tandfonline.com/doi/abs/10.1080/00335637009382977?journalCode=rqjs20

Siroux, V., Agier, L., & Slama, R. (2016a). The exposome concept: A challenge and a potential driver for environmental health research. *European Respiratory Review*, *25*(140), 124–129. https://doi.org/10.1183/16000617.0034-2016

Siroux, V., Agier, L., & Slama, R. (2016b). The exposome concept: A challenge and a potential driver for environmental health research. *European Respiratory Review*, *25*(140), 124–129. https://doi.org/10.1183/16000617.0034-2016

Skinner, B. F. (1989). *Upon further reflection*.

Smith, A. (2017). Record shares of Americans have smartphones, home broadband. *Pew Research Center*. http://www.pewresearch.org/fact-tank/2017/01/12/evolution-of-technology/

Society, N. G. (2011, August 18). *Pollution*. National Geographic Society. http://www.nationalgeographic.org/encyclopedia/pollution/

statista.com. (2018). *App stores: Number of apps in leading app stores 2018*. Statista. https://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores/

Statistics of Health Risks From Eating Fast Food. (2018). https://healthyeating.sfgate.com/statistics-health-risks-eating-fast-food-3290.html

Stiff, J. B. (1986). Cognitive processing of persuasive message cues: A meta-analytic review of the effects of supporting information on attitudes. *Communication Monographs*, *53*(1), 75–89. https://doi.org/10.1080/03637758609376128

Stiff, J. B., & Mongeau, P. A. (2016). *Persuasive Communication, Third Edition*. Guilford Publications.

Teng, S., Khong, K. W., & Goh, W. W. (2015). Persuasive Communication: A Study of Major Attitude-Behavior Theories in a Social Media Context. *Journal of Internet Commerce*, *14*(1), 42–64. https://doi.org/10.1080/15332861.2015.1006515

The Human Exposome Project, B. S.-. (2016). *The Human Exposome Project*. The Human Exposome Project. https://humanexposomeproject.com/

Turner, M. C., Nieuwenhuijsen, M., Anderson, K., Balshaw, D., Cui, Y., Dunton, G., Hoppin, J. A., Koutrakis, P., & Jerrett, M. (2017). Assessing the Exposome with External Measures: Commentary on the State of the Science and Research Recommendations. *Annual Review of Public Health*, *38*(1), 215–239. https://doi.org/10.1146/annurev-publhealth-082516-012802

University of Texas at Austin. (2018). *Air pollution reduces global life expectancy by more than one year*. ScienceDaily. https://www.sciencedaily.com/releases/2018/08/180822112406.htm

US Census Bureau. (2019). *Geography Program*. https://www.census.gov/programs-surveys/geography.html van Tongeren, M., & Cherrie, J. W. (2012). An Integrated Approach to the Exposome. *Environmental Health Perspectives*, *120*(3), a103–a104. https://doi.org/10.1289/ehp.1104719

Vaughan, A., Stevanovic, S., Jafari, M., Rahman, M., Bowman, R. V., Fong, K. M., Ristovski, Z., & Yang, I. A. (2019). The effect of diesel emission exposure on primary human bronchial epithelial cells from a COPD cohort: N-acetylcysteine as a potential protective intervention. *Environmental Research*, *170*, 194–202. https://doi.org/10.1016/j.envres.2018.12.035

Vikraman, S., Fryar, C. D., & Ogden, C. L. (2015). Caloric Intake From Fast Food Among Children and Adolescents in the United States, 2011-2012. *NCHS Data Brief*, *213*, 1–8.

Vrijheid, M. (2014). The exposome: A new paradigm to study the impact of environment on health. *Thorax*, *69*(9), 876–878. https://doi.org/10.1136/thoraxjnl-2013-204949

WHO global report. (2017). *WHO | Overview - Preventing chronic diseases: A vital investment*. WHO. http://www.who.int/chp/chronic_disease_report/part1/en/

Wild, C. P. (2005). Complementing the Genome with an “Exposome”: The Outstanding Challenge of Environmental Exposure Measurement in Molecular Epidemiology. *Cancer Epidemiology and Prevention Biomarkers*, *14*(8), 1847–1850. https://doi.org/10.1158/1055-9965.EPI-05-0456

Wild, C. P. (2012). The exposome: From concept to utility. *International Journal of Epidemiology*, *41*(1), 24–32. https://doi.org/10.1093/ije/dyr236

Zhao, Y. Y., & Pei, Y. S. (2012). Risk evaluation of groundwater pollution by pesticides in China: A short review. *Procedia Environmental Sciences*, *13*, 1739–1747. https://doi.org/10.1016/j.proenv.2012.01.167

Appendix A: The Cost of Chronic Disease

Chronic diseases are the leading drivers of health care costs (Centers for Disease Control and Prevention [CDC], 2018). Heart disease and stroke are ranked among the top causes of death in the U.S. with an estimated 810,000 deaths every year at a cost of $330 billion annually to the healthcare system directly and in indirect costs, such as future productivity loss (American Heart Association, 2017).

Cancer is considered to be the second leading cause of death with more than 1.7 million estimated diagnoses in 2018 and nearly 609,000 deaths (National Cancer Institute, 2018). The cost of cancer treatment is expected to reach $174 billion by 2020 (Mariotto et al., 2011).

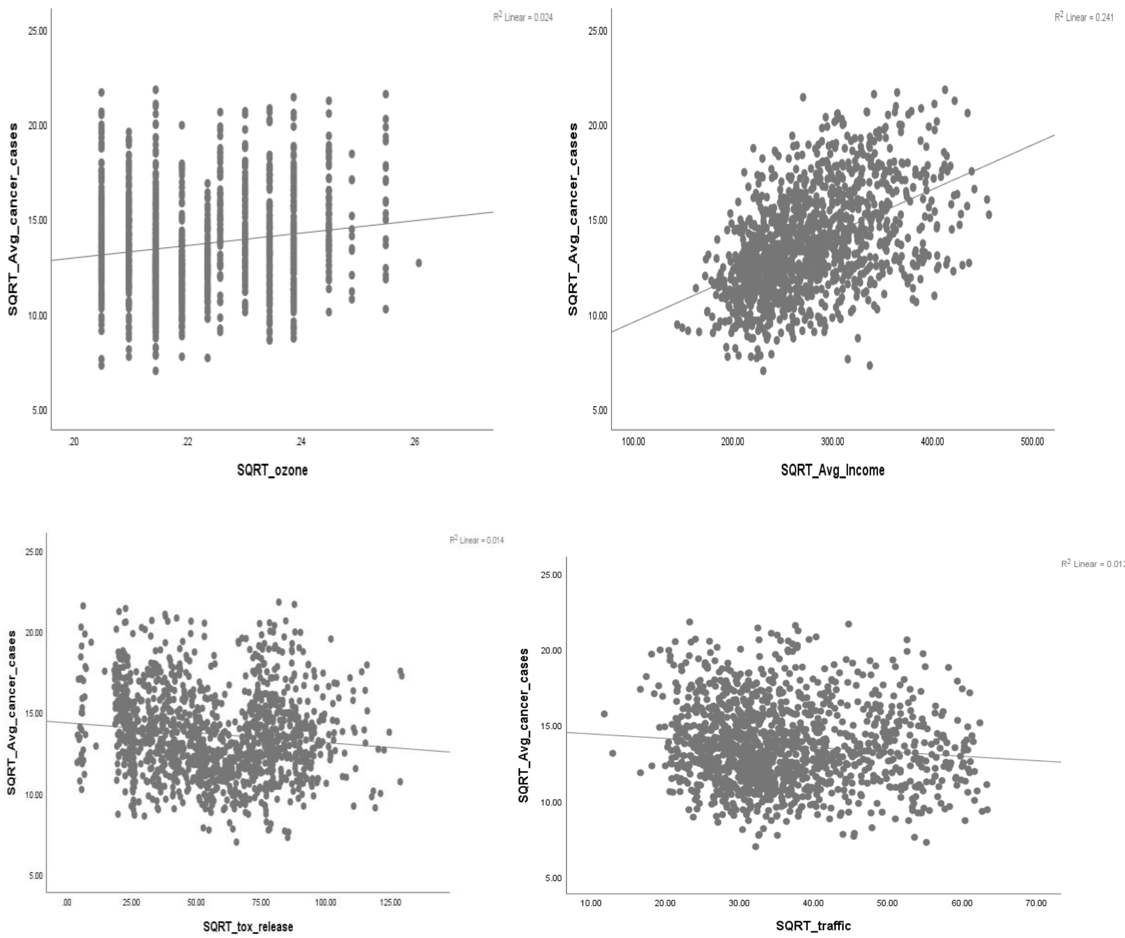
More than 30 million Americans have diabetes and another 84 million adults have prediabetes, which puts them at risk for type 2 (CDC, 2020c). Diabetes treatment cost the US health care system and employers $327 billion in 2017 (American Diabetes Association, 2018).

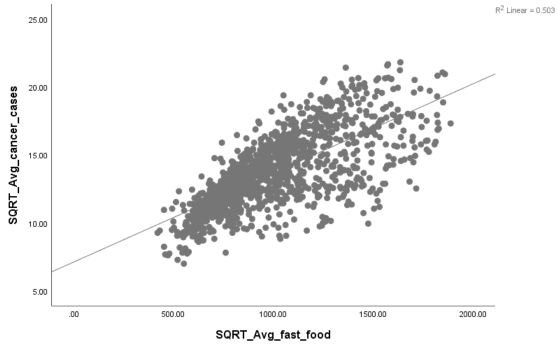
Obesity affects around 1 in 5 children and 1 in 3 adults, which leads to serious outcomes and might cause chronic diseases such as diabetes or cardiovascular diseases (CDC, 2020a). Obesity costs the U.S. health care system $147 billion a year (Finkelstein et al., 2009).

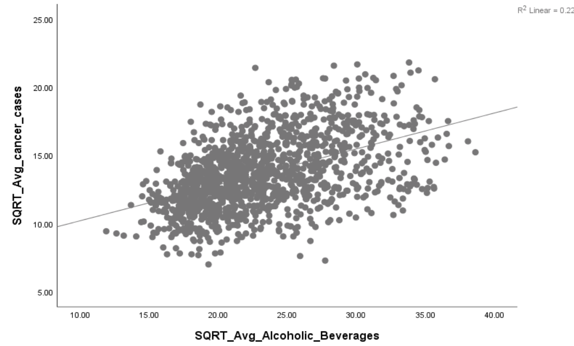
Arthritis affects 54.4 million, or about 1 in 5, adults in the U.S.(Barbour et al., 2017). It is a major cause of work disability and a common cause of chronic pain. Arthritis treatment costs were $304 billion in 2013 (Centers for Disease Control and Prevention, 2020b).

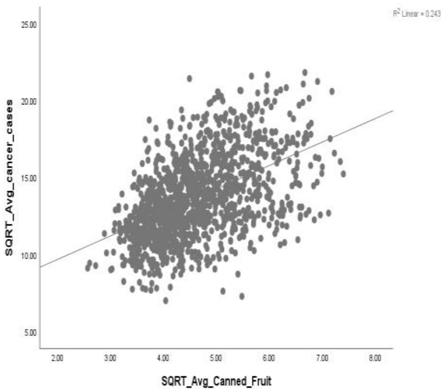
Alzheimer’s disease affects 5.7 million Americans (Alzheimer’s Association, 2018) and costs an estimated $159 billion to $215 billion (Hurd et al., 2013)

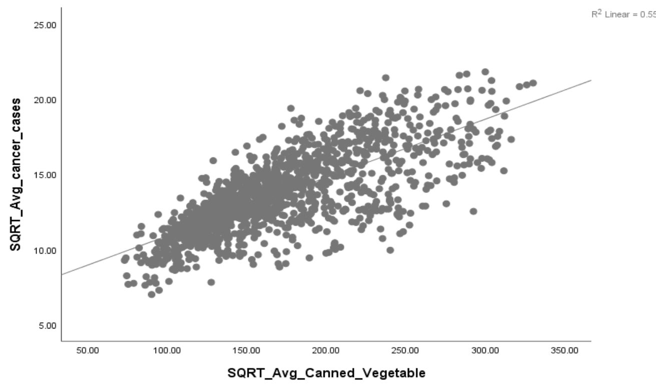
Appendix B: Linear Relationship Between Dependent and Independent Variables

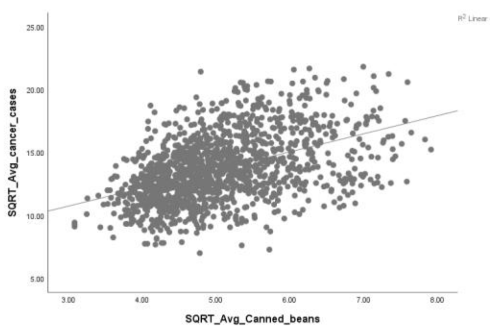


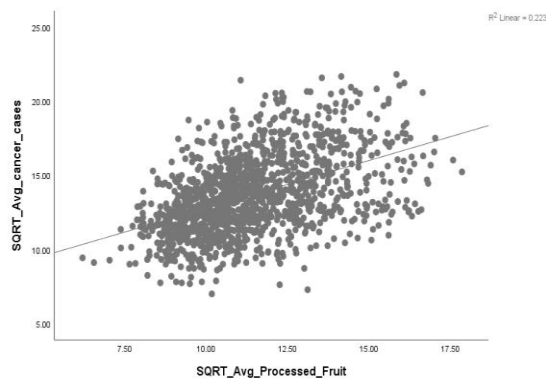

R² Linear = 0.503
SQRT_Avg_cancer_cases
SQRT_Avg_fast_food

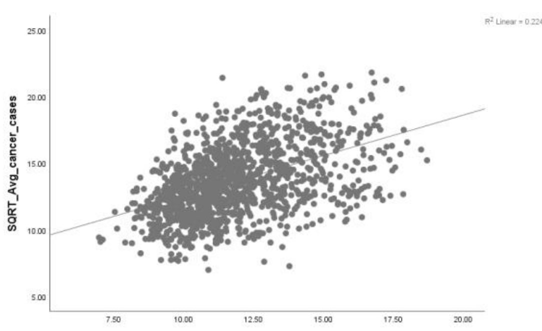

R² Linear = 0.223
SQRT_Avg_cancer_cases
SQRT_Avg_Alcoholic_Beverages

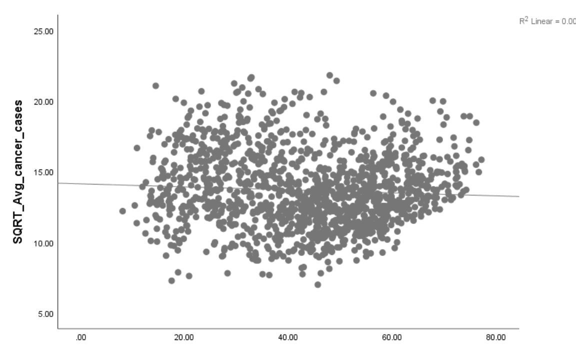

R² Linear = 0.243
SQRT_Avg_cancer_cases
SQRT_Avg_Canned_Fruit

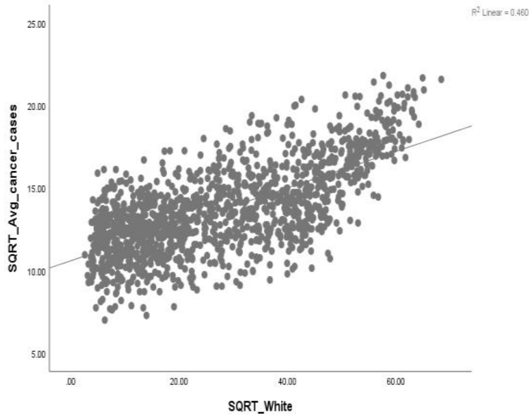

R² Linear = 0.559
SQRT_Avg_cancer_cases
SQRT_Avg_Canned_Vegetable

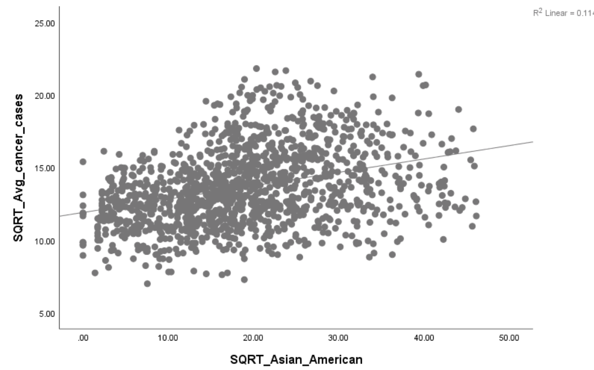

R² Linear = 0.212
SQRT_Avg_cancer_cases
SQRT_Avg_Canned_beans

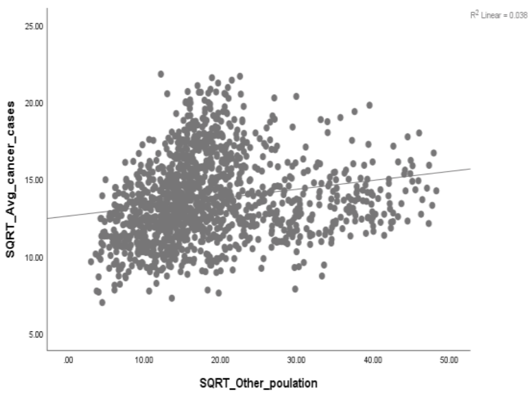

R² Linear = 0.223
SQRT_Avg_cancer_cases
SQRT_Avg_Processed_Fruit

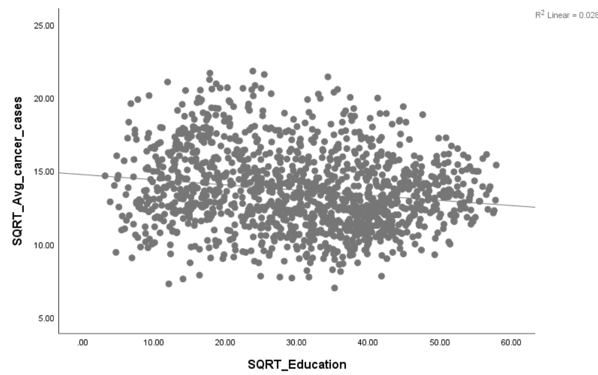

R² Linear = 0.224
SQRT_Avg_cancer_cases
SQRT_Avg_Processed_Vegetables

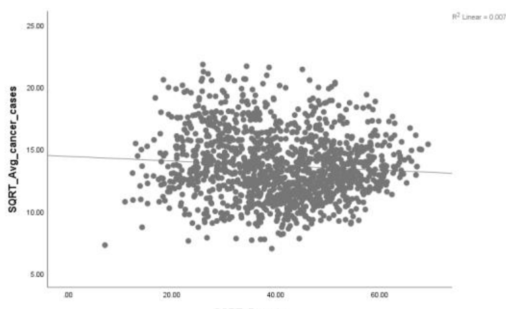

R² Linear = 0.004
SQRT_Avg_cancer_cases
SQRT_Hispanic

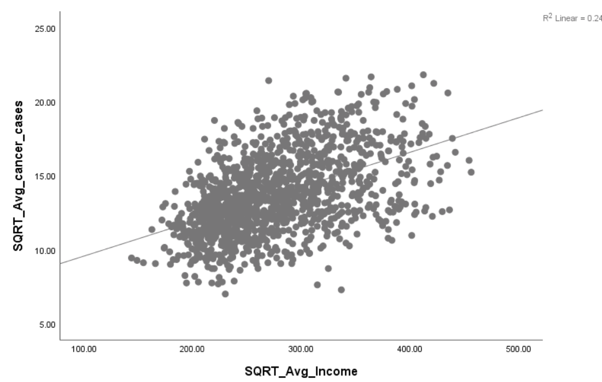

R² Linear = 0.460
SQRT_Avg_cancer_cases
SQRT_White


R² Linear = 0.114
SQRT_Avg_cancer_cases
SQRT_Asian_American


R² Linear = 0.038
SQRT_Avg_cancer_cases
SQRT_Other_poulation


R² Linear = 0.028
SQRT_Avg_cancer_cases
SQRT_Education


R² Linear = 0.007
SQRT_Avg_cancer_cases
SQRT_Poverty


R² Linear = 0.241
SQRT_Avg_cancer_cases
SQRT_Avg_Income

Appendix C: Details of Multicollinearity Analysis

Figure 24

Examining Multicollinearity by Running Pearson Correlation Coefficient Test

		SQRT_ozone	SQRT_diesel	SQRT_tox_release	SQRT_traffic	SQRT_Avg_Alcoholic_Beverage	SQRT_Fast_Food	SQRT_Avg_Canned_Fruit	SQRT_Avg_Canned_Vegetable	SQRT_Avg_Canned_beans	SQRT_Avg_Processed_Fruit	SQRT_Avg_Processed_Vegetables	SQRT_Hispanic	SQRT_White	SQRT_Asian_American	SQRT_Other_poulation	SQRT_Education	SQRT_Poverty	SQRT_Avg_Income
SQRT_ozone	Pearson Correlation	1	-.335**	-.729**	0.048	.110**	.190**	.210**	.131**	.201**	.197**	.203**	-.075**	.299**	.071*	-.285**	-.125**	-.134**	.203**
	Sig. (2-tailed)		0.000	0.000	0.084	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.006	0.000	0.010	0.000	0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_diesel	Pearson Correlation	-.335**	1	.237**	.227**	-.260**	-.373**	-.399**	-.295**	-.371**	-.376**	-.381**	.117**	-.335**	-0.038	-0.018	.216**	.242**	-.399**
	Sig. (2-tailed)	0.000		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.171	0.505	0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_tox_release	Pearson Correlation	-.729**	.237**	1	-.126**	-.240**	-.286**	-.277**	-.243**	-.292**	-.289**	-.290**	.318**	-.403**	0.012	.099**	.302**	.238**	-.270**
	Sig. (2-tailed)	0.000	0.000		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.673	0.000	0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_traffic	Pearson Correlation	0.048	.227**	-.126**	1	-0.033	-0.052	-.068*	-0.049	-.058*	-0.054	-.060*	-0.040	-.055*	-0.001	0.032	-0.020	0.015	-.068*
	Sig. (2-tailed)	0.044	0.000	0.000		0.226	0.058	0.013	0.074	0.037	0.052	0.028	0.151	0.045	0.980	0.243	0.479	0.597	0.013
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Avg_Alcoholic_Beverage	Pearson Correlation	.110**	-.260**	-.240**	-0.033	1	.799**	.774**	.993**	.795**	.795**	.793**	-.458**	.849**	.364**	.081**	-.575**	-.413**	.770**
	Sig. (2-tailed)	0.000	0.000	0.000	0.226		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Fast_Food	Pearson Correlation	.190**	-.373**	-.286**	-0.052	.799**	1	.992**	.830**	.984**	.996**	.992**	-.652**	.766**	.232**	-.136**	-.767**	-.766**	.995**
	Sig. (2-tailed)	0.000	0.000	0.000	0.058	0.000		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Avg_Canned_Fruit	Pearson Correlation	.210**	-.399**	-.277**	-.068*	.774**	.992**	1	.815**	.991**	.994**	.996**	-.619**	.745**	.229**	-.129**	-.741**	-.760**	.998**
	Sig. (2-tailed)	0.000	0.000	0.000	0.013	0.000	0.000		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Avg_Canned_Vegetable	Pearson Correlation	.131**	-.295**	-.243**	-0.049	.993**	.830**	.815**	1	.826**	.826**	.826**	-.477**	.867**	.374**	.073**	-.598**	-.459**	.810**
	Sig. (2-tailed)	0.000	0.000	0.000	0.074	0.000	0.000	0.000		0.000	0.000	0.000	0.000	0.000	0.000	0.008	0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Avg_Canned_beans	Pearson Correlation	.201**	-.371**	-.292**	-.058*	.795**	.984**	.991**	.826**	1	.990**	.998**	-.633**	.744**	.215**	-.114**	-.749**	-.743**	.983**
	Sig. (2-tailed)	0.000	0.000	0.000	0.037	0.000	0.000	0.000	0.000		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Avg_Processed_Fruit	Pearson Correlation	.197**	-.376**	-.289**	-0.054	.795**	.996**	.994**	.826**	.990**	1	.996**	-.628**	.752**	.228**	-.132**	-.745**	-.747**	.992**
	Sig. (2-tailed)	0.000	0.000	0.000	0.052	0.000	0.000	0.000	0.000	0.000		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Avg_Processed_Vegetables	Pearson Correlation	.203**	-.381**	-.290**	-.060*	.793**	.992**	.996**	.826**	.998**	.996**	1	-.636**	.752**	.223**	-.123**	-.752**	-.754**	.992**
	Sig. (2-tailed)	0.000	0.000	0.000	0.028	0.000	0.000	0.000	0.000	0.000	0.000		0.000	0.000	0.000	0.000	0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Hispanic	Pearson Correlation	-.075**	.117**	.318**	-0.040	-.458**	-.652**	-.619**	-.477**	-.633**	-.628**	-.636**	1	-.614**	-.184**	.062*	.916**	.814**	-.617**
	Sig. (2-tailed)	0.006	0.000	0.000	0.151	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.000	0.000	0.023	0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_White	Pearson Correlation	.299**	-.335**	-.403**	-.055*	.849**	.766**	.745**	.867**	.744**	.752**	.752**	-.614**	1	.329**	-.098**	-.703**	-.557**	.745**
	Sig. (2-tailed)	0.000	0.000	0.000	0.045	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.000	0.000	0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Asian_American	Pearson Correlation	.071*	-0.038	0.012	-0.001	.364**	.232**	.229**	.374**	.215**	.228**	.223**	-.184**	.329**	1	-.142**	-.218**	-.122**	.224**
	Sig. (2-tailed)	0.010	0.171	0.673	0.980	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.000	0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Other_poulation	Pearson Correlation	-.285**	-0.018	.099**	0.032	.081**	-.136**	-.129**	.073**	-.114**	-.132**	-.123**	.062*	-.098**	-.142**	1	.120**	.260**	-.142**
	Sig. (2-tailed)	0.000	0.505	0.000	0.243	0.003	0.000	0.000	0.008	0.000	0.000	0.000	0.023	0.000	0.000		0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Education	Pearson Correlation	-.125**	.216**	.302**	-0.020	-.575**	-.767**	-.741**	-.598**	-.749**	-.745**	-.752**	.916**	-.703**	-.218**	.120**	1	.892**	-.739**
	Sig. (2-tailed)	0.000	0.000	0.000	0.479	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Poverty	Pearson Correlation	-.134**	.242**	.238**	0.015	-.413**	-.766**	-.760**	-.459**	-.743**	-.747**	-.754**	.814**	-.557**	-.122**	.260**	.892**	1	-.763**
	Sig. (2-tailed)	0.000	0.000	0.000	0.597	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Avg_Income	Pearson Correlation	.203**	-.399**	-.270**	-.068*	.770**	.995**	.998**	.810**	.983**	.992**	.992**	-.617**	.745**	.224**	-.142**	-.739**	-.763**	1
	Sig. (2-tailed)	0.000	0.000	0.000	0.013	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 24 indicated that significant correlations exist between exposome data. High correlation was detected between Avg_Canned_Fruit and Fast_Food (r = .992, p < .05), Avg_Canned_Vegetable and Avg_Alcoholic_Beverage (r = .993, p < .05), Avg_Canned_Vegetable and Fast_Food ($r = .830$, $p < .05$), Avg_Canned_Vegetable and Avg_Canned_Fruit ($r = .815$, $p < .05$), Avg_Canned_Beans and Fast_Food ($r = .984$, $p < .05$), Avg_Canned_Beans and Avg_Canned_Fruit ($r = .991$, $p < .05$), Avg_Canned_Beans and Avg_Canned_Vegetable ($r = .996$, $p < .05$), Avg_Processed_Fruit and Fast_Food ($r = .826$, $p < .05$), Avg_Processed_Fruit and Avg_Canned_Fruit ($r = .994$, $p < .05$), Avg_Processed_Fruit and Avg_Canned_Vegetable ($r = .826$, $p < .05$), Avg_Processed_Fruit and Avg_Canned_Beans ($r = .990$, $p < .05$), White and Avg_Alcoholic_Beverage ($r = .849$, $p < .05$), White and Avg_Canned_Vegetable ($r = .867$, $p < .05$), Low_Education and Hispanic ($r = .916$, $p < .05$), Poverty and Hispanic ($r = .814$, $p < .05$), Poverty and Education ($r = .892$, $p < .05$), Avg_Income and Fast_Food ($r = .995$, $p < .05$), Avg_Income and Avg_Canned_Fruit ($r = .998$, $p < .05$), Avg_Income and Avg_Canned_Vegetable ($r = .810$, $p < .05$), Avg_Income and Avg_Canned_BeansBeans ($r = .983$, $p < .05$), Avg_Income and Avg_Processed_Fruit ($r = .992$, $p < .05$), and Avg_Income and Avg_Processed_Vegetables ($r = .992$, $p < .05$).

I consolidated all canned variables as a result of the high correlation between food variables (canned and processed) and other variables (Avg_canned_fruit, Agv_canned_vegetable, Avg_canned_Beans) to form one variable named "Avg_canned_food." The same procedure was conducted with processed food (Avg_processed_fruit and Avg_processed_vegetable) to form a new variable named "Avg_processed_food." I calculated the Pearson correlation coefficients with these variables (see Figure 25).

Figure 25

Examining Multiclonality by Running Pearson Correlation Coefficient Test for the Second Time

Correlations

		SQRT_ozone	SQRT_diesel	SQRT_tox_release	SQRT_traffic	SQRT_Avg_Alcoholic_Beverage	SQRT_Fast_Food	SQRT_Hispanic	SQRT_White	SQRT_Asian_American	SQRT_Other_poulation	SQRT_Education	SQRT_Poverty	SQRT_Avg_Income	Avg_Canned_food	Avg_Porcessed_food
SQRT_ozone	Pearson Correlation	1	-.335**	-.729**	0.048	.110**	.190**	-.075**	.299**	.071*	-.285**	-.125**	-.134**	.203**	.134**	.200**
	Sig. (2-tailed)		0.000	0.000	0.084	0.000	0.000	0.006	0.000	0.010	0.000	0.000	0.000	0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_diesel	Pearson Correlation	-.335**	1	.237**	.227**	-.260**	-.373**	.117**	-.335**	-0.038	-0.018	.216**	.242**	-.399**	-.300**	-.379**
	Sig. (2-tailed)	0.000		0.000	0.000	0.000	0.000	0.000	0.000	0.171	0.505	0.000	0.000	0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_tox_release	Pearson Correlation	-.729**	.237**	1	-.126**	-.240**	-.286**	.318**	-.403**	0.012	.099**	.302**	.238**	-.270**	-.246**	-.290**
	Sig. (2-tailed)	0.000	0.000		0.000	0.000	0.000	0.000	0.000	0.673	0.000	0.000	0.000	0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_traffic	Pearson Correlation	0.048	.227**	-.126**	1	-0.033	-0.052	-0.040	-.055*	-0.001	0.032	-0.020	0.015	-.068*	-0.050	-.057*
	Sig. (2-tailed)	0.044	0.000	0.000		0.226	0.058	0.151	0.045	0.980	0.243	0.479	0.597	0.013	0.070	0.038
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Avg_Alcoholic_Beverage	Pearson Correlation	.110**	-.260**	-.240**	-0.033	1	.799**	-.458**	.849**	.364**	.081**	-.575**	-.413**	.770**	.992**	.795**
	Sig. (2-tailed)	0.000	0.000	0.000	0.226		0.000	0.000	0.000	0.000	0.003	0.000	0.000	0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Fast_Food	Pearson Correlation	.190**	-.373**	-.286**	-0.052	.799**	1	-.652**	.766**	.232**	-.136**	-.767**	-.766**	.995**	.840**	.995**
	Sig. (2-tailed)	0.000	0.000	0.000	0.058	0.000		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Hispanic	Pearson Correlation	-.075**	.117**	.318**	-0.040	-.458**	-.652**	1	-.614**	-.184**	.062*	.916**	.814**	-.617**	-.485**	-.633**
	Sig. (2-tailed)	0.006	0.000	0.000	0.151	0.000	0.000		0.000	0.000	0.023	0.000	0.000	0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_White	Pearson Correlation	.299**	-.335**	-.403**	-.055*	.849**	.766**	-.614**	1	.329**	-.098**	-.703**	-.557**	.745**	.868**	.753**
	Sig. (2-tailed)	0.000	0.000	0.000	0.045	0.000	0.000	0.000		0.000	0.000	0.000	0.000	0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Asian_American	Pearson Correlation	.071*	-0.038	0.012	-0.001	.364**	.232**	-.184**	.329**	1	-.142**	-.218**	-.122**	.224**	.371**	.226**
	Sig. (2-tailed)	0.010	0.171	0.673	0.980	0.000	0.000	0.000	0.000		0.000	0.000	0.000	0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Other_poulation	Pearson Correlation	-.285**	-0.018	.099**	0.032	.081**	-.136**	.062*	-.098**	-.142**	1	.120**	.260**	-.142**	.067*	-.128**
	Sig. (2-tailed)	0.000	0.505	0.000	0.243	0.003	0.000	0.023	0.000	0.000		0.000	0.000	0.000	0.015	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Education	Pearson Correlation	-.125**	.216**	.302**	-0.020	-.575**	-.767**	.916**	-.703**	-.218**	.120**	1	.892**	-.739**	-.606**	-.750**
	Sig. (2-tailed)	0.000	0.000	0.000	0.479	0.000	0.000	0.000	0.000	0.000	0.000		0.000	0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Poverty	Pearson Correlation	-.134**	.242**	.238**	0.015	-.413**	-.766**	.814**	-.557**	-.122**	.260**	.892**	1	-.763**	-.471**	-.751**
	Sig. (2-tailed)	0.000	0.000	0.000	0.597	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Avg_Income	Pearson Correlation	.203**	-.399**	-.270**	-.068*	.770**	.995**	-.617**	.745**	.224**	-.142**	-.739**	-.763**	1	.820**	.993**
	Sig. (2-tailed)	0.000	0.000	0.000	0.013	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
Avg_Canned_food	Pearson Correlation	.134**	-.300**	-.246**	-0.050	.992**	.840**	-.485**	.868**	.371**	.067*	-.606**	-.471**	.820**	1	.837**
	Sig. (2-tailed)	0.000	0.000	0.000	0.070	0.000	0.000	0.000	0.000	0.000	0.015	0.000	0.000	0.000		0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
Avg_Porcessed_food	Pearson Correlation	.200**	-.379**	-.290**	-.057*	.795**	.995**	-.633**	.753**	.226**	-.128**	-.750**	-.751**	.993**	.837**	1
	Sig. (2-tailed)	0.000	0.000	0.000	0.038	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

The result of running the Pearson correlation coefficient for the second time indicated a high correlation between new consolidated variables and other variables: Avg_Canned_food and Avg_Alcoholic_Beverage (r = .992, p < .05), Avg_Canned_food and Fast_Food (r = .840, p < .05), Avg_Canned_food and White (r = .868, p < .05), Avg_Canned_food and Avg_Income (r = .820, p < .05), Avg_Processed_food and Fast_Food (r = .995, p < .05), Avg_Processed_food and Avg_Income (r = .993, p < .05), Avg_Processed_food and Avg_Canned_food (r = .837, p < .05). I made the decision to delete the Avg_Canned_food and Avg_Processed_food variables.

Figure 26 illustrates the result of running the Pearson correlation coefficient for the third time. A high correlation existed between White and Avg_Alcoholic_Beverage (r = .849, p < .05), Education and Hispanic (r = .916, p < .05), Poverty and Hispanic (r = .814, p < .05), Poverty and Education (r = .892, p < .05), Avg_Income and Fast_Food (r = .995, p < .05). I made the decision to keep the race variables (Hispanic and White) and Fast_Food and exclude the Avg_Alcoholic_Beverage, Education, Poverty, and Avg_Income variables from the analysis.

Figure 26

Examining Multicollinearity by Running Pearson Correlation Coefficient Test for the Third Time

Correlations

		SQRT_ozone	SQRT_diesel	SQRT_tox_release	SQRT_traffic	SQRT_Avg_Alcoholic_Beverage	SQRT_Fast_Food	SQRT_Hispanic	SQRT_White	SQRT_Asian_American	SQRT_Other_poulation	SQRT_Education	SQRT_Poverty	SQRT_Avg_Income
SQRT_ozone	Pearson Correlation	1	-.335**	-.729**	0.048	.110**	.190**	-.075**	.299**	.071*	-.285**	-.125**	-.134**	.203**
	Sig. (2-tailed)		0.000	0.000	0.084	0.000	0.000	0.006	0.000	0.010	0.000	0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_diesel	Pearson Correlation	-.335**	1	.237**	.227**	-.260**	-.373**	.117**	-.335**	-0.038	-0.018	.216**	.242**	-.399**
	Sig. (2-tailed)	0.000		0.000	0.000	0.000	0.000	0.000	0.000	0.171	0.505	0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_tox_release	Pearson Correlation	-.729**	.237**	1	-.126**	-.240**	-.286**	.318**	-.403**	0.012	.099**	.302**	.238**	-.270**
	Sig. (2-tailed)	0.000	0.000		0.000	0.000	0.000	0.000	0.000	0.673	0.000	0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_traffic	Pearson Correlation	0.048	.227**	-.126**	1	-0.033	-0.052	-0.040	-.055*	-0.001	0.032	-0.020	0.015	-.068*
	Sig. (2-tailed)	0.044	0.000	0.000		0.226	0.058	0.151	0.045	0.980	0.243	0.479	0.597	0.013
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Avg_Alcoholic_Beverage	Pearson Correlation	.110**	-.260**	-.240**	-0.033	1	.799**	-.458**	.849**	.364**	.081**	-.575**	-.413**	.770**
	Sig. (2-tailed)	0.000	0.000	0.000	0.226		0.000	0.000	0.000	0.000	0.003	0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Fast_Food	Pearson Correlation	.190**	-.373**	-.286**	-0.052	.799**	1	-.652**	.766**	.232**	-.136**	-.767**	-.766**	.995**
	Sig. (2-tailed)	0.000	0.000	0.000	0.058	0.000		0.000	0.000	0.000	0.000	0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Hispanic	Pearson Correlation	-.075**	.117**	.318**	-0.040	-.458**	-.652**	1	-.614**	-.184**	.062*	.916**	.814**	-.617**
	Sig. (2-tailed)	0.006	0.000	0.000	0.151	0.000	0.000		0.000	0.000	0.023	0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_White	Pearson Correlation	.299**	-.335**	-.403**	-.055*	.849**	.766**	-.614**	1	.329**	-.098**	-.703**	-.557**	.745**
	Sig. (2-tailed)	0.000	0.000	0.000	0.045	0.000	0.000	0.000		0.000	0.000	0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Asian_American	Pearson Correlation	.071*	-0.038	0.012	-0.001	.364**	.232**	-.184**	.329**	1	-.142**	-.218**	-.122**	.224**
	Sig. (2-tailed)	0.010	0.171	0.673	0.980	0.000	0.000	0.000	0.000		0.000	0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Other_poulation	Pearson Correlation	-.285**	-0.018	.099**	0.032	.081**	-.136**	.062*	-.098**	-.142**	1	.120**	.260**	-.142**
	Sig. (2-tailed)	0.000	0.505	0.000	0.243	0.003	0.000	0.023	0.000	0.000		0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Education	Pearson Correlation	-.125**	.216**	.302**	-0.020	-.575**	-.767**	.916**	-.703**	-.218**	.120**	1	.892**	-.739**
	Sig. (2-tailed)	0.000	0.000	0.000	0.479	0.000	0.000	0.000	0.000	0.000	0.000		0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Poverty	Pearson Correlation	-.134**	.242**	.238**	0.015	-.413**	-.766**	.814**	-.557**	-.122**	.260**	.892**	1	-.763**
	Sig. (2-tailed)	0.000	0.000	0.000	0.597	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Avg_Income	Pearson Correlation	.203**	-.399**	-.270**	-.068*	.770**	.995**	-.617**	.745**	.224**	-.142**	-.739**	-.763**	1
	Sig. (2-tailed)	0.000	0.000	0.000	0.013	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318	1318

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Figure 27

Examining Multicollinearity by Running Pearson Correlation Coefficient Test for the Fourth Time

Correlations

		SQRT_ozone	SQRT_diesel	SQRT_tox_release	SQRT_traffic	SQRT_Fast_Food	SQRT_Hispanic	SQRT_White	SQRT_Asian_American	SQRT_Other_poulation
SQRT_ozone	Pearson Correlation	1	-.335**	-.729**	0.048	.190**	-.075**	.299**	.071*	-.285**
	Sig. (2-tailed)		0.000	0.000	0.084	0.000	0.006	0.000	0.010	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_diesel	Pearson Correlation	-.335**	1	.237**	.227**	-.373**	.117**	-.335**	-0.038	-0.018
	Sig. (2-tailed)	0.000		0.000	0.000	0.000	0.000	0.000	0.171	0.505
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_tox_release	Pearson Correlation	-.729**	.237**	1	-.126**	-.286**	.318**	-.403**	0.012	.099**
	Sig. (2-tailed)	0.000	0.000		0.000	0.000	0.000	0.000	0.673	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_traffic	Pearson Correlation	0.048	.227**	-.126**	1	-0.052	-0.040	-.055*	-0.001	0.032
	Sig. (2-tailed)	0.044	0.000	0.000		0.058	0.151	0.045	0.980	0.243
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Fast_Food	Pearson Correlation	.190**	-.373**	-.286**	-0.052	1	-.652**	.766**	.232**	-.136**
	Sig. (2-tailed)	0.000	0.000	0.000	0.058		0.000	0.000	0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Hispanic	Pearson Correlation	-.075**	.117**	.318**	-0.040	-.652**	1	-.614**	-.184**	.062*
	Sig. (2-tailed)	0.006	0.000	0.000	0.151	0.000		0.000	0.000	0.023
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_White	Pearson Correlation	.299**	-.335**	-.403**	-.055*	.766**	-.614**	1	.329**	-.098**
	Sig. (2-tailed)	0.000	0.000	0.000	0.045	0.000	0.000		0.000	0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Asian_American	Pearson Correlation	.071*	-0.038	0.012	-0.001	.232**	-.184**	.329**	1	-.142**
	Sig. (2-tailed)	0.010	0.171	0.673	0.980	0.000	0.000	0.000		0.000
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318
SQRT_Other_poulation	Pearson Correlation	-.285**	-0.018	.099**	0.032	-.136**	.062*	-.098**	-.142**	1
	Sig. (2-tailed)	0.000	0.505	0.000	0.243	0.000	0.023	0.000	0.000	
	N	1318	1318	1318	1318	1318	1318	1318	1318	1318

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Appendix D: e-fotonovela Script

Page 1–Sarah and Lawrence's Home

Panel 1: A suburban home in Los Angeles. It's the morning, and it's a normal middle-class home: small yard, etc.

LOCATION CAPTION: One day, in a family home in Los Angeles, a couple is about to learn that their child's sickness isn't what it seems.

Panel 2: Inside of the house. We're in the kitchen, and we see Lawrence looking through a medicine cabinet on top of the counter. Sarah, his wife, is dressed for work and coming in from the other room.

SARAH: Oh, honey, you're still home. I thought you'd left with Daniel already. What's going on?

Panel 3: Lawrence shows her the bottle he was looking at – cold medicine. She sighs.

SARAH: He's sick again, I'm guessing.

LAWRENCE: Yep. I've already let the school know.

Panel 4: Lawrence and Sarah look at each other with concern across the table.

LAWRENCE: I'll stay home with him this time.

SARAH: Okay. Let me know if you need anything, but I guess we've got lots of medicine from last time. Love you.

Page 2 – Daniel's bedroom

Panel 1: Daniel's bedroom is scattered with books, games, action figures, art supplies, and other toys appropriate to an 8-year-old kid. He's asleep on his bed. Lawrence is sitting next to him in a chair, on the phone.

LAWRENCE: I understand what you're saying, and I appreciate the school's concern over my son. But we've already been to doctors and they say he's just prone to getting sick. I'm not sure what you want us to do.

Panel 2: Shot of a school counselor, dressed in professional clothing, in her office: bookcases, a big desk, etc.

COUNSELOR: I can't make medical decisions for your family, Mr. Blackwell. I just want to make sure that you know your son's sick days are setting him behind.

Panel 3: Shot of Lawrence, who is reaching over to pull up the covers on his son's bed while holding the phone.

LAWRENCE: He'll catch up. He's a smart kid.

COUNSELOR: (on the phone) His teachers say so, but he's going to start struggling if this keeps up.

Page 3 – Living room

Panel 1: Lawrence and Sarah are sitting on their living room couch. Shot of them facing each other.

SARAH: I'm telling you, this never used to happen before we moved to LA.

LAWRENCE: You're right, but it's got to be a coincidence. I don't know why LA would be any different from Austin. They're both just big sunny cities.

Panel 2: Shot of Sarah, looking very concerned.

SARAH: He's been to the doctor so many times….

LAWRENCE: Maybe some people just get sick.

Panel 3: Daniel comes in through another room, wrapped in a blanket, looking ill.

LAWRENCE: You okay there, son?

DANIEL: Yeah, I just need some water. I don't have to go to the doctor again, do I? Last time, I missed Jake's birthday.

LAWRENCE: You might. I'm sorry.

DANIEL: I wish I could just go play outside.

LAWRENCE: I wish that, too.

PANEL 4: Shot of Sarah looking out the window.

SARAH: Maybe this time we'll get to the bottom of this.

Page 4 – Doctor's office

Panel 1: Lawrence and Sarah in the Doctor's office with Daniel, who is in a patient's gown. The doctor is an older, Asian male.

LOCATION CAPTION: One day later…

LAWRENCE: We're just hoping you can tell us anything new.

DOCTOR: It's hard to say for sure. This just looks like a respiratory infection. But, given how often Daniel's been getting sick, I may have some information that could help you.

Panel 2: Shot of the doctor.

DOCTOR: Something that can make people more prone to illness is environmental exposure – or exposome. There's even a new study linking exposure in LA to cancer.

Panel 3: Shot of Sarah, looking curious – this is a potential answer to her son's sickness.

SARAH: Environmental exposure… you mean, pollution in the air? I hadn't even thought of that.

Panel 4: Shot of the doctor.

DOCTOR: That's part of it, but more specifically, it's related to ozone, diesel, and toxic release from emissions from facilities.

Panel 5: Shot of both Sarah and the doctor, with Lawrence in the background.

SARAH: So something in the air might be harming Daniel.

DOCTOR: It could be. And, it's not just the air. Diet is also a source of exposome. Does Daniel eat fast food? Processed food?

Panel 6: Shot of Lawrence and the Doctor, with Sarah in the background.

LAWRENCE: I mean, of course. We've got busy lives, and no one's perfect. And the kid likes French fries.

DOCTOR: That could be a problem. You'll find exposome in fast food, like French fries, or processed food that's high in sugar and fat, like candy or dessert. These things stand a real chance of causing cancer.

Panel 7: Shot of the doctor handing them a pamphlet.

DOCTOR: The good news is, there's a lot you can do.

Page 5 – An office in Sarah and Lawrence's house, with laptops, bookshelves, and a desk.

Panel 1: Shot of Sarah and Lawrence, each searching on their laptops in the office.

LOCATION CAPTION: Sarah and Lawrence take the doctor's advice and decide to research exposomes.

SARAH: I just can't believe it's so bad here. We've got three factories to the north that are causing air pollution, plus a lot of diesel usage in the area. This neighborhood is an exposome magnet.

LAWRENCE: Take a look at what I found.

Panel 2: Shot of Lawrence's laptop screen, which is opened to a Facebook group about exposomes in High Meadow Neighborhood.

LAWRENCE: We're not the only parents in the area that are concerned.

Panel 3: Shot of Sarah and Lawrence looking at each other.

LAWRENCE: Is there anything we can do about this? Do we have to move to keep Daniel healthy?

SARAH: The materials Dr. Hung gave us had some ideas. Let's try those first.

Page 6 – Implementing strategies

Panel 1: Long shot of the interior of Sarah and Lawrence's living room. Sarah is setting up an air purifier while Lawrence shuts an open window. Daniel is reading on the couch.

CAPTION: Sarah and Lawrence make a plan based on Dr. Hung's materials. They decide to keep windows shut as often as possible to reduce exposure to the open air. They also get an air purifier to make the air indoors clean.

Panel 2: Shot of Lawrence's laptop screen, open to an air quality index page.

CAPTION: They decide to pay attention to the air quality index. If the index is low, they stay inside, and wait for a better-quality day to go outdoors to walk in the park or take Daniel to soccer.

Panel 3: Sarah and Daniel wearing an air mask.

CAPTION: If they do have to go outside on a low-quality day, they wear masks to protect themselves. It becomes part of their daily routine.

Panel 4: Shot of grocery bags full of fruit, vegetables, and whole grains.

CAPTION: The family decides to take Daniel's health into their own hands.

Panel 5: Shot of Lawrence making sandwiches with whole-grain bread and greens.

CAPTION: They pick vegetables, fruits, and whole grains for a scientifically-proven, high-fiber diet.

Panel 6: Daniel at school, with Tupperware containing the sandwich.

CAPTION: Although the family has to eat out from time to time, they keep it to a minimum, and tries to make healthier choices overall.

Page 7 – Results

Panel 1: Shot of the living room, with the air purifier, with Daniel eating a healthy breakfast (eggs, greens) on the dining room table. Sarah is cooking while Lawrence is facing Daniel.

CAPTION: Six months later, Daniel's health improved.

LAWRENCE: Daniel, you about ready to head to school?

DANIEL: Sure dad.

Panel 2: Sarah smiling as they leave.

SARAH: (thinking) He hasn't been sick in four months. I'm glad we can get back to normal again.

Panel 3: The house as Lawrence's car is exiting the driveway.

CAPTION: The end.

Appendix E: Demographic Characteristic Analysis

Figure 28

Gender

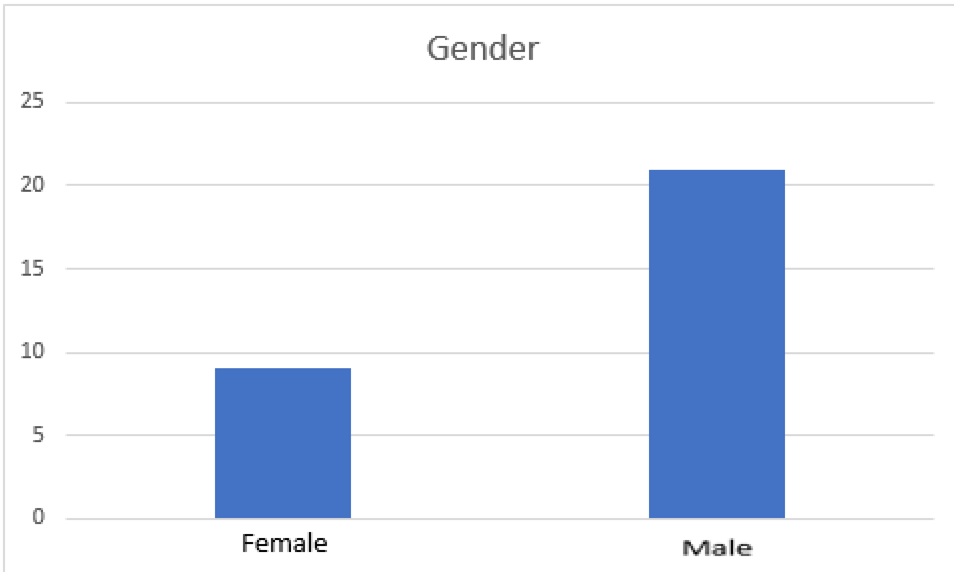


Table 40

Gender

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	9	30.0	30.0	30.0
	2	21	70.0	70.0	100.0
	Total	30	100.0	100.0	

Figure 29

Race

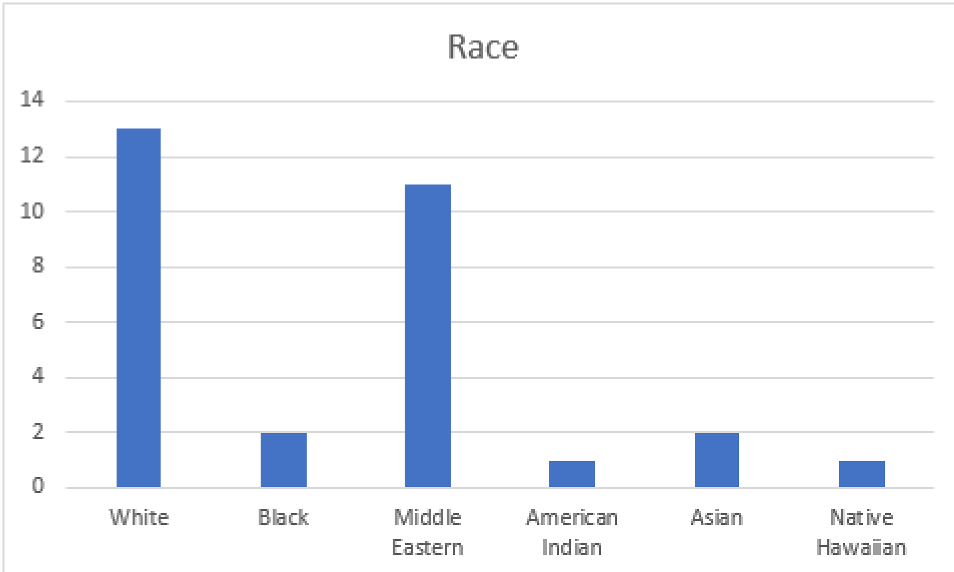


Table 41

Race

		Frequency	Percent	Valid percent	Cumulative percent
Valid	1	13	43.3	43.3	43.3
	2	2	6.7	6.7	50
	3	11	36.7	36.7	86.7
	4	1	3.3	3.3	90
	5	2	6.7	6.7	96.7
	6	1	3.3	3.3	100
	Total	30	100	100	

Figure 30

Marital Status

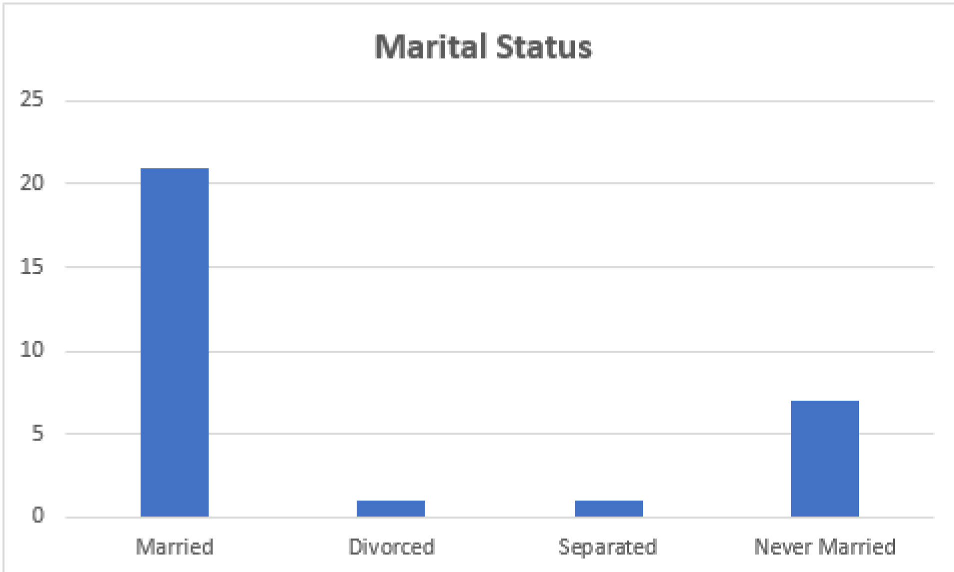


Table 42

Marital Status

		Frequency	Percent	Valid Percent	Cumulative percent
Valid	1	21	70.0	70.0	70.0
	3	1	3.3	3.3	73.3
	4	1	3.3	3.3	76.7
	5	7	23.3	23.3	100.0
	Total	30	100.0	100.0	

Figure 31

Education

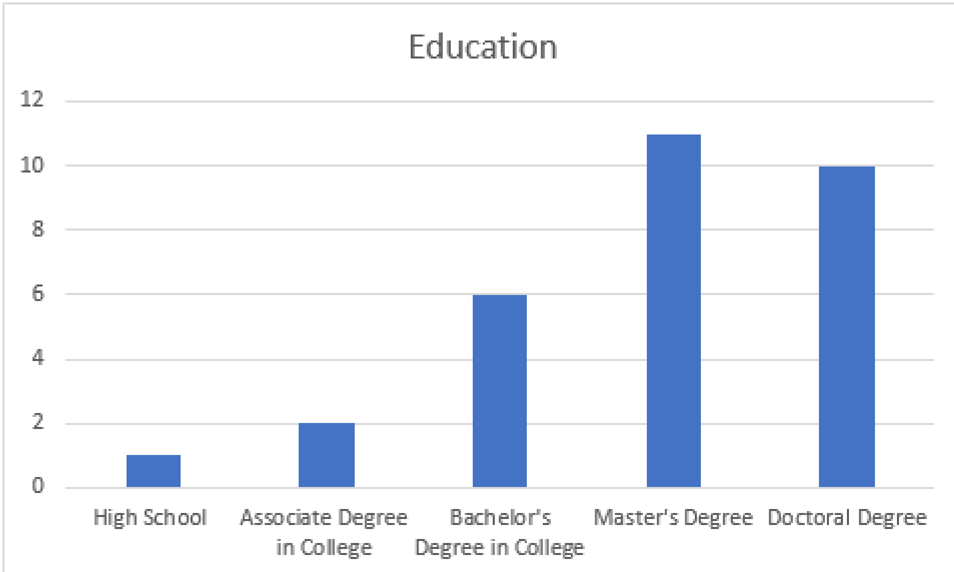


Table 43

Education

		Frequency	Percent	Valid percent	Cumulative percent
Valid	1	1	3.3	3.3	3.3
	2	2	6.7	6.7	10
	3	6	20	20	30
	4	11	36.7	36.7	66.7
	5	10	33.3	33.3	100
	Total	30	100	100	

Figure 32

Income

Table 44

Income

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	3	10.0	10.0	10.0
	1	4	13.3	13.3	23.3
	2	2	6.7	6.7	30.0
	3	1	3.3	3.3	33.3
	4	3	10.0	10.0	43.3
	5	5	16.7	16.7	60.0
	6	5	16.7	16.7	76.7
	7	2	6.7	6.7	83.3
	8	2	6.7	6.7	90.0
	9	3	10.0	10.0	100.0
	Total	30	100.0	100.0	